Quantum metrology for accelerating future searches for fundamental physics including gravitational-wave observation

James W. Gardner

A thesis submitted for the degree of
Doctorate of Philosophy in Physics at
The Australian National University

August 2025

For my parents and grandparents

Declaration

This thesis is an account of research undertaken between June 2022 and August 2025 at the Centre for Gravitational Astrophysics, Research Schools of Physics and Astronomy and Astrophysics, The Australian National University, Canberra, Australia.

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree at any other university.

James W. Gardner

August 2025

Acknowledgements

Hell is empty and all of the devils are here

Ariel (Act 1, Scene 2)

William Shakespeare, The Tempest

Dear all,

I extend my sincerest gratitude to the following people who have made this program worthwhile. I have not tried the impossible task of enumerating you all, but trust that I am better for having known you.

Firstly, I thank my awesome team of supervisors and mentors, Bram, Joe, Lee, Simon, Tuvia, and Yanbei, for what has been a most fun and fulfilling journey these last three years. Thank you for your patience and generosity and for the many lessons that you have taught me, only the least important of which have been about physics.

Secondly, I thank my amazing colleagues and collaborators at ANU and Caltech. Thank you for making a warm environment and showing how science should be done: with passion, humility, and sensitivity.

Finally, I thank my lovely friends and family for their unceasing support. Thank you for making the world more multi-dimensional and ever brighter.

Cheers,
James

P.S.: I quoted Miranda from William Shakespeare's *The Tempest* in my Honours thesis and, since these last three years have felt like a whirlwind, I will now let the other characters speak in the epigraphs of each chapter herein. I thank the Bard for his words.

Abstract

Be not afeard. The isle is full of noises,
Sounds and sweet airs that give delight and hurt not.

CALIBAN (ACT 3, SCENE 2)
WILLIAM SHAKESPEARE, THE TEMPEST

Searches for fundamental physics such as gravitational waves, dark matter, and quantum gravity are now reaching the quantum limit around a century since the invention of quantum mechanics. To push past the fundamental quantum noise and accelerate the discovery of new physics, we need to understand and exploit the interplay of quantum information and quantum metrology. In this thesis, we advocate for a new perspective on the design of gravitational-wave interferometers and other precision waveform sensors, one centred on optimising the sensitivity through studying and manipulating the relevant quantum information-theoretic quantities with practical considerations for loss and classical noise. We demonstrate this approach with a series of papers proposing various measurement schemes, initial quantum states, and quantum control sequences to achieve the fundamental quantum limits of different waveform estimation problems. We suggest that the ultimate potential sensitivity of present experimental designs remains untapped: present in the capability of the device but not yet accessed. In particular, we show that a quantum advantage exists in sensing non-stationary or stochastic waveform signals or those with a wide range of possible frequencies. We discuss the implications of these results and how to realise them for future searches for fundamental physics.

Contents

Chapter 1

Introduction

On a ship at sea: a tempestuous noise
of thunder and lightning heard.

(Act 1, Scene 1)
William Shakespeare, The Tempest

We now introduce the motivation, context, and story of this thesis. Further background material is contained within each paper in the following chapters and the citations therein.

1.1 Motivation

Many fundamental questions about the physical nature of reality remain that we hope to probe this century through precision measurement at the quantum limit. In particular, some of the fundamental questions that concern linear waveform estimation of a continuous signal are as follows.

Firstly, concerning gravitational waves, multimessenger astronomy of binary neutron star mergers promises to resolve the Hubble-Lemaitre tension [1, 2, 3, 4, 5] as well as determine the behaviour of exotic matter at high densities and refine the jet astrophysics of gamma-ray bursts [6, 4]. Meanwhile, observations of black hole mergers provide tests of general relativity and the standard model as well as insights into stellar evolution such as the origin of low-mass black holes [7]. Beyond the audioband of the Laser Interferometric Gravitational-wave Observatory (LIGO) [8], higher and lower frequency gravitational-wave searches promise to study the post-bounce dynamics of core-collapse supernovae [9], search the stochastic gravitational-wave background for primordial black holes as a dark

matter candidate [10, 11], hunt for cosmic domain walls [12], and constrain superradiance from axionic dark matter clouds around black holes [13, 14].

Secondly, for direct detection of dark matter as opposed to the indirect detection via gravitational waves discussed above, the quantum chromodynamics (QCD) axion remains a compelling candidate as it would also solve the strong charge-parity ($\mathcal{CP}$) problem [15, 16]. Searches across the immense parameter space of the axion-photon coupling rate and axion frequency continue to slowly exclude different regions, e.g. experiments involving a microwave cavity in a strong static magnetic field [17, 18, 19, 20, 21, 22]. However, much parameter space remains which wants understanding how to deepen, broaden, or accelerate the search using quantum techniques [23].

Finally, resolving the classicality or quantumness of gravity remains a formidable task but an essential one to reconcile the titans of 20th-century physics: general relativity and quantum mechanics. Furthermore, even if gravity is quantum, determining which theory it obeys at low energies is not direct if we allow violation of general coordinate or local Lorentz invariance, and the possible theories are highly unconstrained at high energies even if we demand local Lorentz invariance [24, 25, 26, 27, 28, 29]. One theory of low-energy "geontropic" quantum gravity predicts particularly large-scale length fluctuations [30, 31] which makes it a prime target for experiments [32, 33, 34, 35] compared to conventional theories which predict much smaller effects [36]. Like stochastic gravitational-wave backgrounds and unresolved axion signals, this geontropic quantum gravity predicts a weak non-deterministic waveform signal quieter than the quantum noise. These different science cases, therefore, all require an understanding of how to optimally sense deterministic and stochastic waveforms.

We emphasise that precision experiments looking for these weak signals are our only way to answer the deepest questions that we have. Enabling and accelerating these searches is thus paramount to our progress in fundamental physics.

1.2 Context

Much remains to be understood, however, about how to best sense these weak signals. The fundamental projection noise from the quantum nature of reality and the weakness of these signals make them hard to detect. This is the role of quantum metrology: to manipulate the available quantum and classical information to extract the best estimates of unknown parameters of interest. Frequency-dependent quantum squeezing of the input quantum state to the LIGO interferometers [37] is a great success of quantum metrology, but even LIGO does not operate in an optimal configuration for different types of signals, e.g. stochastic ones, or from the perspective of quantum control. Although the fundamental quantum limits of deterministic waveform estimation are understood with respect to the final measurement scheme [38, 39] and the initial state [40], we are yet to determine or realise the optimal quantum control [41], non-asymptotic methods [42], or ideal operation in the stochastic regime. Our quantum information theoretic tools, however, remain strong which makes this a promising area for progress. For example, the quantum Fisher information formalism remains a useful and decisive theory [43, 44, 45] and has achieved successes across a wide range of sensing tasks as diverse as, e.g., cold atom interferometry [46] and optical superresolution [47]. The perspective so well exploited thus far in the literature is to study the various bounds on, and gaps between, the information present in the quantum states, classical measurements, and parameter estimates and optimise them over all available experimental axes of inputs, controls, and outputs. Although widely successful, this information-centric approach often suggests experimental designs that are too complicated or impractical and so there is a demand for proposals that remain advantageous in the presence of decent loss and noise and can be implemented with current or near-future technology rather than requiring universal fault-tolerant quantum computation. Embracing this tempered approach, informed by the above context, and applying it wholeheartedly to waveform estimation, e.g., for gravitational-wave detection is the thrust of this thesis.

1.3 Story

The remainder of this thesis consists of five papers followed by a conclusions chapter in accordance with ANU policy for a *Thesis by Compilation*. For the papers in review or preparation, we use the latest version which may differ from the arXiv manuscript. The current publication status of the papers compiled in this thesis is as follows:

1. Chapter 2: Multimessenger astronomy with a Southern-hemisphere gravitational-wave observatory
    (a) Published in PRD [48]: *J. W. Gardner*, L. Sun, S. Borhanian, P. D. Lasky, E. Thrane, D. E. McClelland, and B. J. J. Slagmolen. 2023. Phys. Rev. D, 108:123026.
2. Chapter 3: Achieving the fundamental quantum limit of linear waveform estimation
    (a) Published in PRL [49]: *J. W. Gardner*, T. Gefen, S. A. Haine, J. J. Hope, and Y. Chen. 2024. Phys. Rev. Lett., 132:130801.
3. Chapter 4: Stochastic waveform estimation at the fundamental quantum limit
    (a) Published in PRX Quantum [50]: *J. W. Gardner*, T. Gefen, S. A. Haine, J. J. Hope, J. Preskill, Y. Chen, and L. McCuller. 2025. PRX Quantum, 6:030311.
4. Chapter 5: Lindblad estimation with fast and precise quantum control
    (a) In review in PRX Quantum [51]: *J. W. Gardner*, S. A. Haine, J. J. Hope, Y. Chen, and T. Gefen. 2025. arXiv:2501.03364 [quant-ph].
5. Chapter 6: Bayesian frequency estimation at the fundamental quantum limit
    (a) Submitted to PRX Quantum [52]: *J. W. Gardner*, T. Gefen, E. Payne, S. Direkci, S. M. Vermeulen, S. A. Haine, J. J. Hope, L. McCuller, and Y. Chen. 2025. arXiv:2507.02811 [quant-ph].

The research in the following papers was also performed during this program but is not part of this thesis:

1. Quantum back action evasion with reservoir engineering

   (a) Accepted in PRA [53]: Y. Nishino, *J. W. Gardner*, Y. Chen, and K. Somiya. 2025. arXiv:2505.01013.

2. Achieving the fundamental quantum limit of linear waveform estimation

   (a) In review in npj Quantum Information [54]: Y. Nishino, *J. W. Gardner,* Y. Chen, and K. Somiya. 2025. arXiv:2504.18111.

We now present a concise summary of the coherent narrative that the compiled papers together tell. In a single sentence, this story is as follows:

> We bring new directions from modern quantum information science to accelerate future searches for fundamental physics.

We now outline each chapter and how they connect to reveal this story.

Chapter 2 presents Ref. [48] on using classical Fisher information methods to benchmark different gravitational-wave observatory networks. Before entering the quantum regime for the rest of the thesis, this paper is an example of classical estimation theory. Numerical degeneracies in the classical Fisher information matrix lead to ill-conditioning which this paper handles through rejections. This paper also mentions what multi-messenger astronomy is possible with more sensitive future observatories. This motivates the rest of the thesis which focuses on finding new ways to enhance the sensitivity through the lens of information theory, but now at the fundamental quantum limit.

Chapter 3 presents Ref. [49] on achieving the fundamental quantum limit of waveform estimation with a linear bosonic device. This paper and the ones that follow determine different parts of the landscape of waveform estimation as shown in Fig. 1.1. This paper concerns the case of a deterministic displacement signal in the asymptotic regime of many independent and identically distributed measurements such that the Fisher information is relevant. We examine the pertinent information quantities more carefully than existing literature and discover a gap in the sensitivity that can be made up by performing a different measurement scheme, e.g. at the output of a detuned gravitational-wave detector. This is summarised by the following ethos:

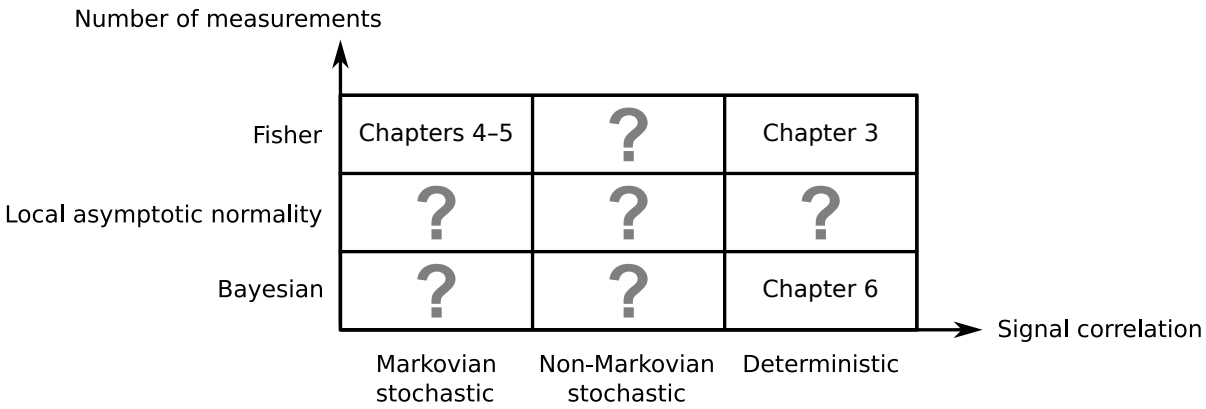

Figure 1.1: Landscape of estimation of the displacement of a bosonic mode/s depending on the number of measurements and the level of correlations of the signal. The associated chapters and papers in this thesis for each regime are shown. The remaining landscape is not fully understood, to the best of our knowledge, but remains under active investigation.

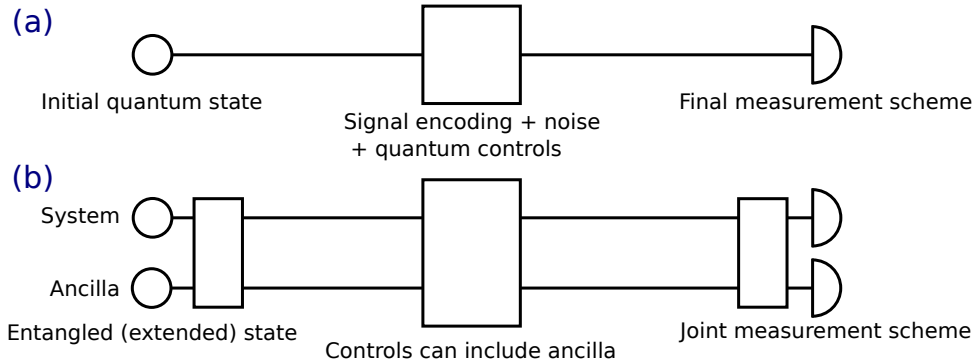

Figure 1.2: Diagram of single shot estimation in the (a) unextended/ancilla-free and (b) extended/ancilla-assisted cases. The boxes represent quantum channels (CPTP maps).

Optimise the relevant quantum information theoretic quantity over the initial quantum state, intervening quantum control, and/or final measurement scheme.

These possible elements of the single-shot estimation process to optimise over are shown in Fig. 1.2 with and without ancilla. In this paper, we only optimised over the measurement but we considered broader optimisation in the subsequent papers, albeit in different regimes as shown in Fig. 1.1. The sensitivity gain (in amplitude units) from our linear measurement is only $\sqrt{2}$ but is much more easily realised than the proposals which follow. In the overall story of this thesis, therefore, this paper serves as an introduction to quantum estimation theory, bridging from existing literature on waveform estimation (e.g. see Refs. [47, 39]) to our later, more radical proposals.

Chapter 4 presents Ref. [50] on waveform estimation of a Markovian stochastic signal in the asymptotic Fisher regime as shown in Fig. 1.1, compared to the deterministic case above. In this paper, we optimise now over both the initial state and the measurement scheme as shown in Fig. 1.2. We demonstrate that orders of magnitude of sensitivity improvement are possible in the right regime of quantum noise, classical noise, loss, and signal strength. However, even with unlimited energy, a finite sensitivity limit exists determined by the loss. To achieve this fundamental quantum limit, however, highly non-Gaussian resources are required which represents a dramatic departure from the linear Gaussian regime of the deterministic case above. Beyond stochastic gravitational-wave detection, this proposal could also accelerate searches for axionic dark matter and quantum gravity, should these non-Gaussian resources become available.

Chapter 5 presents Ref. [51] on stochastic signal sensing but now additionally with fast and precise quantum control intervening in the natural signal and noise evolution of the system as shown in Fig. 1.2. We show that the loss-limited finite sensitivity above can be exceeded by a measure-and-reset sensing strategy, with now unbounded sensitivity in the high energy limit. This demonstrates how optimising over all aspects of the sensing protocol shown in Fig. 1.2 can lead to further improved sensitivity. The results from this paper apply not just to the bosonic waveform sensing considered above but also to Lindblad estimation of Markovian stochastic signals more generally, e.g. to Pauli estimation

of qubits, thus potentially accelerating searches for fundamental physics.

Chapter 6 presents Ref. [52] on measurement scheme optimisation for deterministic waveform sensing once again but now leaving the asymptotic Fisher limit to study the Bayesian regime as shown in Fig. 1.1. In this paper, we show that two measurement schemes that are equally optimal in the Fisher limit can perform differently and both be suboptimal with a finite number of measurements that necessitates a Bayesian approach instead. We consider the single-parameter problem of estimating the frequency of a sinusoid with amplitude and phase known and consider whether a quantum computer is required to reach the fundamental sensitivity limit, as claimed by a recent work [55].

The assumption in the above paper of knowing the amplitude and phase illustrates another axis of the estimation landscape not shown in Fig. 1.1: the number of parameters to estimate. This axis extends from single-parameter estimation through multi-parameter estimation (e.g. as considered in Chapter 3) to many-parameter estimation, also known as quantum learning. We do not explore this axis, like non-Markovian stochastic sensing and the limits in the locally asymptotically normal regime [56]. The vast majority of the estimation landscape remains unknown to the best of our knowledge which motivates current and future work. We survey this and other future work in Chapter 7 as well as draw conclusions. Together, these papers show that modern information-theoretic tools offer new insights into old sensing modalities and inspire a different approach to probing fundamental physics.

In the following Chapters 2–6, we will briefly introduce each paper and connect it to the rest of the thesis, as well as highlight the main results and provide any additional useful context not already provided in the paper itself.

Chapter 2

Multimessenger astronomy with a Southern-hemisphere gravitational-wave observatory

What's past is prologue.

ANTONIO (ACT 2, SCENE 1)

WILLIAM SHAKESPEARE, THE TEMPEST

Here, we present Ref. [48] on benchmarking the multimessenger astronomy capabilities of future networks with and without a Southern-hemisphere gravitational-wave observatory. Although this paper is more disconnected from the other four, it plays two key roles in this thesis: (1) it introduces the classical Fisher information approach to parameter estimation, the quantum version of which pervades the rest of the thesis, and (2) it motivates the ceaseless hunt for increased sensitivity in the rest of the thesis by demonstrating how more powerful detectors enable greater scientific outcomes. In particular, by facilitating further multimessenger astronomy of binary neutron star mergers, we may be able to resolve the Hubble-Lemaitre tension and understand gamma-ray bursts and equations of state of matter at supranuclear densities.

We now point out two main results of this paper independent of the rest of the thesis. The first is a key step in the methods: the rejection handling of numerically ill-conditioned matrices in Sec. III D and Appendix D. This is rather technical but is important to the numerical Fisher analysis as it can otherwise introduce strong biases not present, for example, in a full Bayesian analysis. Our conclusion is that the rejections must

be unified amongst the different networks considered and the geometric and arithmetic improvements in the localisation metric must be considered together. The second main result is the dramatic improvement in this metric shown in Figs. 2–3 and discussed in Sec. IV B, particularly when a next-generation Australian observatory is added to the global network. The subsequently increased rate of joint observations is impressive but has a wide range of uncertainty due to the unconstrained binary neutron star merger rate.

Multimessenger astronomy with a Southern-hemisphere gravitational-wave observatory

James W. Gardner,[1,2,*] Ling Sun,[1] Ssohrab Borhanian,[3] Paul D. Lasky,[4,5] Eric Thrane,[4,5] David E. McClelland,[1] and Bram J. J. Slagmolen[1]

[1]*OzGrav-ANU, Centre for Gravitational Astrophysics, Research Schools of Physics, and of Astronomy and Astrophysics, The Australian National University, Canberra ACT 2601, Australia*
[2]*Walter Burke Institute for Theoretical Physics, The California Institute of Technology, Pasadena, California 91125, USA*
[3]*Theoretisch-Physikalisches Institut, Friedrich-Schiller-Universität Jena, 07743, Jena, Germany*
[4]*School of Physics and Astronomy, Monash University, Melbourne, Victoria 3800, Australia*
[5]*OzGrav: The ARC Centre of Excellence for Gravitational Wave Discovery, Clayton, Victoria 3800, Australia*

(Received 24 August 2023; accepted 15 November 2023; published 14 December 2023)

Joint observations of gravitational waves and electromagnetic counterparts will answer questions about cosmology, gamma-ray bursts, and the behavior of matter at supranuclear densities. The addition of a Southern-hemisphere gravitational-wave observatory to proposed global networks creates a longer baseline, which is beneficial for sky localization. We analyze how an observatory in Australia can enhance the multimessenger astronomy capabilities of future networks. We estimate the number of binary neutron star mergers with joint observations of gravitational waves and kilonova counterparts detectable by the Vera C. Rubin Observatory. First, we consider a network of upgrades to current observatories. Adding an Australian observatory to a three-observatory network (comprising two observatories in the USA and one in Europe) boosts the rate of joint observations from $2.5^{+4.5}_{-2.0}$ yr^{-1} to $5.6^{+10}_{-4.5}$ yr^{-1} (a factor of two improvement). Then, we consider a network of next-generation observatories. Adding a 20 km Australian observatory to a global network of a Cosmic Explorer 40 km in the USA and an Einstein Telescope in Europe only marginally increases the rate from 40^{+71}_{-32} yr^{-1} to 44^{+79}_{-35} yr^{-1} (a factor of 1.1 improvement). The addition of an Australian observatory, however, ensures that at least two observatories are online far more often. When the Cosmic Explorer 40 km is offline for a major upgrade, the Australian observatory increases the joint observation rate from $0.5^{+0.8}_{-0.4}$ yr^{-1} to 38^{+68}_{-30} yr^{-1} (a factor of 82 improvement). When the Einstein Telescope is offline, the joint observation rate increases from $0.2^{+0.3}_{-0.1}$ yr^{-1} to 19^{+34}_{-15} yr^{-1} (a factor of 113 improvement). We sketch out the broader science case for a Southern-hemisphere gravitational-wave observatory.

DOI: 10.1103/PhysRevD.108.123026

I. INTRODUCTION

Approximately 90 gravitational-wave signals have been observed to date from the mergers of black holes and neutron stars [1–3]. These signals have led to many scientific discoveries in cosmology, astrophysics, nuclear physics, and astronomy, but much remains to be discovered in the decades ahead [4–7]. The contributions of the two Laser Interferometric Gravitational-wave Observatories (LIGOs) in the USA [8] and Virgo in Europe [9], were essential to the first detection of a binary neutron star merger [10]. The joint third observing run of the global network also included the Kamioka Gravitational-Wave Observatory (KAGRA) in Japan [11,12]. Looking to the future of the ground-based global network, one additional LIGO observatory is under construction in India [13] and many proposals have been made for upgrades to the existing observatories [14–18]. Next-generation observatories, e.g., Cosmic Explorer (CE) [19] and the Einstein Telescope (ET) [20], are also under development.

The success to date of gravitational-wave astronomy has been achieved through international collaboration and the burgeoning global network of observatories. The addition of a Southern-hemisphere gravitational-wave observatory could provide a long baseline from the Northern-hemisphere observatories to improve sky localization. Several previous studies have proposed Australia as a potential host for a Southern-hemisphere observatory [19,21–23].

We consider how an Australian observatory can enhance the multimessenger astronomy capabilities of the global network. Neutron stars are a promising tool for multimessenger astronomy with gravitational waves. They host one of the most extreme environments in the known Universe, far exceeding densities that can be

*james.gardner@anu.edu.au

studied on Earth. Observing gravitational waves from the merger of binary neutron stars offers a momentary glimpse into the stars' otherwise inaccessible, potentially exotic cores [24–27]. Two binary neutron star mergers, GW170817 [10] and GW190425 [28], have been observed by the current generation of gravitational-wave observatories. The event GW170817 was constrained to a sky-area of 28 deg^2 at a luminosity distance of 40^{+8}_{-14} Mpc by Advanced LIGO and Virgo (Bayesian 90%-credible intervals). A gamma-ray burst was observed by the Fermi Gamma-ray Burst Monitor about 1.7 s after the merger [29]. A subsequent follow-up campaign discovered a bright optical transient in the host galaxy NGC 4993 [30], which was identified as a kilonova, an electromagnetic transient powered by the radioactive decay of heavy nuclei produced in the ejecta, with shock fronts visible in x-ray and radio, e.g., see Refs. [31,32].

The joint gravitational and electromagnetic observation of GW170817 provided invaluable information about astrophysics, dense matter, and cosmology. Combining the distance of the source measured from the gravitational-wave signal and the recession velocity inferred from the redshift measurements from electromagnetic observations, GW170817 was used as a "standard siren" [33] to provide an independent measure of the Hubble constant [34], later refined using radio observations of the jet [35,36]. Population studies of multimessenger observations of binary neutron stars promise a variety of exciting astronomical and astrophysical discoveries including resolving the Hubble-Lemaître tension [22,23,37,38]. According to Ref. [39], $\mathcal{O}(100)$ joint observations will lead to an approximately 1% determination of the Hubble constant. The promise of multimessenger astronomy from $\mathcal{O}(100)$ observations, however, is far broader than just cosmology. For example, it will provide unprecedented insight into the physics and astrophysics of gamma-ray bursts including jet launching and propagation, the formation mechanisms of neutron stars and neutron-star binaries, as well as precision studies on the hot and cold equations of state of matter at supranuclear densities; see, e.g., Refs. [23,40] and the references therein.

The rest of this paper is organized as follows. In Sec. II, we describe different scenarios where an Australian observatory is added to the global network. Then, in Sec. III, we describe our method for estimating the improvement in multimessenger capabilities due to adding the Australian observatory. Finally, we present our results in Sec. IV and conclusions in Sec. V.

II. PROPOSED GRAVITATIONAL-WAVE OBSERVATORIES IN AUSTRALIA

We consider three potential Australian gravitational-wave observatories with sensitivities shown in Fig. 1. We assume that each of these observatories is situated at 26°S and 116°E [41] with the Y-arm of the interferometer at due Northeast; we label this location and orientation as "AU." While we are primarily interested in the possibility of an Australian observatory, the results would be similar using another Southern-hemisphere location with a similar baseline to the Northern-hemisphere observatories.

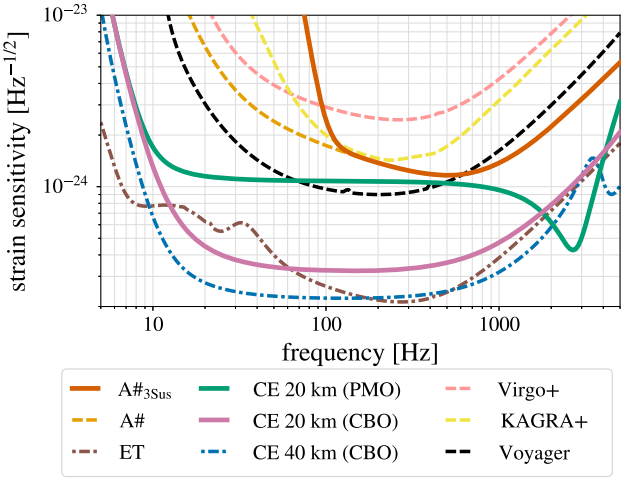

FIG. 1. Strain sensitivity versus frequency for proposed Australian gravitational-wave observatories considered in this work (solid curves) compared to the global context (dashed curves for the upgraded 2G–era and dashed-dotted curves for the XG-era) [43]. Lower values of the strain sensitivity indicate greater sensitivity to gravitational waves. For future observatories, the labels refer only to the displayed sensitivity, not a particular technology used to achieve that sensitivity, e.g., for the compact-binary coalescence optimized (CBO) and postmerger optimized (PMO) designs for CE.

We do not consider all the possible Australian observatories proposed in the literature. The Neutron-star Extreme Matter Observatory (NEMO) [22,23,44] is another proposal for an Australian observatory, however, its science case focuses on 1–4 kHz gravitational-wave physics while we focus on multimessenger astronomy from broadband frequencies. Further discussion about NEMO can be found in Appendix A.

A. Modified A#

The LIGO A# concept [14,45] is a proposed upgrade to the present LIGO's dual-recycled Fabry-Pérot Michelson interferometers. The main upgrades for A# are 100 kg test masses, improved seismic isolation and suspension systems, increased laser input power and quantum squeezing, and improved coating thermal noise. We consider an A# observatory in Australia, which has been modified to reduce the financial cost and mitigate operational complexities for a potentially shorter commissioning time. In particular, this modified version of A# employs the original LIGO A+ test-mass coatings and a triple suspension system for the test-masses [46], which reduces the sensitivity level below

TABLE I. Network scenarios showing the proposed Australian observatories and a selection of possible global environments. Cosmic Explorer (CE) has a variety of possible configurations including those labeled as "CE 20 km" ("CE 40 km") for CE with 20 km (40 km) arms as well as the choice between the compact-binary coalescence optimized (CBO) and postmerger optimized (PMO) designs. We abbreviate the following locations: H for Hanford, USA; L for Livingston, USA; I for Hingoli, India; C for Idaho, USA; and N for New Mexico, USA. Virgo and KAGRA are modeled as being at their present sites. The Einstein Telescope (ET) is assumed to be in Cascina, Italy.

Generation	Australian (AU) observatory	Global network of observatories
Upgraded 2G	$A\#_{3Sus}$	Voyager (H), Voyager (L)
	$A\#_{3Sus}$	Voyager (H), Voyager (L), Virgo+
	$A\#_{3Sus}$	Voyager (H), Voyager (L), Virgo+, KAGRA+
	$A\#_{3Sus}$	Voyager (H), Voyager (L), Virgo+, KAGRA+, Voyager (I)
XG	CE 20 km (CBO)	CE 40 km (CBO; C)
	CE 20 km (CBO)	ET
	CE 20 km (CBO)	CE 40 km (CBO; C), CE 20 km (CBO; N)
	CE 20 km (CBO)	ET, CE 40 km (CBO; C)
	CE 20 km (PMO)	CE 40 km (CBO; C)
	CE 20 km (PMO)	ET
	CE 20 km (PMO)	CE 40 km (CBO; C), CE 20 km (CBO; N)
	CE 20 km (PMO)	ET, CE 40 km (CBO; C)

100 Hz relative to the unmodified A#. We refer to this modified, triple-suspension configuration as $A\#_{3Sus}$.

B. Cosmic Explorer 20 km

We also consider a next-generation Cosmic Explorer (CE) observatory with 20 km arms in Australia [19,47]. The CEs will be new 20 km-long or 40 km-long facilities at new locations. The observatories will utilise larger test masses, improved suspension and isolation systems, increased laser input power and quantum squeezing, and lower loss mirror coatings over the current generation of observatories. There are two possible CE designs for the 20 km facility that we consider for the Australian observatory [19,47,48]. The compact-binary coalescence optimized (CBO) configuration is designed for broadband sensitivity and, therefore, is particularly relevant to our science case. For completeness, we also study the postmerger optimized (PMO) configuration designed for enhanced kilohertz sensitivity. We refer to these configurations as CE 20 km (CBO) and CE 20 km (PMO), respectively.

C. Global network scenarios

We define two eras or "generations" of the global gravitational-wave observatory network to study: future upgrades to current observatories referred to as "upgraded 2G," and new facilities operating with an order of magnitude better strain sensitivity referred to as "XG." We envision an upgraded 2G–era in which LIGO is upgraded to the Voyager design [15] which, among many changes, uses a different laser wavelength and cryogenic silicon test masses, alongside the Advanced Virgo+ (which we label as Virgo+) [16] and KAGRA+ [18] designs. For the XG-era, we consider an Australian CE operating simultaneously with a CE 40 km in the USA [19]. We also examine how an Australian CE would benefit other next-generation global networks including those involving ET [20]. We list various network scenarios in Table I. Previous network studies have examined similar but not identical scenarios [49,50]. For future observatories, we assume fiducial locations for the purpose of calculation as the actual sites are not yet known.

D. Observational duty cycle

To fully assess the benefit of an Australian observatory, it is necessary to consider the network duty cycle—the fraction of time for which the full network is observing. During the third Advanced LIGO observing run, each LIGO achieved an individual duty cycle of 76%–79% [3,51]. The "coincident duty cycle" when both LIGOs were operating simultaneously in a two-observatory network was approximately 58%–62% (assuming that the downtimes of different observatories are uncorrelated). There are a number of reasons that gravitational-wave observatories do not continuously operate during an observing run, including regular maintenance and lock loss from environmental disturbances [51]. These values for the duty cycle, however, do not account for the fact that gravitational-wave observatories regularly undergo long periods of downtime between observing runs for major upgrades and commissioning. Accounting for this downtime between observing runs, the overall effective coincident duty cycle since the construction of Advanced LIGO is approximately 27%. It is important that at least two observatories are operating simultaneously to achieve the sky-localization necessary for electromagnetic follow-up of binary neutron star mergers. For future networks particularly in the XG-era, therefore, we emphasise how a third observatory can be used to maximize the coincident duty cycle when two or more observatories are online.

III. METHOD

We benchmark each network scenario by simulating a universe of sources, predicting how the network observes each source, summarizing the total performance over the population, and addressing known biases. We detail each of these steps in turn.

A. Monte Carlo population

We focus on optical and ultraviolet observations of kilonovae emissions and use the Vera C. Rubin Observatory (Rubin) [52] as our standard. Rubin's typical survey depth for a GW170817-like kilonova given an exposure time of 30 s is 500 Mpc (i.e., redshift ≈ 0.1), assuming the absolute magnitude for the kilonova is roughly -14 in the g-, r-, i-bands. We call sources "in range" within this distance henceforth. Within range, we can still observe other counterparts, e.g., gamma-rays, x-rays, and radio waves. We defer a complete study of the capabilities of multimessenger astronomy beyond 500 Mpc to future work.

We simulate a population of binary neutron star mergers in range. We calculate the redshift distribution of sources in Appendix B using the GWTC-3 local merger rate of 106^{+190}_{-84} $\mathrm{yr}^{-1}\,\mathrm{Gpc}^{-3}$ inferred in Sec. IV A of Ref. [53]. This rate implies 44^{+79}_{-35} yr^{-1} sources in range. The merger rate is only loosely constrained since the width of the 90%-credible interval (the uncertainty range quoted above) is greater than twice the maximum *a posteriori* estimate obtained from the Bayesian parameter estimation. This is the largest uncertainty in our analysis. In comparison, we estimate that the sampling error from the multiyear Monte Carlo population contributes an uncertainty of only $\mathcal{O}(1)$ source per year in the observation rate since we simulate a multiyear population. To mitigate the uncertainty in the merger rate, since it is ultimately only a normalization factor, we study the observation rate normalized to the total number of sources in range per year.

For the other parameters of the population, the right ascension, sine of the declination, polarization angle, and cosine of the inclination angle all follow uniform distributions. The chirp mass, mass ratio, and angular momenta follow a cosmological model detailed in Ref. [54]. We use the `IMRPhenomD_NRTidalv2` waveform (which is an aligned-spin model) from the LIGO Scientific Collaboration Algorithm Library Suite (LALSuite) [55] to model the binary neutron star merger signals.

B. Fisher information analysis

To determine how well a network observes a given source, we use the Fisher information analysis tool GWBENCH [43] written in Python [56] which is similar to other Fisher benchmarking tools [57,58]. We calculate the Cramér-Rao bound on the parameter estimates, specifically the sky-area, by computing the gradient of the measured gravitational-wave strain with respect to each astrophysical parameter. In the single-parameter case [59], the Cramér-Rao bound provides a lower bound on the mean-square-error $\mathrm{E}[(\hat{\theta}-\theta)^2]$ for an estimate $\hat{\theta}$ of a parameter θ. For example, when estimating the true mean θ of a Gaussian random variable $X \sim \mathrm{N}(\theta, \sigma)$, the Cramér-Rao bound of σ^2/N for N observations is achieved by the sample mean $\hat{\theta}$. We provide technical specifications of our computation in Appendix C. We recover the signal-to-noise ratio and sky-area estimates from Ref. [54] which use the same method [61].

Compared to Bayesian analysis, the Fisher information approach is computationally less expensive but requires careful handling of rejection errors discussed later [62]. The reliability of Fisher analysis and specifically the bound on the sky-area compared to Bayesian analysis improves with the network signal-to-noise ratio [63–65]. For non-Gaussian distributions, the Cramér-Rao lower bound can differ significantly from the achievable minimal error. For example, at low signal-to-noise ratios $\lesssim 10$, Fisher analysis systematically underestimates the sky-area. While at signal-to-noise ratios $\gtrsim 25$, Ref. [64] shows that the Fisher and Bayesian sky-area estimates roughly agree at the population level. Although there remains some variance in the individual estimates, there is little systematic bias between the two approaches. In particular, the number of sources with signal-to-noise ratio $\gtrsim 25$ that are localized to within 10 deg^2 is consistent.

C. Follow-up metrics for benchmarking

We study sources satisfying the following conditions:

(1) *A signal-to-noise ratio threshold.* For the Bayesian and Fisher estimates to agree at the population level, we require the network signal-to-noise ratio to be greater than 25. We also examine other 2G-era networks with $\mathrm{A\#}_{\mathrm{3Sus}}$ alongside the LIGOs operating at the A# rather than the Voyager sensitivity. For these networks, however, less than 5% of sources in range have a signal-to-noise ratio greater than 25 even for $\mathrm{A\#}_{\mathrm{3Sus}}$ with five Northern hemisphere observatories. We omit these networks from the results.

(2) *A luminosity distance threshold.* To observe a source with electromagnetic telescopes, it needs to be within their survey depth. We use Rubin's survey depth of 500 Mpc as discussed previously. For this threshold, we use the true luminosity distance to the source from our simulation, although this is not exactly known in practice.

(3) *A sky localization threshold.* To point the electromagnetic telescopes toward the source, the gravitational-wave observation has to sufficiently localize the source as quantified by the 90%-credible sky-area in square degrees. For context, GW170817 was localized to 28 deg^2 [10] and Rubin's field-of-view is 10 deg^2 [52,66]. Tighter localization allows for faster and more precise identification of the host

TABLE II. Thresholds for the loose and tight follow-up metrics used to benchmark the networks for enabling electromagnetic follow-up of binary neutron star mergers.

Follow-up metric [yr^{-1}]	Loose	Tight
Signal-to-noise ratio	≥ 25	≥ 25
Luminosity distance [Mpc]	≤ 500	≤ 500
90%-credible sky-area [deg^2]	≤ 10	≤ 1

galaxy and potentially better multimessenger astronomy. We model Earth's rotation as this assists in localization for observatories with sufficient low-frequency sensitivity, e.g., CE 20 km (CBO). For the sake of choosing a threshold for a simple counting metric, we examine sources localized "loosely" (within 10 deg^2) and "tightly" (within 1 deg^2).

As summarized in Table II, we use the number of joint observations per year that satisfy the above three thresholds as our loose and tight follow-up metrics to benchmark different networks. As discussed previously, we normalise the metrics to the number of sources in range to define the "fractional metrics."

These metrics assume no improvements in the survey depth and localization capabilities of electromagnetic telescopes in the coming decades, e.g., see Refs. [23,67,68]. These advancements would improve the metrics, particularly for XG. Future work should determine the likely impact of these improvements.

D. Numerical rejections

Our Fisher information analysis requires careful handling of numerically ill-conditioned matrices to avoid biasing the metrics. Ill-conditioned matrices, e.g., from nearly edge-on sources, are numerically inaccurate to invert. These sources are rejected from the analysis and not included in the metrics. The rejection rate is higher with fewer observatories in the network. We detail our approach to handling these rejections in Appendix D. To summarize, the fractional follow-up metrics are overestimated by an amount depending on the rejection scenario. The most realistic scenario lies somewhere between the worst-case scenario where the fractional metrics are uniformly reduced by a multiplicative factor of 0.32 (e.g., from 50% to 16%) and the "null-case" scenario where they remain unchanged.

To avoid this ambiguity, we provide a secondary figure-of-merit that is invariant of the merger and rejection rates. We consider the geometric improvement in the follow-up metrics from adding an Australian observatory to a baseline network, i.e., the multiplicative factor of improvement in the number of joint observations. Whether or not we normalize the rates to the number of sources in range for Rubin does not affect this geometric improvement since the same normalization factor appears in both the baseline network's rate and the network with an Australian observatory's rate and thus cancels out. Similarly, any factor from the merger rate or unified rejection rate (i.e., the rejection rate made equal among networks) cancels out. When the baseline network's metric is zero, the geometric improvement is undefined. In such cases, we provide a lower bound derived from the Monte Carlo resolution of our simulated population [69]. We emphasise that the geometric improvement is different from the arithmetic increase in the number of observations since the baseline network's performance is factored in when evaluating the geometric improvement. This makes the geometric improvement sensitive to small but nonzero values of metric from the baseline network. To understand the complete picture, we study both the arithmetic and geometric improvements in the number of joint observations.

IV. RESULTS

We examine how a long-baseline global network with an observatory in Australia improves the joint observation rate quantified by the fractional follow-up metrics.

A. Upgraded 2nd generation (upgraded 2G)

Figure 2(a) shows the arithmetic improvement in the fractional loose follow-up metric when $\mathrm{A\#}_{\mathrm{3Sus}}$ (AU) is added to each upgraded 2G–era global environment. For example, with three Northern hemisphere observatories (Voyager (H), Voyager (L), and Virgo+) we observe 6% of the sources within 500 Mpc that are localized within 10 deg^2 and have a network signal-to-noise ratio above 25. Adding $\mathrm{A\#}_{\mathrm{3Sus}}$ (AU) to this network increases the fraction to 13% of the sources in range. This improvement comes mainly from better sky localization rather than from increasing the signal-to-noise ratio. Figure 3(a) shows the geometric improvement of 2.2 which is invariant of the unified rejection rate.

Overall, Fig. 3(a) shows that $\mathrm{A\#}_{\mathrm{3Sus}}$ (AU) improves the fractional loose follow-up metric by a multiplicative factor of at least 1.1 across all upgraded 2G–era scenarios. The greatest improvement seen in Fig. 2(a) is for the network with two Northern hemisphere observatories [Voyager (H) and Voyager (L)] where adding $\mathrm{A\#}_{\mathrm{3Sus}}$ (AU) improves the metric from approximately 0% to 10% of the sources in range. With $\mathrm{A\#}_{\mathrm{3Sus}}$ (AU), Voyager (H), and Voyager (L), this corresponds to observing $4.5^{+8.1}_{-3.6}$ yr^{-1} sources accounting for the uncertainty in the GWTC-3 merger rate which dominates the other uncertainties in our method such as the rejection scenario and the Monte Carlo error. This rate increases to 11^{+20}_{-9} yr^{-1} with $\mathrm{A\#}_{\mathrm{3Sus}}$ (AU) and five Northern hemisphere observatories such that it would take $\mathcal{O}(9)$ years to observe $\mathcal{O}(100)$ sources.

The upgraded 2G–era results for the tight metric from Table II are shown in Figs. 2(b) and 3(b), indicating sources localized to within 1 deg^2 which can be located even faster and more precisely. With $\mathrm{A\#}_{\mathrm{3Sus}}$ (AU) and five Northern

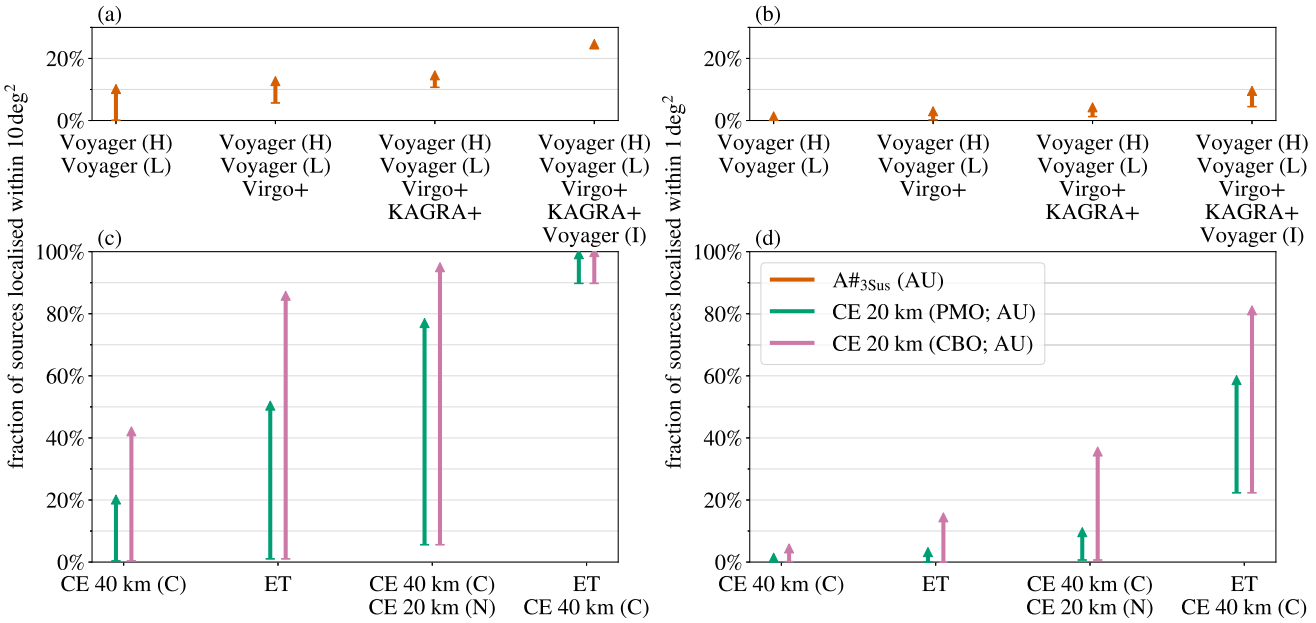

FIG. 2. Fractional follow-up metrics from Table II, i.e., the fraction of sources within 500 Mpc that have a network signal-to-noise ratio greater than 25 and are localized within 10 deg^2 [panels (a) and (c)] or 1 deg^2 [panels (b) and (d)]. The length of each line shows the arithmetic improvement in the expected number of joint observations of gravitational waves and electromagnetic counterparts (normalized to the number of sources in range) from adding an Australian observatory to the baseline global networks in Table I in the upgraded 2G–era [panels (a) and (b)] and XG-era [panels (c) and (d)]. We assume the compact-binary coalescence optimized (CBO) design for the Northern hemisphere CEs but study the CBO and postmerger optimized (PMO) designs for the Australian CE.

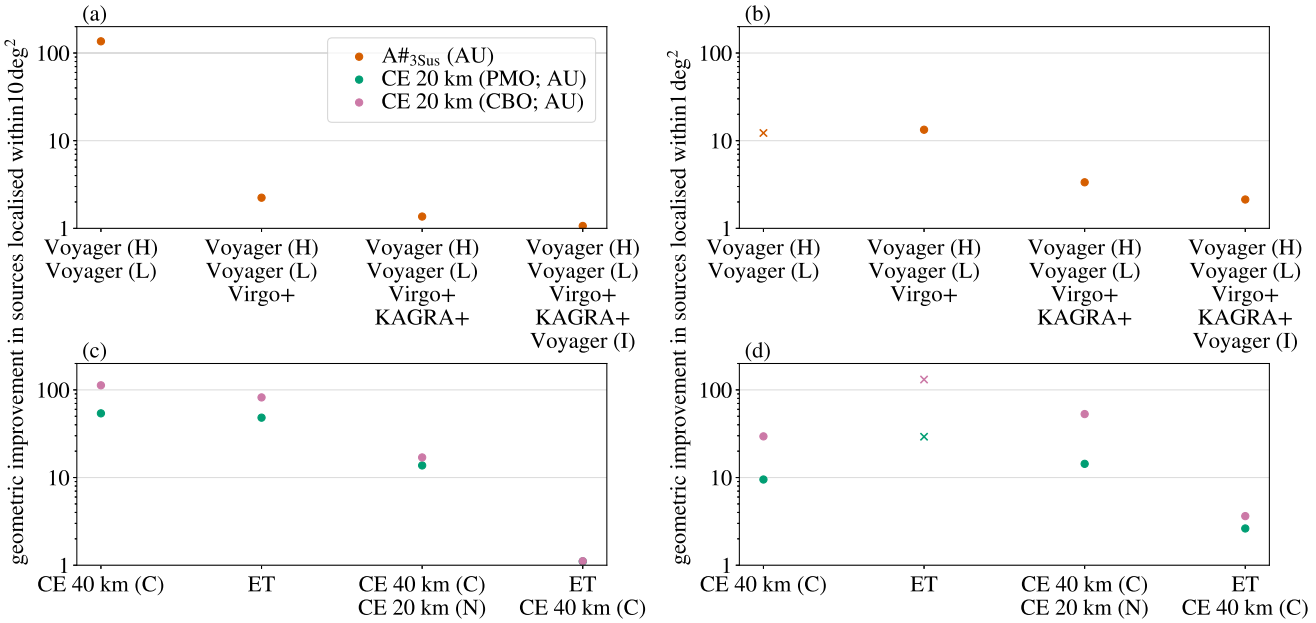

FIG. 3. Geometric improvement in the follow-up metrics from Table II, i.e., the factor of improvement in the number of observations within 500 Mpc that have a network signal-to-noise ratio greater than 25 and are localized within 10 deg^2 [panels (a) and (c)] or 1 deg^2 [panels (b) and (d)]. This shows the multiplicative increase—compared to the arithmetic increase shown in Fig. 2—in the expected number of joint observations of gravitational waves and electromagnetic counterparts from adding an Australian observatory to the baseline global networks in Table I in the upgraded 2G–era [panels (a) and (b)] and XG-era [panels (c) and (d)]. We assume the compact-binary coalescence optimized (CBO) design for the Northern hemisphere CEs, but study the CBO and post-merger optimized (PMO) designs for the Australian CE. When the baseline network observes zero sources, we provide a lower bound on the geometric improvement indicated by ×.

hemisphere observatories, the tight metric is 10% which corresponds to $4.2^{+7.6}_{-3.4}$ yr^{-1}. The geometric improvement with three or more Northern hemisphere observatories is greater for the tight metric than the loose metric shown in Fig. 3(a) although the arithmetic improvement is comparable because the baseline networks' performance is reduced for the tight metric.

B. The next generation (XG)

Figure 2(c) [Fig. 3(c)] quantifies the global observation arithmetic (geometric) improvement in the loose metric from adding CE 20 km (CBO) or CE 20 km (PMO) in Australia. Both configurations enable significant improvement with a geometric improvement of at least 14 for the global environments of CE 40 km (CBO; C); ET; and CE 40 km (CBO; C) and CE 20 km (CBO; N). As expected, the PMO configuration is outperformed here by the CBO configuration which provides a geometric improvement of at least 17 for these environments and an arithmetic improvement of up to 89% of the sources in range [seen for CE 40 km (CBO; C) and CE 20 km (CBO; N)].

The remaining XG global environment of ET and CE 40 km (CBO; C) observes 90% of the sources in range by itself. When ET and CE 40 km (CBO; C) both operate at design sensitivity, therefore, there is little room for improvement in the loose metric using an Australian observatory. For example, adding CE 20 km (CBO; AU) would improve the metric from 40^{+71}_{-32} yr^{-1} to 44^{+79}_{-35} yr^{-1}, a multiplicative factor of only 1.1 because the network then observes 100% of the sources in range. As discussed in Sec. II D, however, the duty cycles make it difficult to maximize the observing time with both ET and CE 40 km (CBO; C) online, and thus the performance of the baseline network is difficult to achieve continuously. This means that an observatory in Australia would significantly improve the amount of time when the global network has two or more XG observatories online. If CE 40 km (CBO; C) [ET] was offline for a major upgrade, then adding CE 20 km (CBO; AU) to ET [CE 40 km (CBO; C)] improves the rate of joint observations from $0.5^{+0.8}_{-0.4}$ yr^{-1} to 38^{+68}_{-30} yr^{-1} [from $0.2^{+0.3}_{-0.1}$ yr^{-1} to 19^{+34}_{-15} yr^{-1}] for a multiplicative factor of 82 [113], as shown in Figs. 2(c) and 3(c). This means that with CE 20 km (CBO; AU), ET, and CE 40 km (CBO; C) it could take less than half a decade to observe $\mathcal{O}(100)$ joint sources even if we do not have all three observatories online at any given time. This is promising for doing cosmology in the XG-era.

Furthermore, the tight metric in Figs. 2(d) and 3(d) shows that either Australian CE configuration improves the tight metric more than the loose metric for the baseline network involving both ET and CE 40 km (CBO; C). Across all scenarios, the tight metric increases by a geometric factor of at least 2.6, leading to even faster and more precise localization. Finally, we reemphasize that there is a multitude of science cases, in addition to the one studied in detail in this paper, which will also benefit from an Australian observatory.

V. CONCLUSIONS

The prospects for multimessenger gravitational-wave astronomy and cosmology with a future long-baseline global network are promising. We focus on localizing binary neutron star mergers to enable electromagnetic follow-up using a gravitational-wave observatory in Australia due to its long baseline to the global network. We define our figure-of-merit as the fraction of binary neutron star mergers available for electromagnetic observation with all channels present, focusing on optical and ultraviolet observations of kilonovae using Rubin. These sources within 500 Mpc have a signal-to-noise ratio above 25 and are localized to within 10 deg^2. In the upgraded 2G–era, a 4 km arm modified A# with A+ coatings and a triple-suspension system in Australia significantly improves the number of well-localized sources for networks with 2–4 other observatories by a multiplicative factor of at least 1.4 to beyond 100. (The improvement is a modest factor of 1.1 when there are five other observatories in the network.) The gain comes mainly from tighter sky localization rather than from increasing the network signal-to-noise ratio. The performance using a 4 km NEMO is similar since it has comparable broadband sensitivity, however, NEMO is designed for a different science case not addressed here. In the XG-era, a broadband-optimized CE 20 km in Australia provides a multiplicative improvement of 82–113 when just one other observatory is operating. The improvement is marginal when an Australian observatory is added to a network with both ET and a CE in the Northern hemisphere where 90% of the sources within 500 Mpc are already jointly observed. As we discuss in Sec. II D, however, it is frequently necessary to shut down observatories for maintenance and upgrades, and so an Australian observatory is essential to maintain two simultaneously operating observatories. An Australian CE 20 km operating alongside ET and a CE in the Northern hemisphere can be expected to observe $\mathcal{O}(100)$ sources within less than half a decade, even if the three observatories are never all online simultaneously.

We defer to future work a comprehensive Bayesian study of the population within and beyond 500 Mpc. Bayesian analysis would eliminate the rejections and allow us to study low signal-to-noise ratio events. Simultaneously, going beyond 500 Mpc would allow us to study multimessenger astronomy without optical or ultraviolet channels. For example, we could study how the inclination angle affects electromagnetic follow-up since jets and gamma-rays provide a selection effect for face-on/face-off sources [35,70–72]. Other avenues of future work include studying network scenarios with a mix of upgraded 2G and XG observatories [45], incorporating

early warnings from space-based gravitational-wave observatories, and constructing broadband metrics that combine several science cases for a variety of sources.

This exploratory study, with a focus on multimessenger astronomy and cosmology, does not address all of the possible science cases or configurations for an Australian observatory. Other designs focusing on science in the kilohertz regime, e.g., NEMO, would enable studies of extremely dense matter with neutron star mergers, exploring hot and cold equations of state at supranuclear densities. Together with the global network, an Australian observatory would enable new science and discoveries. The improved network sensitivity and increased volume of observations would allow for deeper tests of general relativity in extreme gravity regimes, provide unprecedented insights into the populations and evolution of compact objects, and potentially open new windows onto the dark sector of the Universe to explore fundamental physics. The promising results presented in this paper encourage future work in this area.

Our code is available online [73] and was written using resources from Refs. [43,56,74–81].

ACKNOWLEDGMENTS

The authors wish to thank the following people: B. S. Sathyaprakash, R. X. Adhikari, B. P. Schmidt, S. M. Scott, E. Payne, P. Landry and the providers and support team of OzSTAR. This research is supported by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (Project No. CE170100004) as well as ARC LE210100002. J. W. G. and this research are supported by an Australian Government Research Training Program (RTP) Scholarship and also partially supported by the United States of America's National Science Foundation under Award No. PHY-2011968. S. B. is supported by the Deutsche Forschungsgemeinschaft, DFG, Project MEMI Number BE 6301/2-1. P. D. L. is supported by ARC Discovery Project DP220101610. P. D. L. and E. T. are supported by DP230103088. This paper has been assigned LIGO Document Number P2300034.

APPENDIX A: NEMO

NEMO is another proposed configuration of the Australian gravitational-wave observatory. It has 4 km arms and uses a triple suspension system like A#$_{3\mathrm{Sus}}$ but with more advanced internal technology: increased laser power and quantum squeezing, more massive silicon mirrors, and a longer resonant signal-extraction cavity [22]. As shown in Fig. 4(e), NEMO is more sensitive than A#$_{3\mathrm{Sus}}$ in the 1–4 kHz band. NEMO achieves this within the upgraded 2G–era without the significantly higher cost and later start time of the XG observatories. This increased kilohertz sensitivity can enable studying the rich nuclear physics of extreme matter in the intensely hot post-merger phase of binary neutron star coalescences. For example, this would allow us to probe the nuclear equation of state in a

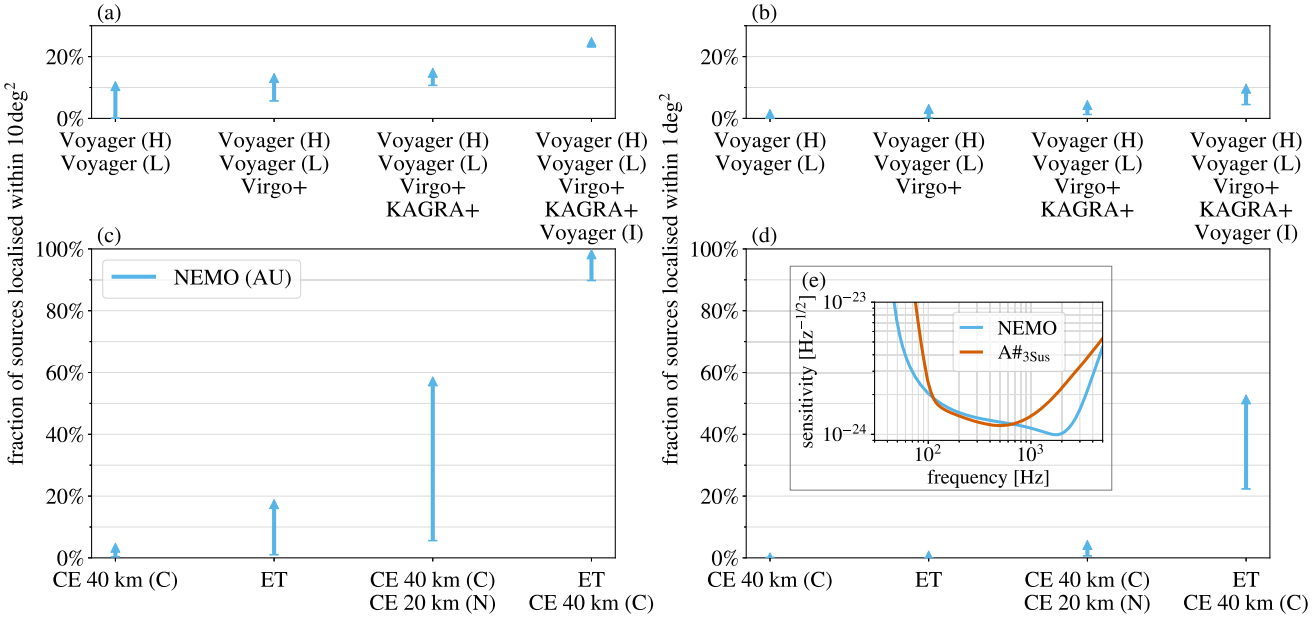

FIG. 4. (Compare these results to Fig. 2.) Arithmetic improvement in the fractional follow-up metrics from adding NEMO in Australia to the baseline global network in Table I in the upgraded 2G–era [panels (a) and (b)] and XG-era [panels (c) and (d)]. The fractional follow-up metrics from Table II are the fraction of sources within 500 Mpc that have a signal-to-noise ratio greater than 25 and are localized within 10 deg^2 [panels (a) and (c)] or 1 deg^2 [panels (b) and (d)]. Panel (e) shows the strain sensitivity of NEMO and A#$_{3\mathrm{Sus}}$ versus frequency.

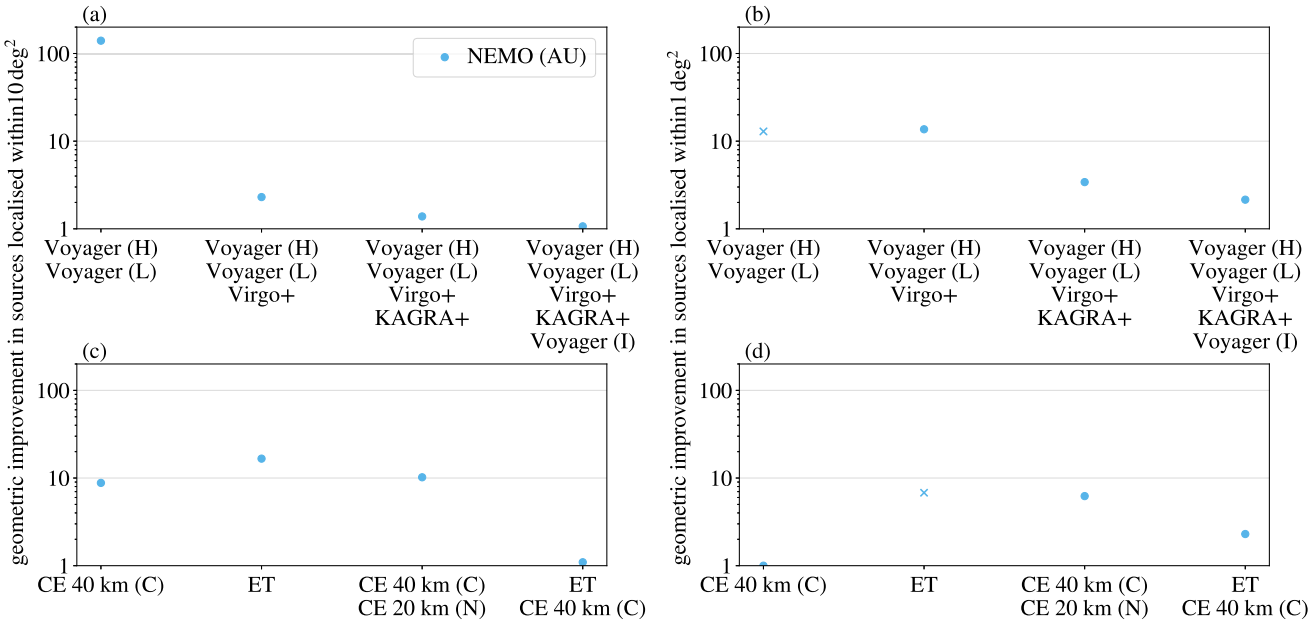

FIG. 5. (Compare these results to Fig. 3.) Geometric improvement in the follow-up metrics from adding NEMO in Australia to the baseline global network in Table I in the upgraded 2G–era [panels (a) and (b)] and XG-era [panels (c) and (d)]. The geometric improvement in the follow-up metrics from Table II is the factor of improvement in the number of sources within 500 Mpc that have a signal-to-noise ratio greater than 25 and are localized within 10 deg^2 [panels (a) and (c)] or 1 deg^2 [panels (b) and (d)]. Whenever the baseline network observes zero sources, we provide a lower bound on the geometric improvement indicated by $\times$.

high-temperature, high-density regime [82,83]. We emphasize that we do not consider this postmerger kilohertz science case here.

Although NEMO is not designed for our multimessenger science case, we still study its performance for completeness. For the upgraded 2G–era, the arithmetic improvement in the metrics using NEMO in Figs. 4(a) and 4(b) compared to $\mathrm{A\#}_{\mathrm{3Sus}}$ in Figs. 2(a) and 2(b) are comparable because their broadband sensitivities are similar [see Fig. 4(e)]. The geometric improvement is also similar between Figs. 5(a), 5(b) and 3(a), 3(b).

For the XG-era, either CE 20 km in Australia in Figs. 2(c) and 2(d) arithmetically improves the metric more than the 4 km arm NEMO in Figs. 4(c) and 4(d). The same is true for the geometric improvement shown in Figs. 5(c)–5(d) compared to Figs. 3(c) and 3(d). CE 20 km (PMO) has a higher integrated signal-to-noise ratio than NEMO. While NEMO's contribution is less than either CE in the XG-era, it would have been operating since the upgraded 2G–era, enabling rich post-merger physics from well before the XG observatories come online. Future work should further consider the trade-offs between the different science cases and eras of the global network.

APPENDIX B: POPULATION DISTRIBUTION IN REDSHIFT

We simulate a multiyear population of sources out to redshift 0.5, although we only present results for the sources within 500 Mpc (redshift 0.1). For greater resolution in redshift, we sample 250,000 sources distributed uniformly in redshift between 0 and 0.5 [54]. We emphasise that this initial population is not cosmologically accurate. But, after analyzing each source, we recover a cosmological distribution. We use the binary neutron star merger rate $R(z)$ at redshift z from Ref. [54]. This uses a Madau-Dickinson star formation rate (that does not account for metallicity) and a comoving volume determined by the *Planck18* cosmology from Ref. [79]. We divide the redshift range into narrow bins with the endpoints of the bins following a geometric progression. We construct the cosmological distribution by first randomly selecting a bin using the merger rate R as described below and then randomly sampling a source in that bin using a uniform distribution. The probability p_i of drawing the ith bin $(z_i, z_i + \Delta_i)$ with width Δ_i is

$$p_i = \frac{R_i \Delta_i}{\sum_j R_j \Delta_j} \approx \frac{\int_{z_i}^{z_i+\Delta_i} R(z)\mathrm{d}z}{\int_0^{0.5} R(z)\mathrm{d}z} \tag{B1}$$

where the merger rate for each bin is sampled at the geometric mean of the endpoints

$$R_i = R(\sqrt{z_i(z_i + \Delta_i)}). \tag{B2}$$

The *multiyear* cosmological population created from this distribution has around 125,000 sources within redshift 0.5 and around 1000 sources within 500 Mpc.

APPENDIX C: COMPUTATION

Using the supercomputer OzSTAR at the Swinburne Supercomputing Facility [84], we processed the 250,000 uniformly distributed sources. The "multinetwork" pipeline of GWBENCH [43] computes the derivatives of the measured signal for each unique observatory only once. Although calculating these derivatives is the slowest part of the computation, the Fisher information approach remains fast overall compared to Bayesian methods. We took advantage of the multinetwork pipeline by grouping similar networks from Table I together in each run of 2048 single-core tasks. For this work, we used roughly five runs with each run taking around a day to complete. (Optimizing our code could likely lead to a significant speed-up.) After processing, we cosmologically resampled the results as described in Appendix B.

APPENDIX D: REJECTIONS

In Sec. III B, we invert the Fisher information matrix to calculate the Cramér-Rao bound [85,86]. This matrix can be numerically ill conditioned and inaccurate to invert [62,63]. In such cases, we reject the source from the analysis, not including it in the cosmological resampling or the results. Figure 6 shows the rejection rate, i.e., the percentage of sources in range that are ill conditioned, for each network scenario. For the single-observatory network CE 40 km (CBO; C), the estimates of the sky position are degenerate and the rejection rate is 98%. (ET comprises an equilateral triangle of three identical observatories at the same site such that the overall effective sensitivity shown in Fig. 1 is a factor of $\frac{2}{3} = (\sqrt{3}\sin(\frac{\pi}{3}))^{-1}$ more sensitive than each individual observatory [20].) For the multiple-observatory networks, the rejection rate is 15%–64% (networks with more observatories tend to have a lower rejection rate). These rates are for the population uniformly distributed in redshift because we reject the sources before the cosmological resampling (see Appendix B). We assume that the rejection rate is the same for the cosmological population and defer to future work eliminating the rejections by using Bayesian analysis instead. Switching to a Bayesian framework would also allow us to study sources with lower signal-to-noise

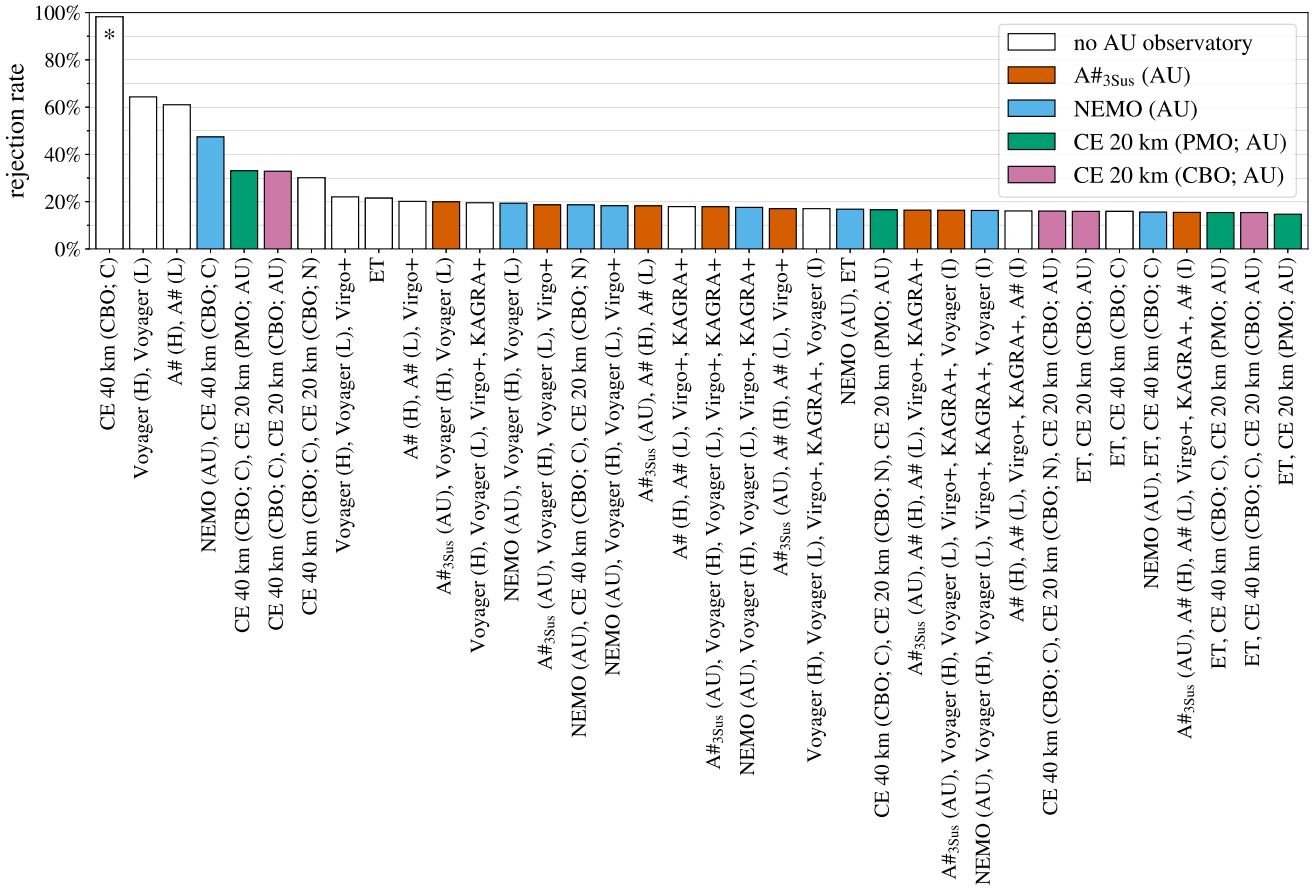

FIG. 6. Rejection rate (i.e., the fraction of sources in range that are numerically ill conditioned) for each upgraded 2G–era and XG-era network shown in Table I in descending order. We also include the other 2G-era networks that use A# instead of Voyager for the LIGOs (although they are omitted from the observation rate results) as well as the networks involving NEMO from Appendix A. The single-observatory network indicated by $*$ is excluded from the unification of the rejections.

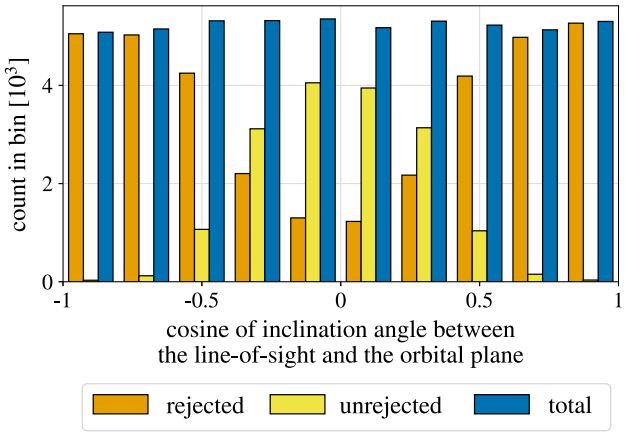

FIG. 7. Histogram of the number of uniformly rejected and unrejected sources in range versus the inclination angle. The initial population (labeled as "total") is sampled from a uniform distribution in the cosine of the inclination angle. There is a bias toward rejecting nearly edge-on sources, corresponding to an inclination angle of close to 0 or π.

ratios because it would no longer be necessary to introduce the threshold of 25 in Sec. III B to ensure that the Fisher and Bayesian estimates agree at the population level. One alternative to eliminate the rejections is to use arbitrary-precision floating-point arithmetic [87,88].

For Fisher information methods, rejections are a perennial issue among other challenges [62,63]. To ensure a fair comparison, we only study sources that are unrejected for all multiple-observatory networks. This "unified" rejection rate is 68%. This guarantees that the cosmological population is the same for each multiple-observatory network. (We include the other 2G-era and NEMO networks but exclude the single-observatory network.) Choosing a smaller set of networks among which to unify the rejections would give different results. But, it is necessary that we are able to compare any two multiple-observatory networks. Unifying the rejections is also necessary to make the geometric improvement invariant of the rejection rate.

We profile what kinds of sources are rejected. The dominant effect is the rejection of nearly edge-on sources shown in Fig. 7. Nearly edge-on sources have a lower signal-to-noise ratio than their counterparts with a more preferred orientation. Our metrics, therefore, may be overestimated in some cases. The "null-case" scenario is that—at the population level—the rejected sources behave the same as the unrejected sources. This would mean that the metrics are unaffected. The worst-case scenario, however, is that none of the rejected sources can be jointly observed. This would reduce the fractional metrics by a multiplicative factor of 0.32.

For example, for $A\#_{3\mathrm{Sus}}$ (AU) added to the network of Voyager (H), Voyager (L), and Virgo+ (see Fig. 2), the null (worst) case scenario is $5.6^{+10}_{-4.5}$ yr^{-1} ($1.8^{+3.2}_{-1.4}$ yr^{-1}). The combined uncertainty interval accounting for both scenarios would be from 0.4 yr^{-1} to 16 yr^{-1}. Since we presently have insufficient evidence to reject the null hypothesis, however, we use the null-case scenario in the results. We defer determining the true rejection scenario or eliminating the rejections to future work.

[1] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt *et al.*, Phys. Rev. X **9**, 031040 (2019).
[2] R. Abbott *et al.*, Phys. Rev. X **11**, 021053 (2021).
[3] R. Abbott *et al.*, Phys. Rev. X **13**, 041039 (2023).
[4] S. Vitale, Science **372**, eabc7397 (2021).
[5] R.-G. Cai, Z. Cao, Z.-K. Guo, S.-J. Wang, and T. Yang, Natl. Sci. Rev. **4**, 687 (2017).
[6] M. C. Miller and N. Yunes, Nature (London) **568**, 469 (2019).
[7] M. Mancarella, N. Borghi, S. Foffa, E. Genoud-Prachex, F. Iacovelli, M. Maggiore, M. Moresco, and M. Schulz, in *Proceedings of 41st International Conference on High Energy Physics* (2022), Vol. 414, p. 127, 10.22323/1.414.0127.
[8] J. Aasi, B. P. Abbott, R. Abbott, T. Abbott, M. R. Abernathy, K. Ackley, C. Adams, T. Adams, P. Addesso *et al.*, Classical Quantum Gravity **32**, 074001 (2015).
[9] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardin *et al.*, Classical Quantum Gravity **32**, 024001 (2015).
[10] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya *et al.*, Phys. Rev. Lett. **119**, 161101 (2017).
[11] T. Akutsu, M. Ando, K. Arai, Y. Arai, S. Araki, A. Araya, N. Aritomi, H. Asada, Y. Aso, S. Atsuta *et al.*, Nat. Astron. **3**, 35 (2019).
[12] B. P. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, V. Adya, C. Affeldt, M. Agathos *et al.*, Living Rev. Relativity **23**, 1 (2020).
[13] M. Saleem, J. Rana, V. Gayathri, A. Vijaykumar, S. Goyal, S. Sachdev, J. Suresh, S. Sudhagar, A. Mukherjee, G. Gaur *et al.*, Classical Quantum Gravity **39**, 025004 (2022).
[14] Report from the LSC Post-O5 Study Group, LIGO Document No. T2200287-v2, 2022.
[15] R. X. Adhikari, K. Arai, A. F. Brooks, C. Wipf, O. Aguiar, P. Altin, B. Barr, L. Barsotti, R. Bassiri, A. Bell *et al.*, Classical Quantum Gravity **37**, 165003 (2020).

[16] D. Bersanetti, B. Patricelli, O. J. Piccinni, F. Piergiovanni, F. Salemi, and V. Sequino, Universe **7**, 322 (2021).
[17] Virgo post-O5 instrument upgrades, LIGO Document No. G2100636-v1, 2021.
[18] T. Akutsu, M. Ando, K. Arai, Y. Arai, S. Araki, A. Araya, N. Aritomi, H. Asada, Y. Aso, S. Bae *et al.*, Prog. Theor. Exp. Phys. **2021**, 05A103 (2020).
[19] M. Evans, R. X. Adhikari, C. Afle, S. W. Ballmer, S. Biscoveanu, S. Borhanian, D. A. Brown, Y. Chen, R. Eisenstein, A. Gruson *et al.*, arXiv:2109.09882.
[20] M. Maggiore, C. Van Den Broeck, N. Bartolo, E. Belgacem, D. Bertacca, M. A. Bizouard, M. Branchesi, S. Clesse, S. Foffa, J. García-Bellido *et al.*, J. Cosmol. Astropart. Phys. 03 (2020) 050.
[21] D. G. Blair, P. Barriga, A. F. Brooks, P. Charlton, D. Coward, J.-C. Dumas, Y. Fan, D. Galloway, S. Gras, D. J. Hosken *et al.*, J. Phys. Conf. Ser. **122**, 012001 (2008).
[22] K. Ackley, V. B. Adya, P. Agrawal, P. Altin, G. Ashton, M. Bailes, E. Baltinas, A. Barbuio, D. Beniwal, C. Blair *et al.*, Publ. Astron. Soc. Aust. **37**, E047 (2020).
[23] N. Sarin and P. D. Lasky, Publ. Astron. Soc. Aust. **39**, e007 (2022).
[24] P. D. Lasky, Publ. Astron. Soc. Aust. **32**, e034 (2015).
[25] L. Baiotti, Prog. Part. Nucl. Phys. **109**, 103714 (2019).
[26] A. Bauswein, N.-U. F. Bastian, D. B. Blaschke, K. Chatziioannou, J. A. Clark, T. Fischer, and M. Oertel, Phys. Rev. Lett. **122**, 061102 (2019).
[27] N. Andersson, Universe **7**, 97 (2021).
[28] B. P. Abbott *et al.*, Astrophys. J. Lett. **892**, L3 (2020).
[29] B. P. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V. Adya *et al.*, Astrophys. J. Lett. **848**, L13 (2017).
[30] D. A. Coulter, R. J. Foley, C. D. Kilpatrick, M. R. Drout, A. L. Piro, B. J. Shappee, M. R. Siebert, J. D. Simon, N. Ulloa, D. Kasen *et al.*, Science **358**, 1556 (2017).
[31] B. P. Abbott *et al.*, Astrophys. J. Lett. **848**, L12 (2017).
[32] I. Andreoni, K. Ackley, J. Cooke, A. Acharyya, J. R. Allison, G. E. Anderson, M. C. B. Ashley, D. Baade, M. Bailes, K. Bannister *et al.*, Publ. Astron. Soc. Aust. **34**, e069 (2017).
[33] D. E. Holz and S. A. Hughes, Astrophys. J. **629**, 15 (2005).
[34] B. P. Abbott *et al.*, Nature (London) **551**, 85 (2017).
[35] K. Hotokezaka, E. Nakar, O. Gottlieb, S. Nissanke, K. Masuda, G. Hallinan, K. P. Mooley, and A. T. Deller, Nat. Astron. **3**, 940 (2019).
[36] K. Mooley, A. Deller, O. Gottlieb, E. Nakar, G. Hallinan, S. Bourke, D. Frail, A. Horesh, A. Corsi, and K. Hotokezaka, Nature (London) **561**, 355 (2018).
[37] E. Pian, Front. Astron. Space Sci. **7**, 609460 (2021).
[38] D. Martynov, H. Miao, H. Yang, F. H. Vivanco, E. Thrane, R. Smith, P. Lasky, W. E. East, R. Adhikari, A. Bauswein *et al.*, Phys. Rev. D **99**, 102004 (2019).
[39] H.-Y. Chen, M. Fishbach, and D. E. Holz, Nature (London) **562**, 545 (2018).
[40] P. D. Lasky, in *Multimessenger Astronomy in Practice* (IOP Publishing, 2021), pp. 9–1 to 9–72.
[41] This is the current site for the Square Kilometre Array Pathfinder in Western Australia at 26°41'49.2"S and 116° 37'51.9"E [42].
[42] S. J. Tingay, R. Goeke, J. D. Bowman, D. Emrich, S. M. Ord, D. A. Mitchell, M. F. Morales, T. Booler, B. Crosse, R. B. Wayth *et al.*, Publ. Astron. Soc. Aust. **30**, e007 (2013).
[43] S. Borhanian, Classical Quantum Gravity **38**, 175014 (2021).
[44] NEMO Noise Budget, LIGO Document No. T2000062-v1, 2021.
[45] The case for dovetailing LIGO observations with next-generation facilities, LIGO Document No. P2300166-v9, 2023.
[46] The sensitivity of A# with A+ coatings is shown in Fig. 19 from Ref. [14]. To model the triple suspension system, we sum the sensitivity curve for A# with A+ coatings in quadrature to an f^{-6} curve intersecting it at 100 Hz. This is such that the new sensitivity is a factor of $\sqrt{2}$ worse at 100 Hz and has an f^{-6} roll-off for frequencies $f < 100$ Hz down to the fundamental mechanical resonance at $\mathcal{O}(1)$ Hz.
[47] M. Evans, A. Corsi, C. Afle, A. Ananyeva, K. Arun, S. Ballmer, A. Bandopadhyay, L. Barsotti, M. Baryakhtar, E. Berger *et al.*, arXiv:2306.13745.
[48] V. Srivastava, D. Davis, K. Kuns, P. Landry, S. Ballmer, M. Evans, E. D. Hall, J. Read, and B. S. Sathyaprakash, Astrophys. J. **931**, 22 (2022).
[49] I. Gupta, C. Afle, K. G. Arun, A. Bandopadhyay, M. Baryakhtar, S. Biscoveanu, S. Borhanian, F. Broekgaarden, A. Corsi, A. Dhani *et al.*, arXiv:2307.10421.
[50] There is not a direct comparison of our results to the recent multimessenger results from Ref. [49] since we consider a different A# sensitivity curve, place an observatory in the Southern hemisphere, and include signal-to-noise ratio and distance thresholds in our sky-localization analysis. The closest immediate comparison that we can make is to the tight localization rates shown in Tables I and VII of Ref. [49] where we agree on the order of magnitudes but not on the less significant digits. This is expected given the $\mathcal{O}(1)$ per year uncertainty from the Monte Carlo simulation.
[51] D. Davis *et al.*, Classical Quantum Gravity **38**, 135014 (2021).
[52] S. J. Thomas, J. Barr, S. Callahan, A. W. Clements, F. Daruich, J. Fabrega, P. Ingraham, W. Gressler, F. Munoz, D. Neill *et al.*, in *Ground-Based and Airborne Telescopes VIII* (SPIE, 2020), Vol. 11445, pp. 68–82.
[53] R. Abbott *et al.*, Phys. Rev. X **13**, 011048 (2023).
[54] S. Borhanian and B. S. Sathyaprakash, arXiv:2202.11048.
[55] LIGO Scientific Collaboration, Free software (GPL) (2018).
[56] G. Van Rossum and F. L. Drake, Jr., *Python Tutorial* (Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands, 1995).
[57] U. Dupletsa, J. Harms, B. Banerjee, M. Branchesi, B. Goncharov, A. Maselli, A. Oliveira, S. Ronchini, and J. Tissino, Astron. Comput. **42**, 100671 (2023).
[58] F. Iacovelli, M. Mancarella, S. Foffa, and M. Maggiore, Astrophys. J. Suppl. Ser. **263**, 2 (2022).
[59] We consider the multiparameter case where the Fisher information matrix provides a bound on the covariance matrix, e.g. see Ref. [60], but for simplicity we only explain the single-parameter case here. Importantly, that the lower bound can be loose outside of the Gaussian regime remains true.

[60] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, England, 2009).
[61] We have found some discrepancies in our reproduction of the results in Ref. [54] for other parameter estimates (e.g., the mass asymmetry and luminosity distance) that are not used anywhere in our analysis. We emphasize that the signal-to-noise ratio and sky-area estimates are consistent with Ref. [54] and that we use the true luminosity distance to define which sources are in range. Although this does not affect any of our presented results, the cause of the discrepancy remains to be investigated further, e.g., to extend our results to other metrics and science cases.
[62] C. L. Rodriguez, B. Farr, W. M. Farr, and I. Mandel, Phys. Rev. D **88**, 084013 (2013).
[63] M. Vallisneri, Phys. Rev. D **77**, 042001 (2008).
[64] R. Magee and S. Borhanian, Astrophys. J. **935**, 139 (2022).
[65] F. Iacovelli, M. Mancarella, S. Foffa, and M. Maggiore, Astrophys. J. **941**, 208 (2022).
[66] The field-of-view does not strictly limit the available sky-area to search over since tiling a given region is common.
[67] E. C. Bellm, S. R. Kulkarni, M. J. Graham, R. Dekany, R. M. Smith, R. Riddle, F. J. Masci, G. Helou, T. A. Prince, S. M. Adams *et al.*, Publ. Astron. Soc. Pac. **131**, 018002 (2018).
[68] Željko Ivezić, S. M. Kahn, J. A. Tyson, B. Abel, E. Acosta, R. Allsman, D. Alonso, Y. AlSayyad, S. F. Anderson, J. Andrew *et al.*, Astrophys. J. **873**, 111 (2019).
[69] The Monte Carlo resolution is the inverse of the total number of sources simulated in range following the cosmological distribution. The simulation finding that the baseline network observes zero sources only implies that it observed fewer sources than the Monte Carlo resolution. This provides a lower bound on the geometric improvement. This bound is not invariant of the unified rejection rate.
[70] T. Y.-Y. Hsiao, T. Goto, T. Hashimoto, D. J. D. Santos, Y. H. V. Wong, S. J. Kim, B. J. R. Raquel, S. C.-C. Ho, B.-H. Chen, E. Kilerci *et al.*, Mon. Not. R. Astron. Soc. **517**, 447 (2022).
[71] G. B. Taylor and R. C. Vermeulen, Astrophys. J. **485**, L9 (1997).
[72] P. J. Easter, P. D. Lasky, and A. R. Casey, arXiv:2106.04064.
[73] J. W. Gardner, CEonlyPony (2023), https://git.ligo.org/jameswalter.gardner/CEonlyPony.
[74] F. Pérez and B. E. Granger, Comput. Sci. Eng. **9**, 90 (2007).
[75] T. Kluyver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier *et al.*, in *Positioning and Power in Academic Publishing: Players, Agents and Agendas* (2016), pp. 87–90.
[76] T. E. Oliphant, *A Guide to NumPy* (Trelgol Publishing, USA, 2006).
[77] J. D. Hunter, Comput. Sci. Eng. **9**, 90 (2007).
[78] Wes McKinney, in *Proceedings of the 9th Python in Science Conference* (2010), pp. 56–61, 10.25080/Majora-92bf1922-012.
[79] A. M. Price-Whelan, P. L. Lim, N. Earl, N. Starkman, L. Bradley, D. L. Shupe, A. A. Patil, L. Corrales, C. E. Brasseur *et al.* (Astropy Collaboration), Astrophys. J. **935**, 167 (2022).
[80] Free Software Foundation, Bash (3.2. 48) [Unix shell program] (2007), https://www.gnu.org/software/bash/.
[81] A. B. Yoo, M. A. Jette, and M. Grondona, in *Job Scheduling Strategies for Parallel Processing* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2003), pp. 44–60.
[82] L. Baiotti, B. Giacomazzo, and L. Rezzolla, Phys. Rev. D **78**, 084033 (2008).
[83] K. Takami, L. Rezzolla, and L. Baiotti, Phys. Rev. D **91**, 064001 (2015).
[84] OzSTAR, https://supercomputing.swin.edu.au/ozstar/. Accessed online on February 3 2023.
[85] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, NJ, 1946).
[86] C. R. Rao, *Information and the Accuracy Attainable in the Estimation of Statistical Parameters* (Springer, New York, 1945).
[87] The mpmath Development Team, MPMATH: A Python library for arbitrary-precision floating-point arithmetic (version 1.3.0) (2023), http://mpmath.org/.
[88] S. Borhanian, GWBENCH (2023), https://gitlab.com/sborhanian/gwbench.

Chapter 3

Achieving the fundamental quantum limit of linear waveform estimation

Me, poor man, my library
Was dukedom large enough.

Antonio (Act 2, Scene 1)
William Shakespeare, The Tempest

We now present Ref. [49] and with it enter the quantum regime of parameter estimation where we must negotiate the quantum projection noise. This paper resolves the origin and attainability of a $\sqrt{2}$ gap in the sensitivity for a detuned gravitational-wave detector that was previously observed [39]. In particular, we determine the optimal measurement to overcome this gap in certain cases. Although detuned gravitational-wave detection was our historical motivation, we close the gap for linear waveform estimation of any deterministic signal in the asymptotic regime, i.e. for signals lying in the top-right of Fig. 1.1. In the larger story of this thesis, this paper is our first instance of optimising the available quantum information over the design of the experiment. This is a motif that will repeat constantly in the following three papers. Here, the relevant quantum information theoretic quantity is the Holevo Cramér-Rao bound (defined formally in Eq. S1) and we optimise it over the final measurement scheme (see Eq. 13 and the Supplemental Material starting from Eq. S23). The measurement scheme that we find, although nonstationary unlike the present DC readout homodyne measurement in LIGO, remains linear and Gaussian unlike the more unconventional non-Gaussian protocols explored in the following three papers. This paper thus represents the first step in this thesis away

from conventional interferometry. It sets up the later, more dramatic departures which assume increasing access and control over the quantum state and thus promise more than the $\sqrt{2}$ sensitivity improvement achieved here.

We now highlight the key results and steps in this paper. The analogy established from Eq. 6 between the waveform estimation problem and the toy model is crucial to this work and the following two papers. Using the linearity of the device, the continuum waveform estimation problem becomes a one or two parameter estimation problem. This toy problem is more readily solved, e.g. in Eqs. 10–11 provide the sensitivity limits achieved by the measurement scheme in Eq. 13. The original waveform sensitivity can then be backed out, e.g. as shown in Fig. 4 for a lossy gravitational-wave interferometer (see also Eq. S45). Realising the optimal nonstationary measurement scheme experimentally is variously difficult depending on the parameter regime as discussed following Eq. 13 and in the Supplemental Material. Even the most ambitious scheme with two detuned squeezed filter cavities, however, is still linear and may be much easier than the highly non-Gaussian protocols explored later in this thesis. This indicates that some of these ideas may take even decades to become possible or necessary, once all other ways of improving the sensitivity classically are exhausted (e.g. increased arm length or circulating optical power for an interferometer).

Achieving the Fundamental Quantum Limit of Linear Waveform Estimation

James W. Gardner,[1,2,*] Tuvia Gefen,[3,†] Simon A. Haine,[4] Joseph J. Hope,[4] and Yanbei Chen[2]
[1]OzGrav-ANU, Centre for Gravitational Astrophysics, Research Schools of Physics, and of Astronomy and Astrophysics, The Australian National University, Canberra ACT 2601, Australia
[2]Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, California 91125, USA
[3]Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA
[4]Department of Quantum Science and Technology and Department of Fundamental and Theoretical Physics, Research School of Physics, The Australian National University, Canberra ACT 0200, Australia

(Received 11 August 2023; revised 23 October 2023; accepted 20 February 2024; published 28 March 2024)

Sensing a classical signal using a linear quantum device is a pervasive application of quantum-enhanced measurement. The fundamental precision limits of linear waveform estimation, however, are not fully understood. In certain cases, there is an unexplained gap between the known waveform-estimation quantum Cramér-Rao bound and the optimal sensitivity from quadrature measurement of the outgoing mode from the device. We resolve this gap by establishing the fundamental precision limit, the waveform-estimation Holevo Cramér-Rao bound, and how to achieve it using a nonstationary measurement. We apply our results to detuned gravitational-wave interferometry to accelerate the search for postmerger remnants from binary neutron-star mergers. If we have an unequal weighting between estimating the signal's power and phase, then we propose how to further improve the signal-to-noise ratio by a factor of $\sqrt{2}$ using this nonstationary measurement.

DOI: 10.1103/PhysRevLett.132.130801

In our efforts to probe fundamental physics, we invariably encounter the quantum limit: the irrevocable statistical nature of our reality [1–3]. This fundamental uncertainty of our measurement devices limits the precision at which we can sense classical signals.

We consider the general problem of estimating a real classical signal $s(t)$ for all times t using a linear quantum device as shown in Fig. 1(a) [4]. The device evolves linearly according to a Hamiltonian $\hat{H}(t) = \hat{H}_0 + \hat{H}_{\rm int}(t)$ with the interaction $\hat{H}_{\rm int}(t) = -s(t)\hat{G}$ [5]. Observables that do not commute with the generator $\hat{G}$ (an internal degree of freedom) respond linearly to $s(t)$. We assume that the device is in a stationary state of $\hat{H}_0$ which is time invariant. In the input-output formalism [6], information about the signal leaks out of the device into the environment imprinted on an outgoing mode of a bosonic field. By measuring this mode, we obtain a classical estimate of the classical signal mediated by the quantum device. The outgoing bosonic mode at each position and time is a harmonic oscillator with canonical quadratures $\hat{x}$ and $\hat{p}$ which obey $[\hat{x}, \hat{p}] = i$ (let $\hbar = 1$ henceforth). Let $\hat{x}_\theta := \cos(\theta)\hat{x} + \sin(\theta)\hat{p}$ for real θ such that the outgoing mode is [7,8]

$$\hat{x}_\theta(t) = \hat{x}_\theta^{(0)}(t) + \int_{-\infty}^{\infty} \mathrm{d}t'\, \chi_{x_\theta G}(t-t')s(t') \tag{1}$$

where the superscript (0) denotes the free evolution under $\hat{H}_0$ and the susceptibility is $\chi_{x_\theta G}(t-t') := i[\hat{x}_\theta^{(0)}(t), \hat{G}^{(0)}(t')]\Theta(t-t')$ with Θ the Heaviside function. In the Fourier domain, Eq. (1) becomes

$$\tilde{x}_\theta(\Omega) = \tilde{x}_\theta^{(0)}(\Omega) + \chi_{x_\theta G}(\Omega)\tilde{s}(\Omega) \tag{2}$$

where the mode at each frequency Ω is displaced by the signal's complex Fourier component $\tilde{s}(\Omega) := \int_{-\infty}^{\infty} \mathrm{d}t\, e^{i\Omega t} s(t)$. Since $s(t)$, $\hat{x}_\theta(t)$, and $\chi_{x_\theta G}(t)$ are real, their Fourier components obey $\tilde{s}(-\Omega) = \tilde{s}^\dagger(\Omega)$ etc. such that it suffices to consider only positive frequencies.

Quantum metrology extends the classical theory of estimating parameters from a probability distribution. In particular, the quantum Cramér-Rao bound (QCRB) [9,10], sets a fundamental precision limit: a lower bound on the variance of unbiased estimation of parameters encoded in a quantum state. This limit only depends on the state itself and not on the measurement scheme. In the real single-parameter case, the QCRB can always be saturated by the optimal measurement if the sample size is large [9].

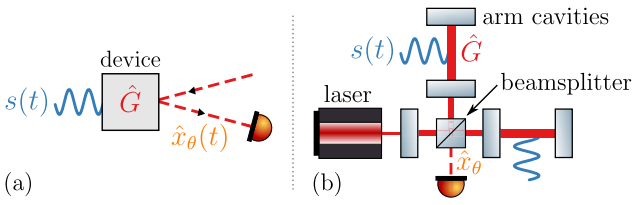

FIG. 1. (a) A linear quantum device coupled to a classical signal. (b) For example, a gravitational-wave interferometer.

Additionally, if the parameter appears as the shift in the mean of a Gaussian state, then the QCRB can be saturated for any sample size. In the multiparameter case, however, reaching the QCRB is not guaranteed. In general, the QCRB can be saturated if and only if the symmetric logarithmic derivatives with respect to the parameters weakly commute [11]. For real parameters s_j encoded as the shift in the mean of a Gaussian state by the unitary transformation $\exp(-i\sum_j s_j \hat{G}_j)$, this condition is equivalent to the generators $\hat{G}_j$ weakly commuting, i.e., $\langle[\hat{G}_j, \hat{G}_k]\rangle = 0, \ \forall \ j,k$.

Here, we want to simultaneously estimate the continuum of independent complex parameters $\tilde{s}(\Omega)$ in Eq. (2), one for each positive frequency Ω.

We assume that the linear device is at the quantum limit such that it is in a pure Gaussian state. The waveform-estimation QCRB S_Q^{wave} [4] sets a lower bound on the error S_{ss} of unbiased estimation of $\tilde{s}(\Omega)$ of

$$\frac{S_{x_\theta x_\theta}(\Omega)}{|\chi_{x_\theta G}(\Omega)|^2} =: S_{ss}(\Omega) \geq S_Q^{\mathrm{wave}}(\Omega) := \frac{1}{S_{GG}(\Omega)} \tag{3}$$

where the power spectral density $S_{z_1 z_2}$ for *stationary* random processes $\hat{z}_1$ and $\hat{z}_2$ is defined as

$$2\pi\delta(\Omega - \Omega')S_{z_1 z_2}(\Omega) = \langle\{\hat{z}_1(\Omega), \hat{z}_2^\dagger(\Omega')\}\rangle. \tag{4}$$

The QCRB depends only on the reciprocal of the fluctuations of the generator S_{GG}.

This can be explained by the generalized uncertainty principle for waveform estimation [12]

$$S_{x_\theta x_\theta}(\Omega)S_{GG}(\Omega) \geq |\chi_{x_\theta G}(\Omega)|^2 + |S_{x_\theta G}(\Omega)|^2 + c(\Omega) \tag{5}$$

where, in the quantum limit, this is an equality and $c(\Omega)$ is a term that vanishes [5].

From Eq. (5), Ref. [5] showed that measuring the stationary complex quadrature $\hat{x}_\theta(\Omega)$ saturates the QCRB in Eq. (3) if and only if it is uncorrelated with the generator, i.e., $S_{x_\theta G} = 0$. Furthermore, Ref. [5] showed that when $S_{x_\theta G} \neq 0, \ \forall \ \theta$ the optimal error S_{ss} is still within a factor of 2 of the QCRB.

It is not necessary *a priori*, however, that we measure a stationary complex quadrature $\hat{x}_\theta(\Omega)$. Equation (3) only applies to such stationary measurements. This leaves several important questions unanswered. What is the QCRB in general and when can it be saturated? If it cannot be saturated, then what is the optimal precision? And, what measurement attains this limit? We will answer these questions and demonstrate our results using gravitational-wave interferometry.

Cosine and sine phases.—At the positive frequency Ω, let $\tilde{s}(\Omega) = \pi T(A + iB)$ where A and B are independent real degrees of freedom and T is the finite integration time. In the time domain, this component of the signal is $A\cos(\Omega t) + B\sin(\Omega t)$ at a given time t where A and B are the cosine and sine phases of the signal $s(t)$ at frequency Ω, respectively.

Our goal of measuring the signal s, therefore, is equivalent to simultaneously estimating A and B at each Ω.

Our weighted figure of merit for the precision at Ω is $\Sigma = 2w\mathrm{Var}[\hat{A}] + 2(1-w)\mathrm{Var}[\hat{B}]$ where $w \in (0,1)$ and $\hat{A}$ and $\hat{B}$ are unbiased estimates of A and B, respectively. (Without loss of generality, we can assume that our weight matrix is diagonal.) The weights may be unequal $w \neq 0.5$ for several reasons. For example, if we want to estimate the signal's power $|\tilde{s}|^2 \propto A^2 + B^2$ more than its phase, then the weights would be unequal because the derivatives of $|\tilde{s}|^2$ depend on A and B. We assume a uniform prior on A and B and distinguish that, while unequal weights indicate how much more is wanted to be known *a posteriori* about A than B, a nonuniform prior would indicate how much more is known *a priori* about A than B [13–15].

Similarly to the signal, we can split the complex quadratures of the outgoing light $\hat{x}_\theta(\Omega)$ into their real and imaginary parts in the frequency domain. Or, equivalently, into their cosine and sine phases in the time domain, e.g., see Refs. [16,17]. These parts are Hermitian but *nonstationary*. In this manner, Eq. (2) becomes

$$\vec{\hat{q}} := \frac{1}{\sqrt{\pi T}} \begin{bmatrix} \mathrm{Re}[\hat{x}(\Omega)] \\ \mathrm{Re}[\hat{p}(\Omega)] \\ \mathrm{Im}[\hat{x}(\Omega)] \\ \mathrm{Im}[\hat{p}(\Omega)] \end{bmatrix} = \vec{\hat{q}}^{(0)} + A\vec{d}_A + B\vec{d}_B. \tag{6}$$

Let $\vec{\chi} = [\chi_{xG}(\Omega), \chi_{pG}(\Omega)]^{\mathrm{T}}$, then the real signal displacements are

$$\vec{d}_A := \sqrt{\pi T}\begin{bmatrix} \mathrm{Re}[\vec{\chi}] \\ \mathrm{Im}[\vec{\chi}] \end{bmatrix}, \qquad \vec{d}_B := \sqrt{\pi T}\begin{bmatrix} -\mathrm{Im}[\vec{\chi}] \\ \mathrm{Re}[\vec{\chi}] \end{bmatrix}. \tag{7}$$

These are orthogonal and have the same Euclidean norm

$$l = \sqrt{\pi T\left(|\chi_{xG}(\Omega)|^2 + |\chi_{pG}(\Omega)|^2\right)}. \tag{8}$$

Since $[\hat{x}(\Omega), \hat{p}(\Omega')] = i2\pi\delta(\Omega - \Omega')$, by using $\mathrm{Re}[z] = \frac{1}{2}(z + z^*)$ and $\mathrm{Im}[z] = (1/2i)(z - z^*)$ we have that $[\vec{\hat{q}}_1, \vec{\hat{q}}_2] = [\vec{\hat{q}}_3, \vec{\hat{q}}_4] = i$ with all other commutators zero ($\vec{\hat{q}}_j$ is the jth element of $\vec{\hat{q}}$). The system at each frequency, therefore, comprises two real displaced harmonic oscillators.

We assume that the noise is stationary such that the complex quadratures $\vec{\hat{x}} = [\hat{x}(\Omega), \hat{p}(\Omega)]^{\mathrm{T}}$ have the 2-by-2 covariance matrix $\frac{1}{2}\langle\{\vec{\hat{x}}_j^{(0)}, \vec{\hat{x}}_k^{(0)}\}\rangle = (V_2)_{jk}$ and the parts $\vec{\hat{q}}$ have the 4-by-4 covariance matrix $\frac{1}{2}\langle\{\vec{\hat{q}}_j^{(0)}, \vec{\hat{q}}_k^{(0)}\}\rangle = (V_2 \oplus V_2)_{jk}$. (Without loss of generality, we assume that $\langle\vec{\hat{x}}^{(0)}\rangle = 0$.) Since the device is linear, distinct frequencies

are uncorrelated. For the moment, we assume that the pure state at each frequency is vacuum, i.e., $V_2 = \mathrm{diag}(\frac{1}{2}, \frac{1}{2})$, and will generalize later.

To measure A and B from the output light in Eq. (6), the naïve optimal unbiased estimates are $\hat{A}_{\text{naïve}} = l^{-2}\vec{d}_A \cdot \vec{\hat{q}}$ and $\hat{B}_{\text{naïve}} = l^{-2}\vec{d}_B \cdot \vec{\hat{q}}$. The variance of each estimate is $\frac{1}{2}l^{-2}$ such that the naïve figure of merit is $\Sigma_{\text{naïve}} = l^{-2}$. [Note that $\Sigma_{\text{naïve}} \propto T^{-1}$ by Eq. (8) such that integrating for longer times reduces the error as expected].

These measurements, however, may not commute since $[\hat{A}_{\text{naïve}}, \hat{B}_{\text{naïve}}] = i\mu l^{-2}$ where

$$\mu = 2\pi T l^{-2}(\mathrm{Re}[\chi_{pG}]\mathrm{Im}[\chi_{xG}] - \mathrm{Re}[\chi_{xG}]\mathrm{Im}[\chi_{pG}]) \quad (9)$$

such that $0 \leq |\mu| \leq 1$. (Without loss of generality, we assume that $\mu \geq 0$.) This means that A and B cannot be simultaneously estimated to attain the naïve figure of merit if $\mu \neq 0$. The displacements in Eq. (7) are generated by their conjugate quadratures $\hat{G}_A$ and $\hat{G}_B$ which obey the same commutation relation such that $\mu = 0$ is equivalent to the weak commutativity condition $\langle[\hat{G}_A, \hat{G}_B]\rangle = 0$. The QCRB for simultaneous estimation of A and B, therefore, can be saturated if and only if $\mu = 0$ which is equivalent to $\exists\, \theta \in \mathbb{R}$ such that $\chi_{x_\theta G} = 0$. In fact, the QCRB is precisely the na-ve figure of merit above such that $\Sigma \geq \Sigma_Q = \Sigma_{\text{naïve}}$. This can be shown from the result that the QCRB with respect to s_j given the unitary transformation $\exp(-is_j\hat{G}_j)$ is $(4\mathrm{Var}[G_j])^{-1}$ [9].

Fundamental precision limit.—If $\mu \neq 0$ such that the QCRB cannot be saturated, then the optimal attainable precision is instead the Holevo Cramér-Rao bound (HCRB) Σ_H which accounts for the commutator of the estimates such that $\Sigma \geq \Sigma_H > \Sigma_Q$ [11,15,18,19]. Since the real parameters A and B appear as the shift in the mean of a pure Gaussian state, the HCRB is saturated by the optimal commuting linear combinations of $\vec{\hat{q}}$ [11]. (This is equivalent to finding the optimal quantum mechanics–free subspace [20].) We calculate the HCRB using the method from Ref. [18] in the Supplemental Material [21].

We show that the ratio of the HCRB Σ_H to the QCRB $\Sigma_Q = l^{-2}$ reduces to single-parameter optimization

$$\frac{\Sigma_H}{\Sigma_Q} = \min_{\phi \in (0,\pi]} \left(\frac{w}{\cos(\phi)^2} + \frac{1-w}{\cos[\phi + \arcsin(\mu)]^2} \right) \quad (10)$$

and we find analytic solutions in certain limits

$$\frac{\Sigma_H}{\Sigma_Q} \overset{\mu=1}{\to} 1 + 2\sqrt{w(1-w)}, \qquad \frac{\Sigma_H}{\Sigma_Q} \overset{w=\frac{1}{2}}{\to} \frac{2}{1+\sqrt{1-\mu^2}}. \quad (11)$$

Figure 2(a) shows that the ratio of the HCRB to the QCRB is at most two which agrees with Ref. [5]. The HCRB increases monotonically with μ and decreases as the weights become less equal as shown in Fig. 2(b). The HCRB reduces to the QCRB for single-parameter estimation at $w = 0, 1$.

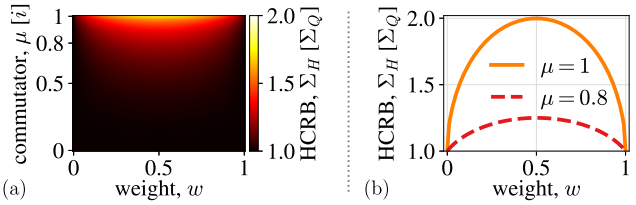

FIG. 2. (a) HCRB versus the commutator and weight. (b) HCRB versus the weight for different commutators.

These results generalize to squeezed states. Let $\hat{S}$ be the conjugate squeezing operator such that $V_2 \mapsto \mathrm{diag}(\frac{1}{2}, \frac{1}{2})$. This unitary transformation does not affect μ or the bounds but does map the signal displacements as $\vec{d} \mapsto (S \oplus S)\vec{d}$ such that

$$l \mapsto l' := \sqrt{\pi T(\|S\mathrm{Re}[\vec{\chi}]\|^2 + \|S\mathrm{Im}[\vec{\chi}]\|^2)}. \quad (12)$$

The general stationary pure Gaussian state case, therefore, is equivalent to the vacuum case with the same μ and w but with $\Sigma_Q = (l')^{-2}$. We emphasize that we only apply $\hat{S}$ mathematically to derive the bounds; it is not required experimentally.

Optimal measurement scheme.—There exists a unique symplectic transformation of the two harmonic oscillators $\vec{\hat{q}} \mapsto \vec{\hat{X}} = (\hat{X}_1, \hat{P}_1, \hat{X}_2, \hat{P}_2)^{\mathrm{T}}$ that maps the normalized displacements as $l^{-1}\hat{d}_A \mapsto \hat{X}_1$ and $l^{-1}\hat{d}_B \mapsto \mu\hat{P}_1 + \sqrt{1-\mu^2}\hat{X}_2$ such that their commutator remains $i\mu$. In this basis, the optimal commuting unbiased estimates are [21]

$$\hat{A} = [l\cos(\bar{\phi})]^{-1}[\cos(\bar{\phi})\hat{X}_1 - \sin(\bar{\phi})\hat{P}_2]$$
$$\hat{B} = \{l\cos[\bar{\phi} + \arcsin(\mu)]\}^{-1}[\cos(\bar{\phi})\hat{X}_2 - \sin(\bar{\phi})\hat{P}_1] \quad (13)$$

where $\bar{\phi}$ is the optimal angle in Eq. (10).

These estimates are two nonstationary quadratures: arbitrary real linear combinations of $\vec{\hat{q}}$. Compare this to the stationary complex quadrature $\hat{x}_\theta(\Omega)$ with the real part $\cos(\theta)\vec{\hat{q}}_1 + \sin(\theta)\vec{\hat{q}}_2$ and imaginary part $\cos(\theta)\vec{\hat{q}}_3 + \sin(\theta)\vec{\hat{q}}_4$. For squeezed states, similarly, the optimal measurement consists of two nonstationary quadratures (mathematically, first apply $\hat{S}$ and then the symplectic transformation).

We propose how to experimentally realize these nonstationary measurements at a given Ω. We expand $\vec{\hat{q}}$ into the time domain using $\mathrm{Re}[\hat{x}_\theta(\Omega)] = \int_{-\infty}^{\infty} dt\, \cos(\Omega t)\hat{x}_\theta(t)$ and $\mathrm{Im}[\hat{x}_\theta(\Omega)] = \int_{-\infty}^{\infty} dt\, \sin(\Omega t)\hat{x}_\theta(t)$. Since the measurements are linear combinations of $\vec{\hat{q}}$, therefore, they are $\hat{A} = \int_{-\infty}^{\infty} dt c_A(t)\hat{x}_{\theta_A(t)}(t)$ and $\hat{B} = \int_{-\infty}^{\infty} dt c_B(t)\hat{x}_{\theta_B(t)}(t)$ for some

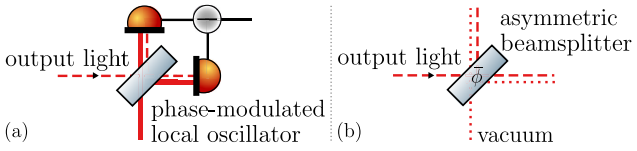

FIG. 3. (a) Phase-modulated balanced homodyne readout. (b) Asymmetric beam splitter with power reflectivity $\cos(\bar{\phi})^2$.

real amplitudes $c_A(t)$ and $c_B(t)$ and phases $\theta_A(t)$ and $\theta_B(t)$ [21]. For example, if $w = 1$ such that we only want to estimate A, then, as shown in Fig. 3(a), we use homodyne readout with a phase-modulated local oscillator with phase $\theta_A(t)$ to obtain the time series $\hat{x}_{\theta_A(t)}(t)$. By integrating this time series multiplied by $c_A(t)$ in postprocessing, we can achieve $\hat{A}$.

Suppose that instead we want to measure both $\hat{A}$ and $\hat{B}$. Although $\hat{A}$ and $\hat{B}$ commute, their integrands $c_A(t)\hat{x}_{\theta_A(t)}(t)$ and $c_B(t)\hat{x}_{\theta_B(t)}(t)$ above may not commute at a given time. This prevents directly performing simultaneous modulated homodyne measurements. If $\mu = 1$ such that the normalized displacements are $\hat{X}_1$ and $\hat{P}_1$, however, then this can be overcome by using an asymmetric beam splitter with reflectivity $\cos(\bar{\phi})^2$ to mix in an ancillary mode (i.e., uncorrelated vacuum) as shown in Fig. 3(b). Then, measuring $\vec{d}_A$ and $\vec{d}_B$ on the two output modes from the beam splitter using modulated homodyne readouts commutes at each time and saturates the HCRB for any w [21]. The added noise from the ancilla is responsible for the gap from the QCRB.

For the general case of any μ and w, we propose a joint homodyne-heterodyne readout scheme. To obtain the individual estimate of $\hat{A}$ above, we integrated the time series $c_A(t)\hat{x}_{\theta_A(t)}(t)$, but information is available at other frequencies too. In particular, the 2Ω Fourier component beats with the time series which oscillates at Ω to produce a linear combination of the quadratures at Ω and 3Ω. This can realize a heterodyne measurement of $\hat{B}$ at Ω [21,22]. The added heterodyne noise at 3Ω can be suppressed by squeezing the output mode using two cascaded, detuned, and narrow band filter cavities—one for each of the upper and lower sidebands at 3Ω—without affecting the estimates or the fundamental limits at Ω. The HCRB, therefore, can be saturated in the narrow band around Ω using a homodyne measurement of $\hat{A}$ and a simultaneous heterodyne measurement of $\hat{B}$.

Gravitational-wave interferometry.—We demonstrate our results for a gravitational-wave interferometer like the Laser Interferometric Gravitational-wave Observatory (LIGO) [23,24] operated in a hypothetical detuned configuration. For simplicity, we model LIGO as a power-recycled Fabry-Pérot Michelson interferometer as shown in Fig. 1(b) with vacuum input into the "dark port" of the beam splitter [25,26]. In our detuned configuration, the 4-km arm cavities with 750 kW of circulating power are detuned away from the input carrier laser frequency of 282 THz by $\Delta = 2\pi \times 3$ kHz [27–33]. We are interested in detecting 1–4 kHz gravitational-wave signals, e.g., from the postmerger remnant of binary neutron-star mergers, to test our theories of extreme matter [34–42]. (Since we focus on the kilohertz response, we ignore quantum radiation pressure noise [3].) Detuning the interferometer makes it resonant at Δ which improves the peak sensitivity without increasing the circulating power [43–45]. We emphasize that operating LIGO in a detuned configuration presents many technical challenges [33] and here we only want to establish the fundamental limit of achievable sensitivity at a given frequency to better evaluate this configuration.

The differential optical mode of the interferometer can be approximated as a single mode in a detuned cavity linearly coupled to the gravitational-wave strain $s(t)$ by $\hat{H}_{\rm int} = gs(t)\hat{x}_{\rm cav}$. Here, g is the effective coupling rate (mediated by free masses in the transverse-traceless gauge) and $\hat{x}_{\rm cav}$ is the amplitude quadrature of the intracavity mode such that $\hat{G} = -g\hat{x}_{\rm cav}$ [46]. The resulting susceptibility is [21]

$$\vec{\chi} = \frac{\sqrt{2\gamma}g}{\Delta^2 + (\gamma - i\Omega)^2}\begin{bmatrix} \Delta \\ -\gamma + i\Omega \end{bmatrix} \tag{14}$$

where $\gamma = 2\pi \times 42$ Hz is the half-width at half-maximum readout rate of the arm cavities. By Eq. (9), $\mu = [2\Delta\Omega/(\gamma^2 + \Delta^2 + \Omega^2)]$ such that the QCRB cannot be saturated for $\Delta \neq 0$ which agrees with Ref. [5].

In Fig. 4, we compare the HCRB versus frequency to the sensitivity using the optimal stationary quadrature (also known as "variational readout") and nonstationary quadrature measurements. For equal weights, the stationary measurement saturates the HCRB such that the gap to the QCRB is insurmountable. For unequal weights, however, our nonstationary measurement is required to saturate the HCRB.

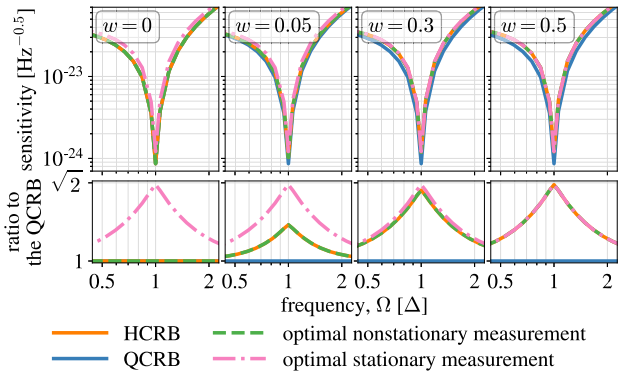

FIG. 4. Strain sensitivity for the detuned LIGO-like interferometer versus frequency for different weights in (top row) effective amplitude spectral density units and (bottom row) ratio to the QCRB.

This unequal weight regime is relevant because, e.g., astrophysically we care more about knowing some parameters of the neutron-star equation of state than others. This can be reduced to having an unequal weighting between the signal's power and phase and, therefore, between A and B. For example, we particularly want to estimate the primary peak of the kilohertz power spectrum to inform our understanding of the equation of state [47]. In the future, with more precise numerical models of the postmerger signal, we may be able to confidently determine the phase of the postmerger signal from a strong enough detection of the inspiral phase. We then only need to estimate the postmerger signal's power which is equivalent to having weights $w = 0$ or 1. In this limiting single-parameter case, we have shown that our nonstationary measurement scheme improves the signal-to-noise ratio by up to a factor of $\sqrt{2}$ at the detuning frequency, an improvement which cannot be surpassed using a different measurement scheme. This corresponds to up to a factor of 2.83 improvement in the volume of the Universe searched for kilohertz signals at the peak frequency, in addition to the gain provided from detuning the interferometer. This could be a significant boost to LIGO's search for kilohertz gravitational waves should the challenges with detuned interferometry be overcome. More realistically, we may instead have partial prior knowledge of the postmerger signal's phase and perform weighted simultaneous estimation of the signal's power and phase. We hypothesize that the sensitivity can still be similarly improved in this regime and defer a detailed study of this application to future work.

Losses limit the possible quantum enhancement of LIGO (where Fig. 4 shows the lossless sensitivity). If we assume optical losses of 100 ppm ($\gamma_l = 2\pi \times 0.3$ Hz) in the arm cavities and $\eta = 0.1$ in the output, then $l \mapsto \sqrt{1-\eta}l$ and $\mu \mapsto \{2\Delta\Omega/[(\gamma+\gamma_l)^2 + \Delta^2 + \Omega^2]\}$ [21]. Since $\gamma_l \ll \gamma \ll \Delta$, μ is unchanged. This implies that the gap from the HCRB to the relevant QCRB, $\Sigma_Q \approx (1-\eta)^{-1}l^{-2}$, is also unchanged and our nonstationary measurement can still achieve up to a factor of $\sqrt{2}$ improvement with losses.

Conclusions.—We have shown how to achieve the fundamental precision limit for the estimation of a classical signal using a linear quantum device. Previous work on linear waveform estimation found an unexplained gap of up to a factor of $\sqrt{2}$ in the signal-to-noise ratio between the optimal stationary quadrature measurement and the QCRB. We showed that this gap stems from the noncommutativity of the na-ve estimates of the cosine and sine phases of the signal at each frequency. This allowed us to establish the fundamental limit of attainable precision and propose how to experimentally realize the optimal nonstationary measurement scheme. We applied these results to the search for postmerger gravitational-wave signals from binary neutron-star mergers using a detuned LIGO-like interferometer. We showed that this nonstationary measurement scheme could significantly increase the volume of the Universe probed for such signals at a given frequency in the unequal weight regime.

Future work could determine the broadband optimal measurement scheme and apply our results to a dual-recycled LIGO-like interferometer with injected squeezed states [48–50] and extend them to other systems, e.g., $\mathcal{PT}$-symmetric interferometers [51–56], axion detectors [57–60], and displacement noise-free interferometers [17,61,62].

Our code is available online [63] and was written using *Mathematica* [64] and PYTHON [65–69].

We thank the following individuals and groups for their advice provided during this research: L. McCuller, S. Kotler, J. Preskill, Z. Mehdi, A. Markowitz, D. Ganapathy, S. Zhou, L. Sun, R. X. Adhikari, the Caltech Y. Chen Quantum Group, and the ANU CGA Squeezer Group. We use component graphics from Ref. [70] with permission. This research is supported by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (Project No. CE170100004). J. W. G. and this research are supported by an Australian Government Research Training Program (RTP) Scholarship and also partially supported by the US NSF Grant No. PHY-2011968. In addition, Y. C. acknowledges the support by the Simons Foundation (Grant No. 568762). T. G. acknowledges funding provided by the Institute for Quantum Information and Matter and the Quantum Science and Technology Scholarship of the Israel Council for Higher Education. S. A. H. acknowledges support through an Australian Research Council Future Fellowship Grant No. FT210100809. This Letter has been assigned LIGO Document No. P2300096.

*james.gardner@anu.edu.au
†tgefen@caltech.edu

[1] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement* (Cambridge University Press, Cambridge, England, 1995).
[2] A. A. Clerk, M. H. Devoret, S. M. Girvin, F. Marquardt, and R. J. Schoelkopf, Rev. Mod. Phys. **82**, 1155 (2010).
[3] S. L. Danilishin and F. Y. Khalili, Living Rev. Relativity **15**, 5 (2012).
[4] M. Tsang, H. M. Wiseman, and C. M. Caves, Phys. Rev. Lett. **106**, 090401 (2011).
[5] H. Miao, R. X. Adhikari, Y. Ma, B. Pang, and Y. Chen, Phys. Rev. Lett. **119**, 050801 (2017).
[6] C. W. Gardiner and M. J. Collett, Phys. Rev. A **31**, 3761 (1985).
[7] R. Kubo, Rep. Prog. Phys. **29**, 255 (1966).
[8] A. Buonanno and Y. Chen, Phys. Rev. D **65**, 042001 (2002).
[9] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[10] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, England, 2009).
[11] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (Springer Science & Business Media, Berlin, 2011).
[12] H. Miao, Phys. Rev. A **95**, 012103 (2017).
[13] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory* (John Wiley & Sons, New York, 2004).
[14] R. D. Gill and B. Y. Levit, Bernoulli **1**, 59 (1995).
[15] M. G. Genoni, M. G. Paris, G. Adesso, H. Nha, P. L. Knight, and M. Kim, Phys. Rev. A **87**, 012107 (2013).
[16] D. Branford, H. Miao, and A. Datta, Phys. Rev. Lett. **121**, 110505 (2018).
[17] T. Gefen, R. Tarafder, R. X. Adhikari, and Y. Chen, Phys. Rev. Lett. **132**, 020801 (2024)..
[18] M. Bradshaw, S. M. Assad, and P. K. Lam, Phys. Lett. A **381**, 2598 (2017).
[19] M. Bradshaw, P. K. Lam, and S. M. Assad, Phys. Rev. A **97**, 012106 (2018).
[20] M. Tsang and C. M. Caves, Phys. Rev. X **2**, 031016 (2012).
[21] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.132.130801 for proofs of the following: Eqs. (10), (11), (13), and (14); the realization of the optimal nonstationary measurements; the results shown in Fig. 4; and the effect of losses.
[22] A. Buonanno, Y. Chen, and N. Mavalvala, Phys. Rev. D **67**, 122005 (2003).
[23] J. Aasi, B. P. Abbott, R. Abbott, T. Abbott, M. R. Abernathy, K. Ackley, C. Adams, T. Adams, P. Addesso *et al.*, Classical Quantum Gravity **32**, 074001 (2015).
[24] A. Buikema, C. Cahillane, G. L. Mansell, C. D. Blair, R. Abbott, C. Adams, R. X. Adhikari, A. Ananyeva, S. Appert, K. Arai *et al.*, Phys. Rev. D **102**, 062003 (2020).
[25] B. J. Meers, Phys. Rev. D **38**, 2317 (1988).
[26] A. Freise and K. Strain, Living Rev. Relativity **13**, 1 (2010).
[27] D. Ganapathy, L. McCuller, J. G. Rollins, E. D. Hall, L. Barsotti, and M. Evans, Phys. Rev. D **103**, 022002 (2021).
[28] R. L. Ward, Ph.D. thesis, California Institute of Technology, 2010.
[29] K. Somiya, Classical Quantum Gravity **29**, 124007 (2012).
[30] O. Miyakawa, R. Ward, R. Adhikari, M. Evans, B. Abbott, R. Bork, D. Busby, J. Heefner, A. Ivanov, M. Smith *et al.*, Phys. Rev. D **74**, 022001 (2006).
[31] A. Buonanno and Y. Chen, Phys. Rev. D **67**, 062002 (2003).
[32] T. Corbitt and N. Mavalvala, J. Opt. B **6**, S675 (2004).
[33] D. Ganapathy, L. McCuller, J. G. Rollins, E. D. Hall, L. Barsotti, and M. Evans, Phys. Rev. D **103**, 022002 (2021).
[34] P. D. Lasky, Pub. Astron. Soc. Aust. **32**, e034 (2015).
[35] L. Baiotti, Prog. Part. Nucl. Phys. **109**, 103714 (2019).
[36] A. Bauswein, N.-U. F. Bastian, D. B. Blaschke, K. Chatziioannou, J. A. Clark, T. Fischer, and M. Oertel, Phys. Rev. Lett. **122**, 061102 (2019).
[37] N. Andersson, Universe **7**, 97 (2021).
[38] M. Shibata, K. Kyutoku, T. Yamamoto, and K. Taniguchi, Phys. Rev. D **79**, 044030 (2009).
[39] C. D. Ott, Classical Quantum Gravity **26**, 063001 (2009).
[40] C. Messenger, K. Takami, S. Gossan, L. Rezzolla, and B. S. Sathyaprakash, Phys. Rev. X **4**, 041004 (2014).
[41] B. P. Abbott *et al.*, Astrophys. J. **875**, 160 (2019).
[42] B. P. Abbott *et al.*, Astrophys. J. Lett. **851**, L16 (2017).
[43] A. F. Brooks, G. Vajente, H. Yamamoto, R. Abbott, C. Adams, R. X. Adhikari, A. Ananyeva, S. Appert, K. Arai, J. S. Areeda *et al.*, Appl. Opt. **60**, 4047 (2021).
[44] M. Evans, S. Gras, P. Fritschel, J. Miller, L. Barsotti, D. Martynov, A. Brooks, D. Coyne, R. Abbott, R. X. Adhikari *et al.*, Phys. Rev. Lett. **114**, 161102 (2015).
[45] L. Barsotti, J. Harms, and R. Schnabel, Rep. Prog. Phys. **82**, 016905 (2019).
[46] H. J. Kimble, Y. Levin, A. B. Matsko, K. S. Thorne, and S. P. Vyatchanin, Phys. Rev. D **65**, 022002 (2001).
[47] K. Topolski, S. D. Tootle, and L. Rezzolla, Astrophys. J. **960**, 86 (2023).
[48] J. Aasi, J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, C. Adams, T. Adams, P. Addesso, R. X. Adhikari *et al.*, Nat. Photonics **7**, 613 (2013).
[49] M. Tse, H. Yu, N. Kijbunchoo, A. Fernandez-Galiana, P. Dupej, L. Barsotti, C. D. Blair, D. D. Brown, S. E. Dwyer, A. Effler *et al.*, Phys. Rev. Lett. **123**, 231107 (2019).
[50] L. McCuller, C. Whittle, D. Ganapathy, K. Komori, M. Tse, A. Fernandez-Galiana, L. Barsotti, P. Fritschel, M. MacInnis, F. Matichard *et al.*, Phys. Rev. Lett. **124**, 171102 (2020).
[51] C. Wang, C. Zhao, X. Li, E. Zhou, H. Miao, Y. Chen, and Y. Ma, Phys. Rev. D **106**, 082002 (2022).
[52] A. Metelmann and A. A. Clerk, Phys. Rev. Lett. **112**, 133904 (2014).
[53] X. Li, M. Goryachev, Y. Ma, M. E. Tobar, C. Zhao, R. X. Adhikari, and Y. Chen, arXiv:2012.00836.
[54] X. Li, J. Smetana, A. S. Ubhi, J. Bentley, Y. Chen, Y. Ma, H. Miao, and D. Martynov, Phys. Rev. D **103**, 122001 (2021).
[55] J. Bentley, H. Nurdin, Y. Chen, X. Li, and H. Miao, arXiv:2211.04016.
[56] J. W. Gardner, M. J. Yap, V. Adya, S. Chua, B. J. J. Slagmolen, and D. E. McClelland, Phys. Rev. D **106**, L041101 (2022).
[57] P. Sikivie, Phys. Rev. Lett. **51**, 1415 (1983).
[58] T. Braine, R. Cervantes, N. Crisosto, N. Du, S. Kimes, L. J. Rosenberg, G. Rybka, J. Yang, D. Bowring, A. S. Chou *et al.*, Phys. Rev. Lett. **124**, 101303 (2020).
[59] M. Malnou, D. A. Palken, B. M. Brubaker, L. R. Vale, G. C. Hilton, and K. W. Lehnert, Phys. Rev. X **9**, 021023 (2019).
[60] D. J. Marsh, Phys. Rep. **643**, 1 (2016).
[61] Y. Chen and S. Kawamura, Phys. Rev. Lett. **96**, 231102 (2006).
[62] Y. Chen, A. Pai, K. Somiya, S. Kawamura, S. Sato, K. Kokeyama, R. L. Ward, K. Goda, and E. E. Mikhailov, Phys. Rev. Lett. **97**, 151103 (2006).
[63] J. W. Gardner, threeeyedFish, 2023, 10.5281/zenodo.10810216.
[64] Wolfram Research, Inc., 2010.
[65] G. Van Rossum and F. L. Drake Jr., *PYTHON Tutorial* (Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands, 1995).
[66] F. Pérez and B. E. Granger, Comput. Sci. Eng. **9**, 21 (2007).

[67] T. Kluyver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier *et al.*, in *Positioning and Power in Academic Publishing: Players, Agents and Agendas* (IOS Press, Amsterdam, 2016), pp. 87–90.

[68] T. E. Oliphant, *A Guide to NumPy* (Trelgol Publishing, USA, 2006).

[69] J. D. Hunter, Comput. Sci. Eng. **9**, 90 (2007).

[70] A. Franzen, 2009, http://www.gwoptics.org/ComponentLibrary/.

Supplemental Material for
Achieving the fundamental quantum limit of linear waveform estimation

James W. Gardner,* Tuvia Gefen,† Simon A. Haine, Joseph J. Hope, and Yanbei Chen
(Dated: October 23, 2023)

In this Supplemental Material, we present the following proofs omitted from the Letter.

HOLEVO CRAMÉR-RAO BOUND

Given a state $\hat{\rho}$ that depends on two real parameters $\vec{s}=(A,B)^{\mathrm{T}}$ and given a weight matrix W (a positive 2-by-2 matrix with $\mathrm{Tr}[\mathrm{W}]=2$), our figure-of-merit is $\Sigma=\mathrm{Tr}[\mathrm{WV}]$ where V is the covariance matrix with respect to $\hat{\rho}$ of the unbiased estimates $\hat{A}$ and $\hat{B}$ of A and B. Without loss of generality, $\mathrm{W}=\mathrm{diag}(2w,2(1-w))$ with $w\in(0,1)$ such that $\Sigma=2w\mathrm{Var}[\hat{A}]+2(1-w)\mathrm{Var}[\hat{B}]$.

The Holevo Cramér-Rao bound (HCRB) Σ_H sets a tight lower bound on this figure-of-merit $\Sigma\geq\Sigma_H$ of [1]

$$\Sigma_H=\min_{\vec{\hat{Y}}\in\mathcal{Y}}(\Sigma_{H,\mathrm{obj}}) \tag{S1}$$

where the objective function is

$$\Sigma_{H,\mathrm{obj}}=\mathrm{Tr}[\mathrm{W}\mathrm{Re}[\mathrm{V}_Y]]+\mathrm{TrAbs}\Big[\sqrt{\mathrm{W}}\mathrm{Im}[\mathrm{V}_Y]\sqrt{\mathrm{W}}\Big] \tag{S2}$$

and $(\mathrm{V}_Y)_{jk}=\mathrm{Tr}\Big[\hat{\rho}\hat{Y}_j\hat{Y}_k\Big]$ is the unsymmeterised covariance matrix with respect to $\hat{\rho}$ of the unbiased error observables $\vec{\hat{Y}}=(\hat{Y}_1,\hat{Y}_2)^{\mathrm{T}}$ which obey the following constraints

$$\mathcal{Y}=\Big\{\vec{\hat{Y}}:\ \hat{Y}_j^\dagger=\hat{Y}_j,\ \mathrm{Tr}\Big[\hat{\rho}\hat{Y}_j\Big]=0,\ \mathrm{Tr}\Big[\hat{Y}_k\partial_{\vec{s}_j}\hat{\rho}\Big]=\delta_{jk}\Big\}. \tag{S3}$$

such that, e.g., $\hat{Y}_1=\hat{A}-A\hat{1}$ and $\hat{Y}_2=\hat{B}-B\hat{1}$. Here, TrAbs of an operator is the sum of the absolute values of its eigenvalues.

Proof of Eq. 10

We calculate the HCRB in Eq. S1 for the vacuum case using the method from Ref. [2]. First, we simplify the problem by using the symplectic transformation discussed in the Letter above Eq. 13. Rearranging Eq. 6,

$$\vec{\hat{q}}=\vec{\hat{q}}^{(0)}+A'(l^{-1}\vec{d}_A)+B'(l^{-1}\vec{d}_B) \tag{S4}$$

where $A'=lA$ and $B'=lB$ and the normalised displacements are $l^{-1}\vec{d}_A$ and $l^{-1}\vec{d}_B$. Then, under the symplectic transformation $\vec{\hat{q}}\mapsto\vec{\hat{X}}$ this becomes $\vec{\hat{X}}=\vec{\hat{X}}^{(0)}+A'\vec{n}_A+B'\vec{n}_B$ where the rotated normalised displacements obey $\vec{n}_A\cdot\vec{\hat{X}}=\hat{X}_1$ and $\vec{n}_B\cdot\vec{\hat{X}}=\mu\hat{P}_1+\sqrt{1-\mu^2}\hat{X}_2$.

In this basis, the state is $\hat{\rho}=|\psi\rangle\langle\psi|$ with $|\psi\rangle=|\alpha_1\rangle\otimes|\alpha_2\rangle$ where $\alpha_1=\frac{1}{\sqrt{2}}(A'+i\mu B')$ and $\alpha_2=\frac{1}{\sqrt{2}}\sqrt{1-\mu^2}B'$. (The displacement operator is $\hat{D}(\alpha)=\exp\big(\alpha\hat{a}^\dagger-\alpha^*\hat{a}\big)$ such that $|\alpha\rangle=\hat{D}(\alpha)|0\rangle$ where $|0\rangle$ is the vacuum state.) We omit a normalisation factor here by assuming that A' and B' are small without loss of generality by a similar argument to that given in Ref. [2]. Then, we linearly approximate the state as

$$|\psi\rangle\approx|\psi_0\rangle+A'|\psi_1\rangle+B'|\psi_2\rangle \tag{S5}$$

where $|\psi_0\rangle=|0\rangle\otimes|0\rangle$, $|\psi_1\rangle=-i\hat{P}_1|0\rangle\otimes|0\rangle$, and

$$|\psi_2\rangle=i\Big(\mu\hat{X}_1-\sqrt{1-\mu^2}\hat{P}_2\Big)|0\rangle\otimes|0\rangle$$

such that

$$\hat{\rho}\approx|\psi_0\rangle\langle\psi_0|+2A'\mathrm{Re}[|\psi_1\rangle\langle\psi_0|]+2B'\mathrm{Re}[|\psi_2\rangle\langle\psi_0|].$$

Here, $|\psi_1\rangle$ and $|\psi_2\rangle$ are unnormalised states.

The QCRB with respect to A' and B' can be calculated to be $\Sigma'_Q=1$ from the symmetric logarithmic derivatives of $\hat{\rho}$ [3]. By a change of variable, this implies that the QCRB with respect to A and B is $\Sigma_Q=l^{-2}$ as stated in the Letter. In the same manner, we first calculate the HCRB Σ'_H with respect to A' and B' and then rescale it to reach the HCRB with respect to A and B as $\Sigma_H=l^{-2}\Sigma'_H=\Sigma_Q\Sigma'_H$.

Using the Gram-Schmidt process, we find an orthonormal basis $\langle e_k|e_j\rangle=\delta_{jk}$ of the $\{|\psi_j\rangle\}_j$ subspace: $|e_0\rangle=|\psi_0\rangle$, $|e_1\rangle=\sqrt{2}|\psi_1\rangle$, and $|e_2\rangle=\sqrt{\frac{2}{1-\mu^2}}(|\psi_2\rangle-i\mu|\psi_1\rangle)$. (In the limit of $\mu=1$, $|e_2\rangle=0$ since the $\{|\psi_j\rangle\}_j$ subspace is only two-dimensional.) Without loss of generality, this subspace contains the support of $\hat{Y}_1$ and $\hat{Y}_2$ in Eq. S1. The constraints in Eq. S3 can then be expanded as

$$\langle e_0|\hat{Y}_j|e_0\rangle=0,\quad j=1,2 \tag{S6}$$

$$2\mathrm{Re}\Big[\langle\psi_k|\hat{Y}_j|e_0\rangle\Big]=\delta_{jk},\quad j,k\in\{1,2\}. \tag{S7}$$

Let $y_{jk}=\langle e_0|\hat{Y}_j|e_k\rangle$ be complex numbers. The constraint in Eq. S7 is then

$$\begin{bmatrix}\frac{1}{\sqrt{2}}&0\\0&\frac{1}{\sqrt{2}}\end{bmatrix}=\begin{bmatrix}\mathrm{Re}[y_{11}]&\sqrt{1-\mu^2}\mathrm{Re}[y_{12}]-\mu\mathrm{Im}[y_{11}]\\\mathrm{Re}[y_{21}]&\sqrt{1-\mu^2}\mathrm{Re}[y_{22}]-\mu\mathrm{Im}[y_{21}]\end{bmatrix}. \tag{S8}$$

* james.gardner@anu.edu.au
† tgefen@caltech.edu

All other independent components of the Hermitian $\hat{Y}_j$ in this basis are zero without loss of generality such that

$$\hat{Y}_1 = \begin{bmatrix} 0 & y_{11} & y_{12} \\ y_{11}^* & 0 & 0 \\ y_{12}^* & 0 & 0 \end{bmatrix}, \quad \hat{Y}_2 = \begin{bmatrix} 0 & y_{21} & y_{22} \\ y_{21}^* & 0 & 0 \\ y_{22}^* & 0 & 0 \end{bmatrix}. \tag{S9}$$

The objective in Eq. S2 is

$$\begin{aligned}\Sigma'_{H,\text{obj}} &= 4\sqrt{w(1-w)}|z| + 2w\Big(|y_{11}|^2 + |y_{12}|^2\Big) \\ &\quad + 2(1-w)\Big(|y_{21}|^2 + |y_{22}|^2\Big)\end{aligned} \tag{S10}$$

where

$$\begin{aligned} z &= y_{21}y_{11}^* + y_{22}y_{12}^* \\ &= \text{Re}[y_{11}]\text{Im}[y_{21}] - \text{Im}[y_{11}]\text{Re}[y_{21}] \\ &\quad + \text{Re}[y_{12}]\text{Im}[y_{22}] - \text{Im}[y_{12}]\text{Re}[y_{22}]. \end{aligned} \tag{S11}$$

By Eq. S8, $\text{Re}[y_{11}] = \frac{1}{\sqrt{2}}$ and $\text{Re}[y_{21}] = 0$. The minimum is achieved with $z = 0$. (Intuitively, the $|z|$ term in Eq. S10 corresponds to the penalty for incompatibility.) This can be achieved with $\text{Re}[y_{12}] = \text{Im}[y_{11}] = 0$ and

$$\begin{aligned} &\text{Re}[y_{22}] = \kappa\cos(\phi), \quad \text{Im}[y_{21}] = \kappa\sin(\phi) \\ &\text{Im}[y_{12}] = \frac{\text{Re}[y_{11}]\text{Im}[y_{21}]}{\text{Re}[y_{22}]} = \frac{1}{\sqrt{2}}\tan(\phi) \end{aligned} \tag{S12}$$

subject to the remaining constraint from Eq. S8 which is $\frac{1}{\sqrt{2}} = \kappa\cos(\nu+\phi)$. Here, $\nu = \arcsin(\mu) \in (0,\pi]$, $\kappa^2 = \text{Re}[y_{22}]^2 + \text{Im}[y_{21}]^2$, and $\phi = \arctan(\text{Im}[y_{21}]/\text{Re}[y_{22}]) \in (0, \frac{\pi}{2}]$. Eq. S10 then becomes

$$\begin{aligned}\Sigma'_{H,\text{obj}} &= 2w\left(\frac{1}{2} + \frac{1}{2}\tan(\phi)^2\right) + 2(1-w)\frac{1}{2\cos(\nu+\phi)^2} \\ &= w\frac{1}{\cos(\phi)^2} + (1-w)\frac{1}{\cos(\nu+\phi)^2}.\end{aligned} \tag{S13}$$

This proves Eq. 10 noting that $\Sigma_H/\Sigma_Q = \Sigma'_H$. This result agrees with numerically optimising the commuting Gaussian measurements.

Proof of Eq. 11

Limit of $\mu = 1$ and any w. — In this limit, the objective function in Eq. S13 becomes

$$\Sigma'_{H,\text{obj}} = \frac{w}{\cos(\phi)^2} + \frac{1-w}{\sin(\phi)^2}. \tag{S14}$$

By the Cauchy-Schwarz inequality, we can find a lower bound on this of

$$\begin{aligned} &1 \cdot \left(\frac{w}{\cos(\phi)^2} + \frac{1-w}{\sin(\phi)^2}\right) \\ &\geq \left(\cos(\phi)\frac{\sqrt{w}}{\cos(\phi)} + \sin(\phi)\frac{\sqrt{1-w}}{\sin(\phi)}\right)^2 \\ &= 1 + 2\sqrt{w(1-w)}. \end{aligned} \tag{S15}$$

Since the Cauchy-Schwarz inequality is tight, we know that this is the optimal value. This proves Eq. 11 in this limit. The HCRB is saturated by the optimal angle $\bar{\phi}$ satisfying

$$\frac{\sqrt{w}}{\cos(\bar{\phi})^2} = \frac{\sqrt{1-w}}{\sin(\bar{\phi})^2} \tag{S16}$$

such that

$$\bar{\phi} = \phi_0 = \arctan\left(\left(\frac{1-w}{w}\right)^{\frac{1}{4}}\right) \tag{S17}$$

or $\bar{\phi} = \pi - \phi_0$.

Limit of $w = \frac{1}{2}$ and any μ. — In this limit, the objective function in Eq. S13 becomes

$$\Sigma'_{H,\text{obj}} = \frac{1}{2\cos(\phi)^2} + \frac{1}{2\cos(\phi+\nu)^2}. \tag{S18}$$

Again, by the Cauchy-Schwarz inequality,

$$\begin{aligned} &\Big(2\cos(\phi)^2 + 2\cos(\phi+\nu)^2\Big)\left(\frac{1}{2\cos(\phi)^2} + \frac{1}{2\cos(\phi+\nu)^2}\right) \\ &\geq \left(\sqrt{2}\cos(\phi)\frac{1}{\sqrt{2}\cos(\phi)} + \sqrt{2}\cos(\phi+\nu)\frac{1}{\sqrt{2}\cos(\phi+\nu)}\right)^2 \\ &= 4. \end{aligned} \tag{S19}$$

such that

$$\Sigma'_{H,\text{obj}} \geq \frac{1}{\cos\left(\frac{\nu}{2}\right)^2} \tag{S20}$$

where we used the fact that, since $\nu \in (0, \frac{\pi}{2}]$,

$$2\cos\left(\frac{\nu}{2}\right)^2 \geq \cos(\phi)^2 + \cos(\phi+\nu)^2. \tag{S21}$$

Since the Cauchy-Schwarz inequality in Eq. S19 and the trigonometric inequality in Eq. S21 are both saturated for $\bar{\phi} = \pi - \frac{\nu}{2}$, the bound on the objective function in Eq. S20 is tight. Furthermore, since $\nu = \arcsin(\mu)$, $\cos\left(\frac{\nu}{2}\right)^2 = \frac{1}{2}\left(1 + \sqrt{1-\mu^2}\right)$ such Eq. 11 is proven in this limit.

The general solution for the optimal angle $\bar{\phi}$ in Eq. 10 for $\mu > 0$ reduces to finding the roots of an order-8 polynomial in $t = \tan\left(\frac{\bar{\phi}}{2}\right)$ by using MATHEMATICA [4]. Since there is no general closed solution for such polynomials, we do not have a closed solution for the HCRB.

OPTIMAL MEASUREMENT SCHEME

Proof of Eq. 13

Here, we prove that the HCRB in Eq. 10 is saturated by measuring $\hat{A}$ and $\hat{B}$ in Eq. 13. It suffices to show that Σ'_H

equals the classical Cramér-Rao bound (CCRB) $\Sigma'_C = \mathrm{Tr}\big[W\mathcal{I}_C^{-1}\big]$ [3] with respect to A' and B' for measurement of $\cos(\bar{\phi})\hat{X}_1 - \sin(\bar{\phi})\hat{P}_2$ and $\cos(\bar{\phi})\hat{X}_2 - \sin(\bar{\phi})\hat{P}_1$. Here, the classical Fisher information matrix $\mathcal{I}_C$ with respect to A' and B' is

$$\mathcal{I}_C = \begin{bmatrix} 2\cos(\bar{\phi})^2 & 0 \\ 0 & 2\Big(\sqrt{1-\mu^2}\cos(\bar{\phi}) - \mu\sin(\bar{\phi})\Big)^2 \end{bmatrix}$$

such that

$$\Sigma'_{\mathrm{C}} = \frac{w}{\cos(\bar{\phi})^2} + \frac{1-w}{\Big(\sqrt{1-\mu^2}\cos(\bar{\phi}) - \mu\sin(\bar{\phi})\Big)^2}. \quad \text{(S22)}$$

By identifying $\cos(\bar{\phi}+\nu)^2$ as the denominator of the second term, therefore, this saturates the HCRB in Eq. 10 since $\bar{\phi}$ was chosen as the optimal angle. This proves Eq. 13.

Proof of the individual realisation

Here, we prove the claim in the Letter that modulated homodyne readout can realise the optimal individual measurement of one of the signal's phases. Let C be the real 2-by-4 matrix of coefficients of the optimal compatible estimates $\hat{A}$ and $\hat{B}$ in Eq. 13 with respect to the original basis of $\vec{\hat{q}}$ such that $(\hat{A},\hat{B})^{\mathrm{T}} = \sqrt{\pi T}\mathrm{C}\vec{\hat{q}}$. To search for an experimental realisation, it is useful to express these estimates in the time domain. Expanding $\vec{\hat{q}}$ in Eq. 6 into the time domain provides that

$$\vec{\hat{q}} = \frac{1}{\sqrt{\pi T}}\int_{-\infty}^{\infty}\mathrm{d}t \begin{bmatrix} \cos(\Omega t)\hat{x}(t) \\ \cos(\Omega t)\hat{p}(t) \\ \sin(\Omega t)\hat{x}(t) \\ \sin(\Omega t)\hat{p}(t) \end{bmatrix} \quad \text{(S23)}$$

such that the estimates are then

$$\begin{bmatrix} \hat{A} \\ \hat{B} \end{bmatrix} = \int_{-\infty}^{\infty}\mathrm{d}t\ \mathrm{C}\begin{bmatrix} \cos(\Omega t)\hat{x}(t) \\ \cos(\Omega t)\hat{p}(t) \\ \sin(\Omega t)\hat{x}(t) \\ \sin(\Omega t)\hat{p}(t) \end{bmatrix}. \quad \text{(S24)}$$

We can show that these correspond to phase and amplitude modulated quadratures. By simplifying Eq. S24,

$$\begin{bmatrix} \hat{A} \\ \hat{B} \end{bmatrix} = \int_{-\infty}^{\infty}\mathrm{d}t \begin{bmatrix} \beta_{11}(t) & \beta_{12}(t) \\ \beta_{21}(t) & \beta_{22}(t) \end{bmatrix}\begin{bmatrix} \hat{x}(t) \\ \hat{p}(t) \end{bmatrix} \quad \text{(S25)}$$

$$= \int_{-\infty}^{\infty}\mathrm{d}t \begin{bmatrix} c_A(t)\hat{x}_{\theta_A(t)}(t) \\ c_B(t)\hat{x}_{\theta_B(t)}(t) \end{bmatrix}. \quad \text{(S26)}$$

Here, the amplitudes c_A and c_B of the time-dependent quadratures are $\vec{c}(t) = (c_A(t), c_B(t))^{\mathrm{T}}$ where, for $j,k = 1,2$, $\vec{c}_j(t) = \sqrt{\beta_{j1}(t)^2 + \beta_{j2}(t)^2}$ for $j,k = 1,2$ and

$$\beta_{jk}(t) = \mathrm{C}_{j,k}\cos(\Omega t) + \mathrm{C}_{j,k+2}\sin(\Omega t).$$

And, the phases θ_A and θ_B are $\vec{\theta}(t) = (\theta_A(t), \theta_B(t))^{\mathrm{T}}$ where $\vec{\theta}_j(t) = \arctan(\beta_{j2}(t)/\beta_{j1}(t))$. (Note that the time dependence is not simply sinusoidal and that these quadratures are defined in the Interaction Frame with respect to $\hat{H}_0$.)

For a sufficiently bright local oscillator $|\alpha(t)| \gg 1$, the difference current $\hat{I}(t)$ from a balanced homodyne readout, as shown in Fig. 3a, satisfies [5]

$$\begin{aligned} \hat{I}(t) &\propto \alpha^*(t)\hat{a}(t) + \alpha(t)\hat{a}^\dagger(t) \\ &= \sqrt{2}|\alpha(t)|\hat{x}_{\arg(\alpha(t))}(t). \end{aligned} \quad \text{(S27)}$$

This means that a phase-modulated local oscillator (with phase $\theta(t) = \arg(\alpha(t))$) can measure a chosen quadrature $\hat{I}(t) \propto \hat{x}_{\theta(t)}(t)$ at each time t such as those in Eq. S26 or, in general, any linear combination of $\vec{\hat{q}}$. This proves the claim in the Letter.

Proof of the joint realisation for $\mu = 1$

Here, we prove the claim in the Letter that for $\mu = 1$ an asymmetric beamsplitter can help realise the optimal simultaneous measurement of the signal's phases. Suppose that we mix the output mode from the device described by $\vec{\hat{q}}$ in Eq. 6 with uncorrelated vacuum $\vec{\hat{q}}^{(0,\mathrm{BS})}$ using an asymmetric beamsplitter with power reflectivity $\cos(\bar{\phi})^2$ as shown in Fig. 3b. The two output modes from the beamsplitter are

$$\begin{aligned} \vec{\hat{q}}^{+} &= \cos(\bar{\phi})\vec{\hat{q}} + \sin(\bar{\phi})\vec{\hat{q}}^{(0,\mathrm{BS})} \\ \vec{\hat{q}}^{-} &= -\sin(\bar{\phi})\vec{\hat{q}} + \cos(\bar{\phi})\vec{\hat{q}}^{(0,\mathrm{BS})}. \end{aligned} \quad \text{(S28)}$$

On these output modes, we can then measure $\vec{d}_A\cdot\vec{\hat{q}}^{+}$ and $\vec{d}_B\cdot\vec{\hat{q}}^{-}$ using a modulated homodyne readout on each beam. These are the analogues of $\hat{A}_{\text{naïve}}$ and $\hat{B}_{\text{naïve}}$, but now they commute due to the ancillary vacuum. The unbiased estimates for A and B, respectively, then are $\cos(\bar{\phi})^{-1}l^{-2}\vec{d}_A\cdot\vec{\hat{q}}^{+}$ and $-\sin(\bar{\phi})^{-1}l^{-2}\vec{d}_B\cdot\vec{\hat{q}}^{-}$. These saturate the HCRB in Eq. 11 for any w. This proves the claim in the Letter. (Note that this realisation works only for $\mu = 1$ because then the denominators in Eq. 10 sum to one which is the energy conservation condition for a beamsplitter.)

Proof of the joint realisation for any μ

Here, we prove the claim in the Letter that for any μ and w a joint homodyne-heterodyne readout scheme can realise the optimal simultaneous measurement of A and B at a given positive frequency Ω. Suppose that we measure $\hat{x}_{\theta_A(t)}(t)$ using modulated homodyne readout and in post processing multiply it by $c_A(t)$ to obtain the timeseries $c_A(t)\hat{x}_{\theta_A(t)}(t)$ for all t. Then, the 2Ω Fourier component of this timeseries is $\int_{-\infty}^{\infty}\mathrm{d}t\ e^{i2\Omega t}c_A(t)\hat{x}_{\theta_A(t)}(t)$. A

linear combination of the real and imaginary parts of this Fourier component with real coefficients λ_1 and λ_2 is then a heterodyne measurement, by Eq. S25, of

$$\hat{q}_{\rm het} = \int_{-\infty}^{\infty} \mathrm{d}t\ (\lambda_1 \cos(2\Omega t) + \lambda_2 \sin(2\Omega t)) \cdot (\beta_{11}(t)\hat{x}(t) + \beta_{12}(t)\hat{p}(t)). \tag{S29}$$

The 2Ω terms and the $\beta_{jk}(t)$ which oscillate at Ω beat to produce components at the difference (Ω) and sum (3Ω) frequencies such that

$$\hat{q}_{\rm het} = \vec{c}_\Omega \cdot \vec{\hat{q}}(\Omega) + \vec{c}_{3\Omega} \cdot \vec{\hat{q}}(3\Omega) \tag{S30}$$

where

$$\vec{c}_\Omega = \frac{\sqrt{\pi T}}{2}\left(\lambda_1 \begin{bmatrix} \mathrm{C}_{11} \\ \mathrm{C}_{12} \\ -\mathrm{C}_{13} \\ -\mathrm{C}_{14} \end{bmatrix} + \lambda_2 \begin{bmatrix} \mathrm{C}_{13} \\ \mathrm{C}_{14} \\ \mathrm{C}_{11} \\ \mathrm{C}_{12} \end{bmatrix}\right) \tag{S31}$$

$$\vec{c}_{3\Omega} = \frac{\sqrt{\pi T}}{2}\left(\lambda_1 \begin{bmatrix} \mathrm{C}_{11} \\ \mathrm{C}_{12} \\ \mathrm{C}_{13} \\ \mathrm{C}_{14} \end{bmatrix} + \lambda_2 \begin{bmatrix} -\mathrm{C}_{13} \\ -\mathrm{C}_{14} \\ \mathrm{C}_{11} \\ \mathrm{C}_{12} \end{bmatrix}\right). \tag{S32}$$

Here, the two vectors of $\vec{c}_\Omega$ in Eq. S31 commute with $\hat{A}$ and are linearly independent of each other and $\hat{A}$. Together with $\hat{A}$, they span the space of all linear combinations of $\vec{\hat{q}}$ that commute with $\hat{A}$. Since $\hat{B}$ commutes with $\hat{A}$, then there always exists real λ_1, λ_2, and λ_3 such that $\hat{B} = \vec{c_\Omega} \cdot \vec{\hat{q}}(\Omega) - \lambda_3 \hat{A}$ and Eq. S30 becomes

$$\hat{q}_{\rm het} = \hat{B} + \lambda_3 \hat{A} + \hat{B}_{\rm het} \tag{S33}$$

where $\hat{B}_{\rm het} = c_{3\vec{\Omega}} \cdot \vec{\hat{q}}(3\Omega)$ is the nonstationary heterodyne noise at 3Ω which commutes with and is statistically independent of $\hat{A}$ and $\hat{B}$. Supposing that this noise can be arbitrarily suppressed, then this shows that measuring $\hat{x}_{\theta_A(t)}(t)$ provides a homodyne estimate of $\hat{A}$ from the DC component and a heterodyne estimate of $\hat{B}$ from the 2Ω component (after subtracting off $\lambda_3\hat{A}$).

It remains to be shown that the noise from $\hat{B}_{\rm het}$ at 3Ω can be squeezed without affecting $\hat{A}$ or $\hat{B}$ at Ω. We have thus far been working in the two-photon formalism [5, 6]. In the one-photon formalism (1PF), however, then $\hat{B}_{\rm het} = \vec{c}_{\rm 1PF} \cdot (\hat{x}'_{\omega_0+3\Omega}, \hat{p}'_{\omega_0+3\Omega}, \hat{x}'_{\omega_0-3\Omega}, \hat{p}'_{\omega_0-3\Omega})^{\rm T}$ where

$$\vec{c}_{\rm 1PF} = \frac{1}{4}\left(\lambda_1 \begin{bmatrix} \mathrm{C}_{11} - \mathrm{C}_{14} \\ \mathrm{C}_{12} + \mathrm{C}_{13} \\ \mathrm{C}_{11} + \mathrm{C}_{14} \\ \mathrm{C}_{12} - \mathrm{C}_{13} \end{bmatrix} - \lambda_2 \begin{bmatrix} \mathrm{C}_{13} + \mathrm{C}_{12} \\ \mathrm{C}_{14} - \mathrm{C}_{11} \\ \mathrm{C}_{13} - \mathrm{C}_{12} \\ \mathrm{C}_{14} + \mathrm{C}_{11} \end{bmatrix}\right)$$

and $\hat{x}'_{\omega_0\pm3\Omega}$ and $\hat{p}'_{\omega_0\pm3\Omega}$ are the canonical Hermitian quadratures at $\omega_0 \pm 3\Omega$ around some carrier frequency ω_0 in the one-photon formalism [5, 6]. To squeeze $\hat{B}_{\rm het}$, therefore, it suffices to squeeze the one-photon quadratures at $\omega_0 \pm 3\Omega$ with the angles of

$$\theta_\pm = \arctan\left(\frac{\lambda_1(\mathrm{C}_{12} \pm \mathrm{C}_{13}) - \lambda_2(\mathrm{C}_{14} \mp \mathrm{C}_{11})}{\lambda_1(\mathrm{C}_{11} \mp \mathrm{C}_{14}) - \lambda_2(\mathrm{C}_{13} \pm \mathrm{C}_{12})}\right). \tag{S34}$$

This could be performed using one-mode squeezing in two detuned cavities, one centered at each of the frequencies $\omega_0 \pm 3\Omega$. Since $\hat{A}$ and $\hat{B}$ are at $\omega_0 \pm \Omega$ in the one-photon formalism, they should be negligibly affected if the two cavities are sufficiently narrowband. In principle, therefore, the joint homodyne-heterodyne measurement of $\hat{A}$ and $\hat{B}$ can saturate the HCRB for any μ and w in the limit of large squeezing and narrowband cavities. This proves the claim in the Letter.

Using the standard Hamiltonian approach, we model the squeezing cavities numerically and confirm this result. Incidentally, assuming that we squeeze $\omega_0 + 3\Omega$ first and then $\omega_0 - 3\Omega$, then we find that it is necessary to adjust the first pump angle by $\mathcal{O}(0.01)$ radians from that inferred from θ_+ in Eq. S34 since there is a slight rotation of the state at $\omega_0 + 3\Omega$ by the detuned cavity centred at $\omega_0 - 3\Omega$. We emphasise that this proposal is lossless and only saturates the HCRB in a narrowband around the chosen frequency Ω. We defer a thorough feasibility study for the general μ case to future work, particularly to address the challenges of achieving narrowband cavities at the table-top scale.

DETUNED GRAVITATIONAL-WAVE INTERFEROMETER

Proof of Eq. 14

We model the LIGO-like interferometer as a detuned cavity as described in the Letter. Let the annihilation operator of the intracavity mode at frequency ω_0 be $\hat{a}_{\rm cav}$ and the continuum of bath modes of the bath modes be $\hat{a}_{\rm bath}(\omega)$ indexed by their frequency ω. These bath modes define the incoming and outgoing modes from the dark port of the interferometer with annihilation operators $\hat{a}_{\rm in}$ and $\hat{a}$ (often called $\hat{a}_{\rm out}$), respectively [7]. The Hamiltonian $\hat{H}$ describing the detuned interferometer is $\hat{H} = \hat{H}_0 + \hat{H}_{\rm int}(t)$ where $\hat{H}_{\rm int}(t) = gs(t)\hat{x}_{\rm cav}$ and

$$\hat{H}_0 = (\omega_0 - \Delta)\hat{a}^\dagger_{\rm cav}\hat{a}_{\rm cav} + \sqrt{2\gamma}\int_{-\infty}^{\infty} \mathrm{d}\omega \left(i\hat{a}^\dagger_{\rm bath}(\omega)\hat{a}_{\rm cav} + \mathrm{H.c.}\right) \tag{S35}$$

where the second term in $\hat{H}_0$ implies the input/output relation $\hat{x}_\theta = -\hat{x}_{{\rm in},\theta} + \sqrt{2\gamma}\hat{x}_{{\rm cav},\theta}$. (The free evolution of the bath modes is also included in $\hat{H}_0$ but we omit it for brevity.) Here, $g^2 = c^{-1}2\omega_0 LP$ where c is the speed of light, L is the arm cavities' length, and P is the circulating power [8]; $\gamma = -c/(4L)\log(1-T)$ is the cavity half-width at half-max bandwidth where T is the input test masses' power transmissivity; and H.c. is the Hermitian conjugate. We emphasise that, in this effective model, the test masses are ignored.

In the Interaction Frame with respect to $\hat{H}_0$, the

Heisenberg-Langevin equations of motion are

$$\begin{aligned}\partial_t \hat{x}_{\rm cav} &= -\gamma \hat{x}_{\rm cav} + \sqrt{2\gamma}\hat{x}_{\rm in} - \Delta \hat{p}_{\rm cav} \\ \partial_t \hat{p}_{\rm cav} &= -\gamma \hat{p}_{\rm cav} + \sqrt{2\gamma}\hat{p}_{\rm in} + \Delta \hat{x}_{\rm cav} - g s(t).\end{aligned} \tag{S36}$$

In the two-photon formalism, Eq. S36 is

$$\begin{aligned}(\gamma - i\Omega)\hat{x}_{\rm cav}(\Omega) &= \sqrt{2\gamma}\hat{x}_{\rm in}(\Omega) - \Delta \hat{p}_{\rm cav}(\Omega) \\ (\gamma - i\Omega)\hat{p}_{\rm cav}(\Omega) &= \sqrt{2\gamma}\hat{p}_{\rm in}(\Omega) + \Delta \hat{x}_{\rm cav}(\Omega) - g\tilde{s}(\Omega)\end{aligned} \tag{S37}$$

Then, Eq. S37 can be linearly solved and the input/output relation used to find that

$$\begin{bmatrix} \hat{x}(\Omega) \\ \hat{p}(\Omega) \end{bmatrix} = \begin{bmatrix} \hat{x}^{(0)}(\Omega) \\ \hat{p}^{(0)}(\Omega) \end{bmatrix} + \vec{\chi}\tilde{s}(\Omega) \tag{S38}$$

such that the susceptibility $\vec{\chi}$ is in Eq. 14 and where the noise is

$$\begin{bmatrix} \hat{x}^{(0)}(\Omega) \\ \hat{p}^{(0)}(\Omega) \end{bmatrix} = \mathrm{M}_{\rm in} \begin{bmatrix} \hat{x}_{\rm in}(\Omega) \\ \hat{p}_{\rm in}(\Omega) \end{bmatrix} \tag{S39}$$

$$\begin{aligned}\mathrm{M}_{\rm in} = &\frac{1}{\Delta^2 + (\gamma - i\Omega)^2} \\ &\cdot \begin{bmatrix} \gamma^2 - \Delta^2 + \Omega^2 & -2\gamma\Delta \\ 2\gamma\Delta & \gamma^2 - \Delta^2 + \Omega^2 \end{bmatrix}\end{aligned} \tag{S40}$$

such that if the incoming mode is vacuum, then the free outgoing mode is also at vacuum.

Proof of results shown in Fig. 4

Here, we prove that for equal weights $w = 0.5$ variational readout is the optimal measurement. By Eq. 8 and Eq. 14, the norm of the displacements is

$$l = \sqrt{2\gamma\pi T}g\sqrt{\frac{(\gamma^2 + \Delta^2 + \Omega^2)}{(\gamma^2 + (\Delta - \Omega)^2)(\gamma^2 + (\Delta + \Omega)^2)}}. \tag{S41}$$

By Eq. 11, then the HCRB for the detuned interferometer with $w = 0.5$ is

$$\Sigma_H = \frac{1}{4g^2\pi T}\frac{\zeta}{\gamma\Delta^2\Omega^2}\left(\gamma^2 + \Delta^2 + \Omega^2 - \sqrt{\zeta}\right)$$

where

$$\zeta = \left(\gamma^2 + (\Delta - \Omega)^2\right)\left(\gamma^2 + (\Delta + \Omega)^2\right). \tag{S42}$$

At the detuning frequency $\Omega = \Delta$, therefore, the gap from the HCRB to the QCRB is $\Sigma_H/\Sigma_Q = \frac{2}{1+G}$ where $G = \frac{\gamma\sqrt{\gamma^2+4\Delta^2}}{\gamma^2+2\Delta^2} \in (0, 1)$. Meanwhile, by Eq. 3, the waveform-estimation QCRB (which is equivalent to Σ_Q up to factors of πT) is

$$S_Q^{\rm wave}(\Omega) = \frac{1}{2g^2}\frac{\left(\gamma^2 + (\Delta - \Omega)^2\right)\left(\gamma^2 + (\Delta + \Omega)^2\right)}{\gamma(\gamma^2 + \Delta^2 + \Omega^2)}. \tag{S43}$$

And, by using MATHEMATICA [4], the optimal stationary measurement satisfies

$$\frac{\min_\theta S_{ss}(\Omega)}{S_Q^{\rm wave}(\Omega)} = \frac{2}{1+G}. \tag{S44}$$

Variational readout, therefore, saturates the HCRB at $\Omega = \Delta$ for $w = 0.5$. Numerically, this also holds at the other frequencies shown in Fig. 4. This proves the claim.

Proof of the effect of losses

Here, we prove the claim in the Letter that the gap from the HCRB to the QCRB is unchanged with loss. It suffices to show that $l \mapsto \sqrt{1-\eta}l$ and $\mu \mapsto \frac{2\Delta\Omega}{(\gamma+\gamma_l)^2+\Delta^2+\Omega^2}$ by the argument in the Letter. We assume that the losses can be modelled as beamsplitters into uncorrelated vacuum. We model the intracavity loss γ_l by adding additional Langevin terms to the frequency domain quadratures in Eq. S37 such that (here, we omit the frequency Ω dependence for brevity)

$$\begin{aligned}(\gamma + \gamma_l - i\Omega)\hat{x}_{\rm cav} &= \sqrt{2\gamma}\hat{x}_{\rm in} + \sqrt{2\gamma_l}\hat{x}_{\rm loss} - \Delta\hat{p}_{\rm cav} \\ (\gamma + \gamma_l - i\Omega)\hat{p}_{\rm cav} &= \sqrt{2\gamma}\hat{p}_{\rm in} + \sqrt{2\gamma_l}\hat{p}_{\rm loss} + \Delta\hat{x}_{\rm cav} - g\tilde{s}\end{aligned} \tag{S45}$$

and the input/output relation remains the same. The resulting noise remains at vacuum, but the susceptibility $\vec{\chi}$ becomes

$$\vec{\chi} = \frac{\sqrt{2\gamma}g}{\Delta^2 + (\gamma + \gamma_l - i\Omega)^2}\begin{bmatrix} \Delta \\ -\gamma - \gamma_l + i\Omega \end{bmatrix} \tag{S46}$$

such that $\mu \mapsto \frac{2\Delta\Omega}{(\gamma+\gamma_l)^2+\Delta^2+\Omega^2}$ by Eq. 9. Then, we also include a loss of $\eta = 0.1$ on the output beam but before the measurement. Since the noise is at vacuum, this simply degrades the signal as $\vec{\chi} \mapsto \sqrt{1-\eta}\vec{\chi}$ such that $l \mapsto \sqrt{1-\eta}l$ by Eq. 8 and μ is unchanged by Eq. 9. This proves the claim.

[1] A. S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory.* Springer Science & Business Media, 2011.

[2] M. Bradshaw, S. M. Assad, and P. K. Lam. 2017. *Phys. Lett. A*, 381(32):2598–2607.

[3] H. M. Wiseman and G. J. Milburn. *Quantum Measurement and Control.* Cambridge University Press, 2009.

[4] Wolfram Research, Inc. 2010.

[5] S. L. Danilishin and F. Y. Khalili. 2012. *Living Rev. Relativ.*, 15(1):5.

[6] C. M. Caves and B. L. Schumaker. 1985. *Phys. Rev. A*, 31(5):3068.

[7] C. W. Gardiner and M. J. Collett. 1985. *Phys. Rev. A*, 31:3761–3774.

[8] X. Li, M. Goryachev, Y. Ma, M. E. Tobar, C. Zhao, R. X. Adhikari, and Y. Chen. 2020. *arXiv:2012.00836.*

Chapter 4

Stochastic waveform estimation at the fundamental quantum limit

We are such stuff
As dreams are made on; and our little life
Is rounded with a sleep.

Prospero (Act 4, Scene 1)
William Shakespeare, The Tempest

We now present Ref. [50] which continues the story of quantum-limited waveform sensing in the asymptotic Fisher limit like the previous paper, but now examines a Markovian stochastic signal rather than the deterministic signal above. In this case, we want to estimate the power spectral density of a Markovian random process rather than the Fourier components of a definite function of time. Moreover, we optimise not just over the measurement scheme but now also over the initial state and find that both should be highly non-Gaussian to achieve the maximum loss-limited sensitivity. In the overall story of the thesis, this is a radical change from the linear Gaussian sensing considered thus far. This further emphasises the different way of thinking about, e.g., operating an interferometer like LIGO. In particular, what state to prepare the incoming bath of modes in at the dark port of the interferometer and what measurements to perform on the outgoing bath of modes. The historical motivation for this work was accelerating searches for holographic or "geontropic" quantum gravity which indicates broadening interest as our story changes from just gravitational-wave detection to other fundamental physics searches.

The summaries of key results on the first page of the paper and in Sec. VII B are already comprehensive so we will only mention here a few points which connect to the rest of the thesis. Firstly, the toy model in Sec. I B is the same as the previous paper which illustrates its continued usefulness. Secondly, the existing result from the literature discussed in Sec. IV A that squeezing and homodyne quadrature measurement is optimal for deterministic sensing (with a tuned interferometer) is important context for understanding that LIGO's present operation and science target cannot be improved through non-Gaussian protocols, but that the science target of stochastic signals can which this paper shows (see Sec. VI B 2). Thirdly, in Sec. V we do not need to use the Holevo Cramér-Rao bound for multiparameter estimation because the weak commutativity condition in Eq. 34 holds here unlike in the previous paper. Finally, the connection between this stochastic waveform estimation problem and the general case of Lindblad estimation in Appendix F was the motivation for the paper in the next chapter.

Stochastic Waveform Estimation at the Fundamental Quantum Limit

James W. Gardner,[1,2,*] Tuvia Gefen,[3,†] Simon A. Haine,[4] Joseph J. Hope,[4] John Preskill,[2,3] Yanbei Chen,[2] and Lee McCuller[3,5]

[1] *OzGrav-ANU, Centre for Gravitational Astrophysics, Research Schools of Physics, and of Astronomy and Astrophysics, The Australian National University, Canberra, Australian Capital Territory 2601, Australia*

[2] *Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, California 91125, USA*

[3] *Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*

[4] *Department of Quantum Science and Technology and Department of Fundamental and Theoretical Physics, Research School of Physics, The Australian National University, Canberra, Australian Capital Territory 0200, Australia*

[5] *Division of Physics, Mathematics and Astronomy, California Institute of Technology, Pasadena, California 91125, USA*

(Received 1 May 2024; revised 9 May 2025; accepted 12 June 2025; published 22 July 2025)

Although measuring the deterministic waveform of a weak classical force is a well-studied problem, estimating a random waveform, such as the spectral density of a stochastic signal field, is much less well understood despite it being a widespread task at the frontier of experimental physics. State-of-the-art precision sensors of random forces must account for the underlying quantum nature of the measurement but the optimal quantum protocol for interrogating such linear sensors is not known. We derive the fundamental precision limit: the extended-channel quantum Cramér-Rao bound. In the experimentally relevant regime in which losses dominate, we prove that non-Gaussian-state preparation and measurement are required to achieve this fundamental limit and we determine numerically the optimal non-Gaussian protocol. We discuss how this scheme could accelerate searches for signatures of quantum gravity, stochastic gravitational waves, and axionic dark matter.

DOI: 10.1103/h91r-4ws9

I. INTRODUCTION

The estimation of the spectral density of a classical process is a ubiquitous task in continuous-variable quantum systems. Examples include searching for excess noise in optical interferometers due to quantum gravity [1–5], probing stochastic gravitational waves with the global network of gravitational-wave observatories [6–10], and hunting for axionic dark matter with microwave cavities [11–14]. The remarkable sensitivity of contemporary devices demands that we contend with quantum noise, the fundamental uncertainty in the state of the device arising from the Heisenberg uncertainty principle. The stochastic signal of interest must be distinguished from the natural fluctuations arising from the measurement of the device.

This task, sometimes called "noise spectroscopy," can be expressed in the language of quantum metrology, the study of estimating parameters encoded in quantum states. We consider a quantum device coupled to a stochastic signal, a classical continuous random variable $y(t)$ at times t. We assume this signal to be a Gaussian stationary random process; assuming that the signal is Gaussian is valid provided that the central limit theorem applies, e.g., that we repeat the experiment many times. Assuming that the signal is a stationary random process allows us to study its power spectral density $S_{yy}(\Omega)$ at each positive frequency Ω. Our goal is to estimate the continuum values of $S_{yy}(\Omega)$ by measuring the state of the quantum device. We want our estimate at each Ω to be unbiased and have the minimal mean-square error (MSE). We will show that this resembles the problem of estimating the uncertainty of a single bosonic mode and so we treat that case extensively.

Much attention has previously been dedicated to studying the fundamental precision limits of estimating the

*Contact author: james.gardner@anu.edu.au
†Contact author: tgefen@caltech.edu

mean values of the signal for various protocols and quantum devices (see, e.g., Refs. [15–18]). Here, we focus instead on estimating the spectral density $S_{yy}(\Omega)$ of the signal rather than its mean value. This is a much less studied problem. It has recently been shown that this scenario is fundamentally different from mean-value estimation in that non-Gaussian measurement techniques are required to obtain optimal estimates [19]. There are still many critical open questions, however, such as the effects of imperfections and the preparation of different initial quantum states.

To guide the reader, we summarize the main results of our work as follows:

(1) In Sec. II, we reduce the noise-estimation problem for a linear quantum device at a fixed frequency Ω to the study of a harmonic oscillator undergoing a random-displacement channel. The task is to estimate the standard deviation σ of the Gaussian distribution of random displacements and, in some cases, to simultaneously estimate the mean displacement μ.

(2) In Sec. III, we review the prior literature on estimating σ in the ideal lossless limit, using an initial vacuum state. A number-resolving measurement, e.g., photon counting, is optimal [19,20]. In contrast, Gaussian measurements, such as homodyne detection, suffer the "Rayleigh curse" when the signal σ is small. The goal of our work is to find the optimal initial state and measurement scheme that achieves the ultimate limit on precision, the extended-channel quantum Fisher information (ECQFI), in the presence of loss. We also consider the simultaneous estimation of μ and σ.

(3) In Sec. IV, we discuss imperfections. We focus on the case of a loss η occurring before the signal is encoded by the random-displacement channel. We assume negligible loss after the encoding and negligible additive classical noise, σ_C. Experimentally, the relevant regime is that of small signals $\sigma^2 \ll 1/2$ and high loss $\eta \gg \sigma^2$.

(4) In Sec. V, we find the optimal protocol for estimating σ. In the lossless case, in which the initial state obeys the energy constraint $\langle \hat{n} \rangle = N$, we show that the ECQFI is saturated by preparing an initial single-mode squeezed vacuum (SMSV) and then, after the encoding, antisqueezing and performing a number measurement. In the lossy case, we show that preparing a two-mode squeezed vacuum (TMSV) state is optimal but only if negligible loss occurs on the ancilla mode. In the experimentally relevant regime of high loss on all modes, numerics indicate that the ECQFI is attainable, without an ancilla, using highly non-Gaussian states and measurements.

(5) In Sec. VI, we discuss how to simultaneously estimate μ and σ. We assume that the initial state is vacuum (which may not be optimal). For separable measurements on M copies of the final state, we construct an adaptive measurement scheme in which we learn about μ via quadrature measurement and then use that information to displace back to the origin and learn about σ via number measurement. To saturate the quantum Fisher information (QFI) for the fixed-vacuum input state, however, a collective measurement on the M modes is required.

(6) Finally, in Sec. VII, we discuss how our results concerning single-mode channel estimation can be leveraged for estimating the power spectral density of a continuously varying signal, we propose experimental implementations to realize the optimal estimation protocols, and we apply our results to searches for signatures of quantum gravity, stochastic gravitational waves, and axionic dark matter.

II. NOISE ESTIMATION OF A LINEAR QUANTUM DEVICE

We now consider estimating $S_{yy}(\Omega)$ at each frequency Ω from measurements of the quantum device. For a linear device, we show that this is equivalent to estimating the random displacement of a harmonic oscillator.

A. Review of linear quantum devices

We consider a quantum device that responds linearly to the continuous stochastic signal $y(t)$ and is coupled to an environment of incoming and outgoing bosonic modes as shown in Fig. 1(a). This is a valid model for many relevant quantum systems. For example, the device could be a gravitational-wave interferometer coupled to a stochastic gravitational-wave signal and to the light entering and exiting the differential port of the Michelson interferometer. In the input-output formalism [21], let the annihilation operator of the outgoing bosonic mode be $\hat{a}(t)$,

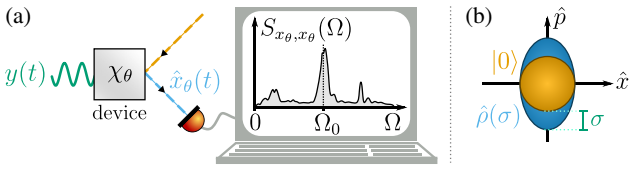

FIG. 1. (a) The spectral density of the outgoing bosonic mode from a linear quantum device coupled to a classical random process $y(t)$, e.g., a gravitational-wave observatory coupled to a stochastic gravitational wave. (b) The phase-plane representation of the analogous single-parameter problem at a particular frequency Ω_0. Given vacuum input $|0\rangle$ with covariance matrix $\mathrm{diag}\left(\frac{1}{2}, \frac{1}{2}\right)$, the final quantum state $\hat{\rho}(\sigma)$ is a squeezed thermal state with covariance matrix $\mathrm{diag}\left(\frac{1}{2}, \frac{1}{2} + \sigma^2\right)$.

satisfying $[\hat{a}(t), \hat{a}^\dagger(t')] = \delta(t-t')$. The canonical quadratures of this mode are $\hat{x}(t) := \frac{1}{\sqrt{2}}[\hat{a}(t) + \hat{a}^\dagger(t)]$ and $\hat{p}(t) := \frac{1}{\sqrt{2}}[-i\hat{a}(t) + i\hat{a}^\dagger(t)]$, which obey $[\hat{x}(t), \hat{p}(t')] = i\delta(t-t')$. The generalized quadrature at angle θ is then defined as $\hat{x}_\theta(t) := \cos(\theta)\hat{x}(t) + \sin(\theta)\hat{p}(t)$ such that $\hat{x}_0(t) = \hat{x}(t)$ and $\hat{x}_{\pi/2}(t) = \hat{p}(t)$. Since the device is linear, these quadratures must obey

$$\hat{x}_\theta(t) = \hat{x}_\theta^{(0)}(t) + \int_{-\infty}^{\infty} \mathrm{d}t'\ \chi_\theta(t-t')y(t'), \tag{1}$$

where $\hat{x}_\theta^{(0)}(t)$ is the free solution (i.e., when no signal is present) and χ_θ is the linear susceptibility of the device [22,23]. In the frequency domain, all of the positive frequencies are independent since the device is linear, such that Eq. (1) becomes

$$\hat{x}_\theta(\Omega) = \hat{x}_\theta^{(0)}(\Omega) + \chi_\theta(\Omega)y(\Omega), \tag{2}$$

where we define the Fourier transform as $y(\Omega) := \int_{-\infty}^{\infty} \mathrm{d}t\, e^{i\Omega t} y(t)$. The frequency-domain quadratures are non-Hermitian but obey a conjugate symmetry $\hat{x}_\theta^\dagger(\Omega) = \hat{x}_\theta(-\Omega)$ and have the commutator $[\hat{x}(\Omega), \hat{p}^\dagger(\Omega')] = i2\pi\delta(\Omega - \Omega')$. Suppose that we measure $\hat{x}_\theta(t)$; then Eq. (2) implies that the observed output-referred power spectral density is

$$S_{x_\theta x_\theta}(\Omega) = S_{x_\theta x_\theta}^{(0)}(\Omega) + |\chi_\theta(\Omega)|^2 S_{yy}(\Omega), \tag{3}$$

where the total output-referred free noise $S_{x_\theta x_\theta}^{(0)}(\Omega)$ includes the quantum noise of the device, which we assume follows a stationary random process. In Eq. (3), the (one-sided) power spectral density $S_{x_\theta x_\theta}(\Omega)$ of the quadratures, which are stationary random processes in time, is defined as follows [24]

$$2\pi\delta(\Omega - \Omega')S_{x_\theta x_\theta}(\Omega) := \langle\{\hat{x}_\theta(\Omega), \hat{x}_\theta^\dagger(\Omega')\}\rangle, \tag{4}$$

where $\{\cdot,\cdot\}$ is the anticommutator and we assume that the quadratures have zero mean. For example, the output-referred power spectral density of the quantum shot noise of the vacuum state is $S_{x_\theta x_\theta}^{(0)}(\Omega) = 1$. Meanwhile, in Eq. (3), the (one-sided) power spectral density of the signal at each frequency is defined as [25]

$$S_{yy}(\Omega) = \lim_{T\to\infty} \frac{2}{T} \mathrm{E}\left[|y_T(\Omega)|^2\right], \tag{5}$$

where $y_T(\Omega) = \int_{-\infty}^{\infty} \mathrm{d}t\, e^{i\Omega t} y_T(t) = \int_0^T \mathrm{d}t\, e^{i\Omega t} y(t)$ is the Fourier transform of the signal $y_T(t) = [\Theta(t) - \Theta(t-T)]y(t)$ windowed to the finite time interval $(0, T)$ for some integration time T and Θ is the Heaviside function. Here, we take the expectation value over different realizations of the continuous random process describing the signal $y(t)$, which we assume has zero mean. Whenever we compute an expectation value with respect to the quantum state, e.g., $\langle\cdot\rangle$ in Eq. (4), we implicitly marginalize over the signal as well. We reserve the notation $\mathrm{E}[\cdot]$ for when we are only taking an expectation value with respect to the signal.

Let us mention a couple of well-known properties of the power spectral density $S_{yy}(\Omega)$. First, the average power P of the signal $y(t)$ over time, i.e., the variance $\mathrm{Var}[y]$ of the continuous random process, is given by

$$P = \frac{1}{2\pi} \int_0^\infty \mathrm{d}\Omega\ S_{yy}(\Omega). \tag{6}$$

And, second, provided that the signal $y(t)$ is ergodic, the power spectral density $S_{yy}(\Omega)$ equals the Fourier transform of the autocorrelation function $C_{yy}(t) = \lim_{T\to\infty}(1/T) \int_{-\infty}^{\infty} \mathrm{d}t'\ y_T(t'-t)y_T(t')$ by the Wiener-Khinchin theorem. These useful properties motivate why the power spectral density is an important quantity to estimate in an experiment.

Returning to the problem at hand, by using Eqs. (4)–(5), we can derive Eq. (3) from Eq. (2) [26]. We have thus far made only the following assumptions about the system: linearity, stationarity, and zero mean. We now make a further assumption about the tuning of the measurement device.

Some of our most sensitive devices of interest are optical interferometers for which, if the resonant optical cavities are all tuned and the Michelson interferometer is held at total destructive interference at the differential port, then a stochastic signal will only appear in one quadrature of the outgoing optical mode [27]. As such, we now restrict our attention to estimating the excess noise in one known quadrature of the outgoing mode. We will later generalize to sensing isotropic excess noise in both quadratures. Without further loss of generality, we assume that the stochastic signal only appears in the $\hat{p}(t)$ quadrature such that Eq. (2) implies that $\hat{x}(\Omega) = \hat{x}^{(0)}(\Omega)$ since $\chi_0 = 0$ and

$$\hat{p}(\Omega) = \hat{p}^{(0)}(\Omega) + \chi_{\frac{\pi}{2}}(\Omega)y(\Omega). \tag{7}$$

The output-referred power spectral density in Eq. (3) observed from measuring $\hat{p}(t)$ is then

$$S_{pp}(\Omega) = S_{pp}^{(0)}(\Omega) + \left|\chi_{\frac{\pi}{2}}(\Omega)\right|^2 S_{yy}(\Omega). \tag{8}$$

We consider running the experiment for a fixed finite integration time T that is long enough such that Eq. (5) implies that $S_{yy}(\Omega) \approx (2/T)\mathrm{E}[|y_T(\Omega)|^2]$ and Eq. (8) becomes

$$S_{pp}(\Omega) \approx S_{pp}^{(0)}(\Omega) + \frac{2}{T}\mathrm{E}\left[\left|\chi_{\frac{\pi}{2}}(\Omega)y_T(\Omega)\right|^2\right]. \tag{9}$$

We want to estimate the signal power $S_{yy}(\Omega)$ at each frequency Ω from measurements of the outgoing mode. We

have up to now considered continuous quadrature measurements. To understand other possible measurements, it is useful to first simplify this continuous-estimation problem.

B. Review of the deterministic case

Let us first review the deterministic case to establish an analogy to a toy model: continuum linear waveform estimation resembles estimating the displacement of a harmonic oscillator at each frequency. In the deterministic case, the signal $y(t)$ is a real-valued function and we want to estimate its Fourier component $y(\Omega)$, a complex number, at each frequency Ω. It is well established that this continuous-estimation problem reduces to a continuum of independent estimation problems, one for each frequency, and the fundamental quantum limits of this task are well understood [15–18]. We will discuss these sensitivity limits later, in Sec. V B. At a given frequency Ω, the canonical deterministic estimation problem in Eq. (7) is to estimate the displacement of $\hat{p}(\Omega)$ by $\chi_{\pi/2}(\Omega)y(\Omega)$ in the presence of noise from $\hat{p}^{(0)}(\Omega)$. Depending on the parametrization of $y(t)$, there are two independent real parameters of $y(\Omega)$ to estimate in general, e.g., its real and imaginary parts, or its amplitude and complex phase. The displacement-estimation problems at different frequencies are independent due to the linearity of the device: the covariance matrix of the quadratures at different frequencies is diagonal. Therefore, continuum-waveform estimation corresponds to estimating the displacements of a continuum of independent harmonic oscillators, each of which can be dealt with separately. The only assumptions that we have made to reach this conclusion are that the device is linear, stationary, zero-mean, and tuned.

C. Canonical noise-estimation problem

In the stochastic case, the signal $y(t)$ is now a continuous real-valued random process with variance given by Eq. (6). At each frequency Ω, we want to estimate $S_{yy}(\Omega)$, the contribution to the average power from that frequency component. Similarly to the deterministic case above, the continuum problem of estimating $S_{yy}(\Omega)$ at each frequency Ω is equivalent to a continuum of independent single-parameter harmonic oscillator estimation problems. In particular, the harmonic oscillators at different frequencies are independent due to linearity and at each frequency the signal appears as a displacement of the harmonic oscillator. Here, however, the displacement of $\hat{p}(\Omega)$ by $\chi_{\pi/2}(\Omega)y(\Omega)$ in Eq. (7) is stochastic, since $y(\Omega)$ is now a complex random variable. In particular, although the absolute value of $y(\Omega)$ is constrained by Eq. (5) in the limit of large T, its complex phase is uniformly distributed on $(0, 2\pi)$ since the signal $y(t)$ is zero-mean and stationary. Thus, we need to determine how the power spectral density $S_{yy}(\Omega)$ relates to the variance of the probability distribution describing these random displacements.

Let us first address a technicality about the frequency-domain quadratures. At a given frequency Ω, the quadratures $\hat{x}(\Omega)$ and $\hat{p}(\Omega)$ are non-Hermitian and comprise a total of four independent real degrees of freedom, whereas a harmonic oscillator has only two independent real degrees of freedom. This can be seen more clearly by decomposing the frequency-domain quadratures into their real and imaginary parts, which we label as $\hat{q}_j(\Omega)$ for $j = 1, 2, 3, 4$ and define as follows [28]:

$$\vec{\hat{q}}(\Omega) = \begin{bmatrix} \hat{q}_1(\Omega) \\ \hat{q}_2(\Omega) \\ \hat{q}_3(\Omega) \\ \hat{q}_4(\Omega) \end{bmatrix} = \sqrt{\frac{2}{T}} \begin{bmatrix} \mathrm{Re}[\hat{x}(\Omega)] \\ \mathrm{Re}[\hat{p}(\Omega)] \\ \mathrm{Im}[\hat{x}(\Omega)] \\ \mathrm{Im}[\hat{p}(\Omega)] \end{bmatrix}. \tag{10}$$

For a finite integration time T, this is equivalent to decomposing the outgoing mode into its cosine and sine phases at frequency Ω, since

$$\vec{\hat{q}}(\Omega) = \sqrt{\frac{2}{T}} \int_0^T \mathrm{d}t \begin{bmatrix} \cos(\Omega t)\hat{x}(t) \\ \cos(\Omega t)\hat{p}(t) \\ \sin(\Omega t)\hat{x}(t) \\ \sin(\Omega t)\hat{p}(t) \end{bmatrix}.$$

This implies that the commutator between $\hat{q}_1(\Omega)$ and $\hat{q}_2(\Omega')$ at two different frequencies Ω and Ω' equals

$$[\hat{q}_1(\Omega), \hat{q}_2(\Omega')] = i\left(\frac{\sin([\Omega - \Omega']T)}{[\Omega - \Omega']T} + \frac{\sin([\Omega + \Omega']T)}{[\Omega + \Omega']T}\right).$$

The second term vanishes, along with all other terms of size $\mathcal{O}(1/[\Omega T])$ henceforth, by assuming that we integrate for a long enough time $T \gg 1/\Omega$ given the lowest frequency of interest Ω. For frequencies that are further apart than the finite frequency resolution of $1/T$, i.e., that satisfy $[\Omega - \Omega']T \gg 1$, the first term also vanishes. However, for unresolved frequencies satisfying $[\Omega - \Omega']T \ll 1$ this commutator is approximately i, since the modes $\cos(\Omega t)$ and $\cos(\Omega' t)$ are not orthogonal on the finite interval $(0, T)$. We thus assume a discrete frequency sampling henceforth that is coarse enough, i.e., satisfying $[\Omega - \Omega']T \gg 1$, such that the harmonic oscillators at different frequencies commute. The number of such frequency bins is proportional to T and we will soon show that the noise and signal within a given bin are independent of T. This means that the total error in estimating, e.g., the power spectral density of a white-noise process will scale as $1/\sqrt{T}$ in amplitude units as expected.

We now focus on solving the estimation problem at a fixed frequency Ω. The four components of $\vec{\hat{q}}(\Omega)$ are Hermitian observables that resemble the quadratures of two harmonic oscillators, since $[\hat{q}_1(\Omega), \hat{q}_2(\Omega)] =$

$[\hat{q}_3(\Omega), \hat{q}_4(\Omega)] = i$, and all other commutators are zero. Since we assume that the signal and noise are zero-mean and stationary, the 4×4 covariance matrix, $\frac{1}{2}\langle\{\hat{q}_j(\Omega), \hat{q}_k(\Omega)\}\rangle$, is block diagonal such that the two harmonic oscillators are independent. For example, the covariance matrix for the vacuum state is $\frac{1}{2}\delta_{jk}$. By Eq. (7), these real and imaginary parts depend on the signal as follows:

$$\vec{\hat{q}}(\Omega) = \vec{\hat{q}}^{\,(0)}(\Omega) + \sqrt{\frac{2}{T}} \begin{bmatrix} 0 \\ \mathrm{Re}\left[\chi_{\frac{\pi}{2}}(\Omega) y(\Omega)\right] \\ 0 \\ \mathrm{Im}\left[\chi_{\frac{\pi}{2}}(\Omega) y(\Omega)\right] \end{bmatrix}. \tag{11}$$

The variances of the real and imaginary parts containing the zero-mean signal are then [29]

$$\begin{aligned} \mathrm{Var}\left[\hat{q}_2(\Omega)\right] &= \mathrm{Var}[\hat{q}_2^{(0)}(\Omega)] + \frac{2}{T}\mathrm{E}\left[\mathrm{Re}\left[\chi_{\frac{\pi}{2}}(\Omega) y(\Omega)\right]^2\right], \\ \mathrm{Var}\left[\hat{q}_4(\Omega)\right] &= \mathrm{Var}[\hat{q}_4^{(0)}(\Omega)] + \frac{2}{T}\mathrm{E}\left[\mathrm{Im}\left[\chi_{\frac{\pi}{2}}(\Omega) y(\Omega)\right]^2\right]. \end{aligned} \tag{12}$$

Let us simplify the excess variance due to the signal, i.e., the second term, starting with the real part as follows:

$$\begin{aligned} &\mathrm{E}\left[\mathrm{Re}\left[\chi_{\frac{\pi}{2}}(\Omega) y(\Omega)\right]^2\right] \\ &\quad = \frac{1}{4}\mathrm{E}\left[\left(\chi_{\frac{\pi}{2}}(\Omega) y(\Omega) + \chi^*_{\frac{\pi}{2}}(\Omega) y^*(\Omega)\right)^2\right] \\ &\quad = \frac{1}{2}\mathrm{E}\left[\left|\chi_{\frac{\pi}{2}}(\Omega) y(\Omega)\right|^2\right], \end{aligned}$$

where we have used the fact that $\mathrm{E}\left[y(\Omega)^2\right] = 0$, since the phase of $y(\Omega)$ is isotropically random. This excess variance in the real part of $\hat{p}(\Omega)$ equals half the excess power spectral density in Eq. (9) and the imaginary part is similar. The excess power spectral density is also equal to $|\chi_{\pi/2}(\Omega)|^2 S_{yy}(\Omega)$ by Eq. (8). Therefore, Eq. (12) may be rewritten as follows:

$$\begin{aligned} \mathrm{Var}\left[\hat{q}_2(\Omega)\right] &= \mathrm{Var}[\hat{q}_2^{(0)}(\Omega)] + \sigma^2, \\ \mathrm{Var}\left[\hat{q}_4(\Omega)\right] &= \mathrm{Var}[\hat{q}_4^{(0)}(\Omega)] + \sigma^2, \end{aligned} \tag{13}$$

where we have defined the common excess variance σ^2 and the gain $G_{py}(\Omega)$ as

$$\sigma^2 := \frac{1}{2} G_{py}(\Omega) S_{yy}(\Omega), \quad G_{py}(\Omega) := |\chi_{\frac{\pi}{2}}(\Omega)|^2. \tag{14}$$

The gain $G_{py}(\Omega)$ is also sometimes called the signal transfer function from y to $\hat{p}$, since it linearly relates the signal $S_{yy}(\Omega)$ to the observed spectral density $S_{pp}(\Omega)$ in Eq. (8) [30]. Since the two harmonic oscillators at each frequency are independent and the noise in each is equal due to stationarity, we can consider them separately to study the fundamental quantum limits on estimating $S_{yy}(\Omega)$ or, equivalently, estimating σ^2. Thus, we can restrict our attention to just the harmonic oscillator corresponding, e.g., to the real part with quadratures $\hat{q}_1(\Omega)$ and $\hat{q}_2(\Omega)$, which we relabel as $\hat{x}$ and $\hat{p}$ below for brevity.

For any initial state, we have shown that estimating the continuum $S_{yy}(\Omega)$ is equivalent to a continuum of independent copies (two for each frequency Ω) of the following canonical single-parameter estimation problem. Given the initial state $\hat{\rho}$ of a harmonic oscillator with canonical quadratures $\vec{\hat{x}} = (\hat{x}, \hat{p})^{\mathrm{T}}$ such that $[\hat{x}, \hat{p}] = i$, we wish to estimate σ^2 from measurements of the final state $\Lambda_\sigma(\hat{\rho})$ after the nonunitary quantum channel Λ_σ that encodes σ. (A quantum channel is a completely positive trace-preserving linear map between density matrices.) The signal-encoding channel Λ_σ represents a random displacement along the $\hat{p}$ quadrature and has the following Kraus representation:

$$\Lambda_\sigma(\hat{\rho}) = \int_{-\infty}^{\infty} \mathrm{d}\alpha \; p(\alpha) \hat{U}_\alpha \hat{\rho} \hat{U}_\alpha^\dagger, \tag{15}$$

where $p(\alpha) \sim \mathcal{N}(0, \sigma^2)$ is the weighting of the different displacement unitaries given by $\hat{U}_\alpha = \exp(i\alpha\hat{x})$, which displaces $\hat{p}$ to $\hat{p} + \alpha$. Here, we have assumed that the zero-mean signal follows a Gaussian random process such that the second moments in Eq. (13) determine the signal. Understanding this toy model of a single harmonic oscillator undergoing random displacements is the main focus of our work. We will solve this toy model in Secs. III–VI, before returning to the original stochastic waveform-estimation problem in Sec. VII.

We now provide two alternative expressions for the encoding channel Λ_σ that will prove useful later. First, in the basis of $\hat{x}$, Λ_σ is the following decoherence channel [31,32]:

$$\langle x|\Lambda_\sigma(\hat{\rho})|x'\rangle = e^{-\frac{1}{2}\sigma^2(x-x')^2} \langle x|\hat{\rho}|x'\rangle. \tag{16}$$

And, second, when $\sigma\hat{x}$ is small, we may expand Eq. (15) to obtain the infinitesimal channel

$$\Lambda_\sigma(\hat{\rho}) \approx \hat{\rho} + \sigma^2\left(\hat{x}\hat{\rho}\hat{x} - \frac{1}{2}\{\hat{x}^2, \hat{\rho}\}\right) + \mathcal{O}(\sigma^4\hat{x}^4), \tag{17}$$

where the odd terms vanish because $p(\alpha)$ is an even function in α.

Let us consider the Gaussian-state case. If the initial state $\hat{\rho}$ is a single-mode Gaussian with a 2×2 covariance matrix $\Sigma_{jk} = \frac{1}{2}\langle\{\hat{\vec{x}}_j^{(0)}, \hat{\vec{x}}_k^{(0)}\}\rangle$ and 2×1 mean vector

$\vec{\mu} = \langle \vec{\hat{x}} \rangle$, then the final state is also Gaussian with the covariance matrix $\Sigma + \mathrm{diag}(0, \sigma^2)$ and the same mean vector $\vec{\mu}$. For example, the vacuum state $|0\rangle\langle 0|$ with covariance matrix $\Sigma_0 = \mathrm{diag}(\frac{1}{2}, \frac{1}{2})$ becomes the squeezed thermal state $\hat{\rho}(\sigma) = \Lambda_\sigma(|0\rangle\langle 0|)$ with covariance matrix $\mathrm{diag}(\frac{1}{2}, \frac{1}{2} + \sigma^2)$, as shown in Fig. 1(b). We thus call Λ_σ a Gaussian-encoding channel, since it sends Gaussian states to Gaussian states.

Our goal now is to solve this canonical noise-estimation problem by understanding the optimal protocol for estimating σ^2. As we will see, despite Λ_σ being a Gaussian-encoding channel, the optimal protocol will turn out to be to prepare non-Gaussian initial states and perform non-Gaussian measurements.

D. Review of Fisher information

We now review our main tools for addressing this single-parameter estimation problem: the concepts of classical and quantum Fisher information (FI) [33–35].

Suppose that a real parameter of interest θ is encoded in a quantum state $\hat{\rho}(\theta)$ and that we estimate θ by performing a given measurement [i.e., a positive operator-valued measure (POVM)] with associated probability distribution $p(x|\theta)$. The minimal MSE $\Delta^2\theta$ of unbiased estimation of θ from M independent and identically distributed measurement results satisfies the *classical Cramér-Rao bound (CCRB)*, $\Delta^2\theta \geq (1/M)[\mathcal{I}_C(\theta)]^{-1}$, where $\mathcal{I}_C(\theta)$ is the *classical Fisher information (CFI)*, given by

$$\mathcal{I}_C(\sigma) = \int_{-\infty}^{\infty} \mathrm{d}x \, \frac{[\partial_\theta p(x|\theta)]^2}{p(x|\theta)}. \tag{18}$$

This bound can be attained by maximum likelihood estimation in the asymptotic limit such that the central limit theorem applies. In general, saturating the CFI may require a parameter that is sufficiently small or has a narrow enough prior, as well as the ability to perform a large number of independent measurements. A useful property of the CFI is that if $p(x|\theta)$ and $p(y|\theta)$ are independent distributions, e.g., describing two separate measurements, then the total CFI from observing one outcome from each distribution is simply the sum of the individual CFIs.

This may not be the optimal measurement, however, for extracting the maximal information about θ from $\hat{\rho}(\theta)$. For single-parameter estimation, the *quantum Fisher information (QFI)* is the CFI maximized over all possible measurements (POVMs), $\mathcal{I}_Q(\theta) = \sup \mathcal{I}_C(\theta)$. In terms of the eigenvalues p_j and eigenvectors $|\phi_j\rangle$ of the final state $\hat{\rho}(\theta) = \sum_j p_j |\phi_j\rangle\langle\phi_j|$, the QFI can be shown to be

$$\mathcal{I}_Q(\theta) = \sum_{j,k} \frac{2}{p_j + p_k} \left| \langle \phi_j | \partial_\theta \hat{\rho}(\theta) | \phi_k \rangle \right|^2$$

where the sum runs over only j and k such that $p_j + p_k > 0$. For example, if a parameter θ is encoded by a unitary $\exp(-i\theta\hat{H})$ applied to a pure state $|\psi\rangle$, then the QFI is $\mathcal{I}_Q(\theta) = 4\mathrm{Var}_{|\psi\rangle}[\hat{H}]$ independent of θ.

The *quantum Cramér-Rao bound (QCRB)*, $\Delta^2\theta(\geq 1/M)[\mathcal{I}_Q(\theta)]^{-1}$, provides the fundamental minimal MSE that can be achieved by maximum likelihood estimation from the M outcomes of the optimal measurement scheme. Similarly, the minimal fractional MSE with respect to a parameter θ is bounded too, $\Delta^2\theta/\theta^2 \geq (1/M)[\theta^2\mathcal{I}_Q(\theta)]^{-1}$. The analog of the additivity of the CFI for independent distributions is that the QFI for a product state $\hat{\rho}_1(\theta) \otimes \hat{\rho}_2(\theta)$ is simply the sum of the individual QFIs.

E. Estimating standard deviation versus variance

For our single-parameter problem, we choose to estimate the standard deviation σ. This is equivalent to estimating the variance σ^2, since $\sigma \geq 0$. The QFIs with respect to σ and σ^2 are related by $\mathcal{I}_Q(\sigma) = 4\sigma^2\mathcal{I}_Q(\sigma^2)$ through the chain rule. Although $\mathcal{I}_Q(\sigma^2)$ may diverge and $\mathcal{I}_Q(\sigma)$ stay finite as $\sigma \to 0$, the limiting behavior of the fractional MSEs is consistent since, in that event, $[\sigma^4\mathcal{I}_Q(\sigma^2)]^{-1} \propto [\sigma^2\mathcal{I}_Q(\sigma)]^{-1} \to \infty$.

III. LOSSLESS VACUUM LIMIT

The task of estimating σ from Λ_σ in Eq. (15), in the absence of loss when the input state is the vacuum, has previously been studied in Refs. [19,20,36]. We briefly discuss this ideal lossless regime here for comparison.

Suppose that we apply Λ_σ to $\hat{\rho} = |0\rangle\langle 0|$ such that the final covariance matrix is $\Sigma = \mathrm{diag}(\frac{1}{2}, \frac{1}{2} + \sigma^2)$. The QFI for a signal encoded solely in the covariance matrix of a single-mode Gaussian state is [37]

$$\mathcal{I}_Q(\sigma) = \frac{\mathrm{Tr}\left[(\Sigma^{-1}\partial_\sigma\Sigma)^2\right]}{2(1+\gamma^2)} + \frac{2(\partial_\sigma\gamma)^2}{1-\gamma^4}, \tag{19}$$

where $\gamma = \det(2\Sigma)^{-\frac{1}{2}}$ is the purity of the Gaussian state. For the vacuum state, Eq. (19) implies that the QFI is

$$\mathcal{I}_Q(\sigma) = \frac{2}{1+\sigma^2} \xrightarrow[\sigma\to 0]{} 2. \tag{20}$$

We will write the important limit of $\lim_{\sigma\to 0}\mathcal{I}_Q(\sigma)$ as $\mathcal{I}_Q(\sigma = 0)$ henceforth. Note that it is nonzero for the QFI.

We want to know what measurement will saturate the QFI in Eq. (20). A homodyne measurement of $\hat{p}$ is a Gaussian measurement for which the CFI with respect to the total standard deviation $\varsigma = \sqrt{\frac{1}{2} + \sigma^2}$ is $\mathcal{I}_C^{\hat{p}}(\varsigma) = 2/\varsigma^2$. This implies that $\mathcal{I}_C^{\hat{p}}(\sigma) = 2\sigma^2/\left(\frac{1}{2} + \sigma^2\right)^2$ such that the CFI for quadrature measurement saturates the QFI in Eq. (20) in the classical regime of $\sigma \gg \frac{1}{\sqrt{2}}$ but vanishes in the quantum regime of $\sigma \to 0$, where $\mathcal{I}_C^{\hat{p}}(\sigma) \approx 8\sigma^2$,

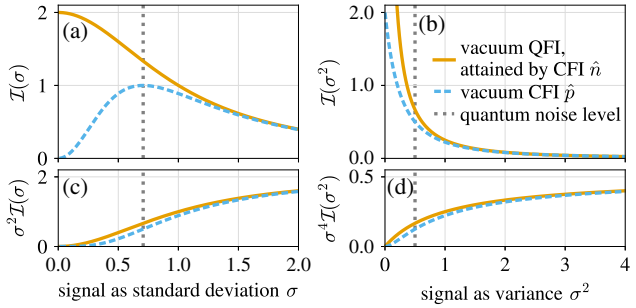

FIG. 2. The Fisher information (FI) with respect to (a) standard deviation σ or (b) variance σ^2 for the single-parameter estimation problem in the lossless vacuum limit. The level of quantum noise for the initial vacuum state, shown by the dotted gray vertical line, is equal to $\mathrm{Var}[\hat{p}] = \frac{1}{2}$. A quadrature measurement $\hat{p}$, with CFI shown in the dashed blue curve, is optimal for signals well above this level. A number-resolving measurement $\hat{n}$, however, is optimal—its CFI attains the QFI shown in the solid orange curve—for all σ including as $\sigma \to 0$. In this limit, the fractional MSE with respect to (c) σ or (d) σ^2 diverges regardless of the measurement scheme, since $\theta^2\mathcal{I}_Q(\theta)$ converges to zero.

as shown in Fig. 2(a). If we estimate σ^2 instead of σ, as shown in Fig. 2(b), then the behavior of the fractional MSE is the same as discussed in Sec. II E and shown in Fig. 2(c). The conventional quadrature measurement, therefore, is highly inefficient in the relevant limit of $\sigma \to 0$. Intuitively, the small change in the variance induced by the weak signal $\sigma^2 \ll \frac{1}{2}$ is masked by the vacuum quantum noise, leading to this vanishing sensitivity.

In comparison, a number-resolving measurement of $\hat{n} = \hat{a}^\dagger\hat{a}$ can be shown to saturate the QFI for all σ. Since $\hat{U}_\alpha|0\rangle = |\alpha'\rangle$ is a coherent state with amplitude $\alpha' = i\alpha/\sqrt{2}$, we can compute the probabilities in the number basis after the encoding in Eq. (15) as

$$p(n) = \int_{-\infty}^{\infty} \mathrm{d}\alpha \ \left|\langle n|\alpha'\rangle\right|^2 p(\alpha) = \frac{(2n)!}{2^{2n}(n!)^2}\frac{\sigma^{2n}}{\left(\sigma^2+1\right)^{n+\frac{1}{2}}}. \tag{21}$$

Then, the CFI for number measurement, calculated as $\mathcal{I}_C^{\hat{n}}(\sigma) = \sum_{n=0}^{\infty}[\partial_\sigma p(n)]^2/p(n)$ by Eq. (18), equals the QFI in Eq. (20), as shown in Fig. 2. Intuitively, if there is no signal, then the vacuum state remains an eigenstate of $\hat{n}$ and no particles are counted. If a particle is counted, then it must come from the signal. This is unlike the quadrature measurement, where some variance in the measurement results remains due to the quantum noise even if there is no signal.

Let us illustrate how much worse quadrature measurement is compared to the optimal number-resolving measurement. If the signal is comparable to the quantum noise, i.e., $\sigma = \frac{1}{\sqrt{2}}$, then the QFI is $\frac{4}{3}$ compared to the quadrature measurement CFI of 1. Instead, if the signal is 10 times (100 times) smaller than the quantum noise in amplitude units, i.e., $\sigma = \frac{1}{10\sqrt{2}}$ ($\sigma = \frac{1}{100\sqrt{2}}$), then the QFI is roughly 51 times (5001 times) greater than the CFI. Correspondingly, suppose that we want to achieve an MSE $\Delta^2\sigma$ below some target value. If $\sigma = \frac{1}{10\sqrt{2}}$, e.g., then it takes only 2% of the number of quadrature measurements required to attain the target value if we instead use number-resolving measurements (recall that the Cramér-Rao Bound on the MSE is $(1/M)[\mathcal{I}(\sigma)]^{-1}$ given M measurements).

This estimation problem is analogous to quantum superresolution in optical imaging [20,38,39] and spectroscopy [40,41]. These resolution problems are characterized by having vanishing signal: $\partial_\theta\hat{\rho} \to 0$ as $\theta \to 0$, where θ is the parameter of interest. This vanishing signal leads to a "Rayleigh curse," where the CFI of the naive measurement, e.g., image-plane photon counting in the case of optical imaging, also vanishes as $\theta \to 0$. The QFI, however, remains positive in this limit and is attained by the CFI of a particular nonstandard measurement, e.g., spatial-mode demultiplexing in the case of optical imaging. Analogously, we define the *Rayleigh curse* to refer to any scenario in which the FI converges to zero in the limit of zero signal. We have seen above that here the Rayleigh curse arises for the CFI of the naive quadrature measurement but can be avoided by performing a number measurement.

IV. LOSS AND CLASSICAL NOISE

Realistically, the quantum state will experience noise channels before and after the encoding such that the noiseless channel Λ_σ becomes the noisy channel Λ'_σ. We restrict our attention to Gaussian noise channels. We want to understand how these imperfections limit our ability to estimate σ and whether the Rayleigh curse can still be avoided to increase QFI in the small-signal limit.

A. Gaussian states

We first consider preparing a Gaussian initial state in the limit of $\sigma \to 0$. For Gaussian noise channels, $\partial_\sigma\Lambda'_\sigma \to 0$ as $\sigma \to 0$ and information about σ is encoded only in the covariance matrix of the output Gaussian state. Whether the Rayleigh curse reappears is addressed by the following claim.

Claim 1. For any Gaussian state such that a parameter σ is encoded only on its covariance matrix $\Sigma(\sigma)$ and $\lim_{\sigma\to 0}\partial_\sigma\Sigma = 0$, then $\mathcal{I}_Q(\sigma = 0) \neq 0$ (i.e., the state does not exhibit the Rayleigh curse) if and only if there exists a symplectic eigenvalue of Σ equal to $\frac{1}{2} + k\sigma^2$ for some constant $k > 0$.

This is the Gaussian version of the quantum resolution criterion in Ref. [40]. The proof of this claim is given in Appendix A. For a single-mode Gaussian state, this claim implies that overcoming the Rayleigh curse is possible only if $\Lambda'_{\sigma=0}(\rho)$ is pure. In that case, the QFI comes only from the purity [the second term in Eq. (19)]: the QFI is nonvanishing in the limit $\sigma \to 0$, and equal to $\mathcal{I}_Q(\sigma = 0) = 2k'$, if and only if the purity is $\gamma = 1 - k'\sigma^2$ for some constant k'. For a multimode Gaussian state in the limit $\sigma \to 0$, at least one of the modes needs to be pure for the QFI not to vanish.

B. Loss channel

We apply Claim 1 to a couple of examples of common Gaussian noise channels, assuming an initial vacuum state. First, consider the "pure" or "cold" loss channel $\Lambda_\eta^{\text{loss}}$, which models, e.g., losses due to coupling with a zero-temperature bath. This can be modeled as a beam-splitter operation with an ancillary vacuum mode that is then traced out as shown in Fig. 3. For a loss $\eta \in (0, 1)$, this channel has the following Kraus representation [42]:

$$\Lambda_\eta^{\text{loss}}(\hat{\rho}) = \sum_{n=0}^{\infty} \hat{K}_n \hat{\rho} \hat{K}_n^\dagger,$$

where the Kraus operators are

$$\hat{K}_n = \frac{1}{\sqrt{n!}} \left(\sqrt{\eta}\hat{a}\right)^n (1-\eta)^{\frac{1}{2}(\hat{a}^\dagger \hat{a} - n)}.$$

Note that here η is the loss and $1 - \eta$ is the efficiency. The effect of this channel on the mean vector and covariance matrix of a single-mode Gaussian state is as follows: $\vec{\mu} \mapsto \sqrt{1-\eta}\vec{\mu}$ and $\Sigma \mapsto (1-\eta)\Sigma + \eta\Sigma_0$, where $\Sigma_0 = \text{diag}\left(\frac{1}{2}, \frac{1}{2}\right)$. If the initial state is a vacuum state, then the final state in the limit of $\sigma \to 0$ is pure and there is no Rayleigh curse. The precise nonzero value of the vacuum QFI depends on whether the loss occurs before or after the encoding. If the loss occurs before the encoding, then the QFI is unchanged from $2/(1 + \sigma^2)$, since the vacuum $|0\rangle$ is stable under the loss channel $\Lambda_\eta^{\text{loss}}$. But if the loss occurs after the encoding, then the QFI degrades to $2(1-\eta)/(1 + (1-\eta)\sigma^2)$, since the state after the encoding is no longer vacuum and therefore is not stable under the loss channel.

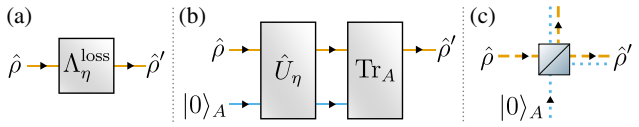

FIG. 3. The quantum state $\hat{\rho}'$ after a loss can be thought of as: (a) the result, $\hat{\rho}' = \Lambda_\eta^{\text{loss}}(\hat{\rho})$, of a nonunitary quantum channel $\Lambda_\eta^{\text{loss}}$; (b) the result of a beam-splitter unitary $\hat{U}_\eta$ with an ancillary vacuum mode $|0\rangle_A$ that is then traced out; or (c) for an optical system, the state at one of the output ports of a fictitious beam splitter with a vacuum mode.

In this work, we will consider different initial states that will not be stable under the loss channel, unlike the vacuum. We will focus on the impact of a known loss η occurring before the encoding, which now will affect the QFI, and assume that no loss occurs after the encoding. We address the effect of noise channels occurring after the encoding in Appendix B 1, where we show that measurement noise can, in theory, be overcome by using a suitable control unitary. We analyze the implications of an unknown loss in Appendix B 2, where we justify neglecting this effect given a vacuum input state.

C. Classical noise channel

We also consider a Gaussian noise channel $\Lambda_{\Sigma_C}^{\text{noise}}$ that models a source of uncorrelated classical noise, e.g., thermal fluctuation processes in optics or a nonzero temperature of microwave resonators. For M modes, this channel is a random-displacement channel

$$\Lambda_{\Sigma_C}^{\text{noise}}(\hat{\rho}) = \int_{\mathbb{C}^M} \mathrm{d}\vec{\alpha}\, p(\vec{\alpha}) \hat{D}(\vec{\alpha}) \hat{\rho} \hat{D}(\vec{\alpha})^\dagger, \tag{22}$$

where $p(\alpha) \sim \mathcal{N}(\mathbf{0}, \frac{1}{2}\Sigma_C)$ for a positive-semidefinite $2M \times 2M$ matrix Σ_C and the M-mode displacement operator is $\hat{D}(\alpha) = \prod_{i=1}^M \exp(\alpha_i \hat{a}_i^\dagger - \alpha_i^* \hat{a}_i)$, where $\hat{a}_i$ is the annihilation operator for the ith mode. This classical noise channel acts on Gaussian states as an additive noise source: $\Sigma \mapsto \Sigma + \Sigma_C$. [Note that $\Lambda_\sigma = \Lambda_{\text{diag}(0,\sigma^2)}^{\text{noise}}$ in Eq. (15).]

For a given single-mode Gaussian state with fixed $\langle \hat{n} \rangle = N$, a nonzero Σ_C will make the final state mixed and always lead to a Rayleigh curse, as $\sigma \to 0$ by Claim 1. For example, if $\Sigma_C = \text{diag}\left(\sigma_x^2, \sigma_p^2\right)$ with vacuum input, then

$$\mathcal{I}_Q(\sigma) = \frac{2\sigma^2 \left(1 + 2\sigma_x^2\right)^2}{\kappa\,(\kappa + 1)}, \tag{23}$$

where

$$\kappa = \sigma^2 + \sigma_x^2 + \sigma_p^2 + 2\sigma_x^2 \left(\sigma^2 + \sigma_p^2\right). \tag{24}$$

As long as one of σ_x and σ_p is nonzero, then $\mathcal{I}_Q(\sigma = 0) = 0$ here. In particular, for the isotropic case $\sigma_C := \sigma_x = \sigma_p$, the Rayleigh curse arises when the signal is dominated by the classical noise, $\sigma \ll \sigma_C$.

Classical noise changes the optimal measurement for the vacuum input-state case, such that number measurement alone no longer saturates the QFI. Instead, the optimal measurement is to squeeze $\Sigma \mapsto \text{diag}\left(e^{-2r}, e^{2r}\right)\Sigma$ with $e^{2r} = 1 + 2\sigma_x^2$ immediately after

the encoding process. The squeezing level is chosen such that $\Sigma = \mathrm{diag}\left(\frac{1}{2}+\sigma_x^2, \frac{1}{2}+\sigma^2+\sigma_p^2\right)$ becomes $\mathrm{diag}\left(\frac{1}{2}, \frac{1}{2}+\kappa\right)$. After this additional squeezing operation, we then perform a number measurement. The gain in the CFI from this additional squeezing operation in the limit of small classical noise compared to direct number measurement is $1+4\sigma_x^2$; hence it is marginal given small σ_x but it is still highly favorable compared to quadrature measurement.

Let us discuss how to determine, e.g., σ_p for a given experiment. The total output-referred free noise $S_{pp}^{(0)}(\Omega)$ in Eq. (3) is the sum of the output-referred quantum noise $S_{pp}^{Q}(\Omega)$ and the output-referred classical noise $S_{pp}^{C}(\Omega)$ such that

$$S_{pp}(\Omega) = S_{pp}^{Q}(\Omega) + S_{pp}^{C}(\Omega) + G_{py}(\Omega)S_{yy}(\Omega). \tag{25}$$

For example, the output-referred quantum noise is $S_{pp}^{Q}(\Omega) = S_{pp}^{Q,\mathrm{vac}}(\Omega) = 1$ if the quantum noise spectrum is dominated by the quantum shot noise from the vacuum, which we denote as $S_{pp}^{Q,\mathrm{vac}}(\Omega)$ (i.e., if the input state is vacuum and the device has no effects, such as quantum radiation pressure noise, that modify the quantum state if there is zero signal). Similarly to how σ is given by Eq. (14), the parallel classical noise is thus $\sigma_p^2 := \frac{1}{2}S_{pp}^{C}(\Omega)$ and, analogously, the perpendicular classical noise is $\sigma_x^2 := \frac{1}{2}S_{xx}^{C}(\Omega)$. (More generally, Σ_C equals $1/2$ times the classical noise contributions to the matrix of cross spectra of $\hat{x}(\Omega)$ and $\hat{p}(\Omega)$.) Furthermore, let us relate, e.g., σ_p to a common set of performance metrics for interferometers: the signal-referred noise spectral densities. These metrics are obtained by dividing Eq. (25) by the gain $G_{py}(\Omega)$ and are given in power units as follows:

$$\begin{aligned} \mathcal{S}(\Omega) &:= \frac{S_{pp}(\Omega)}{G_{py}(\Omega)} = \mathcal{S}^{Q}(\Omega) + \mathcal{S}^{C}(\Omega) + S_{yy}(\Omega), \\ \mathcal{S}^{Q}(\Omega) &:= \frac{S_{pp}^{Q}(\Omega)}{G_{py}(\Omega)}, \quad \mathcal{S}^{C}(\Omega) := \frac{S_{pp}^{C}(\Omega)}{G_{py}(\Omega)}. \end{aligned} \tag{26}$$

Here, $\mathcal{S}^{Q}(\Omega)$ is the signal-referred quantum noise, $\mathcal{S}^{C}(\Omega)$ is the signal-referred classical noise, and $\mathcal{S}^{(0)}(\Omega) = \mathcal{S}^{Q}(\Omega) + \mathcal{S}^{C}(\Omega)$ is the total signal-referred free noise. For example, for gravitational-wave observatories, $\mathcal{S}^{(0)}(\Omega)$ is the strain sensitivity to deterministic gravitational waves. If the total free-noise spectrum is dominated by the quantum shot noise from the vacuum, which is the case for gravitational-wave observatories at kilohertz frequencies, then $\mathcal{S}^{(0)}(\Omega) = \mathcal{S}^{Q,\mathrm{vac}}(\Omega) = 1/G_{py}(\Omega)$, where $\mathcal{S}^{Q,\mathrm{vac}}(\Omega) = S_{pp}^{Q,\mathrm{vac}}(\Omega)/G_{py}(\Omega)$. This provides the following simple relationship to determine the gain $G_{py}(\Omega)$ from existing tuned interferometer designs that meet these assumptions:

$$G_{py}(\Omega) = \mathcal{S}^{Q,\mathrm{vac}}(\Omega)^{-1}. \tag{27}$$

The classical noise may then be determined from the signal-referred spectra, e.g., $\sigma_p^2 := \frac{1}{2}\mathcal{S}^{C}(\Omega)/\mathcal{S}^{Q,\mathrm{vac}}(\Omega)$.

In summary, classical noise limits the QFI for small signals when preparing the vacuum state. Unless stated otherwise, we assume henceforth that the classical noise is negligible, i.e., that the signal is dominant, $\sigma \gg \sigma_C$. This is motivated by certain applications, discussed later in Sec. VII B, for which the search for stochastic signals is limited by imperfections from decoherence—quantum backgrounds—rather than classical backgrounds. We will revisit nonzero classical noise and whether its impact can be avoided by preparing different initial states in Sec. V G.

V. OPTIMAL INITIAL STATE

We now consider the initial state and measurement scheme that comprise the optimal protocol for sensing the signal σ encoded by Λ_σ with and without losses. In particular, we want to know whether entangled resources and collective measurements are necessary.

A. Review of channel quantum Fisher information

Building on Sec. II D, here, we introduce our tools for determining the optimal initial state: the channel QFI and extended-channel QFI [43,44].

Given an initial state $|\psi\rangle$ and a channel Λ_θ that encodes a parameter θ in the final state $\hat{\rho}(\theta) = \Lambda_\theta(|\psi\rangle\langle\psi|)$, then let the QFI with respect to θ be denoted as $\mathcal{I}_Q^{\Lambda_\theta(|\psi\rangle\langle\psi|)}(\theta)$. The QFI from the optimal initial state $|\psi\rangle$ is called the *channel QFI (CQFI)* of Λ_θ and is given by

$$\mathcal{I}_Q^{\Lambda_\theta,\text{no ancilla}}(\theta) = \sup_{|\psi\rangle} \mathcal{I}_Q^{\Lambda_\theta(|\psi\rangle\langle\psi|)}(\theta), \tag{28}$$

where we emphasize that no ancilla is allowed when calculating the CQFI. From the convexity of the QFI, it suffices to optimize over pure initial states.

If we prepare an initial state $|\Psi\rangle$ that is also entangled with some ancilla, e.g., a TMSV state, then we might improve the estimation of θ. Let the joint channel be $\Lambda_\sigma \otimes \Lambda_A$, where Λ_A is some channel that acts on the ancilla. In the noiseless-ancilla case, $\Lambda_A = 1_A$ is the identity. In this ideal case, the CQFI of the joint channel $\Lambda_\sigma \otimes 1_A$ is called the *extended-channel QFI (ECQFI)* of Λ_σ and is given by

$$\mathcal{I}_Q^{\Lambda_\theta}(\theta) = \sup_{|\Psi\rangle} \mathcal{I}_Q^{(\Lambda_\theta\otimes 1_A)(|\Psi\rangle\langle\Psi|)}(\theta). \tag{29}$$

The ECQFI is the maximum amount of information about the parameter θ that can be extracted after the channel Λ_θ acts on the quantum device. Note that the inequality $\mathcal{I}_Q^{\Lambda_\theta}(\theta) \geq \mathcal{I}_Q^{\Lambda_\theta,\text{no ancilla}}(\theta)$ between the ECQFI and the

CQFI always holds. If, as in some cases, $\mathcal{I}_Q^{\Lambda_\theta}(\theta)$ is strictly larger than $\mathcal{I}_Q^{\Lambda_\theta,\text{no ancilla}}(\theta)$, then entanglement with an ancilla is a required resource for optimal signal extraction [45].

Realistically, however, we expect that the ancilla should also experience some loss such that $\Lambda_A = \Lambda_{\eta_A}^{\text{loss}}$, where η_A is some ancilla loss, rather than $\Lambda_A = 1_A$. The CQFI of this joint channel $\Lambda_\sigma \otimes \Lambda_{\eta_A}^{\text{loss}}$ then represents the maximum information feasibly available. The key questions to ask are then whether the CQFI and ECQFI of $\Lambda_\sigma \otimes \Lambda_{\eta_A}^{\text{loss}}$ are equal and, similarly, whether the CQFI and ECQFI of Λ_σ are equal. If the latter is true, then all four of the CQFI and ECQFI of $\Lambda_\sigma \otimes \Lambda_{\eta_A}^{\text{loss}}$ and the CQFI and ECQFI of Λ_σ are equal and the ultimate sensitivity limit can be achieved without using an ancilla. This will turn out to be the case in our work (see Sec. V F below).

When exploring these limits, we can also impose other physical restrictions. For example, since the unconstrained ECQFI $\mathcal{I}_Q^{\Lambda_\theta}(\theta)$ might be unbounded at $\theta = 0$ in the lossless case, we can constrain the average energy $\langle \hat{H} \rangle \propto \langle \hat{n} \rangle + \frac{1}{2}$ of the initial state of the harmonic oscillator. Let the ECQFI in Eq. (29) constrained to initial states $|\Psi\rangle$ with $\langle \hat{n} \rangle = N$ average occupation number per mode be denoted $\mathcal{I}_Q^{\Lambda_\theta,N}(\theta)$. (Note that while $\langle \hat{n} \rangle \leq N$ may be a more natural constraint, the bounds that we find are always nondecreasing in N such that it suffices to consider $\langle \hat{n} \rangle = N$.)

We need to therefore calculate the ECQFI of our channel Λ_σ with respect to σ to determine the optimal initial state and whether entanglement is a required resource, with and without noise channels on the system and ancilla.

B. Review of the deterministic case

Before proceeding to our case of a random-displacement channel, we briefly review the case of a deterministic displacement channel to establish the similarities and differences between the two cases.

Consider a deterministic displacement of the state $\hat{\rho}$ by μ along $\hat{p}$ such that we want to estimate μ from measurements of $\hat{U}_\mu \hat{\rho} \hat{U}_\mu^\dagger$, where $\hat{U}_\mu = \exp(i\mu\hat{x})$. In the lossless case, if $\hat{\rho} = |\psi\rangle\langle\psi|$ is pure, then the QFI is $\mathcal{I}_Q(\mu) = 4\text{Var}_{|\psi\rangle}[\hat{x}]$. Since $\text{Var}_{|\psi\rangle}[\hat{x}]$ can be made arbitrarily large, the unconstrained ECQFI is unbounded. As such, we constrain the initial state $|\psi\rangle$ to have N average occupation number per mode, $\langle \hat{n} \rangle = N$. The maximum value of $\text{Var}_{|\psi\rangle}[\hat{x}]$ given the constraint of $\langle \hat{n} \rangle = N$ is attained by an SMSV state, in which case $\text{Var}_{|\psi\rangle}[\hat{x}]$ is equal to $\xi_N := N + \frac{1}{2} + \sqrt{N(N+1)}$. Note that $\xi_N \to 2N$ as $N \to \infty$. The ECQFI is then

$$\mathcal{I}_Q^{\hat{U}_\mu,N}(\mu) = 4\xi_N \xrightarrow[N\to\infty]{} 8N, \tag{30}$$

which is achieved by preparing an SMSV state and, e.g., measuring the quadrature $\hat{p}$. Another optimal measurement is to instead antisqueeze after the encoding and then perform a number measurement $\hat{n}$. We emphasize that this asymptotic scaling with N is the fundamental limit for lossless deterministic displacement sensing. The Heisenberg limit of N^2 for phase estimation cannot be achieved for displacements.

Realistically, however, the state will experience some losses such that the total channel is not unitary. Suppose that the state encounters a loss η before the encoding such that the total channel is $\Lambda'_\mu(\rho) = \hat{U}_\mu \Lambda_\eta^{\text{loss}}(\hat{\rho}) \hat{U}_\mu^\dagger$. In this lossy case, Ref. [18] has shown that an SMSV state is still the optimal initial state and attains the following ECQFI given the constraint of $\langle \hat{n} \rangle = N$:

$$\mathcal{I}_Q^{\Lambda'_\mu,N}(\mu) = \frac{4\xi_N}{\eta\,(2\xi_N - 1) + 1} \xrightarrow[N\to\infty]{} \frac{2}{\eta}. \tag{31}$$

The high-energy limit of the deterministic ECQFI is then bounded by $2/\eta$, which cannot be surpassed using any initial state. For example, this result is known in the gravitational-wave-observatory context where, to sense deterministic gravitational waves in the presence of optical losses, it is optimal to inject a squeezed optical state into the differential port of the Michelson interferometer and perform a quadrature measurement [46], as is presently done [47–50].

C. Lossless case

We now return to studying the ECQFI of the random-displacement channel Λ_σ with respect to σ. We prove the following claim about the optimal protocol.

Claim 2. In the lossless case, the ECQFI constrained to initial states with $\langle \hat{n} \rangle = N$ per mode is

$$\mathcal{I}_Q^{\Lambda_\sigma,N}(\sigma) = \frac{4}{2\sigma^2 + \xi_N^{-1}}, \tag{32}$$

which is saturated by preparing an initial SMSV state and performing a number-resolving measurement after first antisqueezing. Entanglement with an ancilla is not a required resource.

The proof of Claim 2 is given in Appendix C 2. We use an established technique of bounding the ECQFI using a sequence of purifications of Λ_σ by Uhlmann's theorem [18,43,44]. We find the following upper bound on the ECQFI:

$$\mathcal{I}_Q^{\Lambda_\sigma,N}(\sigma) \leq \frac{4}{2\sigma^2 + \text{Var}[\hat{x}]^{-1}}, \tag{33}$$

where $\text{Var}[\hat{x}]$ is calculated with respect to the initial state. Minimizing Eq. (33) over the initial state is equivalent to

optimizing $\mathrm{Var}[\hat{x}]$ given the constraint of $\langle\hat{n}\rangle = N$, which is attained by an SMSV state. By calculating the QFI using Eq. (19), we observe that an SMSV state is optimal for any σ and saturates the ECQFI in Eq. (32).

Claim 2 implies that the behavior of the ECQFI depends on the ratio between N and $1/4\sigma^2$. For $N \ll 1/4\sigma^2$, $\mathcal{I}_Q^{\Lambda_\sigma,N}(\sigma)$ grows linearly as $8N$ similarly to Eq. (30) but if $N \gg 1/4\sigma^2$, then $\mathcal{I}_Q^{\Lambda_\sigma,N}(\sigma)$ converges to $2/\sigma^2$.

We comment on the limit of vanishingly small signals $\sigma \to 0$ such that the upper bound in Eq. (33) becomes $4\mathrm{Var}[\hat{x}]$ and is tight for any initial state in this limit [51]. (Curiously, this equals the QFI for deterministic displacements in Sec. V B.) In Fig. 4(a), for this small-signal limit of $\sigma \to 0$, we compare the ECQFI attained by an SMSV state versus the average number $\langle\hat{n}\rangle$ to the QFI of other states, such as coherent, TMSV, Fock, Schrödinger's cat, and finite-energy Gottesman-Kitaev-Preskill (GKP) grid states [52]. (We review finite-energy GKP states in Appendix H.) In the high-energy limit of $\langle\hat{n}\rangle \to \infty$; cat states also attain the ECQFI of $8N$; TMSV, finite-energy GKP states, and Fock only grow as $4N$; and coherent states remain at the vacuum level of 2, since displacements commute with the encoding channel Λ_σ. In this limit of $\sigma \to 0$, an optimal measurement of the final state $\hat{\rho}$ is to project it onto the initial state $|\Psi\rangle\langle\Psi|$, which can be implemented by an "echo protocol" [53]. For example, it suffices to directly perform a number-resolving measurement of the SMSV in this vanishing-signal limit of $\sigma \to 0$. For arbitrary $\sigma > 0$, however, we need to first antisqueeze and then perform a number-resolving measurement as stated in Claim 2.

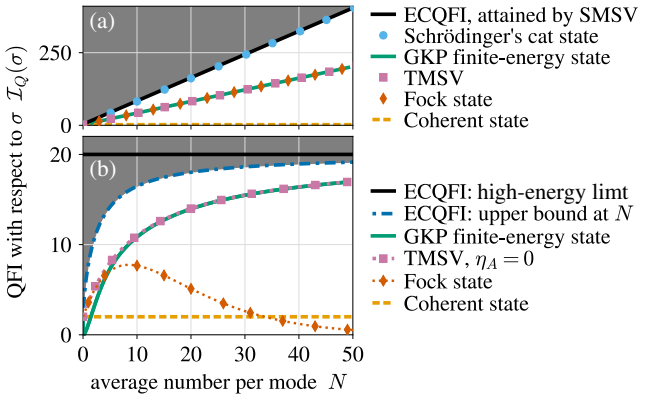

FIG. 4. The QFI from preparing different initial states, indicated in the legend, versus the initial average occupation number per mode in (a) the lossless case and (b) the case of a loss of $\eta = 0.1$ occurring before the encoding with $\sigma = 10^{-3}$. We compare the QFI from preparing each initial state to the ultimate precision limit, the extended-channel QFI (ECQFI). The shaded gray region is thus inaccessible [but the upper bound at N is loose in (b)]. In the lossy case, the SMSV and the TMSV with $\eta_A > 0$ exhibit the Rayleigh curse, i.e., QFI $\to 0$ as $\sigma \to 0$, and the Schrödinger's cat state has the same QFI as a coherent state for approximately $N > 3$.

D. Lossy case

Any actual experiment will experience loss that will dramatically change the ECQFI and optimal initial state compared to the lossless case. Consider using the initial state that is optimal in the lossless case, an SMSV state. Given a loss η occurring before the encoding, then the total channel becomes $\Lambda'_\sigma = \Lambda_\sigma \circ \Lambda_\eta^{\mathrm{loss}}$, where we assume that the classical noise is negligible. By Eq. (19), the QFI for an SMSV state in the high-energy limit of $N \to \infty$ is

$$\mathcal{I}_Q(\sigma) = \frac{8\sigma^2}{\left(\eta + 2\sigma^2\right)^2}. \tag{34}$$

In the small-loss limit, $\eta \ll \sigma^2$, the QFI is approximately $2/(\eta + \sigma^2)$, which will be shown to be optimal at any σ^2. In the experimentally relevant loss-dominated regime of $0 < \sigma^2 \ll \eta$, the SMSV performs poorly. The QFI for an SMSV vanishes as $\sigma \to 0$: it suffers from the Rayleigh curse and performs even worse than vacuum, as shown in Fig. 5. This raises the question: What is the ECQFI in the presence of significant loss, $\eta \gg \sigma^2$, and which initial states saturate this bound?

We prove the following claim about the optimal protocol in the lossy case.

Claim 3. The ECQFI given a loss η before the encoding is given by

$$\mathcal{I}_Q^{\Lambda'_\sigma}(\sigma) = \frac{2}{\eta + \sigma^2}, \tag{35}$$

which is attained by preparing a TMSV state with noiseless ancilla ($\eta_A = 0$) in the high-energy limit of $N \to \infty$.

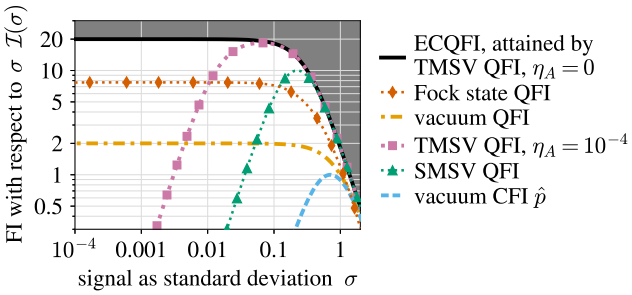

FIG. 5. The FI versus the standard deviation σ for different initial states, indicated in the legend, with a loss of $\eta = 0.1$ occurring before the encoding. The high-energy limit of the ECQFI and the squeezed states is shown. The shaded gray region is inaccessible. For large σ, e.g., $\sigma \sim \eta$, we calculate the QFI for the Fock state $|8\rangle$ numerically using a truncated Hilbert space of dimension 50.

The proof of Claim 3 is given in Appendix C 3, where we use the same method as the lossless case but with different purifications, inspired by the deterministic case [18]. The optimal measurement of a TMSV state with noiseless ancilla is to perform number-resolving measurements of the two modes after first applying antisqueezing and, in general, a beam splitter [37]. Suppose that we instead perform number-resolving measurements directly on the two modes of a TMSV state with average number per mode $n \in \mathbb{Z}$ without first antisqueezing. Then, the CFI does not attain the TMSV QFI and instead equals the lower QFI of the equivalent Fock state $|n\rangle$ given below.

We have also determined the following upper bound on the ECQFI for a given finite $\langle \hat{n} \rangle = N$:

$$\mathcal{I}_Q^{\Lambda'_\sigma, N} < \frac{4}{2(\eta + \sigma^2) + (1-\eta)\xi_N^{-1}}, \tag{36}$$

but this upper bound is not tight for $\eta > 0$ and a fixed finite N. For example, a TMSV state with noiseless ancilla ($\eta_A = 0$) does not saturate this upper bound for finite N, as shown in Fig. 4(b).

For finite $\sigma > 0$, if the loss η_A on the ancilla mode before the encoding is small, $\eta_A \ll \sigma^2, \eta$, then a TMSV state saturates the ECQFI in Eq. (35) in the high-energy limit of $N \to \infty$, as shown in Appendix D. Experimentally, however, this requirement is likely too stringent to probe the small signals of interest. In the realistic regime of $\sigma^2 \ll \eta, \eta_A$, a TMSV state does not saturate the ECQFI.

In the $\sigma \to 0$ limit, all squeezed Gaussian states suffer the Rayleigh curse by Claim 1, provided that loss occurs on every mode (i.e., with fixed $\eta, \eta_A > 0$). The only Gaussian states with nonvanishing QFI are then the coherent states, which still have a QFI of $2/(1+\sigma^2)$, since they remain coherent after the loss before the encoding. This raises the question of whether it is possible to attain the ECQFI in the realistic regime of $\sigma^2 \ll \eta, \eta_A$ by using non-Gaussian states. We address this question in Sec. V F below and show numerically that indeed it appears to be possible.

E. Limit of small signals

Before moving to discuss our numerical results with non-Gaussian states, let us gain some more understanding of the small-signal $\sigma \to 0$ limit. In the lossless case, we have already observed that the QFI in this limit is $4\mathrm{Var}[\hat{x}]$ [51]. Here, we want to understand the lossy case in which the state is mixed before the channel is applied. In Appendix E, we prove the following general claim.

Claim 4. Given an initial state $\hat{\rho}$ and the random unitary channel

$$\Lambda_\sigma(\hat{\rho}) = \int_{-\infty}^{\infty} \mathrm{d}\theta \; p(\theta) \, \hat{U}_\theta \hat{\rho} \hat{U}_\theta^\dagger, \tag{37}$$

where $p(\theta) \sim \mathcal{N}(0, \sigma^2)$ and $\hat{U}_\theta$ is unitary, then the QFI is

$$\mathcal{I}_Q(\sigma = 0) = 4\langle \hat{H} \hat{\Pi}_\perp \hat{H} \rangle,$$

where $\hat{H} = i\hat{U}_0^\dagger \dot{\hat{U}}_0$ is the Hermitian generator at $\sigma = 0$, $\hat{\Pi}_\perp$ is the projection operator onto the null space of $\hat{\rho}$, and the expectation value is calculated with respect to $\hat{\rho}$, i.e., $\langle \hat{O} \rangle = \mathrm{Tr}[\hat{\rho}\hat{O}]$. An optimal measurement that attains the QFI is measurement of $\hat{U}_0^\dagger \hat{\Pi}_\perp \hat{U}_0$.

This result means that $\hat{H}$ needs to map some of $\hat{\rho}$ into its null space to obtain information about σ in the limit of $\sigma \to 0$. In particular, if $\hat{\rho}$ is full rank, then it suffers the Rayleigh curse.

For our random-displacement channel, given a pure initial state $|\psi\rangle$, then $\mathcal{I}_Q(\sigma = 0) = 4\langle \hat{x} \hat{\Pi}_\perp \hat{x} \rangle$, where $\hat{\rho} = \Lambda_\eta^{\mathrm{loss}}(|\psi\rangle\langle\psi|)$ is the state after the loss but before the encoding. This result is useful when $\mathcal{I}_Q(\sigma = 0)$ is nonzero. For example, it implies that, for a Fock state $|N\rangle$, the QFI is $\mathcal{I}_Q(\sigma = 0) = 2(1-\eta)^N(N+1)$ and is attained by number measurement.

What is the optimal initial state in this limit of $\sigma \to 0$? The ECQFI is not well defined unless we specify the order of limits since, by Eq. (29),

$$\lim_{\sigma \to 0} \mathcal{I}_Q^{\Lambda_\theta}(\theta) \neq \sup_{|\Psi\rangle} \lim_{\sigma \to 0} \mathcal{I}_Q^{(\Lambda_\theta \otimes 1_A)(|\Psi\rangle\langle\Psi|)}(\theta).$$

Nevertheless, we claim that $\lim_{\sigma \to 0} \mathcal{I}_Q^{\Lambda_\theta}(\theta)$ is the relevant quantity of interest since, in practice, the signal is small but finite $0 < \sigma \ll 1$ and we can only search for signals above the classical noise floor. In the following subsection, we discuss a family of initial states that are numerically optimal for a fixed small but finite $0 < \sigma \ll 1$.

An open question about this limit of $\sigma \to 0$ is whether the only single-mode pure initial states that are finite rank after the loss $\Lambda_\eta^{\mathrm{loss}}$ are either finite superpositions of coherent states or bounded in the Fock basis.

We discuss a generalization of Claim 4 in Appendix F. Also, in Appendix G, we give an example of how the upper bound on the ECQFI, which is analogous to Eq. (32), can be loose for random unitary channels acting on finite-dimensional systems.

F. Non-Gaussian states

We now explore whether there exist non-Gaussian states of the probe that can outperform the Gaussian states and saturate the ECQFI in Eq. (35) in the relevant high-loss regime $\sigma^2 \ll \eta, \eta_A$.

We start by analyzing the QFI with Fock states. As shown above, the QFI with a Fock state $|N\rangle$ is $\mathcal{I}_Q(\sigma = 0) = 2(1-\eta)^N(N+1)$. The optimal Fock state $|N\rangle$ is thus $N \approx -(1/\log(1-\eta)) - 1$, where the Bose enhancement factor $(N+1)$ balances with the loss factor $(1-\eta)^N$

to achieve a QFI of $\mathcal{I}_Q(\sigma=0) \approx -2/e(1-\eta)\log(1-\eta)$. In the limit of small η, $\mathcal{I}_Q(\sigma=0) \approx 2/e\eta$, which misses the ECQFI by roughly a factor of e, or a penalty of roughly 4.3 dB. This is shown in Fig. 4(b) for $\eta = 0.1$, for which the optimal Fock QFI is $\mathcal{I}_Q(\sigma=0) = 7.75$ at $N = 8, 9$, while the ECQFI is $2/\eta = 20$. That is, for a loss of 10%, we only need to prepare a Fock state of eight particles to get within 4.1 dB of the ultimate limit. This may be achievable experimentally in the microwave regime as, e.g., Fock states of up to 100 photons have been generated in superconducting microwave cavities [54] and Fock states of four photons have already been used in axion searches [55], as we discuss later. The key takeaway here is that for $\sigma^2 \ll \eta, \eta_A$, Fock states outperform all Gaussian states but do not attain the ECQFI in Eq. (35), although they may be a good first step experimentally.

To attain the ECQFI for $0 < \sigma^2 \ll \eta, \eta_A$, we now consider finite-energy GKP states. We review these states in Appendix H; in particular, the $|\mathrm{GKP}_\Delta\rangle$ family of finite-energy states given by a superposition of displaced SMSV states of width $\propto \Delta$ within a Gaussian window of width $\propto \Delta^{-1}$. To calculate the QFI with respect to σ numerically, we model $|\mathrm{GKP}_\Delta\rangle$ in the Fock basis of a truncated Hilbert space. For the lossless case and $\Delta \ll 1$, the overall variance of the pure initial state $|\mathrm{GKP}_\Delta\rangle$ is $\mathrm{Var}[\hat{x}] = \mathrm{Var}[\hat{p}] = \frac{1}{2}\Delta^{-2}$, which equals the average occupation number $\langle\hat{n}\rangle = N$ [56] such that the QFI for small σ is $4N$, as shown in Fig. 4(a). GKP states are promising for quantum computing and error-correction applications and much engineering is being done to produce them reliably [57]. For example, controlled displacements with a qubit can produce any quantum state of a bosonic mode, at least in principle [58]. Fock states, binomial-code states, and GKP states have been prepared experimentally using this method [59]. To saturate the QFI of the GKP state above, a non-Gaussian measurement is once again required. We numerically determine the Fock-basis coefficients of the optimal set of orthogonal states to project onto for $\sigma > 0$. These states are more complicated than number or momentum eigenstates. We do not know how these states relate to the GKP states or what they represent physically, but we suspect that realizing this measurement will be as challenging as producing the GKP states themselves. We defer determining how to realize this measurement to future work.

For the lossy case, e.g., a loss of $\eta = 0.1$ occurring before the encoding, the QFI from preparing finite-energy GKP states $|\mathrm{GKP}_\Delta\rangle$ converges numerically to the QFI of a TMSV state with noiseless ancilla ($\eta_A = 0$) for $N > 10$, i.e., $\Delta < 0.2$, as shown in Fig. 4(b). At higher energies, the QFI increases toward the ECQFI. For example, with $\eta = 0.1$ and $\sigma = 10^{-3}$, $|\mathrm{GKP}_\Delta\rangle$ attains a QFI above 19, within 95% of the ECQFI of 20, using states with $\langle\hat{n}\rangle = N > 100$ in a truncated Hilbert space of dimension $\mathcal{O}(1000)$. Note that we have only considered the small-signal σ behavior here, since it is the most relevant regime, and not the large-σ behavior shown in Fig. 5 for the other initial states.

We conjecture that preparing finite-energy GKP states of higher average number, which will require more peaks and a larger truncated Hilbert space, can get arbitrarily close to the ECQFI for any fixed $\sigma^2 \ll \eta$. We expect the convergence to be slow given that a TMSV state with noiseless ancilla ($\eta_A = 0$) only converges asymptotically. This conjecture is based on the above numerics and, heuristically, the connection between the GKP and TMSV infinite-energy states discussed below. It also would be interesting to understand the performance in the limit of $\sigma \to 0$.

Moreover, we conjecture that the ECQFI for a fixed finite $\langle\hat{n}\rangle = N$ is saturated by preparing a TMSV state with noiseless ancilla for all N. Numerically, we have searched for different non-Gaussian single-mode states that perform better than a TMSV state with noiseless ancilla with the same large $\langle\hat{n}\rangle = N$ but have not found any. We describe our numerical methods in Appendix I. Briefly, we have found that sparse superpositions of finitely many Fock states also outperform Fock states and approach the ECQFI at high energies. For example, we have found a sparse state $|\psi_{\mathrm{num.}}\rangle = \sum_{j=0}^{23} c_j |20j\rangle$ with $\langle\hat{n}\rangle = 158.9$ and a QFI of 18.4, within 9% of the ECQFI of 20 for $\sigma^2 = 10^{-6}$ and $\eta = 0.1$. Intuitively, e.g., the signal trajectory from $|0\rangle$ to $|1\rangle$ dominates the loss trajectory from $|20\rangle$ to $|1\rangle$ for finite signals $\sigma^2 = 10^{-6}$ and $\eta = 0.1$. These sparse states are similar to optimized binomial quantum error-correcting codes [60].

Finally, we remark that the high numerical performance of the finite-energy GKP states for sensing a random displacement in the presence of loss is intriguing, as the GKP infinite-energy state was originally designed for the correction of random-displacement noise [52]. Both the GKP infinite-energy state and the TMSV infinite-energy state can be used to form error-correction codes that are sensitive to random displacement signals along $\hat{p}$ yet protected against random-displacement noise along $\hat{x}$. (We discuss this further in Appendix D.) In the next subsection, we will show that this property makes these states resilient to classical noise along $\hat{x}$. It would be interesting to understand how this property is related to their performance in the presence of loss.

G. Classical noise case

We now address the case of significant classical noise such that the total channel is $\Lambda'_\sigma = \Lambda_\sigma \circ \Lambda^{\mathrm{noise}}_{\Sigma_C}$ for a given classical noise matrix Σ_C in Eq. (22). Losses are not present unless otherwise noted. In Appendix J, we prove that the ECQFI and optimal initial state depend on Σ_C as follows.

First, suppose that the classical noise is confined to the same quadrature, $\hat{p}$, as the signal, i.e., $\Sigma_C = \mathrm{diag}\left(0, \sigma_p^2\right)$ with $\sigma_p > 0$. Then, the Rayleigh curse is unavoidable for $\sigma \ll \sigma_p$ and the noiseless ECQFI cannot be recovered, since there is no way to distinguish the signal and the noise. For a fixed finite $\sigma > 0$, the ECQFI is

$$\mathcal{I}_Q^{\Lambda'_\sigma, N}(\sigma) = \frac{4\sigma^2}{(\sigma^2 + \sigma_p^2)\left[2(\sigma^2 + \sigma_p^2) + \xi_N^{-1}\right]}, \tag{38}$$

which reduces to Eq. (23) in the vacuum case ($N = 0$) with $\sigma_x = 0$. The ECQFI in Eq. (38) is attained for a given $\langle \hat{n} \rangle = N$ by preparing the appropriate SMSV state and performing the noiseless optimal measurement—antisqueezing followed by a number-resolving measurement. In the case of parallel classical noise σ_p and loss η, in the high-energy limit the ECQFI becomes

$$\mathcal{I}_Q^{\Lambda'_\sigma}(\sigma) = \frac{2\sigma^2}{(\eta + \sigma^2 + \sigma_p^2)(\sigma^2 + \sigma_p^2)}, \tag{39}$$

such that, in the experimentally relevant regime of $\sigma^2 \ll \sigma_p^2 \ll \eta$, the ECQFI approaches $2\sigma^2/\eta\sigma_p^2$, which exhibits the Rayleigh curse. In comparison, the QFI from preparing an SMSV state in the high-energy limit is

$$\mathcal{I}_Q(\sigma) = \frac{8\sigma^2}{\left[\eta + 2\left(\sigma^2 + \sigma_p^2\right)\right]^2}, \tag{40}$$

which can be attained by performing the above scheme—antisqueezing followed by a number-resolving measurement. It can also be attained by performing a quadrature measurement, but only in the high-energy limit, as otherwise quadrature measurement is suboptimal for an SMSV state at any given finite N. The ratio of the ECQFI in Eq. (39) to the QFI from preparing an SMSV state in Eq. (40) is thus

$$\mathfrak{g} := \frac{\mathcal{I}_Q^{\Lambda'_\sigma}(\sigma)}{\mathcal{I}_Q(\sigma)} = \frac{\left(\eta + 2\varsigma^2\right)^2}{4\varsigma^2\left(\eta + \varsigma^2\right)}, \tag{41}$$

which is a function of η/ς^2, where $\varsigma^2 := \sigma^2 + \sigma_p^2$ and $\mathfrak{g} \geq 1$ by definition of the ECQFI. If $\eta \ll \varsigma^2$, then this ratio $\mathfrak{g}$ approaches one and preparing an SMSV state is close to optimal. In the above experimentally relevant regime of $\sigma^2 \ll \sigma_p^2 \ll \eta$, however, $\eta \gg \varsigma^2$ and $\mathfrak{g}$ approaches $\eta/4\sigma_p^2$, which is large. When $\mathfrak{g}$ is large, i.e., the classical noise is sufficiently small compared to the loss, then preparing a TMSV state with noiseless ancilla or preparing non-Gaussian states can outperform an SMSV state. If the loss on the ancilla can be made smaller than the classical noise on the system, i.e., $\eta_A \ll \sigma_p^2 \ll \eta$, then a TMSV state with lossy ancilla can also outperform an SMSV state. Developing low-loss quantum memories may thus be a viable direction to pursue experimentally as an alternative to non-Gaussian-state preparation, as the required measurement scheme for the TMSV state is simply Gaussian operations (beam-splitter and squeezing unitaries) followed by photodetection.

Second, suppose that the classical noise is confined to the quadrature opposite to the signal, i.e., $\Sigma_C = \mathrm{diag}\left(\sigma_x^2, 0\right)$, with $\sigma_x > 0$. The effect of this opposite quadrature noise is fundamentally different from the effect of noise in the same quadrature as the signal; unlike the same quadrature case, noise in the opposite quadrature can be overcome in the limit of large N and does not necessarily lead to a Rayleigh curse. For fixed finite values of $\sigma, \sigma_x > 0$, the noiseless ECQFI in Eq. (32) can be recovered by preparing an SMSV state in the high-energy limit of $N \to \infty$. Intuitively, the classical noise can be squeezed after the encoding while simultaneously antisqueezing the signal. A given SMSV state with fixed finite $\langle \hat{n} \rangle = N$, however, still exhibits the Rayleigh curse as $\sigma \to 0$, such that an SMSV state may not be optimal for a given $\langle \hat{n} \rangle = N$ and σ. In comparison, there exist finite-energy states, e.g., a TMSV state with noiseless ancilla, that do not exhibit the Rayleigh curse and that recover their respective noiseless QFI in the high-energy limit, e.g., $\lim_{N\to\infty} \lim_{\sigma\to 0} \mathcal{I}_Q(\sigma) = 4N$ for a TMSV state.

Finally, suppose that the classical noise appears in both quadratures, i.e., $\Sigma_C = \mathrm{diag}\left(\sigma_x^2, \sigma_p^2\right)$ with $\sigma_x, \sigma_p > 0$. For example, suppose that the classical noise is isotropic with $\Sigma_C = \mathrm{diag}\left(\sigma_C^2, \sigma_C^2\right)$ and $\sigma_C > 0$. Then, the Rayleigh curse is unavoidable, as $\sigma \to 0$ with σ_C fixed and the noiseless ECQFI cannot be recovered. The ECQFI with $\sigma_x = 0$ in Eq. (38), however, can be recovered for a fixed finite $\sigma > 0$ by preparing an SMSV state in the high-energy limit despite $\sigma_x > 0$. In the isotropic case, this means simply replacing σ_p by σ_C in Eq. (38).

VI. SIMULTANEOUS ESTIMATION OF THE MEAN AND VARIANCE

We have focused so far on estimating the variance of a random process. In contrast, most work on quantum metrology up to now has been dedicated to estimating the mean value of a signal rather than its variance. Formally, estimating the deterministic encoding of a signal is equivalent to mean estimation in the zero-variance limit. Here, we unify these two efforts as we consider the optimal simultaneous estimation of the mean and the variance. We emphasize that we only consider the case in which the initial state is vacuum such that any loss occurring before the encoding can be ignored.

A. Review of quantum multiparameter estimation

We now introduce the tools of quantum multiparameter estimation, in contrast to single-parameter estimation as reviewed in Sec. II D. In the multiparameter case, a vector of real parameters $\vec{\theta}$ is encoded in a quantum state $\hat{\rho}(\vec{\theta})$. Given M independent and identically distributed measurements of the probability distribution $p(x|\vec{\theta})$, the CCRB provides a tight lower bound on the covariance matrix, $\Sigma_{\vec{\theta}}$, of the estimators of $\vec{\theta}$ as $\Sigma_{\vec{\theta}} \geq (1/M)[\mathcal{I}_C(\vec{\theta})]^{-1}$, where matrix inequalities $A \geq B$ are interpreted henceforth as $A - B$ being positive semidefinite and the classical Fisher information matrix (CFIM) is

$$\mathcal{I}_C(\vec{\theta})_{i,j} = \int_{-\infty}^{\infty} \mathrm{d}x \, \frac{\partial_{\theta_i} p(x|\vec{\theta}) \partial_{\theta_j} p(x|\vec{\theta})}{p(x|\vec{\theta})}.$$

As in the single-parameter case, the CCRB is asymptotically saturable with maximal likelihood estimators and the CFIMs of two independent observations sum: if $p(x|\vec{\theta})$ and $p(y|\vec{\theta})$ are independent distributions, then the total CFIM from observing one outcome from each is the sum of the two CFIMs.

The QCRB provides a further lower bound on the covariance matrix of the estimators given any measurement strategy,

$$\Sigma_{\vec{\theta}} \geq \frac{1}{M}[\mathcal{I}_C(\vec{\theta})]^{-1} \geq \frac{1}{M}[\mathcal{I}_Q(\vec{\theta})]^{-1}.$$

Here, the quantum Fisher information matrix (QFIM) is defined as

$$[\mathcal{I}_Q(\vec{\theta})]_{ij} = \frac{1}{2}\mathrm{Tr}[\hat{\rho}\{\hat{L}_i, \hat{L}_j\}],$$

where the Hermitian operator $\hat{L}_i$ is the symmetric logarithmic derivative (SLD) of $\hat{\rho}$ with respect to θ_i, which is defined implicitly by the equation $\partial_{\theta_i}\hat{\rho} = \frac{1}{2}\{\hat{L}_i, \hat{\rho}\}$. Given the spectral decomposition of $\hat{\rho} = \sum_j p_j |\phi_j\rangle\langle\phi_j|$, the SLDs are given by

$$\hat{L}_i = \sum_{k,l} \frac{2\langle\phi_k|\partial_{\theta_i}\hat{\rho}|\phi_l\rangle}{p_k + p_l} |\phi_k\rangle\langle\phi_l|,$$

where the sum runs over only k, l such that $p_k + p_l > 0$. Similarly to the single-parameter case, the QFIM for a product state $\hat{\rho}_1(\theta) \otimes \hat{\rho}_2(\theta)$ is simply the sum of the individual QFIMs.

Unlike the single-parameter case, however, this bound is not saturable *a priori*. A necessary and sufficient condition for the asymptotic saturability of the QFIM is the weak-commutativity condition [61]

$$\mathrm{Tr}\Big[\hat{\rho}[\hat{L}_i, \hat{L}_j]\Big] = 0, \quad \forall i, j. \tag{42}$$

If this weak-commutativity condition holds, then a joint (i.e., collective) measurement of many copies of $\hat{\rho}$ may be necessary to asymptotically saturate the QFIM. (If this condition does not hold, then the Holevo CRB would instead be required to find a tight bound [17,62].)

B. QCRB for simultaneous estimation

We now consider the problem of estimating μ and σ from the encoding in Eq. (15) but with $p(\alpha) \sim \mathcal{N}(\mu, \sigma^2)$. We assume that the initial state is the vacuum state, such that it becomes the displaced squeezed thermal state with $\vec{\mu} = (0, \mu)$ and $\Sigma = \mathrm{diag}\left(\frac{1}{2}, \frac{1}{2} + \sigma^2\right)$. The QFIM about parameters $\vec{\theta}$ encoded in a single-mode Gaussian state is given by a generalization of Eq. (19) [63]

$$[\mathcal{I}_Q(\vec{\theta})]_{jk} = \frac{\mathrm{Tr}\left[\Sigma^{-1}[\partial_{\theta_j}\Sigma]\Sigma^{-1}[\partial_{\theta_k}\Sigma]\right]}{2(1+\gamma^2)} + \frac{2[\partial_{\theta_j}\gamma][\partial_{\theta_k}\gamma]}{1-\gamma^4} + [\partial_{\theta_j}\vec{\mu}]^{\mathrm{T}}\Sigma^{-1}[\partial_{\theta_k}\vec{\mu}]. \tag{43}$$

Let $\vec{\theta} = (\mu, \sigma)^{\mathrm{T}}$. Then the QFIM for the vacuum state is

$$\mathcal{I}_Q(\mu, \sigma) = \begin{bmatrix} \frac{2}{1+2\sigma^2} & 0 \\ 0 & \frac{2}{1+\sigma^2}, \end{bmatrix} \tag{44}$$

such that the QCRB per independent measurement is $\Delta^2\mu = \Delta^2\sigma = \frac{1}{2}$ for $\sigma \ll 1$. We emphasize here that the M measurements must be independent and identically distributed, which means that the stochastic signal must be uncorrelated between the M measurements, i.e., all off-diagonal components are zero in the covariance matrix of the M experiments in time or space. This assumption will be relevant later when we discuss the application to searching for a stochastic gravitational-wave background.

This bound is the best we could expect for: the estimation variances of σ and μ given by it are the same as in each respective single-parameter case. If the QCRB is tight, therefore, we can estimate the two parameters simultaneously with a precision for each that is identical to its respective single-parameter case.

We check the saturability of the QCRB using the weak-commutativity condition in Eq. (42). By calculating the SLDs for the Gaussian channel [37,64], we find that Eq. (42) holds for the vacuum case. Our remaining task is to determine the optimal asymptotic measurement scheme that saturates the QCRB and whether it must be a joint measurement acting collectively on multiple copies of the state.

C. Separate measurements

For the vacuum case, while a number-resolving measurement of $\hat{n}$ is optimal for estimating σ alone (in the $\mu = 0$ case) and μ alone (in the $\sigma = 0$ case), it is inefficient for estimating the two parameters simultaneously.

Analogously to Eq. (21), the probability of a number-measurement outcome n is

$$p(n) = \int_{-\infty}^{\infty} \mathrm{d}\alpha \ \left|\langle n|\alpha'\rangle\right|^2 p(\alpha)$$
$$= \frac{e^{\frac{-\mu^2}{2(\sigma^2+1)}}\sigma^{2n}}{n!\left(\sigma^2+1\right)^{n+\frac{1}{2}}}\left[\mu^{2n}\frac{(\sigma^{2n}-[4(\sigma^2+1)]^n)}{2^n\sigma^{4n}(\sigma^2+1)^n}\right.$$
$$\left.+n!\mathcal{L}_n^{(-\frac{1}{2})}\left(-\frac{2\mu^2}{4\sigma^2(\sigma^2+1)}\right)\right],$$

where $\mathcal{L}_n^{(\alpha)}(x)$ is the generalized Laguerre polynomial. In the relevant limit of $\mu,\sigma \ll 1$, the probabilities are $p(0) \approx 1-a$ and $p(1) \approx a$, where $a = \frac{1}{2}(\mu^2+\sigma^2)$. These probabilities are degenerate with respect to μ and σ such that only $\sqrt{\mu^2+\sigma^2}$ can be estimated and μ cannot be distinguished from σ. Indeed, the CFIM for a number measurement is

$$\mathcal{I}_C^{\hat{n}}(\mu,\sigma) = \frac{2}{\left(\mu^2+\sigma^2\right)\left(1-\mu^2+\sigma^2\right)}\begin{bmatrix}\mu^2 & \mu\sigma\\ \mu\sigma & \sigma^2,\end{bmatrix} \tag{45}$$

which is singular, and the only nonzero eigenvalue corresponds to the parameter $\sqrt{\mu^2+\sigma^2}$.

For simultaneous estimation of μ and σ, therefore, we need to add a measurement that will break this degeneracy. This can be accomplished by adding a quadrature measurement. In the limit of $\mu,\sigma \ll 1$, quadrature measurement is optimal for estimating μ yet provides no information about σ since it is Rayleigh cursed, as can be observed from the CFIM

$$\mathcal{I}_C^{\hat{p}}(\mu,\sigma) = \begin{bmatrix}\frac{2}{1+2\sigma^2} & 0\\ 0 & \frac{8\sigma^2}{\left(1+2\sigma^2\right)^2}\end{bmatrix} \xrightarrow[\sigma\to 0]{} \begin{pmatrix}2 & 0\\ 0 & 0\end{pmatrix}.$$

Performing a combination of number and quadrature measurements should thus allow for simultaneous estimation. Suppose that we perform many measurements of which the proportion of number measurements is p for some $p \in (0,1)$. For this protocol, the CCRB per measurement is

$$\Sigma = \left[p\mathcal{I}_C^{\hat{n}} + (1-p)\mathcal{I}_C^{\hat{p}}\right]^{-1}$$
$$= \frac{1}{2(1-p)}\begin{pmatrix}1 & -\frac{\mu}{\sigma}\\ -\frac{\mu}{\sigma} & \frac{1-p}{p}+\frac{\mu^2}{p\sigma^2},\end{pmatrix}$$

such that the minimum MSEs are $\Delta^2\mu = 1/2(1-p)$ and $\Delta^2\sigma = 1/2p + \mu^2/2(1-p)p\sigma^2$. The QCRB per measurement of $\Delta^2\mu = \Delta^2\sigma = \frac{1}{2}$ is, therefore, not saturated by this separate measurement strategy.

Intuitively, $\hat{p}$ provides an optimal measurement of μ and $\hat{n}$ provides an optimal measurement of σ for a sufficiently small μ. These observables, however, only weakly commute, since $\left\langle[\hat{p},\hat{n}]\right\rangle = -i\left\langle\hat{x}\right\rangle = 0$, such that these measurements cannot be performed simultaneously. Since only one of them is performed each time, $\Delta^2\mu$ suffers from a factor-of-$(1-p)$ gap from the QCRB and, similarly, $\Delta^2\sigma$ suffers from a gap of p. Although $\Delta^2\sigma$ suffers an additional uncertainty of $\mu^2/2p(1-p)\sigma^2$, this can be suppressed asymptotically by reducing μ adaptively, i.e., by estimating μ and then nulling by displacing the state back toward the vacuum using the estimate of μ before the number measurement. We describe a suitable adaptive protocol further in Appendix K.

In various cases, the figure of merit is a weighted sum of the variances, i.e., $\mathrm{Tr}[W\Sigma]$, where $W = 2\,\mathrm{diag}(1-w,w)$, with $w \in (0,1)$, is the weight matrix. Informally, the weight w represents how much more we care about estimating σ than μ or vice versa. We thus want to know the optimal proportion of number measurements p for a given w to minimize this figure of merit:

$$\mathrm{Tr}[W\Sigma] = \frac{(-2pw+p+w)\sigma^2+w\mu^2}{(1-p)p\sigma^2}.$$

The optimal proportion can be shown to be

$$p = \frac{w}{w+\sqrt{\frac{w\left[w\mu^2+(1-w)\sigma^2\right]}{\mu^2+\sigma^2}}},$$

such that the optimal figure of merit is

$$\mathrm{Tr}[W\Sigma] = 1 + \frac{2\mu^2 w}{\sigma^2} + 2\frac{\sqrt{\mu^2+\sigma^2}}{\sigma^2}\sqrt{w\left[w\mu^2+(1-w)\sigma^2\right]},$$

which does not saturate the QCRB of $\mathrm{Tr}[W\mathcal{I}_Q^{-1}] = 1$, except in the limit of $w = 0$, which corresponds to estimating μ by solely performing quadrature measurements. This is due to two penalty terms. The first penalty term, $2\mu^2 w/\sigma^2$, stems from nonzero μ and can be asymptotically removed with an adaptive protocol. The second penalty term is fundamental and appears for any $w \neq 0,1$. In the limit $\mu \to 0$, which is attainable with an adaptive protocol, the optimal p is $p = \sqrt{w}/(\sqrt{w}+\sqrt{1-w})$ and the optimal figure of merit is $1+2\sqrt{w(1-w)}$, which attains the QCRB for $w = 0,1$. The worst-case scenario corresponds to equal weights, i.e., $w = 0.5$, where the optimal proportion of number and quadrature measurements is also equal, i.e., $p = 0.5$, and the figure of merit is $\mathrm{Tr}[W\Sigma] = 2+2\mu^2/\sigma^2$, since $\Delta^2\mu = 1$ and $\Delta^2\sigma = 1+2\mu^2/\sigma^2$.

Numerical results of the adaptive and nonadaptive schemes are shown in Fig. 6, where we plot the MSE in

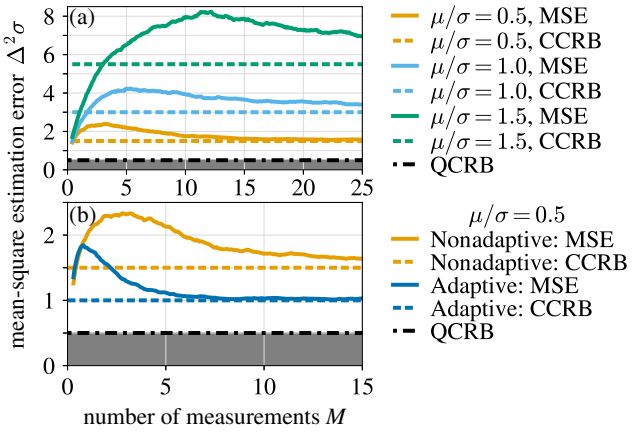

FIG. 6. The numerical mean-square estimation error (MSE) compared to the Cramér-Rao bounds (CRBs) for simultaneous estimation of μ and σ versus the number of independent measurements M. The shaded gray region is inaccessible with the input state fixed as the vacuum state. (a) The nonadaptive-separate-measurements scheme with $\sigma = 0.05$ and different ratios of μ/σ. (b) Adaptive separate measurements are more precise than the nonadaptive strategy with $\sigma = 0.05$ and $\mu/\sigma = 0.5$ but neither reach the QCRB asymptotically as $M \to \infty$.

estimating σ. Denoting the M outcomes of the quadrature measurement and number measurement as $\{p_i\}_{i=1}^M$ and $\{n_i\}_{i=1}^M$ respectively, the estimator of μ is $\tilde{\mu} = (1/M)\sum_i p_i$ and the estimator of σ is $\varsigma = \sqrt{(2/M)\sum_i n_i - \tilde{\mu}^2}$. The distribution of ς is not Gaussian and hence it only converges to the CCRB of $\Delta^2\sigma = 1$ asymptotically, as shown in Fig. 6.

While we have shown that there is no combination of number and quadrature measurements that can saturate the QCRB, we would need to calculate the Nagaoka-Hayashi bound to prove that no other separate measurement scheme can attain the QCRB [65,66].

D. Joint measurements

Suppose that we perform joint measurements on the M independent and identical copies of the final state $\hat{\rho} = \Lambda_\sigma(|0\rangle\langle 0|)$. This collective state $\hat{\rho}^{\otimes M}$ is an M-mode Gaussian state with $2M$-by-1 mean vector $\vec{\mu} = (0, \mu, 0, \mu, \cdots, 0, \mu)^{\mathrm{T}}$ and $2M \times 2M$ covariance matrix $\Sigma = \bigoplus_{j=1}^M \mathrm{diag}\left(\frac{1}{2}, \frac{1}{2} + \sigma^2\right)$ in the basis $(\hat{x}_1, \hat{p}_1, \cdots, \hat{x}_M, \hat{p}_M)^{\mathrm{T}}$. This M-mode state can be transformed to the (anti)symmetric basis such that $\vec{\mu} = (\sqrt{M}\mu, 0, 0, 0, \cdots, 0, 0)^{\mathrm{T}}$, where the symmetric mode is listed first and Σ is unchanged. We emphasize that the elements of this (anti)symmetric basis are orthogonal and commute. This shows that measuring the $\hat{p}$ quadrature of the symmetric mode is a sufficient statistic for μ. The remaining $M-1$ antisymmetric modes can then be used to measure σ performing a number measurement on each mode. Absorbing the M factors from the CRBs, the CFIM for this joint measurement is $2\,\mathrm{diag}(M, M-1)$, which saturates the QFIM in Eq. (44) of $2M\,\mathrm{diag}(1,1)$ in the asymptotic limit of $M \to \infty$.

It is therefore possible to saturate the QFIM if we can perform joint measurements on M independent and identical copies of the final state. These M copies can correspond to different experiments distributed in space or time. For uncorrelated experiments distributed in space, the required transformation to symmetric and antisymmetric modes could be done using a sequence of beam-splitter unitaries; while for uncorrelated experiments distributed in time, the transformation may be done using quantum memories, as further discussed in Sec. VII A.

This protocol is an instance of the following general statement for any quantum system.

Claim 5. Given an initial pure state $|\psi\rangle$ and the random unitary channel

$$\Lambda_{\sigma,\mu}(|\psi\rangle\langle\psi|) = \int_{-\infty}^{\infty} \mathrm{d}\theta\; p\,(\theta)\; \hat{U}_\theta |\psi\rangle\langle\psi| \hat{U}_\theta^\dagger, \qquad (46)$$

where $p(\theta) \sim \mathcal{N}(\mu, \sigma^2)$ and $\hat{U}_\theta = \exp(-i\theta\hat{H})$, then the QFI with respect to either $\mu \to 0$ or $\sigma \to 0$ equals $4\mathrm{Var}[\hat{H}]$ and the QFIM is simultaneously saturable with a joint measurement in the asymptotic limit of $M \to \infty$.

The proof of Claim 5 is given in Appendix L. The key idea is the same as in our particular case above: while the parameter μ displaces the state in the symmetric subspace of the M copies, σ takes the state out of the symmetric subspace. The QFI with respect to μ is saturated by measuring a suitable basis in the symmetric subspace, while the QFI with respect to σ is saturated by measuring the projection onto the antisymmetric subspace. Similar protocols to saturate the QCRB by projecting onto the symmetric and antisymmetric subspaces have previously been proposed for superresolution in imaging [67] and estimating phase diffusion in qubits [68].

Lastly, we remark that preparing collective initial states could improve the joint measurement of the mean and variance in the presence of losses. The optimal collective strategy is to prepare the symmetric mode in an SMSV state to sense the mean μ and to prepare the antisymmetric modes in the optimal non-Gaussian states to sense the standard deviation σ, e.g., GKP states.

VII. STOCHASTIC WAVEFORM ESTIMATION, IMPLEMENTATION, AND APPLICATION

We now solve the continuous-estimation problem of determining $S_{yy}(\Omega)$ at each frequency Ω. By the chain rule from σ to $S_{yy}(\Omega)$ using Eq. (14), the QFI with respect to

$S_{yy}(\Omega)$ is

$$\mathcal{I}_Q[S_{yy}(\Omega)] = \frac{G_{py}(\Omega)^2}{8\sigma^2}\mathcal{I}_Q(\sigma). \tag{47}$$

where we include a factor of 2 to account for adding the information from the independent real and imaginary parts of the mode at Ω.

By Eqs. (39) and (47), the ECQFI for sensing $S_{yy}(\Omega)$ in the presence of an input loss η and signal-referred parallel classical noise $\mathcal{S}^C(\Omega)$ from Eq. (26) is thus equal to

$$\mathcal{I}_Q^{\Lambda'_\sigma}[S_{yy}(\Omega)] = \left[S_{yy}(\Omega) + \mathcal{S}^C(\Omega) + 2\eta\mathcal{S}^{Q,\mathrm{vac}}(\Omega)\right]^{-1} \cdot [S_{yy}(\Omega) + \mathcal{S}^C(\Omega)]^{-1}, \tag{48}$$

where we have used Eq. (27) to relate the gain $G_{py}(\Omega)$ to the signal-referred quantum shot noise from the vacuum, $\mathcal{S}^{Q,\mathrm{vac}}(\Omega)$. For example, if the initial state is vacuum, then the QFI with respect to $S_{yy}(\Omega)$ is given above by setting $\eta = 1$ and is attained by performing a number-resolving measurement at each frequency. To attain the ECQFI in Eq. (48) at each frequency Ω and increase the QFI beyond the vacuum case by roughly a factor of $1/\eta$ in the vanishing-signal limit, we need to instead prepare non-Gaussian initial states and perform non-Gaussian measurements independently in the $\cos(\Omega t)$ and $\sin(\Omega t)$ signal components of Eq. (10). This is experimentally demanding, as we discuss in detail later.

Let us compare the ECQFI in Eq. (48) to the QFI from preparing an SMSV state. In the limit where the signal and classical noise are small compared to the quantum shot noise, i.e., $S_{yy}(\Omega), \mathcal{S}^C(\Omega) \ll \eta\mathcal{S}^{Q,\mathrm{vac}}(\Omega)$, the ECQFI in Eq. (48) may be approximated as

$$\mathcal{I}_Q^{\Lambda'_\sigma}[S_{yy}(\Omega)] \approx \left\{2\eta\mathcal{S}^{Q,\mathrm{vac}}(\Omega)[S_{yy}(\Omega) + \mathcal{S}^C(\Omega)]\right\}^{-1}.$$

In comparison, by Eqs. (40) and (47), the QFI from preparing an SMSV state is attained in the high-energy limit by performing a quadrature measurement and is given by

$$\mathcal{I}_Q[S_{yy}(\Omega)] = [S_{yy}(\Omega) + \mathcal{S}^C(\Omega) + \eta\mathcal{S}^{Q,\mathrm{vac}}(\Omega)]^{-2},$$

which is approximately $\mathcal{I}_Q[S_{yy}(\Omega)] \approx [\eta\mathcal{S}^{Q,\mathrm{vac}}(\Omega)]^{-2}$ in the shot-noise–dominated limit. (For example, in the vacuum-injection case, the CFI from quadrature measurement is given by setting $\eta = 1$.) The ratio between the ECQFI and the QFI from preparing an SMSV state is thus large and given by [cf. Eq. (41)]

$$\frac{\mathcal{I}_Q^{\Lambda'_\sigma}[S_{yy}(\Omega)]}{\mathcal{I}_Q[S_{yy}(\Omega)]} \approx \frac{\eta\mathcal{S}^{Q,\mathrm{vac}}(\Omega)}{2\left[S_{yy}(\Omega) + \mathcal{S}^C(\Omega)\right]},$$

such that experiments where the signal and classical noise are small compared to the quantum shot noise can be significantly accelerated by preparing non-Gaussian states compared to preparing an SMSV state.

In principle, there is a continuum of independent parameters $S_{yy}(\Omega)$ to estimate the spectrum but, in practice, we are often only interested in distinguishing spectra from a family $S_{yy}(\Omega|\vec{\theta})$ parametrized by a finite set of real parameters, $\vec{\theta}$. For example, below, in Sec. VII B, we will examine the single-parameter case in which the spectrum $S_{yy}(\Omega|\alpha) = \alpha\Phi(\Omega)$ is solely determined by the unknown scale parameter α, since the morphology $\Phi(\Omega)$ is known. Additionally, given a limited measurement bandwidth $\Delta\Omega$ and a total measurement interval ΔT, it is only possible to make a finite number of measurements $M = \Delta\Omega\Delta T/\pi$. The task then becomes to estimate $\vec{\theta}$ from the measurement outcomes $\{S_{yy}(\Omega_j|\vec{\theta})\}_{j=1}^M$ at the frequencies $\{\Omega_j\}_{j=1}^M$ such that the QFIM is

$$[\mathcal{I}_Q(\vec{\theta})]_{ij} = \sum_{k=1}^{M}\mathcal{I}_Q[S_{yy}(\Omega_k|\vec{\theta})]\frac{\partial S_{yy}(\Omega_k|\vec{\theta})}{\partial\theta_i}\frac{\partial S_{yy}(\Omega_k|\vec{\theta})}{\partial\theta_j}, \tag{49}$$

where, since the modes at each frequency are independent of each other by linearity, the total QFIM is the sum of the individual QFIMs. Note that the optimal measurement at each frequency to attain the QFIM depends on the initial state.

We may generalize the choice of temporal basis. We have so far assumed that we measure the state of the Fourier component at a given frequency, equivalent to measuring the cosine and sine functions in the time domain as given by Eq. (10). We generalize this description, as the Fourier basis is expressed in a formal limit that is not experimentally accessible. Instead, we may measure the state in an arbitrary temporal basis $\{w_j\}_{j=1}^M$ of orthogonal functions such that $\int_{-\infty}^{\infty}\mathrm{d}t\, w_j(t)w_k^*(t) = \delta_{jk}$. The temporal basis can be chosen to fit the given signal model and, in particular, each temporal mode, w_j, can be associated with a parameter of interest, θ_j. The state of each temporal mode is described by the density matrix $\hat{\rho}_k := \iint_{-\infty}^{\infty}\mathrm{d}\tau\mathrm{d}\tau'\, w_k(\tau)\hat{\rho}(\tau,\tau')w_k^*(\tau')$ and the operators that act on this state may be similarly constructed, e.g., the annihilation operator is $\hat{a}_k := \int_{-\infty}^{\infty}\mathrm{d}\tau\, w_k(\tau)\hat{a}(\tau)$ [69]. Here, $\hat{\rho}(\tau,\tau')$ is the density matrix of a single transverse mode of the outgoing bosonic field where the wave parameter $\tau = t - cz$ depends on time t and the distance along the spatial propagation axis z, along which the mode propagates at speed c. For example, this state could represent a paraxial Gaussian beam of light or a wire acting as a transmission line. Note that we still assume that the noise is stationary.

Using a temporal measurement basis is a significant departure from the deterministic displacement case, in which we record the time series from quadrature measurement and can then choose any temporal basis *post*

hoc using classical processing. Here, we instead directly perform, e.g., a number measurement of each temporal mode, which is highly advantageous compared to quadrature measurement for sensing stochastic signals. The measurement comes at the cost, however, of not being able to later process the data classically to study a different temporal basis. Implementations of the independent preparation of non-Gaussian states and performance of non-Gaussian measurements of each temporal mode must be developed for optimal stochastic waveform estimation. We discuss two possible pathways below—resonant filters and cooperative quantum memories.

A. Potential experimental implementations

We first consider the case of preparing the vacuum state and measuring in the Fourier basis. The QFI for an input vacuum state at a given frequency is saturated by number measurement at that frequency. This measurement can be achieved using a resonant filter to extract a band of Fourier modes followed by a number-resolving measurement. For example, in the optical domain, this may be done using resonating filter cavities and low-background photodetectors [24,70]. This implementation, however, can demand a high number of resonant filters. While in principle we only require one filter per parameter θ, in practice we may require additional filters, e.g., to remove background classical noise. Optically, implementing narrow-band filter cavities may be difficult because long round-trip lengths and low scattering loss are required. Furthermore, implementing the measurement in an arbitrary temporal basis or, e.g., the non-Gaussian measurement required for a GKP finite-energy input state may be challenging to do all optically.

A more promising pathway is to instead prepare the initial states in a dedicated ancilla and then couple the states from that ancilla to the appropriate temporal basis of the incoming bosonic mode of the main device. The time reversal of the coupling process can then load the outgoing bosonic mode from the device into an ancilla to be measured. A key example of this protocol would be to use the two-photon Raman transitions of atoms [71] to transmit and receive states in a given temporal basis. To accomplish this in the future with atom-based experimental platforms, we outline the following five requirements:

(R1) preparing non-Gaussian bosonic initial states using the electronic states of atoms or ensembles of atoms
(R2) implementing the optimal non-Gaussian projective measurements using atomic states
(R3) multiplexing such preparation and measurement procedures across many such atomic ensembles
(R4) achieving high cooperativity couplings that minimize transmission loss into and out of these ancillas
(R5) creating long-lived or distantly distributed memories

In general, an optimal stochastic waveform search would demand simultaneous implementation of (R1)–(R4), which have some competing requirements. (R1) and (R2) are feasible with various platforms and schemes, e.g., preparing Fock states with efficient projective readout [72–78] and coherent population trapping [79,80] of atomic states. (R3) is feasible using lattice traps [81] or tweezer arrays [82]. (R4) is feasible using the Purcell effect via cavity enhancement or nanophotonics [83]. Finally, (R5) is required to realize the optimal joint measurement of the associated mean and variance from Sec. VI D using many independent and identical copies of the state, which may correspond to experiments distributed distantly in space (e.g., observatory networks jointly estimating common deterministic and uncorrelated stochastic signals) or in time (e.g., rare events with collective deterministic and uncorrelated stochastic properties). (R5) has been demonstrated in many kinds of systems [84–86] but its integration with the other requirements is an ongoing effort in the field of emerging quantum technologies.

While, in the above discussion, we emphasize optical sensors and atomic memories, microwave systems are similarly promising. They can leverage superconducting cavities and nonlinear junctions [87–91] to form qubits that can efficiently produce non-Gaussian states [92] across multiplexed devices [93] for fundamental physics and sensing applications [94]. The need for low classical noise suggests that the greatest benefit will be for microwave sensors operating in the $\gg$200-MHz range that can search for signals above the thermal black-body radiation of dilution refrigerators [95], i.e., in the $\sigma \gg \sigma_C$ regime [96].

B. Applications

We now consider fundamental-physics applications that highlight the metrological advantages of non-Gaussian-state preparation and measurement for sensing stochastic signals.

1. Quantum gravity

The Verlinde-Zurek theory of quantum gravity predicts large-scale "geontropic" length fluctuations that scale with the holographic surface area of a given volume of spacetime [1,2]. This signal is predicted to manifest as excess noise in the phase quadrature $\hat{p}$ of an optical Michelson interferometer that scales linearly with the arm length. This scenario is described by Eq. (3) with $\theta = \pi/2$, where $S_{yy}(\Omega) = \alpha\Phi(\Omega)$ for some scale factor α to be estimated, and known signal morphology $\Phi(\Omega)$ (e.g., the fiducial spectrum from Ref. [2]) such that $\sigma^2 = \frac{1}{2}\alpha G_{py}(\Omega)\Phi(\Omega)$ by Eq. (14). Previous observations have constrained α to be less than one [4]. The future Gravity from the

Quantum Entanglement of Space-Time (GQuEST) optical interferometry experiment is proposed to further constrain α [3,70] (see also Ref. [5]).

In GQuEST, the classical noise background is projected to be 7 orders of magnitude below the quantum shot noise, $\sigma_C^2 \sim 10^{-7} \ll \frac{1}{2}$, while the optical loss is projected to be $\eta, \eta_A \sim 0.1$, limited by resonating filter cavities outside the interferometer. (Contributions to the optical loss from inside the interferometer are projected to be as low as 10^{-4}.) For signals below the classical noise background, i.e., $\sigma \ll \sigma_C$, this situation is described by Eq. (39), which exhibits the Rayleigh curse. We search instead for signals above the classical noise background, i.e., $\sigma_C \ll \sigma \ll \eta, \eta_A$. In this loss-dominated regime, preparing squeezed states or performing quadrature measurements introduces the Rayleigh curse. By Eqs. (47) and (49), the integrated QFI with respect to α is

$$\mathcal{I}_Q(\alpha) = \sum_{j=1}^{M} \frac{G_{py}(\Omega_j)\Phi(\Omega_j)}{4\alpha} \mathcal{I}_Q(\sigma).$$

For example, if the initial state is a vacuum state, then, by Eq. (48) with $\eta = 1$, the QFI is

$$\mathcal{I}_Q(\alpha) = \sum_{j=1}^{M} \frac{G_{py}(\Omega_j)\Phi(\Omega_j)}{\alpha[2 + \alpha G_{py}(\Omega_j)\Phi(\Omega_j)]}, \tag{50}$$

which can be attained by number measurement at each frequency Ω_j. For the vacuum case, the results from Sec. III thus indicate that photon counting accelerates the accrual of information by $1/4\sigma^2$, compared to quadrature measurement at each frequency. This acceleration reduces the required observing time or number of independent measurements to reach a given confidence level by the same factor. In comparison, if an SMSV state is prepared, for which quadrature measurement is optimal in the high-energy limit, then the acceleration compared to the vacuum case with quadrature measurement is a factor of $1/\eta^2$ by Eq. (34). Thus, photon counting beats squeezing for this application of stochastic signal estimation since $4\sigma^2 \ll \eta$.

The constraint on α can be further improved by preparing non-Gaussian states. The ECQFI in Eq. (35) indicates that if we prepare the optimal states and perform the optimal measurements, then the accrual of information about α could be accelerated by $1/(4\sigma^2\eta)$ compared to the vacuum case with quadrature measurement, or by $1/\eta$ compared to the vacuum case with number measurement. Thus, optimal state preparation and measurement, if implemented through the aforementioned emerging quantum technology, could significantly accelerate this fundamental-physics application.

2. Stochastic gravitational waves

Gravitational-wave searches may also benefit from our results. The global network of gravitational-wave observatories [7] presently consists of the Laser Interferometric Gravitational-wave Observatory (LIGO) [8], Virgo [9], and the Kamioka Gravitational-Wave Observatory (KAGRA) [10]. Each of these observatories operates as a Michelson interferometer with two ports—the common and differential ports. The common port corresponds to the common mode of the interferometer, i.e., the symmetric mode of the two arms, and a large coherent state is injected into it. The differential port corresponds to the differential (antisymmetric) mode of the interferometer. Information about the gravitational wave is encoded in the output of the differential port: a passing gravitational wave causes small displacements along the phase quadrature of the quantum states reflecting from the output differential port. Presently, these observatories are optimized to estimate the mean value (i.e., signal amplitude) of the gravitational wave by injecting highly squeezed states into the differential port and performing quadrature measurements of the output mode [50]. This is the optimal strategy for sensing deterministic displacements, as discussed in Sec. V B. The gravitational-wave strain sensitivity as an amplitude spectral density is proportional to the inverse square root of the FI.

Here, we consider instead operating a gravitational-wave observatory as a variance (i.e., signal-power) sensor for the stochastic signal from a single stochastic source or an incoherent background of many sources [6,97]. The present strategy outlined above, however, suffers the Rayleigh curse when estimating small stochastic signals: LIGO's present loss of $\eta \leq 0.23$ [50] imposes a detection horizon of $\sigma^2 > \eta/2 \approx 0.12$ by Eq. (34), beyond which it suffers the Rayleigh curse. The possible improvement from using the optimal stochastic sensing protocol, which we discuss below, in sensing astrophysical parameters $\vec{\theta}$ for a given signal model can be calculated from Eq. (49). We focus on astrophysical applications at kilohertz frequencies since there, unlike at low frequencies, LIGO's present classical noise is a factor of 4 below the vacuum shot noise [50], i.e., $\sigma_p^2 \sim 0.13$, and noise models predict that it could be made as low as $\sigma_p^2 \sim 0.05$ in the future [98]. The resulting classical noise horizon ($\sigma^2 > \sigma_p^2$) suffered by the optimal protocol is thus up to 2.3 times further than the loss horizon ($\sigma^2 > \eta/2$) suffered by preparing SMSV states [99]. We focus on detecting stochastic signals within this additional quantum-enhanced range.

The optimal stochastic sensing protocol involves preparing non-Gaussian states at the differential port and performing non-Gaussian measurements. (We assume that we still prepare a large coherent state at the common port.) In particular, we consider photon counting directly in the temporal basis of the gravitational-wave signal templates,

$\{w_k\}_{k=1}^{\infty}$, as follows [3]. A compact binary coalescence strain signal $y(t) = h(t)$ may be decomposed as

$$h(t) = h_{\text{det-LF}}(t) + h_{\text{det-HF}}(t) + h_{\text{stoc}}(t),$$

where $h_{\text{det-LF}}(t)$ and $h_{\text{det-HF}}(t)$ are low-frequency (premerger) and high-frequency (merger and post-merger) deterministic terms, respectively, and $h_{\text{stoc}}(t)$ is a stochastic term. This stochastic term has an associated temporal correlation function $\langle h_{\text{stoc}}(t) h_{\text{stoc}}(t')\rangle = H(t,t')$ which is *a priori* nonstationary. If we assume Gaussianity, then, by the Karhunen-Loéve theorem, we can further decompose this stochastic process into a sum of orthonormal temporal modes, $h_{\text{stoc}}(t) = \sum_{k=1}^{\infty} \sigma_k n_k w_k(t)$, where $\{n_k\}_{k=1}^{\infty}$ are independent unit-normal-distributed random variables and $\{\sigma_k\}_{k=1}^{\infty}$ are the parameters of interest. Preparing non-Gaussian states and photon counting in the basis $\{w_k\}_{k=1}^{\infty}$ could accelerate the search for stochastic sources beyond LIGO's present loss horizon discussed above. Note that this is distinct from the "event-stacking" technique [100–102] for aggregating information on the *deterministic* component of post-merger signals, $h_{\text{det-HF}}(t)$, from multiple binary-neutron-star coalescences [103].

It is also possible to operate a gravitational-wave observatory to simultaneously estimate the mean and variance as a test for unmodeled physics. In this scenario, the temporal basis to measure in is chosen beforehand and both the deterministic and stochastic components of the signal in each basis mode are unknown. For example, there are degrees of freedom in our models of binary neutron stars that are unconstrained by current data, and possibly additional unmodeled physics that would make the source either intrinsically stochastic or have unpredictable variations over the astrophysical population. Also, for black-hole events, we could accelerate tests for a stochastic departure from general relativity using a large ensemble of observed binary black-hole events. In either case, by performing joint measurements to estimate σ from the observation of many events with similar μ with a single observatory, we could perform stronger assumption-free tests for unmodeled physics. We assume here that the stochastic component is uncorrelated between the M different events in time.

To observe μ and σ simultaneously and optimally from a single event, we would need a network of M spatially separated observatories. The stochastic component would be uncorrelated between distant observatories because, e.g., the overlap reduction function between the two LIGO sites is vanishing at kilohertz frequencies [97]. As discussed in Sec. VII A, the challenges in this spatial case are in sending, storing, and receiving the quantum states from each observatory to prepare joint states and perform collective measurements. In Sec. VI D, we have assumed asymptotically large M to find the optimal collective measurement. Here, the number of observatories M is instead likely small such that we do not know what the optimal collective measurement is [104]. Our proposed collective measurement, however, still attains the QFI with respect to μ and is a factor of $(M-1)/M$ away from attaining the QFI with respect to σ. Let us compare this to the local strategy in Sec. VI C, which is a factor of 2 away from the QFI of both μ and σ: if $M = 2$ ($M = 3$) [$M = 4$], then our collective measurement is a factor of 2 more sensitive to μ and simultaneously a factor of 1 $\left(\frac{4}{3}\right)$ $\left[\frac{3}{2}\right]$ more sensitive to σ than the local strategy.

3. Axionic dark matter

Axions are a hypothetical wavelike dark-matter candidate that could also solve the strong $\mathcal{CP}$-problem [11,12]. More generally, if axionlike particles exist, then it is predicted that they should weakly interact with photons at the coupling rate $g_{a\gamma\gamma}$. To search for axions, therefore, we want to estimate $g_{a\gamma\gamma}$ at each frequency, where each frequency corresponds to a possible mass for the axion. Many experiments involving a microwave cavity in the presence of a static magnetic field are searching for axions (see, e.g., Refs. [13,14,55,105–107]). Since the coherence time of the axion is predicted to be short, the resulting displacement of the microwave cavity mode would be stochastic and without a preferred phase on longer timescales.

At a given frequency, this transformation is canonically the symmetric case of the additive Gaussian noise channel $\Lambda_{\Sigma_C}^{\text{noise}}$ in Eq. (22), with $\Sigma_C = \text{diag}(\sigma^2, \sigma^2)$ [108]. This is a two-dimensional random-displacement channel with

$$\Lambda_{\sigma}^{\text{2D}}(\hat{\rho}) \approx \hat{\rho} + \sigma^2 \left(\hat{x}\hat{\rho}\hat{x} + \hat{p}\hat{\rho}\hat{p} - \frac{1}{2}\{\hat{x}^2 + \hat{p}^2, \hat{\rho}\} \right). \quad (51)$$

We want to estimate σ since it is proportional to the coupling constant $g_{a\gamma\gamma}$ at the given frequency (i.e., axion mass), because the excess microwave power is proportional to $g_{a\gamma\gamma}^2$ up to the local dark-matter density, magnetic field strength, and cavity parameters [13,14,57,106]. Optimizing the QFI with respect to σ thus maximizes the axion scan rate, since the scan rate is proportional to the FI integrated over all frequencies [108].

Suppose that the initial pure state encounters loss before the encoding such that the total channel is $\Lambda'_{\sigma} = \Lambda_{\sigma}^{\text{2D}} \circ \Lambda_{\eta}^{\text{loss}}$. The ECQFI of this noisy channel with respect to σ, with energy constraint $\langle \hat{n} \rangle = N$, is known to be [108]

$$\mathcal{I}_Q^{\Lambda'_{\sigma}, N}(\sigma) = \frac{4\left(\eta + N\left(\eta + 2\sigma^2\right) + \sigma^2\right)}{\left(\eta + \sigma^2\right)\left(N\left(\eta + 2\sigma^2\right) + \sigma^2 + 1\right)} \xrightarrow[N\to\infty]{} \frac{4}{\eta + \sigma^2},$$

which is attained by a TMSV state with noiseless ancilla ($\eta_A = 0$) in the high-energy limit. (This limit is simply twice the one-dimensional case in Claim 3.) If $\eta_A \neq 0$,

however, then we show in Appendix D that the TMSV QFI vanishes for $\sigma^2 \ll \eta_A$. Moreover, by Claim 1, all squeezed states [109,110] are Rayleigh cursed, i.e., the QFI with respect to σ converges to zero in the limit of $\sigma \to 0$, provided that loss occurs on every mode. In general, we expect to operate in this loss-dominated regime experimentally, $\sigma^2 \ll \eta, \eta_A$, since the excess microwave power is small compared to the cavity loss [109]. In this loss-dominated regime, can we prepare non-Gaussian states instead?

Preparing Fock states for axion searches has been previously demonstrated [55] but without awareness of the relation between the optimal Fock state and the ultimate sensing limit. We establish this relation here. By Appendix F 2, the QFI in the limit of $\sigma \to 0$ is

$$\mathcal{I}_Q(\sigma = 0) = 4\left(\langle \hat{x}\hat{\Pi}_\perp \hat{x}\rangle + \langle \hat{p}\hat{\Pi}_\perp \hat{p}\rangle\right), \tag{52}$$

where $\hat{\Pi}_\perp$ projects onto the null space of the state before Λ_σ^{2D}. For a Fock state $|N\rangle$, Eq. (52) implies that the QFI is $\mathcal{I}_Q(\sigma = 0) = 4(1-\eta)^N(N+1)$. The optimal N depends on the loss η as discussed above, e.g., $N = 8$ for $\eta = 0.1$. The optimal Fock state is better than the Gaussian states but is again roughly a factor of e (4.3 dB) away from the ECQFI. For an input vacuum or Fock state, the optimal measurement is photon counting [95,107]. (Note that here, while we include a large loss before the encoding, we assume that no loss occurs after the encoding.)

Numerically, we observe that preparing finite-energy GKP states can achieve a QFI of at least 36, which is at least 90% of the ECQFI of 40 for $\sigma = 10^{-3}$ and $\eta = 0.1$. Similarly to the one-dimensional case, we conjecture that higher-energy finite-energy GKP states can converge to the ECQFI and are the optimal single-mode states for a given large $\langle \hat{n}\rangle = N$ and small $\sigma \ll \eta$. This means that non-Gaussian states beyond Fock states could further accelerate the search for axionic dark matter.

VIII. CONCLUSIONS AND OUTLOOK

We now draw our main conclusions, summarize our contributions to the literature, discuss the experimental practicality of our work, and look to the future of stochastic waveform estimation.

A. Conclusions

We have found the optimal protocol for stochastic waveform estimation using a linear quantum device. We have simplified the problem to the single-variable estimation of the excess noise in each temporal mode. For realistic losses, all Gaussian protocols exhibit the Rayleigh curse and fail to attain the ultimate precision limit. Instead, we have shown numerically that it is optimal to prepare non-Gaussian states such as finite-energy GKP states. For small signals above the classical noise floor, this non-Gaussian protocol outperforms all Gaussian protocols by orders of magnitude. We have also shown that a joint non-Gaussian measurement protocol is optimal for simultaneously estimating the mean and variance of a stochastic signal. Finally, we have demonstrated how our results may be applied to enhance searches for geontropic fluctuations from quantum gravity, stochastic gravitational-wave signals, and axionic dark matter.

B. Summary of results

For ease of reference, we now summarize our technical contributions to the literature in Table I and as follows:

(1) In Claim 1, we have proved a sufficient and necessary condition for overcoming the Rayleigh curse with Gaussian states. This general result applies beyond our specific encoding channel of random displacements and the context of stochastic waveform estimation. In Sec. V D, we have applied this result to prove that all squeezed states with noisy ancilla are cursed for sensing random displacements in the presence of loss on all modes. This establishes that to beat the vacuum state in the limit of vanishing signal, we must prepare non-Gaussian states.

(2) In Claims 2 and 3 we have calculated the ultimate limit, the ECQFI, on sensing random displacements in the lossless and lossy cases, respectively. We have used a purification-based approach given in Appendix C to accomplish this. In the lossless case, we have found that preparing an SMSV state is optimal given a constraint on the average energy per mode of the initial state; whereas, in the lossy case, we have found that preparing a TMSV state with noiseless ancilla is optimal in the high-energy limit.

(3) In Claim 4, we have calculated the QFI and optimal measurement for any random unitary channel acting on a Hilbert space, not just for random displacements of a bosonic mode, in the limit of vanishing signal. We have generalized this result in Appendix F to sensing the weak decay rate from any Lindbladian master equation.

(4) In Sec. V F, we have analyzed the performance of non-Gaussian states. We have proved that preparing Fock states and directly photon counting beats the

TABLE I. Summary of QFIs for different noise channels and initial states for single-parameter estimation of the variance. The QFIM for multiparameter estimation of the mean and variance for vacuum input is given in Eq. (44).

	Initial quantum state		
Noise channel	Vacuum	Gaussian	Non-Gaussian
Lossless	Eq. (20)	Claim 2	Sec. V C
Lossy	Sec. IV B	Claim 1	Claim 3
Classical noise	Eq. (23)	Eqs. (38) and (40)	Eq. (39)

vacuum state but misses the ECQFI by roughly 4.3 dB. We have shown numerically that finite-energy GKP states approach the ECQFI without requiring a noiseless ancilla like a TMSV state. The optimal non-Gaussian state is likely not unique, as we have also found a variational ansatz, sparse superpositions of finitely many Fock states, using numerical methods given in Appendix I. These sparse states are similar to binomial error correcting codes, outperform the Fock states, and approach the ECQFI if the energy of the input state is unconstrained. In Appendix I, we discuss that finding the optimal state is a biconvex-optimization problem and that the alternating convex search, used previously to efficiently find the optimal state, does not work in this case.

(5) In Sec. V G, we have considered sensing random displacements in the presence of classical noise. We have divided the classical noise into its components parallel and perpendicular to the signal. In the presence of parallel classical noise and no loss, we have shown that preparing an SMSV state is optimal. In the additional presence of loss, we have shown that preparing a GKP state instead is optimal. Meanwhile, in the presence of perpendicular classical noise and no loss, we have shown that preparing a GKP state is not only optimal but can recover the ultimate limit from the case with no classical noise. The above results for GKP states also hold for TMSV states but only if we assume noiseless ancilla, i.e., that the ancilla is free from classical noise and loss.

(6) Finally, in Sec. VI, we have considered the case in which the mean is also unknown, such that the mean and variance of the probability distribution behind the random-displacement signal need to be simultaneously estimated. We have considered a vacuum input state and shown that a number-resolving measurement cannot distinguish the mean from the variance of the signal. Instead, we have shown that a joint measurement of many independent copies of the final state is required to saturate the QFIM, while an adaptive protocol using separate measurements remains a factor of 2 below the QFIM. We have proved in Claim 5 that a similar collective measurement saturates the QFIM for any random unitary channel.

C. Experimental practicality

We now address the practicality of our work. As we discuss in Sec. VII A, there are several challenges in preparing the non-Gaussian states and implementing the non-Gaussian measurements that we propose. Thus, these protocols may take some time to be realized. We believe, however, that there are clear paths to addressing these challenges and that the time that it may require to overcome them does not diminish the importance of our results.

Preparing the optimal Fock state and photon counting can outperform all Gaussian states and get us within roughly 4.3 dB of the ultimate sensing limit, the ECQFI, as we have discussed in Sec. V F. For example, the optimal Fock state contains eight photons given a loss of 10%. This is likely achievable for stochastic waveform estimation in the microwave domain. In superconducting microwave cavities, Fock states have already been generated with up to 100 photons [54] and, in axion searches, the benefits of Fock states of four photons have already been shown [55]. This work adds valuable context to the prior experimental results, establishing where their results stand with respect to the ECQFI for stochastic waveform estimation.

To attain the ultimate limit and beat the Fock states by up to roughly 4.3 dB, we need to prepare more complicated non-Gaussian states such as GKP states or sparse superpositions of finitely many Fock states (which are similar to binomial-code states). Progress to this end is also being made experimentally in the microwave regime, motivated by using these states in quantum computing and error-correction applications. As we have discussed in Sec. V F, GKP states and binomial-code states have been prepared experimentally in a superconducting microwave cavity using controlled displacements with a qubit [59]. Alternatively, as we have discussed in Sec. IV C, it may be possible to reach the ultimate limit by preparing TMSV states and performing Gaussian operators followed by photodetection provided that the loss on the ancilla can be engineered to be sufficiently small compared to the classical noise and the signal.

Finally, we wish to make an analogy to the development of squeezing for gravitational-wave observatories. Four decades have passed from when squeezing was first proposed for LIGO [111] to today, where we measure 5 dB of frequency-dependent squeezing [112]. It may take decades more to produce the 10-dB broadband noise reduction stated for the third generation of gravitational-wave observatories [113,114]. The original squeezing proposal helped catalyze the development of LIGO and gravitational-wave astronomy and led to the maturity of squeezing today. Stochastic waveform estimation provides alternative and parallel science cases to deterministic waveform estimation that are relevant to both particle physics and astrophysics. This work has established the goals and benchmarks to gauge past and future demonstrations of stochastic waveform estimation. This is a valuable contribution given the statistical subtleties in sensing random signals and in systematically defining background noise and degradation processes. While the full extent of our results may not be immediately realized experimentally, many of the preliminary state-preparation and readout technologies have been demonstrated and

our results motivate and provide context to their future development. Our work informs the design of future experiments to maximally benefit from state preparation and optimally extract information.

D. Outlook

There are many related open questions to consider in the future. While we have considered the asymptotic limit and maximum likelihood estimation, it remains to be understood whether we can obtain faster convergence to the fundamental limit or better performance when restricted to a small number of measurements. It would also be interesting to understand the relation with the Bayesian-estimation problem in which, using the notation from Eq. (15), α is instead being estimated and $p(\alpha)$ is its prior distribution. We note that there is no immediate relation between this Bayesian problem and the problem of estimating σ that we consider. For example, in the vacuum case with a Gaussian prior [115,116], quadrature measurement $\hat{p}$ remains asymptotically optimal for estimating deterministic displacements, whereas here $\mathcal{I}_C^{\hat{p}}(\sigma = 0) = 0$. It remains to be understood whether a subtler connection exists between this Bayesian problem and our estimation problem.

Further open questions include the impact of different noise channels, determining the optimal states for fixed finite $\langle \hat{n} \rangle = N$, and the complete estimation of a non-Gaussian stochastic waveform since, e.g., estimating the fourth-order cumulant of a quantum state is of interest to testing nonclassicality [117] and quantum gravity [118–120].

ACKNOWLEDGMENTS

We thank the following people for their advice provided during this research: Rana Adhikari, Evan Hall, Thakur Giriraj Hiranandani, Konrad Lehnert, Katarzyna Macieszczak, Ian MacMillan, Patrick Meyers, Haixing Miao, Swadha Pandey, John Teufel, Mankei Tsang, Sander Vermeulen, Chris Whittle, and Sisi Zhou. We also thank the Caltech Chen Quantum Group and the ANU CGA Squeezer Group. In Fig. 1, we use component graphics from Ref. [121], with permission. The computations presented here were conducted in the Resnick High Performance Computing Center, a facility supported by Resnick Sustainability Institute at the California Institute of Technology. This research is supported by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (Project No. CE170100004). J.W.G. and this research are supported by an Australian Government Research Training Program (RTP) Scholarship and also partially supported by U.S. National Science Foundation (NSF) Grant No. PHY-2011968. In addition, Y.C. acknowledges the support by the Simons Foundation (Award No. 568762) and the NSF Grant No. PHY-2309231. T.G. acknowledges funding provided by the Institute for Quantum Information and Matter and the Quantum Science and Technology Scholarship of the Israel Council for Higher Education. S.A.H. acknowledges support through Australian Research Council Future Fellowship Grant No. FT210100809. J.P. acknowledges support from the U.S. Department of Energy (DOE), Office of Science, Office of Advanced Scientific Computing Research (Grants No. DE-NA0003525 and No. DE-SC0020290), the U.S. DOE, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator, and the NSF (Grant No. PHY-1733907). The Institute for Quantum Information and Matter is an NSF Physics Frontiers Center. L.M.'s photon-counting effort on the GQuEST project is funded in part by the Heising-Simons Foundation through Grant No. 2022-3341. This paper has been assigned LIGO Document No. P2400069.

DATA AVAILABILITY

The data that supports our results uses *Mathematica* [122] and PYTHON [123–131] and is openly available online [132].

APPENDIX A: PROOF OF CLAIM 1

Here, we prove Claim 1 by combining a known result about the eigenvalues of the density matrix and the Williamson decomposition of a Gaussian state [133].

Since $\partial_\sigma \hat{\rho} \to 0$ as $\sigma \to 0$, the optimal measurement basis in this limit is the eigenbasis of the density matrix $\hat{\rho}$ such that the QFI is nonvanishing if and only if there exists an eigenvalue of $\hat{\rho}$ proportional to σ^2 [40]. We therefore need to find a condition on the symplectic eigenvalues of Σ that is equivalent to this condition on the eigenvalues of $\hat{\rho}$. For an M-mode Gaussian state, the Williamson decomposition has the eigenvalues $\prod_{i=1}^{M} \lambda_{k_i}$ and eigenstates $\{\bigotimes_{i=1}^{M} \hat{S}(r_i)\,|k_i\rangle\}_{k_1 \cdots k_M}$, where $\hat{S}(r)$ is the single-mode squeezing operator and $|k_i\rangle$ is a Fock state (potentially of a combination of the original modes, e.g., a TMSV state). The *symplectic eigenvalues* associated with the ith mode are $\frac{1}{2} + \bar{n}_i$, such that the corresponding eigenvalues are $\lambda_{k_i} = (1/(1+\bar{n}_i))(\bar{n}_i/(1+\bar{n}_i))^{k_i}$ with degeneracy 2. Since the QFI is not Rayleigh cursed if and only if there exists some k_i such that $\lambda_{k_i} \propto \sigma^2$ [40], therefore, the QFI is nonvanishing if and only if there exists some mode i such that $\bar{n}_i \propto \sigma^2$ or, equivalently, a symplectic eigenvalue equal to $\frac{1}{2} + k\sigma^2$ for some constant k (not to be confused with k_i).

We remark that for a single-mode Gaussian state, this result can be seen directly from the QFI expression: if $\partial_\sigma \Sigma \to 0$, then the first term in Eq. (19) also vanishes. The second term in Eq. (19), however, is nonzero if and only if $\gamma = 1 - \beta\sigma^2$, which is equivalent to the above condition on the symplectic eigenvalues.

Here, we have assumed that the noisy channel is analytic at $\sigma = 0$ such that we may expand it in a Taylor series. This will usually be the case. For nonanalytic noisy channels, however, this result can be extended: the QFI is nonvanishing if and only if there exists a symplectic eigenvalue equal to $\frac{1}{2} + k\sigma^l$ with $1 < l \leq 2$ [40].

APPENDIX B: OTHER NOISE CHANNELS

We have focused on the case of a noise channel with known parameters occurring before the encoding channel. In this appendix, we briefly discuss noise channels occurring after the encoding channel and explore an example of what happens when the parameters of the noise channel are unknown.

1. Noise occurring after the encoding

Noise occurring after the encoding can be divided into two types: (1) noise associated with or "baked into" the acquisition of the signal or (2) noise associated with the later measurement. In the former case, it is not possible to insert an arbitrary parameter-independent "control" channel between the encoding and noise channels to mitigate the noise. In the latter case of "measurement noise," however, it is possible to entirely overcome the noise by implementing a suitable control channel between the encoding and measurement noise channels.

Let us illustrate this point that measurement noise can be overcome for sensing σ in the case of preparing an SMSV initial state with $\langle \hat{n} \rangle = N$. Note that in the noiseless case, the optimal measurement after the encoding channel is to first antisqueeze and then perform a number-resolving measurement. Here, we assume that we still antisqueeze but then perform a different, fixed, POVM that can be decomposed into a measurement noise channel followed by a number-resolving projective measurement. In, e.g., an optical system, our measurement apparatus is a photodetector and we consider two relevant noise models: detection loss and dark counts. Photodetection with detection loss is modeled as a loss channel $\Lambda^{\text{loss}}_{\eta_{\text{meas}}}$ followed by a projective number measurement. Photodetection with dark counts is modeled as the noise channel $\Lambda^{\text{dark}}_{N_{\text{meas}}}$ followed by a projective number measurement, where $\Lambda^{\text{dark}}_{N_{\text{meas}}}$ is defined as

$$\Lambda^{\text{dark}}_{N_{\text{meas}}}(\hat{\rho}) = \sum_{n=0}^{\infty} \hat{K}_n \hat{\rho} \hat{K}_n^\dagger, \quad \hat{K}_n = \sum_{m=0}^{\infty} \sqrt{p_{\text{th}}(m)} |n+m\rangle\langle n|,$$

where $p_{\text{th}}(m) = (1/(1+N_{\text{meas}}))\,(N_{\text{meas}}/(1+N_{\text{meas}}))^m$ is the number distribution of a thermal state with average number N_{meas}. The results of measuring $\hat{n}$ on $\Lambda^{\text{dark}}_{N_{\text{meas}}}(\hat{\rho})$ are equivalent to convolving the results of measuring $\hat{n}$ on $\hat{\rho}$ with p_{th}. Intuitively, $\Lambda^{\text{dark}}_{N_{\text{meas}}}$ adds m particles to the quantum state with probability $p_{\text{th}}(m)$ independent of the state. In comparison, it is unsuitable to model dark counts as an isotropic classical noise channel $\Lambda^{\text{noise}}_{\Sigma_{C,\text{meas}}}$ with $\Sigma_{C,\text{meas}} = \text{diag}(\sigma_{C,\text{meas}}, \sigma_{C,\text{meas}})$, since $\Lambda^{\text{noise}}_{\Sigma_{C,\text{meas}}}$ both adds and subtracts particles at a rate that depends on the quantum state due to Bose enhancement.

We first consider the impact of measurement noise in the absence of any additional control channels. In the noiseless case, given an initial SMSV state, the QFI is $\mathcal{I}_Q(\sigma) = 4\xi_N$ in the limit of $\sigma \ll 1$. In this limit, almost all of the CFI from the projective number measurement after antisqueezing comes from the $\hat{n} = 1$ single-particle detection probability, $p\,(1) \approx \xi_N \sigma^2$. In the case of measurement loss η_{meas}, e.g., detection loss for an optical system, the total channel before the projective number measurement is $\Lambda'_\sigma = \Lambda^{\text{loss}}_{\eta_{\text{meas}}} \circ \Lambda_{\text{antisqz}} \circ \Lambda_\sigma$, where Λ_{antisqz} is the antisqueezing unitary channel for the given N. The single-particle detection probability is now $p\,(1) \approx (1-\eta_{\text{meas}})\,\xi_N \sigma^2$, such that the CFI from the subsequent projective number measurement falls to $\mathcal{I}^{\hat{n}}_C(\sigma) = (1-\eta_{\text{meas}})\,4\xi_N$ but the Rayleigh curse is not introduced. This is unlike loss occurring before the antisqueezing operation, which does introduce the Rayleigh curse. In contrast, the Rayleigh curse will arise in the case of the dark-count-measurement noise channel. In this case, the total channel before the projective number measurement is $\Lambda'_\sigma = \Lambda^{\text{dark}}_{N_{\text{meas}}} \circ \Lambda_{\text{antisqz}} \circ \Lambda_\sigma$. The convolved probability distribution of number measurements is then $p\,(0) = (1-\xi_N\sigma^2)/(1+N_{\text{meas}})$ and $p\,(n) = N_{\text{meas}}^{n-1}(N_{\text{meas}} + \xi_N\sigma^2)/(1+N_{\text{meas}})^{n+1}$ for $n \geq 1$ such that the CFI is $\mathcal{I}^{\hat{n}}_C(\sigma) = 4\xi_N^2\sigma^2/\left(1-\xi_N\sigma^2\right)\left(\xi_N\sigma^2 + N_{\text{meas}}\right)$, which vanishes in the limit of $\sigma \to 0$ for fixed $\xi_N, N_{\text{meas}} > 0$.

We now show that there exists, in theory, a unitary that allows us to recover the noiseless QFI. After the antisqueezing operation, applying any unitary channel Λ_{swap} that swaps the Fock states $|1\rangle$ and $|k\rangle$ for some $k \gg 1$ and stabilizes the vacuum $|0\rangle$ (up to a phase) will recover the noiseless QFI in the limit of large enough k as proved below. Intuitively, this control unitary channel Λ_{swap} amplifies the signal to make it more tolerant to loss and distinguish it from dark counts. In the case of measurement loss, e.g., detection loss for an optical system, the total channel before the projective number measurement is now $\Lambda'_\sigma = \Lambda^{\text{loss}}_{\eta_{\text{meas}}} \circ \Lambda_{\text{swap}} \circ \Lambda_{\text{antisqz}} \circ \Lambda_\sigma$. The probability of not detecting zero particles is now $\left(1-\eta^k_{\text{meas}}\right)\xi_N\sigma^2$, keeping only the relevant σ^2 terms in the limit of $\sigma \ll 1$, such that the CFI from the subsequent projective number measurement is $\mathcal{I}^{\hat{n}}_C(\sigma) = 4\xi_N\left(1-\eta^k_{\text{meas}}\right)$, which recovers the noiseless QFI of $4\xi_N$ in the limit of $k \to \infty$. In the case of the dark-count-measurement noise, the total channel before the projective number measurement is now $\Lambda'_\sigma = \Lambda^{\text{dark}}_{N_{\text{meas}}} \circ \Lambda_{\text{swap}} \circ \Lambda_{\text{antisqz}} \circ \Lambda_\sigma$. The probability of detecting k or more particles after this channel is

$$p(\hat{n} \geq k) = \xi_N\sigma^2\left[1 - \left(\frac{N_{\text{meas}}}{N_{\text{meas}}+1}\right)^k\right] + \left(\frac{N_{\text{meas}}}{N_{\text{meas}}+1}\right)^k$$

in the limit of $\xi_N\sigma^2 \ll 1$. The CFI from detecting k or more particles is then

$$\mathcal{I}_C^{\hat{n}} = \frac{4\xi_N^2\sigma^2\left[1-\left(\frac{N_{\text{meas}}}{N_{\text{meas}}+1}\right)^k\right]^2}{\xi_N\sigma^2\left[1-\left(\frac{N_{\text{meas}}}{N_{\text{meas}}+1}\right)^k\right]+\left(\frac{N_{\text{meas}}}{N_{\text{meas}}+1}\right)^k},$$

for fixed $\sigma, N, \bar{n} > 0$. Since, for fixed $\xi_N, \sigma, N_{\text{meas}}$ and $\varepsilon > 0$, $\exists K$ such that $(N_{\text{meas}}/(N_{\text{meas}}+1))^k < \varepsilon$ for $k > K$, the CFI from number measurement recovers the noiseless QFI of $4\xi_N$ in the limit of $k \to \infty$. We defer to future work to find possible ways to implement suitable control unitary channels such as Λ_{swap}.

While we have assumed a specific fixed POVM above, e.g., a noisy photodetector, the results hold more generally [134–136]. Any fixed POVM that is modeled as a loss or dark-counts channel followed by a projective measurement will have the above limitations. For any such POVM, there exists an analogous control unitary channel that can overcome the measurement noise, where the particular control unitary channel needed may depend on the POVM [135,136].

In the case of a deterministic displacement channel, it is known that a phase-sensitive amplifier after the encoding can mitigate subsequent measurement loss in the high-gain limit. This is sometimes called a "Caves's amplifier" in the context of gravitational-wave observatories [111]. In our case, however, a phase-sensitive amplifier after the encoding would introduce the Rayleigh curse and should be avoided.

2. Unknown loss

Here, we discuss the implications of not precisely knowing the loss η that the state $\hat{\rho}$ experiences via the loss channel $\Lambda_\eta^{\text{loss}}$ introduced in Sec. IV B. We illustrate this for the input vacuum case and an indeterminate loss occurring after the encoding but before the measurement.

Suppose that the loss η follows some probability distribution $p(\eta)$. By the central limit theorem, asymptotically, this distribution approaches $p(\eta) \sim \mathcal{N}(\mu_\eta, \sigma_\eta^2)$ with some mean μ_η and standard deviation σ_η. The state after this indeterminate loss is

$$\Lambda_{\mu_\eta,\sigma_\eta}^{\text{indet.}}(\hat{\rho}) = \int_0^1 \mathrm{d}\eta\ \Lambda_\eta^{\text{loss}}(\hat{\rho})p(\eta).$$

We want to know whether this will introduce the Rayleigh curse in the relevant regime of $\sigma_\eta \ll 1$ where the loss is not precisely known but is well constrained.

The nth moment of a weighted average of distributions is the weighted average of their nth moments. The first moment is zero for the vacuum case. The second moment after $\Lambda_\eta^{\text{loss}}$ is linear in η such that the second moment after $\Lambda_{\mu_\eta,\sigma_\eta}^{\text{indet.}}$ equals that after $\Lambda_{\mu_\eta}^{\text{loss}}$, e.g., the signal term is $(1-\mu_\eta)\sigma^2$. The loss indeterminacy σ_η does not affect the first two moments, so, it must be a non-Gaussian perturbation of the third, fourth, and higher-order moments.

For the vacuum case, the final state after $\Lambda_{\mu_\eta,\sigma_\eta}^{\text{indet.}} \circ \Lambda_\sigma$ has the second moment $\langle\hat{p}^2\rangle = \varsigma^2$ and fourth moment $\langle\hat{p}^4\rangle = 3\varsigma^4 + 3\sigma^2\sigma_\eta^2$, where $\varsigma^2 := \frac{1}{2} + (1-\mu_\eta)\sigma^2$. The fourth-order cumulant $3\sigma^2\sigma_\eta^2$ indicates that the distribution has slightly fatter tails than a Gaussian. Numerically, the CFI from number measurement shows that this perturbation is negligible when μ_η is known and σ_η is small. For example, an indeterminacy of $\sigma_\eta/\mu_\eta = 10\%$ leads to an $\mathcal{O}(10^{-4})$ fractional change in the CFI with respect to σ. By convexity, this implies that the QFI also changes negligibly. If we also do not know μ_η, then the CFIM with respect to μ_η and σ is not singular and thus they can be estimated simultaneously.

3. Coarse-grained number-resolving measurements

Realistic measurements are unable to resolve infinitely many different particle numbers. We now consider a coarse-grained number-resolving measurement that can only distinguish between zero counts ($|0\rangle\langle 0|$) and nonzero counts ($\hat{I} - |0\rangle\langle 0|$), but cannot distinguish between, e.g., one or two particles. Suppose that the initial state is vacuum and there is no classical noise. Then, by Eq. (21), the probabilities of these two outcomes are $p(n=0) = 1/\sqrt{\sigma^2+1}$ and $p(n \geq 1) = 1 - p(n=0)$. By Eq. (18), the CFI is thus

$$\mathcal{I}_C(\sigma) = \frac{1}{\left(\sigma^2+1\right)^{5/2}} + \frac{1}{\left(\sigma^2+1\right)^2},$$

which achieves the QFI in Eq. (20) in the vanishing-signal limit of $\sigma \to 0$ but not generally for fixed $\sigma > 0$. This is because all of the information in the vanishing-signal limit comes from the single-particle probability $p(n=1)$, as can be seen from the summand of the CFI in Eq. (18),

$$\frac{[\partial_\sigma p(n)]^2}{p(n)} = \frac{(2n)!\,\sigma^{2n-2}\left(\sigma^2-2n\right)^2}{4^n(n!)^2\left(\sigma^2+1\right)^{n+\frac{5}{2}}},$$

which vanishes in the limit of $\sigma \to 0$ for $n \neq 1$ but saturates the QFI for $n = 1$.

Similarly, for a Fock state $|N\rangle$, all of the information comes from the $p(n = N+1)$ probability in the vanishing-signal limit. For example, this means that for a loss of $\eta = 0.1$, if we prepare the optimal $|8\rangle$, then we need to at least distinguish nine or more particles from eight or fewer particles to have nonzero sensitivity. This is more experimentally difficult than achieving the coarse-grained zero-or-nonzero measurement above. Alternatively, we could

first perform a unitary that maps, e.g., $|N\rangle \mapsto |0\rangle$ and $|N+1\rangle \mapsto |1\rangle$ and then use the coarse-grained zero-or-nonzero measurement above, although implementing such a unitary may be challenging.

More generally, the fact that coarse-grained binary measurements, i.e., those with only two outcomes, suffice to saturate the QFI in the vanishing-signal limit appears to be a general property of Rayleigh-cursed problems. As shown in Refs. [38,40], all the information about σ is encoded in eigenspaces of $\hat{\rho}(\sigma)$, the eigenvalues of which go as σ^2. Hence, the QFI is saturated by a binary measurement that differentiates between $\hat{\rho}(\sigma = 0)$ and the orthogonal eigenspace(s) that go as σ^2. In particular, Claim 1 implies that for Rayleigh-cursed Gaussian states, the optimal measurement corresponds to a symplectic operation followed by a coarse-grained number-resolving measurement. For example, the optimal measurement for a TMSV state corresponds to a symplectic operation that diagonalizes $\hat{\rho}(\sigma = 0)$ followed by a coarse-grained number-resolving measurement of one of the modes.

However, for more complex scenarios such as GKP states and for finite values of σ, binary measurements may not be sufficient and the effect of a realistic counting model is unknown. As we have explained in Sec. V F, we do not know if or how the required measurement for GKP states relates to the process to generate them. This means that we do not know whether a number-resolving measurement is required to implement the measurement for the GKP state, let alone what the effects of the realistic model of counting might be. We leave this to future work once the GKP measurement is first understood.

APPENDIX C: PURIFICATIONS AND THE ECQFI

To prove Claims 2 and 3 about the ECQFI, we first review the established method that we use to optimize over the initial states.

1. Review of purifications

Consider an initial pure state $|\psi\rangle$ and a nonunitary channel Λ_θ such that the final state is $\hat{\rho}(\theta) = \Lambda_\theta(|\psi\rangle\langle\psi|)$. The channel Λ_θ can be purified (also called dilated) to a unitary process $\hat{U}_\theta$ acting on $|\psi\rangle \otimes |\varphi\rangle$ such that the final state is $\hat{\rho}(\theta) = \mathrm{Tr}_A[\hat{U}_\theta(|\psi\rangle \otimes |\varphi\rangle)]$ for all $|\psi\rangle$ and θ. Note that the ancilla $|\varphi\rangle$ is independent of $|\psi\rangle$ and θ but depends on the purification $\hat{U}_\theta$ chosen. The purification of Λ_θ to $\hat{U}_\theta$ should not be confused with the purification of a mixed state or with the ECQFI discussed below.

The choice of purification, however, is not unique. By Uhlmann's theorem for the quantum fidelity, the QFI for a fixed initial state $|\psi\rangle$ is the infimum of the QFI over all possible purifications $\hat{U}_\theta$ of Λ_θ [43,44,108,137]

$$\mathcal{I}_Q^{\Lambda_\theta(|\psi\rangle\langle\psi|)}(\theta) = \inf_{\hat{U}_\theta} \mathcal{I}_Q^{\hat{U}_\theta(|\psi\rangle\otimes|\varphi\rangle)}(\theta). \tag{C1}$$

Here, the QFI of the unitary process $\hat{U}_\theta$ is $4\mathrm{Var}[\hat{H}]$ for all θ, where $\hat{H} = -i\hat{U}_\theta^\dagger \hat{U}_\theta$ is the Hermitian generator of local displacements in θ and the variance is calculated with respect to the pure initial state $|\psi\rangle \otimes |\varphi\rangle$ [138]. Using this fact, then Eq. (C1) becomes

$$\mathcal{I}_Q^{\Lambda_\theta(|\psi\rangle\langle\psi|)}(\theta) = 4 \inf_{\hat{U}_\theta} \mathrm{Var}_{|\psi\rangle\otimes|\varphi\rangle}[\hat{H}],$$

where $|\varphi\rangle$ and $\hat{H}$ are determined by the purification $\hat{U}_\theta$ of Λ_θ.

The CQFI of Λ_θ in Eq. (28), which optimizes over the initial state, is then given by

$$\mathcal{I}_Q^{\Lambda_\theta,\text{no ancilla}}(\theta) = 4 \sup_{|\psi\rangle} \inf_{\hat{U}_\theta} \mathrm{Var}_{|\psi\rangle\otimes|\varphi\rangle}[\hat{H}].$$

Exchanging the order of maximization and minimization above results in an upper bound on the CQFI of Λ_θ. Remarkably, in Ref. [139] it has been shown that this upper bound is exactly the ECQFI of Λ_θ in Eq. (29); hence

$$\mathcal{I}_Q^{\Lambda_\theta}(\theta) = 4 \inf_{\hat{U}_\theta} \sup_{|\psi\rangle} \mathrm{Var}_{|\psi\rangle\otimes|\varphi\rangle}[\hat{H}]. \tag{C2}$$

Thus, to attain the ECQFI, it suffices to find a purification $\hat{U}_\theta$ for which the following upper bound is tight:

$$\mathcal{I}_Q^{\Lambda_\theta}(\theta) \leq 4 \sup_{|\psi\rangle} \mathrm{Var}_{|\psi\rangle\otimes|\varphi\rangle}[\hat{H}]. \tag{C3}$$

To find such a purification, we will use the following technique of introducing a "hiding" unitary $\hat{U}_{\text{hide}}$ [18]. Suppose that we have a purification $\hat{U}_{\text{enc.}}$ of Λ_θ. We can generate a family of related purifications $\hat{U}_\theta = \hat{U}_{\text{hide}}\hat{U}_{\text{enc.}}$ by performing any θ-dependent unitary transformation $\hat{U}_{\text{hide}}$ on the ancilla, since the ancilla is traced out to recover $\hat{\rho}(\theta)$. Intuitively, this hiding unitary $\hat{U}_{\text{hide}}$ is meant to remove (or hide) the excess information about θ present in the ancilla after $\hat{U}_{\text{enc.}}$ to make Eq. (C3) tight.

2. Proof of Claim 2

Here, we prove Claim 2 about the lossless case by using the above purification method with a hiding unitary.

We purify the channel Λ_σ by introducing an ancillary mode 2 prepared in a vacuum state, such that the resulting encoding unitary is a conditional displacement $\hat{U}_{\text{enc.}} = \exp(-i\sigma\hat{H}_{\text{enc.}})$ with the generator $\hat{H}_{\text{enc.}} = \sqrt{2}\hat{x}_1\hat{x}_2$. We also include a hiding unitary $\hat{U}_{\text{hide}} = \exp(-i\sigma\hat{H}_{\text{hide}})$ such that the overall purification is $\hat{U}_\sigma = \hat{U}_{\text{hide}}\hat{U}_{\text{enc.}}$. Here, $\hat{H}_{\text{hide}} = \frac{1}{2}g\left(\hat{x}_2\hat{p}_2 + \hat{p}_2\hat{x}_2\right)$ is a squeezing Hamiltonian and g is a real parameter over which to minimize to obtain a tight upper bound in Eq. (C3).

The Hermitian generator for the overall purification $\hat{U}_\sigma$ is

$$\begin{aligned}\hat{H} &= \hat{H}_{\text{enc.}} + \hat{U}_{\text{enc.}}^\dagger \hat{H}_{\text{hide}} \hat{U}_{\text{enc.}}. \\ &= \left(\sqrt{2} + \sqrt{2} g\sigma\right) \hat{x}_1 \hat{x}_2 + \frac{g}{2}\left(\hat{x}_2 \hat{p}_2 + \hat{p}_2 \hat{x}_2\right).\end{aligned}$$

The QFI of $\hat{U}_\sigma |\psi\rangle \otimes |\varphi\rangle$ is therefore

$$4\text{Var}[\hat{H}] = 2\left(\sqrt{2} + \sqrt{2} g\sigma\right)^2 \langle \hat{x}_1^2 \rangle + 2g^2, \tag{C4}$$

where we have assumed that $\langle \hat{x}_1 \rangle = 0$ without loss of generality. Here, since mode 2 is in a vacuum state, we have also used the fact that

$$\text{Var}\left[\hat{x}_2 \hat{p}_2 + \hat{p}_2 \hat{x}_2\right] = 2. \tag{C5}$$

Minimizing the QFI in Eq. (C4) over g, we derive the upper bound to the ECQFI in Eq. (33). This equals the QFI from preparing an SMSV state by Eq. (19).

3. Proof of Claim 3

Here, we prove Claim 3 about the lossy case by using a similar purification method to the lossless case above. Our choice of hiding unitary to handle the loss channel was inspired by the proof of Eq. (31) for the deterministic case [18].

We purify the total noisy channel $\Lambda'_\sigma = \Lambda_\sigma \circ \Lambda_\eta^{\text{loss}}$ as follows. The loss channel $\Lambda_\eta^{\text{loss}}$ is purified to the following beam-splitter unitary between mode 1 and some ancillary mode 3 that starts in vacuum:

$$\begin{aligned}\hat{U}_\eta &= \exp[i \arcsin(\sqrt{\eta})(\hat{a}_1^\dagger \hat{a}_3 + \hat{a}_3^\dagger \hat{a}_1)] \\ &= \exp[i \arcsin(\sqrt{\eta})(\hat{x}_1 \hat{x}_3 + \hat{p}_1 \hat{p}_3)].\end{aligned}$$

The encoding channel Λ_σ is again purified to $\hat{U}_{\text{enc.}} = \exp(-i\sigma \hat{H}_{\text{enc.}})$ with $\hat{H}_{\text{enc.}} = \sqrt{2}\hat{x}_1 \hat{x}_2$ and mode 2 starting in vacuum. We also include two hiding unitaries: $\hat{U}_{\text{hide},1} = \exp(-i\sigma \hat{H}_{\text{hide},1})$ with $\hat{H}_{\text{hide},1} = g_1 \hat{x}_2 \hat{x}_3$, which removes the excess information due to the purification of the loss, and $\hat{U}_{\text{hide},2} = \exp(-i\sigma \hat{H}_{\text{hide},2})$ with $\hat{H}_{\text{hide},2} = \frac{1}{2} g_2 (\hat{x}_2 \hat{p}_2 + \hat{p}_2 \hat{x}_2)$, which again removes the excess information due to the purification of the random displacement. Here, g_1 and g_2 are parameters to be minimized over again to obtain a tight bound.

The overall unitary is thus

$$\hat{U}_\sigma = \hat{U}_{\text{hide},2} \hat{U}_{\text{hide},1} \hat{U}_{\text{enc.}} \hat{U}_\eta,$$

the Hermitian generator of which is

$$\begin{aligned}\hat{H} = \hat{U}_\eta^\dagger \hat{U}_{\text{enc.}}^\dagger \hat{U}_{\text{hide},1}^\dagger \hat{H}_{\text{hide},2} \hat{U}_{\text{hide},1} \hat{U}_{\text{enc.}} \hat{U}_\eta \\ + \hat{U}_\eta^\dagger \hat{U}_{\text{enc.}}^\dagger \hat{H}_{\text{hide},1} \hat{U}_{\text{enc.}} \hat{U}_\eta + \hat{U}_\eta^\dagger \hat{H}_{\text{enc.}} \hat{U}_\eta,\end{aligned}$$

where there is no term associated with $\hat{H}_\eta$, since $\hat{U}_\eta$ is independent of σ. This expression then simplifies to

$$\begin{aligned}\hat{H} = \frac{g_2}{2}\left(\hat{x}_2 \hat{p}_2 + \hat{p}_2 \hat{x}_2\right) \\ + (g_2\sigma + 1)\left(\sqrt{2\eta} + g_1\sqrt{1-\eta}\right)\hat{x}_3 \hat{x}_2 \\ + (g_2\sigma + 1)\left(\sqrt{2(1-\eta)} - g_1\sqrt{\eta}\right)\hat{x}_1 \hat{x}_2.\end{aligned}$$

The QFI of $\hat{U}_\sigma |\psi\rangle \otimes |\varphi\rangle$ is therefore

$$4\text{Var}[\hat{H}] = 2g_2^2 + f\ (g_1)\ (g_2\sigma + 1)^2$$

where, by Eq. (C5),

$$\begin{aligned}f\ (g_1) = \left(\sqrt{2\eta} + g_1\sqrt{1-\eta}\right)^2 \\ + 2\langle \hat{x}_1^2 \rangle \left(\sqrt{2(1-\eta)} - g_1\sqrt{\eta}\right)^2.\end{aligned}$$

Minimizing over the free parameters of the hiding unitaries, g_1 and g_2, we find that

$$\begin{aligned}4\text{Var}[\hat{H}] &= \min_{g_1, g_2}\left[2g_2^2 + f\ (g_1)\ (g_2\sigma + 1)^2\right] \\ &= \min_{g_2}\left[2g_2^2 + \frac{4\langle \hat{x}_1^2 \rangle}{(1-\eta) + 2\eta\langle \hat{x}_1^2 \rangle}(g_2\sigma + 1)^2\right] \\ &= \frac{4}{2(\eta + \sigma^2) + (1-\eta)\langle \hat{x}_1^2 \rangle^{-1}}.\end{aligned}$$

Using the optimal value of $\langle \hat{x}_1^2 \rangle$ for a given constraint $\langle \hat{n} \rangle = N$ leads to the upper bound of Eq. (36). This optimal value of $\langle \hat{x}_1^2 \rangle$ for a fixed N would correspond to the state after the loss but before the encoding being an SMSV state, but there is no such state that becomes an SMSV state after a loss. Note that while, e.g., Schrödinger's cat states approach $\langle \hat{x}_1^2 \rangle = 2N$ for large N too, the relation $\langle \hat{x}_1^2 \rangle < \xi_N$ holds for any fixed N and any non-SMSV state. This bound, therefore, is not tight for a fixed finite N, as preparing an SMSV state does not saturate it. In the limit of $N \to \infty$, however, this bound proves the ECQFI in Eq. (35), which is saturated asymptotically by TMSV (see Appendix D) and thus is tight.

APPENDIX D: TWO-MODE SQUEEZED VACUUM

Here, we analyze the QFI from preparing a TMSV state given different noise models (loss and classical noise) and different encoding channels (one-dimensional and two-dimensional random-displacement channels).

In Ref. [37], it has been shown that the QFI of an M-mode Gaussian state with the parameter encoded in the

covariance matrix Σ is given by the following formula:

$$\mathcal{I}_Q(\sigma) = 2\mathrm{Tr}\left[(\partial_\sigma \Sigma)(4\Sigma \otimes \Sigma - \omega \otimes \omega)^{-1}(\partial_\sigma \Sigma)\right], \quad \text{(D1)}$$

where, if Σ is written in the quadrature basis of $(\hat{x}_1, \cdots, \hat{x}_M, \hat{p}_1, \cdots, \hat{p}_M)^{\mathrm{T}}$, then the commutator matrix is given by $\omega = \left[\begin{smallmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{smallmatrix}\right]$, where $\mathbb{1}$ is the $M \times M$ identity matrix. This result generalizes Eq. (19) for single-mode Gaussian states.

For the one-dimensional random-displacement channel in Eq. (15) preceded by a loss channel, the QFI of a TMSV state with noisy ancilla $\eta_A > 0$ in the high-energy limit $N \to \infty$ and with $\sigma \ll 1$ is given by Eq. (D1) as

$$\mathcal{I}_Q(\sigma) = \frac{2(1-\eta_A)\sigma^2(\eta+\eta_A-2\eta\eta_A)^2}{\xi + \sigma^2(\eta+\eta_A-2\eta\eta_A)^3}, \quad \text{(D2)}$$

where

$$\xi = 2(1-\eta)\eta\eta_A\left(\eta^2(2(\eta_A-1)\eta_A+1) - 2\eta\eta_A^2 + \eta_A^2\right).$$

Attaining this QFI requires first antisqueezing and then, in general, applying a beam splitter followed by number-resolving measurements of the two modes [37]. For any $\sigma > 0$ and with noiseless ancilla ($\eta_A = 0$), the QFI in Eq. (D2) converges to the ECQFI in Eq. (35) in the high-energy limit. For example, in Fig. 4, we use the full form of the QFI for any σ, η, η_A, and N, which is too verbose to provide here but is found with the same method.

In comparison, for the two-dimensional random-displacement channel in Eq. (51) preceded by a loss channel, the high-energy QFI for a TMSV state with noisy ancilla is given by Eq. (D1) as

$$\mathcal{I}_Q(\sigma) = \frac{4(1-\eta_A)\sigma^2}{(\eta+\sigma^2)\left[(1-\eta)\eta_A + (1-\eta_A)\sigma^2\right]}. \quad \text{(D3)}$$

For $0 < \eta_A, \sigma^2 \ll \eta$, the two-dimensional encoding TMSV QFI in Eq. (D3) with ancilla loss η_A is approximately twice the one-dimensional encoding TMSV QFI in Eq. (D2) with ancilla loss $\eta_A/2$. In either case, the TMSV QFI attains the ECQFI at high energy if $\eta_A \ll \sigma^2, \eta$ but vanishes if $\sigma^2 \ll \eta, \eta_A$.

We now briefly analyze the perpendicular classical noise case discussed in Appendix J. Consider the one-dimensional random-displacement channel preceded by a perpendicular classical noise channel that adds σ_x to the system and $\sigma_{x,A}$ to the ancilla. Let us first consider the case of noiseless ancilla, i.e., $\sigma_{x,A} = 0$. Interestingly, for any $\langle\hat{n}\rangle = N$ per mode and any amount of perpendicular noise $\sigma_x > 0$, the QFI is

$$\mathcal{I}_Q(\sigma = 0) = \frac{8N(N+1)}{2N+1} \xrightarrow[N\to\infty]{} 4N, \quad \text{(D4)}$$

which does not exhibit the Rayleigh curse. Intuitively, for any N, there exists a symplectic transformation of $\hat{\rho}$, the state after the encoding channel, to a pair of modes such that one of them is completely decoupled from the perpendicular noise but nevertheless contains some information about the signal. An optimal measurement, therefore, would be to apply this symplectic transformation to $\hat{\rho}$ and then perform a number measurement of the perpendicular-noise–free mode. In the quadrature basis of $(\hat{x}_S, \hat{x}_A, \hat{p}_S, \hat{p}_A)^{\mathrm{T}}$, where the system S and ancilla A are entangled, the covariance matrix of $\hat{\rho}$ is

$$\Sigma = \frac{1}{2}\mathbb{1} + \begin{bmatrix} \sigma_x^2 + N & \sqrt{N(N+1)} \\ \sqrt{N(N+1)} & N \end{bmatrix} \oplus \begin{bmatrix} \sigma^2 + N & -\sqrt{N(N+1)} \\ -\sqrt{N(N+1)} & N \end{bmatrix}.$$

After performing a suitable antisqueezing operation, the inverse of the two-mode squeezing unitary that prepares the TMSV state from vacuum, the covariance matrix becomes

$$\Sigma = \frac{1}{2}\mathbb{1} + \frac{1}{2}\sigma_x^2 \begin{bmatrix} N+1 & -\sqrt{N(N+1)} \\ -\sqrt{N(N+1)} & N \end{bmatrix} \oplus \frac{1}{2}\sigma^2 \begin{bmatrix} N+1 & \sqrt{N(N+1)} \\ \sqrt{N(N+1)} & N \end{bmatrix}.$$

The collective mode with annihilation operator $\hat{a}_{\mathrm{dec}} = (1/\sqrt{2N+1})\left(\sqrt{N}\hat{a}_S + \sqrt{N+1}\hat{a}_A\right)$, therefore, is completely decoupled from the perpendicular noise (σ_x^2) yet remains coupled to the signal (σ^2). The covariance matrix of this perpendicular-noise–free mode is

$$\Sigma_{\mathrm{dec}} = \frac{1}{2}\mathbb{1} + \mathrm{diag}\left(0,\ 2\sigma^2\frac{N(N+1)}{2N+1}\right),$$

such that a number measurement of this mode attains the QFI in Eq. (D4). This protocol is reminiscent of the idea of displacement-noise–free interferometry in which there exist modes unaffected by displacement noise but that nevertheless contain information about the signal [140–142].

We now consider the case of noisy ancilla, i.e., $\sigma_{x,A} > 0$. Here, however, the QFI as $\sigma \to 0$ exhibits the Rayleigh curse by Claim 1 for a fixed value of N, since, for all i, $\bar{n}_i = c_i(\sigma_x, \sigma_{x,A}) + \mathcal{O}(\sigma^2)$ for some $c_i(\sigma_x, \sigma_{x,A}) > 0$ (see Appendix A). For a fixed finite signal $\sigma > 0$ and any fixed σ_x and $\sigma_{x,A}$, the QFI in Eq. (D1) attains the noiseless ECQFI of $2/\sigma^2$ in the high-energy limit of $N \to \infty$ particles per mode.

Finally, we elaborate on the relation between a TMSV state with noiseless ancilla and GKP states in the high-energy limit. The infinite-energy TMSV pure state is the

following unnormalized state:

$$\int_{-\infty}^{\infty} \mathrm{d}x\ |\hat{x}_S = x\rangle|\hat{x}_A = x\rangle = \int_{-\infty}^{\infty} \mathrm{d}p\ |\hat{p}_S = p\rangle|\hat{p}_A = -p\rangle,$$

where $|\hat{x}_j = x_0\rangle$ ($|\hat{p}_j = p_0\rangle$) is the position (momentum) eigenstate of mode j at x_0 (p_0). We observe that deterministic or random displacements along $\hat{p}_S$ keep the state inside the code space of $\left\{|\hat{x}_S = x\rangle|\hat{x}_A = x\rangle\right\}_x$, where deterministic (random) displacement induces rotation (dephasing) inside this subspace. In comparison, deterministic or random displacements along $\hat{x}_S$ take the state outside of this code space and can be detected by measuring $\hat{x}_S - \hat{x}_A$. This is similar to the GKP infinite-energy state $|\mathrm{GKP}_{\mathrm{ideal}}\rangle$ from Appendix H, for which the code space is $\left\{\sum_{j=-\infty}^{\infty}|\hat{p} = 2j\sqrt{\pi} + p_0\rangle\right\}_{p_0}$. Deterministic (random) displacements along $\hat{p}$ induce rotation (dephasing) inside this subspace, while displacements along $\hat{x}$ take the state outside of this code space and can be detected by the relevant "syndrome" measurement of $\hat{x}$ mod $2\sqrt{\pi}$ [52].

APPENDIX E: PROOF OF CLAIM 4

Here, we prove Claim 4 by a Taylor expansion and again using the fact that, in the limit of small signals $\sigma \to 0$, an optimal measurement is to project onto the eigenbasis.

In general, a unitary $\hat{U}_\theta$ may be expanded around $\theta = 0$ to obtain the following approximation up to $\mathcal{O}(\theta^3)$:

$$\hat{U}_\theta \approx \hat{U}_0 + \theta\dot{\hat{U}}_0 + \frac{1}{2}\theta^2\ddot{\hat{U}}_0 \tag{E1}$$

where the unitarity condition $\hat{U}_\theta^\dagger\hat{U}_\theta = 1$ implies that

$$0 \approx \dot{\hat{U}}_0^\dagger\hat{U}_0 + \hat{U}_0^\dagger\dot{\hat{U}}_0 \tag{E2}$$

$$0 \approx \ddot{\hat{U}}_0^\dagger\hat{U}_0 + \hat{U}_0^\dagger\ddot{\hat{U}}_0 + 2\dot{\hat{U}}_0^\dagger\dot{\hat{U}}_0. \tag{E3}$$

Let $\hat{H} := i\hat{U}_0^\dagger\dot{\hat{U}}_0$ and $\hat{Z} := -\hat{U}_0^\dagger\ddot{\hat{U}}_0$ such that $\hat{Z} = \hat{H}^2 + 2i\hat{B}$, where $\hat{H}^2 = \mathrm{Re}[\hat{Z}] = \dot{\hat{U}}_0^\dagger\dot{\hat{U}}_0$ and $\hat{B} = \frac{1}{2}\mathrm{Im}[\hat{Z}] = (i/4)(\dot{\hat{U}}_0^\dagger\ddot{\hat{U}}_0 - \ddot{\hat{U}}_0^\dagger\dot{\hat{U}}_0)$.

Using this expansion around $\theta = 0$, the random unitary channel Λ_σ in Eq. (37) is approximately

$$\Lambda_\sigma(\hat{\rho}) \approx \hat{U}_0\hat{\rho}\hat{U}_0^\dagger + \sigma^2\left(\dot{\hat{U}}_0\hat{\rho}\dot{\hat{U}}_0^\dagger + \frac{1}{2}(\ddot{\hat{U}}_0\hat{\rho}\hat{U}_0^\dagger + \hat{U}_0\hat{\rho}\ddot{\hat{U}}_0^\dagger)\right).$$

Without loss of generality, we can study $\hat{U}_0^\dagger\Lambda_\sigma(\hat{\rho})\hat{U}_0$, since $\hat{U}_0^\dagger$ does not depend on the parameter θ and thus will not affect the QFI if applied to the state after the channel. The final state then becomes

$$\begin{aligned}\hat{U}_0^\dagger\Lambda_\sigma(\hat{\rho})\hat{U}_0 &\approx \hat{\rho} + \sigma^2\left(\hat{H}\hat{\rho}\hat{H} - \frac{1}{2}(\hat{Z}\hat{\rho} + \hat{\rho}\hat{Z}^\dagger)\right)\\ &= \hat{\rho} - i[\sigma^2\hat{B}, \hat{\rho}] + \sigma^2\left(\hat{H}\hat{\rho}\hat{H} - \frac{1}{2}\{\hat{H}^2, \hat{\rho}\}\right).\end{aligned} \tag{E4}$$

This is an approximation to the master-equation evolution of $\hat{\rho}$ by the Hamiltonian $\sigma^2\hat{B}$ and the Lindbladian jump operator $\hat{H}$ with decay rate σ^2 in the limit of $\sigma^2 T \ll 1$, where T is the total evolution time. In the additive case of $\hat{U}_\theta = \exp(-i\theta\hat{H})$, $\hat{B} = 0$ such that $\hat{Z} = \hat{H}^2$ and there is no Hamiltonian evolution. Moreover, Λ_σ is then a decoherence channel in the eigenbasis of $\hat{H}$ similarly to Eq. (16) for $\hat{H} = \hat{x}$.

Since $\partial_\sigma(\hat{U}_0^\dagger\Lambda_\sigma(\hat{\rho})\hat{U}_0) \to 0$ as $\sigma \to 0$, the optimal measurement (after $\hat{U}_0^\dagger$) is projection $\hat{\Pi}$ onto the support of $\hat{\rho}$, i.e., $\hat{\Pi} = \sum_j|\psi_j\rangle\langle\psi_j|$ in the eigenbasis $\{|\psi_j\rangle\}_j$ of $\hat{\rho}$ [40,51]. In particular, since $\mathrm{Tr}[\hat{\Pi}[\sigma^2\hat{B}, \hat{\rho}]] = 0$, we may ignore the Hamiltonian evolution in the limit of $\sigma \to 0$. By Eq. (E4), the probability to remain in the support of $\hat{\rho}$ is

$$p = \mathrm{Tr}[\hat{\Pi}\hat{U}_0^\dagger\Lambda_\sigma(\hat{\rho})\hat{U}_0] = 1 + \sigma^2\left(\langle\hat{H}\hat{\Pi}\hat{H}\rangle - \langle\hat{H}^2\rangle\right)$$

and the probability to leave is $1 - p = \sigma^2\langle\hat{H}\hat{\Pi}_\perp\hat{H}\rangle$, where $\hat{\Pi}_\perp = 1 - \hat{\Pi}$ is the projection onto the null space of $\hat{\rho}$. Here, the expectation values are calculated with respect to $\hat{\rho}$. Since the QFI equals the CFI of the optimal measurement,

$$\mathcal{I}_Q(\sigma = 0) = 4\langle\hat{H}\hat{\Pi}_\perp\hat{H}\rangle, \tag{E5}$$

and if $\hat{\rho}$ is pure, then the QFI is

$$\mathcal{I}_Q(\sigma = 0) = 4\mathrm{Var}[\hat{H}], \tag{E6}$$

which is equal to the QFI for the deterministic unitary case $\hat{U}_\sigma\hat{\rho}\hat{U}_\sigma^\dagger$. This correspondence between the deterministic and random cases does not hold for $\sigma > 0$. A similar observation for nonunitary channels in the limit of $\sigma \to 0$ has been made previously in Ref. [36].

APPENDIX F: ESTIMATING A WEAK DECAY RATE

Here, we consider the general scenario of estimating a weak decay rate. This is inspired by the fact that Eq. (E4) is analogous to a short-time solution of a master equation, where the jump operator is $\hat{H}$ and σ^2 is equivalent to the decay rate multiplied by the evolution time. We consider

the cases of a single jump operator and multiple jump operators sharing the same weak decay rate. This problem has previously been studied in Ref. [36], in the case of an initial pure state.

Suppose that a quantum state $\hat{\rho}(t)$ evolves in time t under the following Lindbladian master equation:

$$\partial_t\hat{\rho}(t)=\gamma\mathcal{L}_{\hat{Y}}[\hat{\rho}(t)],\quad \mathcal{L}_{\hat{Y}}[\hat{\rho}(t)]=\hat{Y}\hat{\rho}(t)\hat{Y}^\dagger-\frac{1}{2}\{\hat{Y}^\dagger\hat{Y},\hat{\rho}(t)\},$$

where $\gamma\geq 0$ is the decay rate, $\hat{Y}$ is the (potentially non-Hermitian) jump operator, and we seek to estimate $\sqrt{\gamma}$. In the short-time or weak-decay limit of $\gamma Tc\ll 1$, where T is the total evolution time and $c=\mathrm{Tr}[\hat{Y}\hat{\rho}\hat{Y}^\dagger]$. Then, the evolution of an initial state $\hat{\rho}$ is approximated by the channel Λ_γ, with the following Kraus representation:

$$\Lambda_\gamma(\hat{\rho})=\hat{K}_0\hat{\rho}\hat{K}_0^\dagger+\hat{K}_1\hat{\rho}\hat{K}_1^\dagger,$$

where the Kraus operators are

$$\hat{K}_0:=1-\frac{\gamma T}{2}\hat{Y}^\dagger\hat{Y}\approx\cos\left(\sqrt{\gamma T\hat{Y}^\dagger\hat{Y}}\right),$$
$$\hat{K}_1:=\sqrt{\gamma T}\hat{Y}\approx\sin\left(\sqrt{\gamma T}\hat{Y}\right),$$

such that

$$1=\hat{K}_0^\dagger\hat{K}_0+\hat{K}_1^\dagger\hat{K}_1+\mathcal{O}\left((\gamma T)^2\right).$$

Since the total evolution time T is known, we can measure the final state $\Lambda_\gamma(\hat{\rho})$ to estimate the weak decay rate $\sqrt{\gamma}$. Since $\partial_{\sqrt{\gamma}}\Lambda_\gamma\to 0$ as $\sqrt{\gamma}\to 0$, then the QFI with respect to $\sqrt{\gamma}$ is [40,51]

$$\mathcal{I}_Q(\sqrt{\gamma}=0)=4T\langle\hat{Y}^\dagger\hat{\Pi}_\perp\hat{Y}\rangle,\tag{F1}$$

where $\hat{\Pi}_\perp$ projects onto the null space of the state $\hat{\rho}=\sum_j p_j|\phi_j\rangle\langle\phi_j|$ before Λ_γ. If $\hat{\rho}=|\phi\rangle\langle\phi|$ is pure, then Eq. (F1) becomes

$$\mathcal{I}_Q(\sqrt{\gamma}=0)=4T\left(\langle\hat{Y}^\dagger\hat{Y}\rangle-|\langle\hat{Y}\rangle|^2\right),\tag{F2}$$

which resembles the formula for the variance $\mathrm{Var}[Z]=\mathrm{E}\left[|Z|^2\right]-|\mathrm{E}[Z]|^2$ of a complex random variable Z. If $\hat{Y}$ is Hermitian, then $\mathcal{I}_Q(\sqrt{\gamma}=0)=4T\mathrm{Var}[\hat{Y}]$, which is equivalent to Eq. (E5) for $\gamma T=\sigma^2$ and $\hat{Y}=\hat{H}$. If instead $\hat{\rho}$ is mixed, then Eq. (F1) implies that the Rayleigh curse will be avoided if and only if $\hat{\rho}$ is partial rank and the jump operator $\hat{Y}$ transitions some of $\hat{\rho}$ into its null space. $\hat{\rho}$ may be mixed, e.g., because Λ_γ is preceded by another noise channel $\Lambda_{\gamma'}$ with its own decay rate γ', which might not be small, and jump operator $\hat{Y}'$ such that the total channel is $\Lambda_\gamma\circ\Lambda_{\gamma'}$. It is perhaps more natural, however, to consider these two processes occurring simultaneously within the following Lindbladian master equation:

$$\partial_t\hat{\rho}(t)=\gamma\mathcal{L}_{\hat{Y}}[\hat{\rho}(t)]+\gamma'\mathcal{L}_{\hat{Y}'}[\hat{\rho}(t)].\tag{F3}$$

We consider an example of this scenario below.

1. Example: Qubit noise channels

Consider a qubit simultaneously undergoing amplitude damping (loss) and dephasing such that the Lindbladian master equation is thus

$$\partial_t\hat{\rho}(t)=\gamma_{\mathrm{deph}}\mathcal{L}\left(\hat{\sigma}_z\right)+\gamma_{\mathrm{loss}}\mathcal{L}\left(\hat{\sigma}_-\right).$$

After evolving for a time T, the individual solutions to the amplitude damping and dephasing channels are, respectively,

$$\Lambda_{\mathrm{deph}}\left(\hat{\rho}\right)=\begin{bmatrix}\rho_{1,1}&\rho_{1,0}e^{-\gamma_{\mathrm{deph}}T}\\ \rho_{1,0}^*e^{-\gamma_{\mathrm{deph}}T}&\rho_{0,0}\end{bmatrix},$$
$$\Lambda_{\mathrm{loss}}\left(\hat{\rho}\right)=\begin{bmatrix}\rho_{1,1}e^{-\gamma_{\mathrm{loss}}T}&\rho_{1,0}e^{-\gamma_{\mathrm{loss}}T/2}\\ \rho_{1,0}^*e^{-\gamma_{\mathrm{loss}}T/2}&\rho_{0,0}+\rho_{1,1}\left(1-e^{-\gamma_{\mathrm{loss}}T}\right)\end{bmatrix},$$

where $\rho_{j,k}=\langle j|\hat{\rho}|k\rangle$. Even though $[\hat{\sigma}_z,\hat{\sigma}_-]\neq 0$, these channels commute such that the solution to the Lindbladian master equation is given by $\Lambda_{\mathrm{deph}}\circ\Lambda_{\mathrm{loss}}\left(\hat{\rho}\right)=\Lambda_{\mathrm{loss}}\circ\Lambda_{\mathrm{deph}}\left(\hat{\rho}\right)$.

This means that estimating a weak loss $\sqrt{\gamma_{\mathrm{loss}}}\ll 1$ is unaffected by dephasing, since the optimal initial state for sensing Λ_{loss} is $|1\rangle\langle 1|$, which is stabilized by Λ_{deph}. By Eq. (F2), the QFI with respect to $\sqrt{\gamma_{\mathrm{loss}}}$ is thus $\mathcal{I}_Q(\sqrt{\gamma_{\mathrm{loss}}})=4T$.

On the other hand, estimating a weak dephasing $\sqrt{\gamma_{\mathrm{deph}}}\ll 1$ is strongly affected by loss. (We assume that $\gamma_{\mathrm{loss}}>0$ is fixed and known.) Eq. (F1) implies that the QFI for any initial state is Rayleigh cursed, since the state after Λ_{loss} is either full rank or $|0\rangle\langle 0|$, which has zero QFI. The QFI with respect to $\sqrt{\gamma_{\mathrm{deph}}}$ therefore vanishes in the limit of $\gamma_{\mathrm{deph}}\to 0$ for any single-qubit initial state.

2. Multiple jump operators

Suppose instead that the same weak decay rate γ is common to m jump operators $\{\hat{Y}_j\}_{j=1}^m$ such that the master equation is

$$\partial_t\hat{\rho}(t)=\gamma\sum_{j=1}^m\mathcal{L}_{\hat{Y}_j}[\hat{\rho}(t)].$$

Then, the evolution of $\hat{\rho}$ is approximated in the limit of $\gamma Tmc\ll 1$, where $c=\max_j\mathrm{Tr}[\hat{Y}_j\hat{\rho}\hat{Y}_j^\dagger]$, by the following

channel:

$$\Lambda_\gamma(\hat{\rho}) = \sum_{j=1}^{m} \left(\hat{K}_{0,j}\,\hat{\rho}\,\hat{K}_{0,j}^\dagger + \hat{K}_{1,j}\,\hat{\rho}\,\hat{K}_{1,j}^\dagger \right),$$

where the Kraus operators are

$$\hat{K}_{0,j} := \frac{1}{\sqrt{m}} - \frac{\gamma T\sqrt{m}}{2}\hat{Y}_j^\dagger \hat{Y}_j,$$
$$\hat{K}_{1,j} := \sqrt{\gamma T}\hat{Y}_j.$$

By a similar argument to the $m = 1$ case above, then

$$\mathcal{I}_Q(\sqrt{\gamma} = 0) = 4T\sum_{j=1}^{m}\langle \hat{Y}_j^\dagger \hat{\Pi}_\perp \hat{Y}_j\rangle. \tag{F4}$$

For example, in Eq. (51), $\gamma T = \sigma^2$, $\hat{Y}_1 = \hat{x}$, and $\hat{Y}_2 = \hat{p}$, such that Eq. (F4) implies Eq. (52).

APPENDIX G: FINITE-DIMENSIONAL SYSTEMS

Here, we consider a random unitary channel acting on a finite-dimensional system. Given a d-dimensional quantum system, the random unitary channel Λ_σ is defined analogously to Eq. (37) as $\Lambda_\sigma(\hat{\rho}) = \int_{-\infty}^{\infty} \mathrm{d}\theta\, p(\theta)\hat{U}_\theta \hat{\rho}\hat{U}_\theta^\dagger$, where $p(\theta) \sim \mathcal{N}(0,\sigma^2)$ and $\hat{U}_\theta$ is a θ-dependent unitary that acts on this system. We discuss the differences between finite-dimensional systems and the infinite-dimensional harmonic oscillator system that we study. In particular, we give an example where the analogous upper bound on the ECQFI to Eq. (32) can be loose for $\sigma > 0$.

For finite-dimensional systems, the optimal initial pure state in the limit of $\sigma \to 0$ is an equal superposition of the states corresponding to the smallest and largest eigenvalues of $\hat{H}$ to maximize $4\mathrm{Var}[\hat{H}]$. If a loss precedes the channel, then the mixed state $\hat{\rho}$ before the encoding needs to instead optimize $\langle \hat{H}\hat{\Pi}_\perp \hat{H}\rangle$.

For $\sigma > 0$, however, the same state may no longer be optimal, even in the lossless case. For example, consider sensing the phase diffusion of a qubit, modeled as a random rotation channel Λ_σ with $\hat{U}_\theta = \exp(-i\theta\hat{H})$ and $\hat{H} = \frac{1}{2}\hat{\sigma}_z$. The optimal initial state for $\sigma > 0$, even allowing for entangled resources, is, e.g., $|\uparrow_x\rangle$ with a QFI of $\sigma^2/(e^{\sigma^2} - 1)$, as this saturates the ECQFI as proved below. In comparison, the naive upper bound on the ECQFI, analogous to Claim 2, is $2/(\sigma^2 + 2)$, which is loose for $\sigma > 0$, unlike for the random bosonic displacement channel.

The ECQFI for the random rotation channel Λ_σ can be found as follows. The channel has the following Kraus operators:

$$\hat{K}_1 = \sqrt{\frac{1}{2}\left(1 + e^{-\frac{1}{2}\sigma^2}\right)}\hat{1}, \quad \hat{K}_2 = \sqrt{\frac{1}{2}\left(1 - e^{-\frac{1}{2}\sigma^2}\right)}\hat{\sigma}_z,$$

such that an upper bound on the ECQFI is [139]

$$\mathcal{I}_Q^{\Lambda_\sigma}(\sigma) \le 4\left\|\sum_j \dot{\hat{K}}_j^\dagger \dot{\hat{K}}_j\right\| = \frac{\sigma^2}{e^{\sigma^2} - 1}.$$

This implies that $|\uparrow_x\rangle$ is optimal and attains the ECQFI.

APPENDIX H: REVIEW OF FINITE-ENERGY GKP STATES

Here, we review the families of finite-energy states that we study, which approximate the ideal GKP state in the high-energy limit.

The ideal GKP grid state is an unnormalized infinite-energy single-mode state consisting of an equal superposition of infinitely many evenly spaced position eigenstates, e.g.,

$$|\mathrm{GKP}_{\mathrm{ideal}}\rangle \propto \sum_{j=-\infty}^{\infty} |\hat{x} = 2j\sqrt{\pi}\rangle = \sum_{j=-\infty}^{\infty} |\hat{p} = 2j\sqrt{\pi}\rangle,$$

where $|\hat{x} = x_0\rangle$ ($|\hat{p} = p_0\rangle$) is the position (momentum) eigenstate at x_0 (p_0) [52]. There are many different continuous families of normalized finite-energy states that approximate the same ideal state in the high-energy limit [143]. One such family, $|\mathrm{GKP}_\Delta\rangle$, is defined as a superposition with a Gaussian window function of $\mathcal{N}(0, \frac{1}{2}\Delta^{-2})$ displaced squeezed single-mode states, with variance $\frac{1}{2}\Delta^2$ and evenly spaced by $2\sqrt{\pi}$, as follows:

$$|\mathrm{GKP}_\Delta\rangle \propto \sum_{j=-\infty}^{\infty} \exp\left(-2\pi\Delta^2 j^2\right)\hat{U}_{2j\sqrt{\pi}}|\mathrm{SMSV}_{\frac{1}{2}\Delta}\rangle,$$

where $\hat{U}_\mu = \exp(-i\mu\hat{p})$ translates $\hat{x}$ to $\hat{x} + \mu$ and the position-basis wave function of the SMSV state is

$$|\mathrm{SMSV}_{\frac{1}{2}\Delta}\rangle = \frac{1}{\pi^{\frac{1}{4}}\sqrt{\Delta}}\int_{-\infty}^{\infty}\mathrm{d}x\ \exp\left(-\frac{x^2}{2\Delta^2}\right)|\hat{x} = x\rangle.$$

This family approximates the ideal GKP state, $|\mathrm{GKP}_{\mathrm{ideal}}\rangle$, in the limit of $\Delta \to 0$. Numerically, for a fixed but small $\Delta > 0$, $|\mathrm{GKP}_\Delta\rangle$ can be approximated by a finite sum over $j = -J, \cdots, J$, where some large integer J is chosen such that $2\pi\Delta^2 J^2 \gg 1$.

The Wigner function of the pure initial state $|\mathrm{GKP}_\Delta\rangle$ is shown in Fig. 7(a), where the marginals fit the Gaussian envelope of $\mathcal{N}(0, \frac{1}{2}\Delta^{-2})$. The mixed state after a loss,

$\hat{\rho} = \Lambda_\eta^{\text{loss}}(|\text{GKP}_\Delta\rangle\langle\text{GKP}_\Delta|)$, is shown in Fig. 7(b), where, visually, the subvacuum peaks grow and move toward the origin. Solely for the purpose of plotting, we calculate these Wigner functions using a different family of finite-energy GKP states $|\text{GKP}'_\varepsilon\rangle$, defined as the normalized result of applying the nonunitary operator $\exp(-\varepsilon\hat{n})$ to the ideal GKP state [143]. This approximation is valid since, when $\frac{1}{24}\Delta^3$ is negligible compared to $\frac{1}{2}\Delta$, then $|\text{GKP}_\Delta\rangle \approx |\text{GKP}'_{\varepsilon=\Delta^2}\rangle$, e.g., this holds for the value of $\varepsilon = 0.01$ used in Fig. 7. We use this family $|\text{GKP}'_\varepsilon\rangle$ here since the Wigner function is more efficient to calculate than in the Fock-basis representation shown in Figs. 7(c)–7(d) [144]. While, at low energies, the different families of finite-energy GKP states diverge, e.g., the behavior of $|\text{GKP}_\Delta\rangle$ as $N \to 0$ seen in Fig. 4(b) may be different from $|\text{GKP}'_{\varepsilon=\Delta^2}\rangle$, we are principally interested in the high-energy limit where they agree.

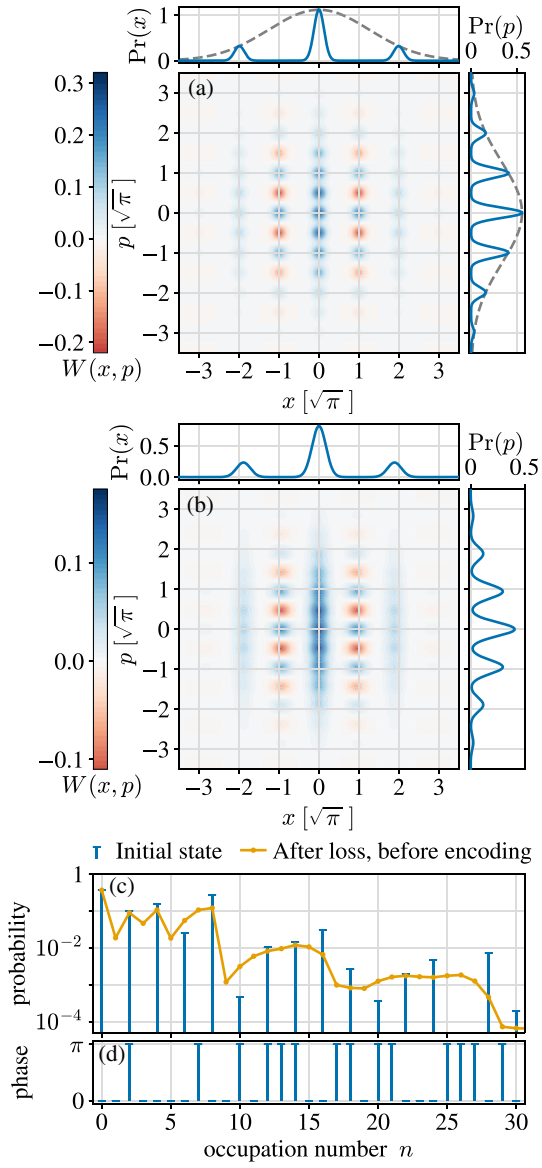

FIG. 7. The GKP finite-energy state for $\Delta^2 \approx 0.1$: Wigner-function and quadrature probability distributions of (a) the pure initial state and (b) the state after a loss of $\eta = 0.1$ before the encoding. The Gaussian envelope of the marginals for the pure initial state is shown. Visually, the loss channel enlarges each subvacuum peak and pulls it toward the origin. (c) Fock-basis probabilities on a logarithmic scale shown up to $n = 30$ before and after the loss, in a truncated Hilbert space of dimension 100. (d) Fock-basis complex phases of the pure initial state. The coefficients are all real.

APPENDIX I: NUMERICAL METHODS

Here, we discuss the numerical methods mentioned in Sec. V F that we use to search for the optimal initial single-mode state in the presence of loss. We first discuss our brute-force approach and then an alternative approach that we have experimented with.

1. Brute-force approach

While we have observed numerically that preparing finite-energy GKP states approaches the ECQFI for $0 < \sigma^2 \ll \eta, \eta_A$ in the limit of large $\langle\hat{n}\rangle = N$, we do not know if this is the only such family of non-Gaussian states. Moreover, while we know that the upper bound in Eq. (36) is loose for a fixed finite $\langle\hat{n}\rangle = N$, we do not know whether a TMSV state with noiseless ancilla is optimal. We search here numerically for non-Gaussian states that might perform better.

We consider sparse superpositions of finitely many Fock states. For simplicity, we consider superpositions equally spaced in the number basis, i.e., $|\psi\rangle = \sum_{k=0}^{K-1} c_k|mk\rangle$ for some spacing m and number of peaks K. For successively larger truncations of the Hilbert space in the number basis, we fix different values of m and K and numerically optimize the basis coefficients $\{c_k\}_{k=0}^{K-1}$ to maximize the QFI using gradient-descent and particle-swarm optimization methods. This is a brute-force approach to finding the optimal state.

We show an example of one such state in Fig 8, which has been found numerically within a truncated Hilbert space of dimension 490. This state, $|\psi_{\text{num.}}\rangle = \sum_{j=0}^{23} c_j|20j\rangle$, where $m = 20$ and $K = 24$, has a QFI of 18.4 for $\langle\hat{n}\rangle = 158.9$, which is within 9% of the ECQFI of 20 for $\sigma^2 = 10^{-6}$ and $\eta = 0.1$. (This value of the QFI appears stable under small perturbations in σ and η.) It is important here that the signal and loss levels are fixed and finite, since the intuition for this sparse state is that, e.g., the transition to $|1\rangle$ is dominated by the signal trajectory from $|0\rangle$ rather than the loss trajectory from $|20\rangle$, as shown in Fig. 8(c). If the loss was larger or the signal was smaller, then we would expect a larger separation m to be necessary. The optimal measurement for this state

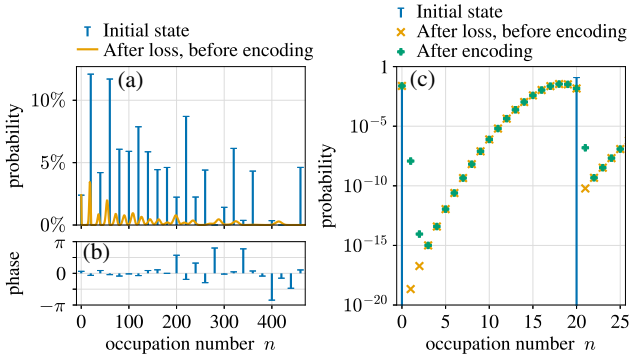

FIG. 8. The Fock-basis (a) probabilities and (b) complex phases of the non-Gaussian initial state $|\psi_{\text{num.}}\rangle$ found by numerical optimization. The probabilities after the loss of $\eta = 0.1$ but before the encoding are a sum of binomial distributions. (c) An example of two neighboring components. For $\sigma = 10^{-3}$, the transition to $|1\rangle$ is dominated by the signal trajectory $|0\rangle \mapsto |1\rangle$ rather than the loss trajectory $|20\rangle \mapsto |1\rangle$. However, this is not the case for all neighboring components of $|\psi_{\text{num.}}\rangle$.

is a superposition of number-resolving measurements, i.e., projections onto different linear combinations of the Fock states, similar to the generalized parity measurements for binomial codes [60]. We conjecture that optimizing over superpositions of unequally spaced but still sparsely separated Fock states in larger truncations of the Hilbert space can get arbitrarily close to the ECQFI for any $\sigma^2 \ll \eta$.

2. Biconvex optimization

Here, we describe this task as a biconvex-optimization problem and discuss how, in principle, it might be solved without using a brute-force approach.

The general problem of finding the optimal protocol, i.e., the initial state and measurement scheme, for sensing a parameter θ encoded by a quantum channel Λ_θ may be formulated as follows. This problem reduces to biconvex optimization of the function $f(\hat{\rho}, X)$, given by [145,146]

$$f(\hat{\rho}, X) := -\text{Tr}\Big[\hat{\rho}\Big(-\Lambda_\theta^\dagger(X^2) + \dot{\Lambda}_\theta^\dagger(X)\Big)\Big],$$

where $\Lambda_\theta^\dagger$ is the conjugate channel of Λ_θ in the Heisenberg-picture sense, $\dot{\Lambda}_\theta^\dagger = \mathrm{d}\Lambda_\theta^\dagger/\mathrm{d}\theta$, and $X \in \mathcal{L}^2(\mathcal{H})$ on a Hilbert space $\mathcal{H}$ with finite dimension (e.g., on a truncated bosonic Hilbert space). Here, the biconvexity of f means that f is convex with respect to $\hat{\rho}$ if X is held constant and with respect to X if $\hat{\rho}$ is held constant, but not necessarily with respect to both at once. The global minimum of f exists and is the negative of the CQFI in Eq. (28),

$$\mathcal{I}_Q^{\Lambda_\theta,\text{no ancilla}}(\theta) = -\inf_X \inf_{\hat{\rho}} f(\hat{\rho}, X), \tag{I1}$$

where the order of the infima may be exchanged and the negative sign is included as a convention such that f is biconvex rather than biconcave. If X is the SLD of $\Lambda_\theta(\hat{\rho})$ with respect to θ, then $f(\hat{\rho}, X) = -\mathcal{I}_Q^{\Lambda_\theta(\hat{\rho})}$ is the negative of the QFI given the initial state $\hat{\rho}$. This biconvex formulation can also be applied to optimize the CQFI subject to a convex constraint. For example, the average-energy–constrained CQFI is given by

$$\mathcal{I}_Q^{\Lambda_\theta,N,\text{no ancilla}}(\theta) = -\inf_X \inf_{\hat{\rho}} f(\hat{\rho}, X) \quad \text{subject to } \text{Tr}\big[\hat{n}\hat{\rho}\big] = N. \tag{I2}$$

This is a biconvex-optimization problem since the set of density matrices that satisfy this average-energy constraint of $\text{Tr}\big[\hat{n}\hat{\rho}\big] = N$ is convex.

In Ref. [145], these biconvex-optimization problems have been proposed to be solved locally using an alternating convex search (ACS) [147]. The ACS proceeds from a given starting point $\hat{\rho}_0$ by finding the corresponding SLD X_0 that solves the convex problem with $|\psi_0\rangle$ held constant, then finds the corresponding $\hat{\rho}_1$ that solves the convex problem with X_0 held constant, and so on. Although the ACS is monotonic and converges, it is only a local search and is not guaranteed to find the global minimum, although different random starting points $\hat{\rho}_0$ can be used to explore the nonconvex geometry [147].

In theory, a deterministic global optimization program (GOP) exists for solving biconvex problems [148]. The GOP, however, only converges in exponential time in the number of confounding variables. Here, this would be proportional to the dimension of the truncated Hilbert space squared in the worst case. Although methods exist to reduce the number of confounding variables in particular cases [149], the GOP is impractical in general. This limitation of the GOP and success of the ACS has previously been discussed in the context of other biconvex-optimization problems in quantum information science (see, e.g., Refs. [136,146,150–152]).

While the ACS has proved successful in these other applications, e.g., it typically converges very fast to optimal values from arbitrary initial states, it appears to be less efficient for our problem of sensing σ from Λ_σ. We have implemented this method for calculating both the unconstrained CQFI in Eq. (I1) and the average-energy–constrained CQFI in Eq. (I2). In the unconstrained CQFI case, we have observed a strong dependence on the starting point. Starting from the vacuum state, the ACS does not converge to or surpass the QFI of the finite-energy GKP states even after 10 000 iterations. This method also does not significantly improve upon the finite superpositions of Fock states from the brute-force approach when started from them. We observe improved performance when running the ACS with random pure states. (We sample the real parts of the number-basis coefficients from a

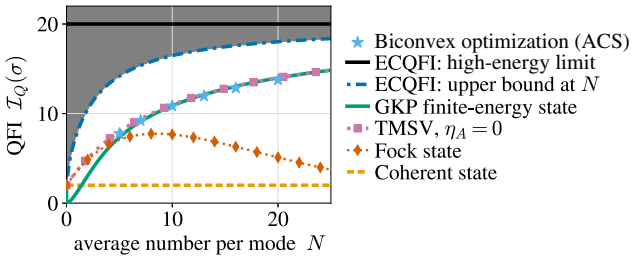

FIG. 9. The QFI versus the initial average occupation number per mode for the states obtained from biconvex optimization via the ACS with the average energy-constraint of Eq. (I2) after 500 iterations and starting from random pure states. The remaining curves corresponding to the QFI from different states are the same as the lossy case in Fig. 4(b).

Gaussian distribution with zero mean and unit variance and then normalize the state. The imaginary parts are zero.) However, the ACS still fails to converge to the QFI of the TMSV and GKP states at finite average energy. Determining what aspects of the geometry of our Rayleigh-cursed bosonic problem limit the ACS is left to future work.

Implementing the ACS for the average energy–constrained CQFI case in Eq. (I2) leads to considerably improved results: the ACS reaches values that are $\geq 98\%$ of the QFI of a finite-energy TMSV after 500 iterations for different N. These results are shown in Fig. 9. This provides strong numerical evidence for our conjecture that the finite-energy CQFI corresponds to the finite-energy TMSV QFI. These results also suggest that the failure of the ACS in the unconstrained CQFI case may be due to the bosonic nature of this problem: we observe that the unconstrained ACS on the truncated Hilbert space does not converge to the QFI of the finite-energy TMSV or GKP states, but requires an average-energy constraint that is much smaller than the truncation does. The states that the ACS finds are not obviously related to the TMSV or GKP states. We defer understanding the full space of states with similar QFI at the same average energy to future work.

APPENDIX J: CLASSICAL NOISE AND THE OPTIMAL INITIAL STATE

Here, we prove the results discussed in Sec. V G about the optimal initial state in the presence of significant classical noise but no loss.

Suppose that the encoding channel is preceded by the classical noise channel $\Lambda^{\text{noise}}_{\Sigma_C}$ in Eq. (22) for a given classical noise covariance matrix Σ_C such that the total channel is $\Lambda'_\sigma = \Lambda_\sigma \circ \Lambda^{\text{noise}}_{\Sigma_C}$. We consider the following three cases in turn: $\Sigma_C = \text{diag}\left(0, \sigma_p^2\right)$, where the random displacements from the classical noise are parallel to (i.e., in the same quadrature, $\hat{p}$, as) those from the signal; $\Sigma_C = \text{diag}\left(\sigma_x^2, 0\right)$, where the classical noise is perpendicular (i.e., in the opposite quadrature, $\hat{x}$) to the signal; and $\Sigma_C = \text{diag}\left(\sigma_x^2, \sigma_p^2\right)$, where the classical noise is in both quadratures, which includes the isotropic case $\Sigma_C = \text{diag}\left(\sigma_C^2, \sigma_C^2\right)$. Whether the Rayleigh curse arises depends on Σ_C as shown below.

1. Parallel classical noise

For $\Sigma_C = \text{diag}\left(0, \sigma_p^2\right)$, we prove that the Rayleigh curse is unavoidable and that the optimal initial state for a given $\langle \hat{n} \rangle = N$ is to prepare an SMSV state.

We first note that the total channel corresponds to a combined classical noise channel:

$$\Lambda'_\sigma = \Lambda_\sigma \circ \Lambda^{\text{noise}}_{\text{diag}\left(0,\sigma_p^2\right)} = \Lambda^{\text{noise}}_{\text{diag}\left(0,\sigma^2+\sigma_p^2\right)}, \tag{J1}$$

where $\Sigma_C = \text{diag}(0, \sigma^2 + \sigma_p^2)$ adds the signal and noise in quadrature since they are uncorrelated.

Equation (J1) holds immediately for Gaussian states, but let us prove it for any initial state $\hat{\rho}$ by direct calculation:

$$\begin{aligned}\Lambda'_\sigma(\hat{\rho}) &= \int_{\mathbb{R}^2} \mathrm{d}\alpha_1 \mathrm{d}\alpha_2 \, p_{\sigma^2}(\alpha_2) \, p_{\sigma_p^2}(\alpha_1) \, \hat{U}_{\alpha_2} \hat{U}_{\alpha_1} \hat{\rho} \hat{U}_{\alpha_1}^\dagger \hat{U}_{\alpha_2}^\dagger \\ &= \int_{\mathbb{R}^2} \mathrm{d}\alpha_1 \mathrm{d}\beta \, p_{\sigma^2}(\beta - \alpha_1) \, p_{\sigma_p^2}(\alpha_1) \, \hat{U}_\beta \hat{\rho} \hat{U}_\beta^\dagger,\end{aligned}$$

where $p_{\varsigma^2}(\alpha) \sim \mathcal{N}(0, \varsigma^2)$, $\beta = \alpha_1 + \alpha_2$, and $\hat{U}_\alpha = \exp(i\alpha\hat{x})$. By recognizing the convolution between two Gaussian distributions, then this equals

$$\begin{aligned}&= \int_{\mathbb{R}} \mathrm{d}\beta \, p_{\sigma^2+\sigma_p^2}(\beta) \, \hat{U}_\beta \hat{\rho} \hat{U}_\beta^\dagger \\ &= \Lambda^{\text{noise}}_{\text{diag}\left(0,\sigma^2+\sigma_p^2\right)}(\hat{\rho}).\end{aligned}$$

By the chain rule, Eq. (J1) implies that the QFI for a given initial state $\hat{\rho}$ is

$$\mathcal{I}_Q^{\Lambda'_\sigma(\hat{\rho})}(\sigma) = \frac{\sigma^2}{\sigma^2 + \sigma_p^2} \mathcal{I}_Q^{\Lambda_\sigma(\hat{\rho})}\left(\sigma = \sqrt{\sigma^2 + \sigma_p^2}\right). \tag{J2}$$

Hence, the ECQFI $\mathcal{I}_Q^{\Lambda'_\sigma}(\sigma)$ for a given value of σ_p is proportional to the noiseless ECQFI $\mathcal{I}_Q^{\Lambda_\sigma}(\sigma)$ evaluated at the signal with variance $\sigma^2 + \sigma_p^2$, which implies Eq. (38). Since preparing an SMSV state is the optimal initial state that attains the noiseless ECQFI for a fixed $\langle \hat{n} \rangle = N$, it is also the optimal initial state in the presence of parallel classical noise. The conversion factor of $\sigma^2/(\sigma^2 + \sigma_p^2)$, however, will introduce the Rayleigh curse for $\sigma \ll \sigma_p$ and mean that the noiseless ECQFI cannot be recovered regardless of the initial state.

Before moving on to discuss the other possible cases of classical noise, we analyze what happens when both parallel classical noise and loss are present. Let the total channel be

$$\Lambda'_\sigma = \Lambda_\sigma \circ \Lambda^{\text{noise}}_{\text{diag}\left(0,\sigma_p^2\right)} \circ \Lambda^{\text{loss}}_\eta.$$

By combining Eqs. 35 and (J2), the ECQFI is then given in Eq. (39). By a similar argument to that made below Eq. (J2), the optimal initial states are the same as for the case with loss but no classical noise.

2. Perpendicular classical noise

For $\Sigma_C = \text{diag}(\sigma_x, 0)$, we prove that, for a fixed finite signal $\sigma > 0$, the noiseless ECQFI can be recovered in the high-energy limit and there exist finite-energy states that do not exhibit the Rayleigh curse as $\sigma \to 0$.

Suppose that we prepare an SMSV state with $\langle \hat{n} \rangle = N$, then, given perpendicular classical noise, the QFI calculated from Eq. (19) is

$$\begin{aligned}\mathcal{I}_Q(\sigma) &= \frac{4}{(2\xi_N)^{-2}\left(\frac{\sigma_x}{\sigma}\right)^2 \frac{\left(\frac{1}{2}\sigma_x^2+\xi_N\right)}{\left(\sigma_x^2+\xi_N\right)^2} + \xi_N^{-1} + 2\sigma^2} \\ &\xrightarrow[N\to\infty]{} \frac{4}{\xi_N^{-1} + 2\sigma^2}.\end{aligned}$$

For a fixed finite signal $\sigma > 0$, the above limit as $N \to \infty$ implies that the noiseless ECQFI in Eq. (32) is recovered for large enough N such that $\sigma\xi_N \gg \sigma_x$. Note that the optimal measurement for an SMSV state in the noiseless case has been to first antisqueeze and then perform a number-resolving measurement. Here, intuitively, this antisqueezing operation squeezes the perpendicular noise such that its effect becomes vanishing in the high-energy limit.

Note that we assume a particular order of limits above: $\lim_{\sigma\to 0} \sup_{|\Psi\rangle} \mathcal{I}_Q(\sigma)$. As discussed in Sec. V E, this is the relevant order in practice because we are interested in small but finite signals $\sigma > 0$. For completeness, we now briefly discuss the opposite order of limits: $\sup_{|\Psi\rangle} \lim_{\sigma\to 0} \mathcal{I}_Q(\sigma)$. While the family of SMSV states is optimal as $N \to \infty$ for a fixed finite $\sigma > 0$, a given SMSV state with fixed $\langle \hat{n} \rangle = N$ exhibits the Rayleigh curse as $\sigma \to 0$ by Claim 1.

Unlike the parallel classical noise case, however, the Rayleigh curse can be avoided using different initial states. For example, we show in Appendix D that preparing a TMSV state with noiseless ancilla, i.e., no noise channels on the ancilla, with a fixed $\langle \hat{n} \rangle = N$ particles per mode yields a QFI of

$$\mathcal{I}_Q(\sigma = 0) = \frac{8N(N+1)}{2N+1} \xrightarrow[N\to\infty]{} 4N, \tag{J3}$$

which is independent of the amount of perpendicular classical noise. This recovers the noiseless QFI of $4N$ for a TMSV state in the high-energy limit, but not the noiseless ECQFI of $4\xi_N \approx 8N$ at $\sigma = 0$. Numerical analysis suggests that finite-energy GKP states are also resilient to perpendicular classical noise.

We note that the above analysis, while promising, does not include the relevant real-world case of perpendicular noise jointly in the presence of optical loss. There, the degradation of the squeezed state from the loss may prevent the complete removal of perpendicular noise.

3. Classical noise in both quadratures

For $\Sigma_C = \text{diag}\left(\sigma_x^2, \sigma_p^2\right)$ with $\sigma_x, \sigma_p > 0$, we combine the above results to show that the Rayleigh curse is inevitable, the noiseless ECQFI cannot be recovered, and that an SMSV state remains the optimal initial state for a fixed $\sigma > 0$ in the high-energy limit. These results apply, e.g., to the isotropic case of $\Sigma_C = \text{diag}\left(\sigma_C^2, \sigma_C^2\right)$ with $\sigma_C > 0$.

The random-displacement channels $\Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,0)}$ and $\Lambda^{\text{noise}}_{\text{diag}(0,\sigma_p^2)}$ commute with each other and obey the following relation:

$$\begin{aligned}\Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,\sigma_p^2)} &= \Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,0)} \circ \Lambda^{\text{noise}}_{\text{diag}(0,\sigma_p^2)} \\ &= \Lambda^{\text{noise}}_{\text{diag}(0,\sigma_p^2)} \circ \Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,0)},\end{aligned}$$

such that, by Eq. (J1), the total channel is

$$\Lambda'_\sigma = \Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,\sigma^2+\sigma_p^2)}.$$

Hence, similarly to Eq. (J2), the QFI for a given initial state $\hat{\rho}$ is

$$\mathcal{I}_Q^{\Lambda'_\sigma(\hat{\rho})}(\sigma) = \frac{\sigma^2}{\sigma^2+\sigma_p^2} \mathcal{I}_Q^{\left(\Lambda_\sigma \circ \Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,0)}\right)(\hat{\rho})}\left(\sigma = \sqrt{\sigma^2+\sigma_p^2}\right).$$

The optimal initial state here, therefore, is the same as for the perpendicular case in which the total channel is $\Lambda_\sigma \circ \Lambda^{\text{noise}}_{\text{diag}(\sigma_x^2,0)}$. For a fixed finite $\sigma > 0$, this implies that preparing an SMSV state is optimal in the high-energy limit but does not recover the noiseless ECQFI due to the $\sigma^2/(\sigma^2+\sigma_p^2)$ factor. In the limit of $\sigma \ll \sigma_p$, the same factor leads to an unavoidable Rayleigh curse.

When we restrict ourselves to study signals above a classical noise floor, e.g., in Sec. V E, we mean signals σ above an isotropic classical noise σ_C, i.e., for $\sigma \gg \sigma_C$, since all protocols exhibit the Rayleigh curse below the noise floor, i.e., for $\sigma \ll \sigma_C$.

APPENDIX K: ADAPTIVE PROTOCOL

Here, we detail an adaptive protocol for our separable measurements scheme (quadrature and number measurements) that converges to a factor of 2 away from the QFIM bound, as shown in Fig. 6(b).

Algorithmically, the adaptive measurement protocol proceeds as follows:

$k \leftarrow 1$
while $k \leq M$ **do**
  $p_k \leftarrow$ outcome of quadrature measurement $\hat{p}$
  Displace quadrature $\hat{p}$ by $-\frac{p_k}{k}$
  $n_k \leftarrow$ outcome of number measurement $\hat{n}$
  $k \leftarrow k+1$
end while

Let us denote $\langle\hat{p}\rangle$ after the kth step as μ_k. Since $p_k = \mu_{k-1} + n_k$, where n_k is a Gaussian random variable $n_k \sim \mathcal{N}\left(0, \frac{1}{2}\right)$, then, by induction, $\mu_k = -(1/k)\sum_{i=1}^{k} n_i$. This implies that $\Delta\mu_k = 1/\sqrt{2k}$, i.e., the "size" of μ_k is equal to the statistical uncertainty in estimating μ after k.

After all of the measurements are performed, we then estimate μ_i with $\tilde{\mu}_i = \mu_i - \mu_M$ and estimate σ with $\varsigma = \sqrt{(2/M)\sum_i n_i - (1/M)\sum_i \tilde{\mu}_i^2}$.

APPENDIX L: PROOF OF CLAIM 5

Here, we prove Claim 5. We first calculate the QFI and then explicitly construct the joint measurement, which is asymptotically optimal.

In the limit of $\sigma \to 0$, the QFI with respect to σ for a pure initial state is $4\mathrm{Var}[\hat{H}]$ by Claim 4. We can assume, without loss of generality, that $\mu \ll 1$. Then, the QFI with respect to μ can be shown to also be $4\mathrm{Var}[\hat{H}]$ by expanding the channel $\Lambda_{\sigma,\mu}$ in Eq. (46) to second order in μ and σ, since higher orders will not contribute to the QFI.

We now construct a joint measurement the CFIM of which saturates these QFIs simultaneously. Without loss of generality, we assume that $\langle\hat{H}\rangle = 0$ and let $V := \mathrm{Var}[\hat{H}]$ for the initial pure state $|\psi\rangle$. Then, $|\phi_1\rangle := (1/\sqrt{V})\hat{H}|\psi\rangle$ and $|\phi_2\rangle := (1/\sqrt{l})(\hat{H}^2 - V)|\psi\rangle$ are orthogonal to $|\psi\rangle$, where l is a normalization factor.

The final state $\Lambda_{\sigma,\mu}(|\psi\rangle\langle\psi|)$ can be expanded up to second order in μ and σ using this orthonormal set of states $\{|\psi\rangle, |\phi_1\rangle, |\phi_2\rangle\}$ as

$$\begin{pmatrix} 1-\left(\mu^2+\sigma^2\right)V & -i\mu\sqrt{V} & -\frac{1}{2}\left(\mu^2+\sigma^2\right)l \\ i\mu\sqrt{V} & \left(\mu^2+\sigma^2\right)V & 0 \\ -\frac{1}{2}\left(\mu^2+\sigma^2\right)l & 0 & 0 \end{pmatrix},$$

such that we can restrict attention to the two-dimensional subspace of $|\psi\rangle$ and $|\phi_1\rangle$.

We want to consider joint measurements on M copies of the final state $\Lambda_{\sigma,\mu}(|\psi\rangle\langle\psi|)$. Consider the following collective states:

$$|e_i\rangle := |\psi\rangle^{\otimes(i-1)}|\phi_1\rangle|\psi\rangle^{\otimes(M-i)}.$$

Let the symmetric state be $|e_s\rangle := (1/\sqrt{M})\sum_i |e_i\rangle$ and let the antisymmetric states be denoted as $\{|e_{a,i}\rangle\}_{i=1}^{M-1}$. In this collective basis, the final state is

$$\begin{aligned}\Lambda_{\sigma,\mu}(|\psi\rangle\langle\psi|)^{\otimes M} \approx \left[1 - MV(\mu^2+\sigma^2)\right]|\psi\rangle^{\otimes M}\langle\psi|^{\otimes M} \\ - i\mu\sqrt{MV}\left(|e_s\rangle\langle\psi|^{\otimes M} - |\psi\rangle^{\otimes M}\langle e_s|\right) \\ + V\left(\sigma^2+\mu^2 M\right)|e_s\rangle\langle e_s| \\ + \sigma^2 V\sum_{i=1}^{M-1}|e_{a,i}\rangle\langle e_{a,i}|.\end{aligned}$$

We now calculate the CFIM obtained with projective measurement onto $(1/\sqrt{2})\left(|\psi\rangle^{\otimes M} + i|e_s\rangle\right)$ and $\left\{|e_{a,i}\rangle\right\}_{i=1}^{M-1}$ and show that it is asymptotically optimal. The probability to be projected onto the antisymmetric subspace is $\sigma^2 V(M-1)$, which provides the information about σ, while the probability to be projected onto $(1/\sqrt{2})\left(|\psi\rangle^{\otimes M} + i|e_s\rangle\right)$ is

$$p = \frac{1}{2}\left(1 - \sigma^2(M-1)V\right) + \sqrt{MV}\mu,$$

which provides the information about μ but no information about σ in the limit of $\sigma \to 0$. The CFIM with respect to μ and σ in the limit of $\sigma \to 0$ is

$$\mathcal{I}_C = \begin{pmatrix} 4MV & 0 \\ 0 & 4(M-1)V \end{pmatrix},$$

which saturates the QFIM and implies that this joint protocol is asymptotically optimal.

[1] E. P. Verlinde and K. M. Zurek, Observational signatures of quantum gravity in interferometers, Phys. Lett. B **822**, 136663 (2021).

[2] D. Li, V. S. H. Lee, Y. Chen, and K. M. Zurek, Interferometer response to geontropic fluctuations, Phys. Rev. D **107**, 024002 (2023).

[3] L. McCuller, Single-photon signal sideband detection for high-power Michelson interferometers, arxiv:2211.04016 [hep-ex, physics:physics, physics:quant-ph].

[4] A. Chou, H. Glass, H. R. Gustafson, C. Hogan, B. L. Kamai, O. Kwon, R. Lanza, L. McCuller, S. S. Meyer, J. Richardson, C. Stoughton, R. Tomlin, and R. Weiss, The holometer: An instrument to probe Planckian quantum geometry, Class. Quantum Grav. **34**, 065005 (2017).

[5] S. M. Vermeulen, L. Aiello, A. Ejlli, W. L. Griffiths, A. L. James, K. L. Dooley, and H. Grote, An experiment for

observing quantum gravity phenomena using twin tabletop 3D interferometers, Class. Quantum Grav. **38**, 085008 (2021).
[6] J. D. Romano and Neil J. Cornish, Detection methods for stochastic gravitational-wave backgrounds: A unified treatment, Living Rev. Relativ. **20**, 2 (2017).
[7] B. P. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, V. Adya, C. Affeldt, M. Agathos *et al.*, Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA, Living Rev. Relativ. **23**, 1 (2020).
[8] J. Aasi *et al.*, Advanced LIGO, Class. Quantum Grav. **32**, 074001 (2015).
[9] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardin *et al.*, Advanced Virgo: A second-generation interferometric gravitational wave detector, Class. Quantum Grav. **32**, 024001 (2015).
[10] T. Akutsu, M. Ando, K. Arai, Y. Arai, S. Araki, A. Araya, N. Aritomi, H. Asada, Y. Aso, S. Atsuta *et al.*, KAGRA: 2.5 generation interferometric gravitational wave detector, Nat. Astron. **3**, 35 (2019).
[11] J. E. Kim and G. Carosi, Axions and the strong *CP* problem, Rev. Mod. Phys. **82**, 557 (2010).
[12] K. Choi, S. H. Im, and C. S. Shin, Recent progress in the physics of axions and axion-like particles, Annu. Rev. Nucl. Part. Sci. **71**, 225 (2021).
[13] L. J. Rosenberg and K. A. Van Bibber, Searches for invisible axions, Phys. Rep. **325**, 1 (2000).
[14] P. W. Graham, I. G. Irastorza, S. K. Lamoreaux, A. Lindner, and K. A. van Bibber, Experimental searches for the axion and axion-like particles, Annu. Rev. Nucl. Part. Sci. **65**, 485 (2015).
[15] M. Tsang, H. M. Wiseman, and C. M. Caves, Fundamental quantum limit to waveform estimation, Phys. Rev. Lett. **106**, 090401 (2011).
[16] H. Miao, R. X. Adhikari, Y. Ma, B. Pang, and Y. Chen, Towards the fundamental quantum limit of linear measurements of classical signals, Phys. Rev. Lett. **119**, 050801 (2017).
[17] J. W. Gardner, T. Gefen, S. A. Haine, J. J. Hope, and Y. Chen, Achieving the fundamental quantum limit of linear waveform estimation, Phys. Rev. Lett. **132**, 130801 (2024).
[18] C. L. Latune, B. M. Escher, R. L. de Matos Filho, and L. Davidovich, Quantum limit for the measurement of a classical force coupled to a noisy quantum-mechanical oscillator, Phys. Rev. A **88**, 042112 (2013).
[19] S. Ng, S. Z. Ang, T. A. Wheatley, H. Yonezawa, A. Furusawa, E. H. Huntington, and M. Tsang, Spectrum analysis with quantum dynamical systems, Phys. Rev. A **93**, 042121 (2016).
[20] M. Tsang, Quantum noise spectroscopy as an incoherent imaging problem, Phys. Rev. A **107**, 012611 (2023).
[21] C. W. Gardiner and M. J. Collett, Input and output in damped quantum systems: Quantum stochastic differential equations and the master equation, Phys. Rev. A **31**, 3761 (1985).
[22] R. Kubo, The fluctuation-dissipation theorem, Rep. Prog. Phys. **29**, 255 (1966).
[23] A. Buonanno and Y. Chen, Signal recycled laser-interferometer gravitational-wave detectors as optical springs, Phys. Rev. D **65**, 042001 (2002).
[24] S. L. Danilishin and F. Y. Khalili, Quantum measurement theory in gravitational-wave detectors, Living Rev. Relativ. **15**, 1 (2012).
[25] A. V. Oppenheim and G. C. Verghese, *Signals, Systems & Inference* (Pearson London, London, 2017).
[26] One possible proof of Eq. (3) uses the identity $2\pi\delta(\Omega-\Omega')|_{\Omega=\Omega'} = \lim_{T\to\infty} T$, which follows from $2\pi\delta(a) = \int_{-\infty}^{\infty} \mathrm{d}t\, e^{iat} = \lim_{T\to\infty}\int_0^T \mathrm{d}t\, e^{iat} = \lim_{T\to\infty}\delta_T(a)$, where $\delta_T(a) = (e^{iaT}-1)/(ia)$ is a finite-time approximation to the delta function obeying $\delta_T(0) = T$.
[27] C. Bond, D. Brown, A. Freise, and K. A. Strain, Interferometer techniques for gravitational-wave detection, Living Rev. Relativ. **19**, 1 (2016).
[28] The approximation that $\delta(\Omega-\Omega')|_{\Omega=\Omega'} \approx T/(2\pi)$ for large but finite integration time T provides the correct $\sqrt{2/T}$ normalization in Eq. (10) to satisfy the canonical commutation relations compared to the $\sqrt{1/(\pi T)}$ normalization in Ref. [17], which incorrectly assumes that $\delta(\Omega-\Omega')|_{\Omega=\Omega'} \approx T$.
[29] We consider the variance of the parts $\hat{q}_j(\Omega)$ here since they are nonstationary, such that the power spectral density in Eq. (4) is not defined.
[30] The value of $G_{py}(\Omega)$ in Eq. (14) is equal to twice the value of $G(\Omega)$ from Ref. [70] since we use a different normalization for the input-output relation.
[31] W. H. Zurek, Decoherence, einselection, and the quantum origins of the classical, Rev. Mod. Phys. **75**, 715 (2003).
[32] W. H. Zurek, in *Quantum Decoherence: Poincaré Seminar 2005* (Birkhäuser Basel, Basel, 2007), pp. 1–31..
[33] S. L. Braunstein and C. M. Caves, Statistical distance and the geometry of quantum states, Phys. Rev. Lett. **72**, 3439 (1994).
[34] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2009).
[35] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Quantum metrology with nonclassical states of atomic ensembles, Rev. Mod. Phys. **90**, 035005 (2018).
[36] M. Takeoka, M. Ban, and M. Sasaki, Unambiguous quantum-state filtering, Phys. Rev. A **68**, 012307 (2003).
[37] A. Monras, Phase space formalism for quantum estimation of Gaussian states, arxiv:1303.3682 [quant-ph].
[38] M. Tsang, R. Nair, and X.-M. Lu, Quantum theory of superresolution for two incoherent optical point sources, Phys. Rev. X **6**, 031033 (2016).
[39] S. Zhou and L. Jiang, Modern description of Rayleigh's criterion, Phys. Rev. A **99**, 013808 (2019).
[40] T. Gefen, A. Rotem, and A. Retzker, Overcoming resolution limits with quantum sensing, Nat. Commun. **10**, 4992 (2019).
[41] S. L. Mouradian, N. Glikin, E. Megidish, K.-I. Ellers, and H. Haeffner, Quantum sensing of intermittent stochastic signals, Phys. Rev. A **103**, 032419 (2021).

[42] Y.-x. Liu, Ş. K. Özdemir, A. Miranowicz, and N. Imoto, Kraus representation of a damped harmonic oscillator and its application, Phys. Rev. A **70**, 042308 (2004).

[43] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology, Nat. Phys. **7**, 406 (2011).

[44] J. Kolodynski and R. Demkowicz-Dobrzanski, Efficient tools for quantum metrology with uncorrelated noise, New J. Phys. **15**, 073043 (2013).

[45] R. Demkowicz-Dobrzański and L. Maccone, Using entanglement against noise in quantum metrology, Phys. Rev. Lett. **113**, 250801 (2014).

[46] R. Demkowicz-Dobrzański, K. Banaszek, and R. Schnabel, Fundamental quantum interferometry bound for the squeezed-light-enhanced gravitational wave detector GEO 600, Phys. Rev. A **88**, 041802 (2013).

[47] J. Aasi *et al.*, Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light, Nat. Photonics **7**, 613 (2013).

[48] M. Tse *et al.*, Quantum-enhanced Advanced LIGO detectors in the era of gravitational-wave astronomy, Phys. Rev. Lett. **123**, 231107 (2019).

[49] L. McCuller, C. Whittle, D. Ganapathy, K. Komori, M. Tse, A. Fernandez-Galiana, L. Barsotti, P. Fritschel, M. MacInnis, F. Matichard, K. Mason, N. Mavalvala, R. Mittleman, Haocun Yu, M. E. Zucker, and M. Evanss, Frequency-dependent squeezing for Advanced LIGO, Phys. Rev. Lett. **124**, 171102 (2020).

[50] LIGO O4 Detector Collaboration, D. Ganapathy, W. Jia, M. Nakano, V. Xu, N. Aritomi, T. Cullen, N. Kijbunchoo, S. E. Dwyer, A. Mullavey *et al.*, Broadband quantum enhancement of the LIGO detectors with frequency-dependent squeezing, Phys. Rev. X **13**, 041021 (2023).

[51] W. Górecki, A. Riccardi, and L. Maccone, Quantum metrology of noisy spreading channels, Phys. Rev. Lett. **129**, 240503 (2022).

[52] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, Phys. Rev. A **64**, 012310 (2001).

[53] K. A. Gilmore, M. Affolter, R. J. Lewis-Swan, D. Barberena, E. Jordan, A. M. Rey, and J. J. Bollinger, Quantum-enhanced sensing of displacements and electric fields with two-dimensional trapped-ion crystals, Science **373**, 673 (2021).

[54] X. Deng, S. Li, Z.-J. Chen, Z. Ni, Y. Cai, J. Mai, L. Zhang, P. Zheng, H. Yu, C.-L. Zou *et al.*, Quantum-enhanced metrology with large Fock states, Nat. Phys. **20**, 1874 (2024).

[55] A. Agrawal, A. V. Dixit, T. Roy, S. Chakram, K. He, R. K. Naik, D. I. Schuster, and A. Chou, Stimulated emission of signal photons from dark matter waves, Phys. Rev. Lett. **132**, 140801 (2024).

[56] Q. Zhuang, J. Preskill, and L. Jiang, Distributed quantum sensing enhanced by continuous-variable error correction, New J. Phys. **22**, 022001 (2020).

[57] A. J. Brady, A. Eickbusch, S. Singh, J. Wu, and Q. Zhuang, Advances in bosonic quantum error correction with Gottesman-Kitaev-Preskill codes: Theory, engineering and applications, Prog. Quantum Electron. **93**, 100496 (2024).

[58] S. Krastanov, V. V. Albert, C. Shen, C.-L. Zou, R. W. Heeres, B. Vlastakis, R. J. Schoelkopf, and L. Jiang, Universal control of an oscillator with dispersive coupling to a qubit, Phys. Rev. A **92**, 040303 (2015).

[59] A. Eickbusch, V. Sivak, A. Z. Ding, S. S. Elder, S. R. Jha, J. Venkatraman, B. Royer, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, Fast universal control of an oscillator with weak dispersive coupling to a qubit, Nat. Phys. **18**, 1464 (2022).

[60] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, New class of quantum error-correcting codes for a bosonic mode, Phys. Rev. X **6**, 031006 (2016).

[61] A. Monras and F. Illuminati, Measurement of damping and temperature: Precision bounds in Gaussian dissipative channels, Phys. Rev. A **83**, 012315 (2011).

[62] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (Springer Science & Business Media, Pisa, 2011).

[63] O. Pinel, P. Jian, N. Treps, C. Fabre, and D. Braun, Quantum parameter estimation using general single-mode Gaussian states, Phys. Rev. A **88**, 040102 (2013).

[64] Y. Gao and H. Lee, Bounds on quantum multiple-parameter estimation with Gaussian state, Eur. Phys. J. D **68**, 347 (2014).

[65] H. Nagaoka, in *Asymptotic Theory of Quantum Statistical Inference: Selected Papers* (World Scientific, 2005), pp. 100–112.

[66] L. O. Conlon, J. Suzuki, P. K. Lam, and S. M. Assad, Efficient computation of the Nagaoka-Hayashi bound for multiparameter estimation with separable measurements, npj Quantum Inf. **7**, 110 (2021).

[67] M. Parniak, S. Borówka, K. Boroszko, W. Wasilewski, K. Banaszek, and R. Demkowicz-Dobrzański, Beating the Rayleigh limit using two-photon interference, Phys. Rev. Lett. **121**, 250503 (2018).

[68] Y. L. Len, Multiparameter estimation for qubit states with collective measurements: A case study, New J. Phys. **24**, 033037 (2022).

[69] If the temporal-basis modes $w(\tau)$ are complex, then their Fourier transforms will have independent positive and negative frequency components, unlike the real basis of cosine and sine functions in Eq. (10). Instead, the real and imaginary parts of each complex temporal mode should be studied, where each part reduces to the canonical noise-estimation problem of a single harmonic oscillator.

[70] S. M. Vermeulen, T. Cullen, D. Grass, I. A. O. MacMillan, A. J. Ramirez, J. Wack, B. Korzh, V. S. H. Lee, K. M. Zurek, C. Stoughton, and L. McCuller, Photon-counting interferometry to detect geontropic space-time fluctuations with GQuEST, Phys. Rev. X **15**, 011034 (2025).

[71] A. V. Gorshkov, A. André, M. D. Lukin, and A. S. Sørensen, Photon storage in Λ-type optically dense atomic media. I. Cavity model, Phys. Rev. A **76**, 033804 (2007).

[72] K. R. Brown, K. M. Dani, D. M. Stamper-Kurn, and K. B. Whaley, Deterministic optical Fock-state generation, Phys. Rev. A **67**, 043818 (2003).
[73] J. M. Geremia, Deterministic and nondestructively verifiable preparation of photon number states, Phys. Rev. Lett. **97**, 073601 (2006).
[74] J. Hastrup and U. L. Andersen, Protocol for generating optical Gottesman-Kitaev-Preskill states with cavity QED, Phys. Rev. Lett. **128**, 170503 (2022).
[75] S. Chen, Y.-A. Chen, T. Strassel, Z.-S. Yuan, B. Zhao, J. Schmiedmayer, and J.-W. Pan, Deterministic and storable single-photon source based on a quantum memory, Phys. Rev. Lett. **97**, 173004 (2006).
[76] S. E. Thomas, L. Wagner, R. Joos, R. Sittig, C. Nawrath, P. Burdekin, I. Maillette de Buy Wenniger, M. J. Rasiah, T. Huber-Loyola, S. Sagona-Stophel, S. Höfling, M. Jetter, P. Michler, I. A. Walmsley, S. L. Portalupi, and P. M. Ledingham, Deterministic storage and retrieval of telecom light from a quantum dot single-photon source interfaced with an atomic quantum memory, Sci. Adv. **10**, eadi7346 (2024).
[77] P. Farrera, G. Heinze, B. Albrecht, M. Ho, M. Chávez, C. Teo, N. Sangouard, and H. de Riedmatten, Generation of single photons with highly tunable wave shape from a cold atomic ensemble, Nat. Commun. **7**, 13556 (2016).
[78] O. Katz and O. Firstenberg, Light storage for one second in room-temperature alkali vapor, Nat. Commun. **9**, 2074 (2018).
[79] K. Bergmann, H. Theuer, and B. W. Shore, Coherent population transfer among quantum states of atoms and molecules, Rev. Mod. Phys. **70**, 1003 (1998).
[80] N. V. Vitanov, A. A. Rangelov, B. W. Shore, and K. Bergmann, Stimulated Raman adiabatic passage in physics, chemistry, and beyond, Rev. Mod. Phys. **89**, 015006 (2017).
[81] I. Bloch, Ultracold quantum gases in optical lattices, Nat. Phys. **1**, 23 (2005).
[82] A. M. Kaufman and K.-K. Ni, Quantum science with optical tweezer arrays of ultracold atoms and molecules, Nat. Phys. **17**, 1324 (2021).
[83] A. Reiserer and G. Rempe, Cavity-based quantum networks with single atoms and optical photons, Rev. Mod. Phys. **87**, 1379 (2015).
[84] K. Hammerer, A. S. Sørensen, and E. S. Polzik, Quantum interface between light and atomic ensembles, Rev. Mod. Phys. **82**, 1041 (2010).
[85] K. Heshami, D. G. England, P. C. Humphreys, P. J. Bustard, V. M. Acosta, J. Nunn, and B. J. Sussman, Quantum memories: Emerging applications and recent advances, J. Mod. Opt. **63**, 2005 (2016).
[86] S.-H. Wei, B. Jing, X.-Y. Zhang, J.-Y. Liao, C.-Z. Yuan, B.-Y. Fan, C. Lyu, D.-L. Zhou, Y. Wang, G.-W. Deng *et al.*, Towards real-world quantum networks: A review, Laser Photonics Rev. **16**, 2100219 (2022).
[87] R. Acharya, S. Brebels, A. Grill, J. Verjauw, Ts. Ivanov, D. Perez Lozano, D. Wan, J. Van Damme, A. M. Vadiraj, M. Mongillo, B. Govoreanu, J. Craninckx, I. P. Radu, K. De Greve, G. Gielen, F. Catthoor, and A. Potočnik, Multiplexed superconducting qubit control at millikelvin temperatures with a low-power cryo-CMOS multiplexer, Nat. Electron. **6**, 900 (2023).
[88] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation, Phys. Rev. A **69**, 062320 (2004).
[89] L. Chen, H.-X. Li, Y. Lu, C. W. Warren, C. J. Križan, S. Kosen, M. Rommel, S. Ahmed, A. Osman, J. Biznárová, A. F. Roudsari, B. Lienhard, M. Caputo, K. Grigoras, L. Grönberg, J. Govenius, A. F. Kockum, P. Delsing, J. Bylander, and G. Tancredi, Transmon qubit readout fidelity at the threshold for quantum error correction without a quantum-limited amplifier, npj Quantum Inf. **9**, 26 (2023).
[90] M. H. Devoret and R. J. Schoelkopf, Superconducting circuits for quantum information: An outlook, Science **339**, 1169 (2013).
[91] D. P. DiVincenzo, The physical implementation of quantum computation, Fortschritte Phys. **48**, 771 (2000).
[92] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, L. Frunzio, M. Mirrahimi, and M. H. Devoret, Quantum error correction of a qubit encoded in grid states of an oscillator, Nature **584**, 368 (2020).
[93] A. A. Diringer, E. Blumenthal, A. Grinberg, L. Jiang, and S. Hacohen-Gourgy, Conditional-NOT displacement: Fast multioscillator control with a single qubit, Phys. Rev. X **14**, 011055 (2024).
[94] R. Moretti, H. A. Corti, D. Labranca, F. Ahrens, G. Avallone, and D. Babusci, Design and simulation of a transmon qubit chip for axion detection, IEEE Trans. Appl. Supercond. **34**, 1 (2024).
[95] S. K. Lamoreaux, K. A. van Bibber, K. W. Lehnert, and G. Carosi, Analysis of single-photon and linear amplifier detectors for microwave cavity dark matter axion searches, Phys. Rev. D **88**, 035020 (2013).
[96] If measuring signals below the classical noise floor, then we expect the regime $\eta \gg 4\sigma_C^2$ to be also relevant; see the discussion surrounding Eq. (39) for the one-dimensional case.
[97] A. I. Renzini, B. Goncharov, A. C. Jenkins, and P. M. Meyers, Stochastic gravitational-wave backgrounds: Current detection efforts and future prospects, Galaxies **10**, 34 (2022).
[98] A. Buikema *et al.*, Sensitivity and performance of the Advanced LIGO detectors in the third observing run, Phys. Rev. D **102**, 062003 (2020).
[99] This argument assumes a uniform source density, which is valid for the current generation of gravitational-wave observatories. Future observatories are proposed to have much further reach [113,114], such that this assumption is no longer valid and a more detailed astrophysical model is required.
[100] K. Chatziioannou, J. A. Clark, A. Bauswein, M. Millhouse, T. B. Littenberg, and N. Cornish, Inferring the post-merger gravitational wave emission from binary neutron star coalescences, Phys. Rev. D **96**, 124035 (2017).
[101] A. W. Criswell, J. Miller, N. Woldemariam, T. Soultanis, A. Bauswein, K. Chatziioannou, M. W. Coughlin, G. Jones, and V. Mandic, Hierarchical Bayesian method for constraining the neutron star equation of state with

an ensemble of binary neutron star postmerger remnants, Phys. Rev. D **107**, 043021 (2023).
[102] A. Sasli, N. Karnesis, and N. Stergioulas, Exploring the potential for detecting rotational instabilities in binary neutron star merger remnants with gravitational wave detectors, Phys. Rev. D **109**, 043045 (2024).
[103] In the deterministic event-stacking case [100–102], there are threshold effects where the CCRB cannot be saturated [153,154]. This behavior arises from the nonlinear process of frequency estimation at a low signal-to-noise ratio and can be expressed as the discrepancy between the classical Cramér-Rao and Barankin bounds [155,156]. Since finding the frequency of excess power is a stochastic estimation problem, studying the quantum analogs of the Barankin bound [157] may be necessary for understanding the behavior here at low signal-to-noise ratios.
[104] Finding the optimal measurement for fixed finite M would require instead calculating the Nagaoka-Hayashi bound [65,66].
[105] R. Cameron, G. Cantatore, A. Melissinos, G. Ruoso, Y. Semertzidis, H. Halama, D. Lazarus, A. Prodell, F. Nezrick, C. Rizzo *et al.*, Search for nearly massless, weakly coupled particles by optical techniques, Phys. Rev. D **47**, 3707 (1993).
[106] N. Du, N. Force, R. Khatiwada, E. Lentz, R. Ottens, L. Rosenberg, G. Rybka, G. Carosi, N. Woollett, and D. Bowring *et al.*, Search for invisible axion dark matter with the axion dark matter experiment, Phys. Rev. Lett. **120**, 151301 (2018).
[107] A. V. Dixit, S. Chakram, K. He, A. Agrawal, R. K. Naik, D. I. Schuster, and A. Chou, Searching for dark matter with a superconducting qubit, Phys. Rev. Lett. **126**, 141302 (2021).
[108] H. Shi and Q. Zhuang, Ultimate precision limit of noise sensing and dark matter search, npj Quantum Inf. **9**, 27 (2023).
[109] H. Zheng, M. Silveri, R. Brierley, S. Girvin, and K. Lehnert, Accelerating dark-matter axion searches with quantum measurement technology, arxiv:1607.02529 [hep-ph].
[110] K. M. Backes *et al.*, A quantum enhanced search for dark matter axions, Nature **590**, 238 (2021).
[111] C. M. Caves, Quantum-mechanical noise in an interferometer, Phys. Rev. D **23**, 1693 (1981).
[112] L. McCuller, C. Whittle, D. Ganapathy, K. Komori, M. Tse, A. Fernandez-Galiana, L. Barsotti, P. Fritschel, M. MacInnis, and F. Matichard, Frequency-dependent squeezing for Advanced LIGO, Phys. Rev. Lett. **124**, 171102 (2020).
[113] D. Reitze *et al.*, Astro2020: Decadal Survey on Astronomy and Astrophysics, APC white papers, in *Bulletin of the American Astronomical Society* (2019), Vol. 51.
[114] M. Maggiore, C. van den Broeck, N. Bartolo, E. Belgacem, D. Bertacca, M. A. Bizouard, M. Branchesi, S. Clesse, S. Foffa, J. García-Bellido, S. Grimm, J. Harms, T. Hinderer, S. Matarrese, C. Palomba, M. Peloso, A. Ricciardone, and M. Sakellariadou, Science case for the Einstein telescope, J. Cosmol. Astropart. Phys. **2020**, 050 (2020).
[115] R. D. Gill and B. Y. Levit, Applications of the van Trees inequality: A Bayesian Cramér-Rao bound, Bernoulli **1**, 59 (1995).
[116] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part IV* (John Wiley & Sons, Incorporated, New York, 2002).
[117] A. Bednorz and W. Belzig, Fourth moments reveal the negativity of the Wigner function, Phys. Rev. A **83**, 052113 (2011).
[118] R. Howl, V. Vedral, D. Naik, M. Christodoulou, C. Rovelli, and A. Iyer, Non-Gaussianity as a signature of a quantum theory of gravity, PRX Quantum **2**, 010325 (2021).
[119] S. A. Haine, Searching for signatures of quantum gravity in quantum gases, New J. Phys. **23**, 033020 (2021).
[120] Z. Mehdi, J. J. Hope, and S. A. Haine, Signatures of quantum gravity in the gravitational self-interaction of photons, Phys. Rev. Lett. **130**, 240203 (2023).
[121] A. Franzen, Componentlibrary, 2009. http://www.gwoptics.org/ComponentLibrary/.
[122] Wolfram Research, Inc., *Mathematica* 8.0 (2010). https://www.wolfram.com.
[123] G. Van Rossum and F. L. Drake, PYTHON *3 Reference Manual* (CreateSpace, Scotts Valley, California, 2009).
[124] F. Pérez and B. E. Granger, IPython: A system for interactive scientific computing, Comput. Sci. Eng. **9**, 21 (2007).
[125] T. Kluyver, B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, C. Willing, and Jupyter Development Team, in *Positioning and Power in Academic Publishing: Players, Agents and Agendas* (IOS Press, 2016), pp. 87–90.
[126] C. R. Harris *et al.*, Array programming with NumPy, Nature **585**, 357 (2020).
[127] J. D. Hunter, MATPLOTLIB: A 2D graphics environment, Comput. Sci. Eng. **9**, 90 (2007).
[128] Wes McKinney, in *Proceedings of the 9th* PYTHON *in Science Conference* (SciPy, Austin, Texas, 2010), pp. 56–61.
[129] P. Virtanen *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in PYTHON, Nat. Methods **17**, 261 (2020).
[130] J. Johansson, P. Nation, and F. Nori, QuTiP 2: A PYTHON framework for the dynamics of open quantum systems, Comput. Phys. Commun. **184**, 1234 (2013).
[131] N. Killoran, J. Izaac, N. Quesada, V. Bergholm, M. Amy, and C. Weedbrook, STRAWBERRY FIELDS: A software platform for photonic quantum computing, Quantum **3**, 129 (2019).
[132] J. W. Gardner, *pleasantPheasant* (2023). https://git.ligo.org/jameswalter.gardner/pleasantpheasant.
[133] J. Williamson, On the algebraic problem concerning the normal forms of linear dynamical systems, Am. J. Math. **58**, 141 (1936).
[134] S. A. Haine, Using interaction-based readouts to approach the ultimate limit of detection-noise robustness for quantum-enhanced metrology in collective spin systems, Phys. Rev. A **98**, 030303 (2018).
[135] Y. L. Len, T. Gefen, A. Retzker, and J. Kołodyński, Quantum metrology with imperfect measurements, Nat. Commun. **13**, 6971 (2022).

[136] S. Zhou, S. Michalakis, and T. Gefen, Optimal protocols for quantum metrology with noisy measurements, PRX Quantum **4**, 040305 (2023).
[137] A. Uhlmann, The "transition probability" in the state space of a $*$-algebra, Rep. Math. Phys. **9**, 273 (1976).
[138] S. Boixo, S. T. Flammia, C. M. Caves, and J. M. Geremia, Generalized limits for single-parameter quantum estimation, Phys. Rev. Lett. **98**, 090401 (2007).
[139] A. Fujiwara and H. Imai, A fibre bundle over manifolds of quantum channels and its application to quantum statistics, J. Phys. A **41**, 255304 (2008).
[140] S. Kawamura and Y. Chen, Displacement-noise-free gravitational-wave detection, Phys. Rev. Lett. **93**, 211103 (2004).
[141] Y. Chen, A. Pai, K. Somiya, S. Kawamura, S. Sato, K. Kokeyama, R. L. Ward, K. Goda, and E. E. Mikhailov, Interferometers for displacement-noise-free gravitational-wave detection, Phys. Rev. Lett. **97**, 151103 (2006).
[142] T. Gefen, R. Tarafder, R. X. Adhikari, and Y. Chen, Quantum precision limits of displacement noise-free interferometers, Phys. Rev. Lett. **132**, 020801 (2024).
[143] I. Tzitrin, J. E. Bourassa, N. C. Menicucci, and K. K. Sabapathy, Progress towards practical qubit computation using approximate Gottesman-Kitaev-Preskill codes, Phys. Rev. A **101**, 032315 (2020).
[144] J. E. Bourassa, N. Quesada, I. Tzitrin, A. Száva, T. Isacsson, J. Izaac, K. K. Sabapathy, G. Dauphinais, and I. Dhand, Fast simulation of bosonic qubits via Gaussian functions in phase space, PRX Quantum **2**, 040315 (2021).
[145] K. Macieszczak, Quantum Fisher information: Variational principle and simple iterative algorithm for its efficient computation, arxiv:1312.1356 [quant-ph].
[146] Á. Lukács, R. Trényi, T. Vértesi, and G. Tóth, Optimizing local Hamiltonians for the best metrological performance, arXiv:2206.02820.
[147] J. Gorski, F. Pfeuffer, and K. Klamroth, Biconvex sets and optimization with biconvex functions: A survey and extensions, Math. Methods Oper. Res. **66**, 373 (2007).
[148] C. A. Floudas and V. Visweswaran, A global optimization algorithm (GOP) for certain classes of nonconvex NLPs—I. Theory, Comput. Chem. Eng. **14**, 1397 (1990).
[149] C. A. Floudas, *Deterministic Global Optimization: Theory, Methods and Applications* (Springer Science & Business Media, New York, 2013), Vol. 37.
[150] K. Noh, V. V. Albert, and L. Jiang, Quantum capacity bounds of Gaussian thermal loss channels and achievable rates with Gottesman-Kitaev-Preskill codes, IEEE Trans. Inf. Theory **65**, 2563 (2018).
[151] G. Tóth and T. Vértesi, Quantum states with a positive partial transpose are useful for metrology, Phys. Rev. Lett. **120**, 020506 (2018).
[152] K. Macieszczak, M. Fraas, and R. Demkowicz-Dobrzański, Bayesian quantum frequency estimation in presence of collective dephasing, New J. Phys. **16**, 113002 (2014).
[153] D. Rife and R. Boorstyn, Single tone parameter estimation from discrete-time observations, IEEE Trans. Inf. Theory **20**, 591 (1974).
[154] A. Steinhardt and C. Bretherton, in *ICASSP '85. IEEE International Conference on Acoustics, Speech, and Signal Processing* (Tampa, FL, USA, 1985), Vol. 10, pp. 1273–1276.
[155] L. Knockaert, The Barankin bound and threshold behavior in frequency estimation, IEEE Trans. Signal Process. **45**, 2398 (1997).
[156] E. Chaumette, Jé. Galy, A. Quinlan, and P. Larzabal, A new Barankin bound approximation for the prediction of the threshold region performance of maximum likelihood estimators, IEEE Trans. Signal Process. **56**, 5319 (2008).
[157] M. Gessner and A. Smerzi, Hierarchies of frequentist bounds for quantum metrology: From Cramér-Rao to Barankin, Phys. Rev. Lett. **130**, 260801 (2023).

Chapter 5

Lindblad estimation with fast and precise quantum control

Now I will believe that there are unicorns...

Sebastian (Act 3, Scene 3)

William Shakespeare, The Tempest

We now present Ref. [51] which is closely related to the previous paper but now answers what the addition of optimal fast and precise quantum control can do to aid sensing a stochastic waveform or, more generally, a Lindblad decay rate, still remaining in the asymptotic Fisher limit. This represents the culmination of the optimisation process showcased in the previous two papers, building up from just choosing the measurement, then adding different initial states, and finally now also changing the natural signal-less evolution of the sensing device. We show that a simple measure-and-reset strategy, as shown in Fig. 1b, is optimal provided that the right initial state is chosen (see Theorem 1). While this result may lead to dramatically increased sensitivity, e.g. the loss-limited sensitivity to stochastic waveforms is now unbounded as shown in Fig. 6, this proposal requires an extreme level of control over the experiment. This continues the trend from the previous two papers from practical near-future improvements to far-future radical changes in architecture to reach levels of precision never-before-seen. This is a tradeoff that is not just seen within this thesis but also broadly in the literature. This paper is the last in the stochastic sensing series as we will next explore a different dimension of the landscape of waveform estimation as shown in Fig. 1.1.

We now emphasise the main steps and results from this paper. Firstly, the existing

result in Eq. 6 is what allows us to study just the infinitesimal behaviour of the quantum channel in time (see Eq. 12) and prove the first part of Theorem 1, our main result, in Appendix A. Secondly, the proof of Theorem 1 also relies on the vanishing signal limit (see Eq. 9) and the order of these two limits matters as explained in Sec. II E. Thirdly, the properties of the optimal measure-and-reset strategy discussed in Sec. C are quite different to the optimal protocol for the deterministic case. Finally, while we focused above on the stochastic waveform estimation in Sec. VI, the Lindblad estimation results also apply to fermionic sensing tasks as explored in Sec. IV G and Sec. V.

Lindblad estimation with fast and precise quantum control

James W. Gardner [1,2,*] Simon A. Haine [3] Joseph J. Hope [3] Yanbei Chen [2] and Tuvia Gefen [4,5,†]

[1] *OzGrav-ANU, Centre for Gravitational Astrophysics, Research Schools of Physics, and of Astronomy and Astrophysics, The Australian National University, Canberra, ACT 2601, Australia*

[2] *Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, California 91125, USA*

[3] *Department of Quantum Science and Technology and Department of Fundamental and Theoretical Physics, Research School of Physics, The Australian National University, Canberra, ACT 0200, Australia*

[4] *Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*

[5] *Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem 91904, Givat Ram, Israel*

(Dated: August 1, 2025)

Enhancing precision sensors for stochastic signals using quantum techniques is a promising emerging field of physics. Estimating a weak stochastic waveform is the core task of many fundamental physics experiments including searches for stochastic gravitational waves, quantum gravity, and axionic dark matter. Simultaneously, noise spectroscopy and characterisation, e.g. estimation of various decay mechanisms in quantum devices, is relevant to a broad range of fundamental and technological applications. We consider the ultimate limit on the sensitivity of these devices for Lindblad estimation given any quantum state, fast and precise control sequence, and measurement scheme. We show that it is optimal to rapidly projectively measure and re-initialise the quantum state. We develop optimal protocols for a wide range of applications including stochastic waveform estimation, spectroscopy with qubits, and Lindblad estimation.

A wide variety of fundamental physics searches and emerging technologies involve sensing a weak stochastic signal: From stochastic waveform estimation [1] for searching for gravitational waves [2–4], quantum gravity [5–7], or axionic dark matter [8–11]; to spectroscopy [12, 13], nano-scale nuclear magnetic resonance [14, 15], and noise characterisation in quantum sensors and devices [16–20]. A common challenge across these different applications is how to differentiate between the weak stochastic signal and the noise in the system: The noise can completely mask the weak stochastic signal and significantly degrade the sensitivity. It is unclear whether this degraded sensitivity is fundamental to sensing a stochastic signal in the presence of noise, or whether it can be overcome with quantum-enhanced techniques. We address this problem by determining the optimal quantum metrological strategy to sense a weak stochastic signal in the presence of noise.

We now explain how we model this question as a Lindblad estimation problem. Under the Born-Markov approximation, the time evolution of the density matrix $\hat{\rho}$ describing our quantum sensor is given by the following Lindblad master equation [21–25]:

$$\dot{\hat{\rho}} = -i[\hat{H},\hat{\rho}] + \gamma_1 \sum_{j=1}^{K} \mathcal{L}_{j,(s)}(\hat{\rho}) + \sum_{k=K+1}^{K+N} \gamma_k \mathcal{L}_{k,(n)}(\hat{\rho}) \quad (1)$$

where $\hat{H}$ is the Hamiltonian and $\hbar = 1$. The signal that we want to estimate is the common decay rate γ_1 of K signal Lindblad jump operators $\{\hat{L}_{j,(s)}\}_{j=1}^{K}$ each of which determines a Lindbladian superoperator as follows:

$$\mathcal{L}_{j,(s)}(\hat{\rho}) = \hat{L}_{j,(s)}\hat{\rho}\hat{L}_{j,(s)}^{\dagger} - \frac{1}{2}\{\hat{L}_{j,(s)}^{\dagger}\hat{L}_{j,(s)},\hat{\rho}\}.$$

The signal γ_1, however, is much weaker than the noise described by the other N decay rates $\{\gamma_k\}_{k=K+1}^{K+N}$ and noise Lindblad jump operators $\{\hat{L}_{k,(n)}\}_{k=K+1}^{K+N}$ which similarly determine the noise Lindbladian superoperators $\{\mathcal{L}_{k,(n)}\}_{k=K+1}^{K+N}$ in Eq. 1. We want to determine how best to extract the information about the signal γ_1 encoded over time in the quantum state $\hat{\rho}$ by preparing different initial states, applying different control operations during the time evolution, and performing different measurements at the end.

We want to determine the fundamental precision limit of this Lindblad estimation problem. This implies that we need to study the sequential metrological strategy with fast and precise control as shown in Fig. 1a. Here, the control operations that we apply to the quantum state can be arbitrary and continuously performed. This fast and precise control limit is hierarchically the most powerful metrological strategy as it can simulate all other strategies including, e.g., the measure-and-reset strategy shown in Fig. 1b or the parallel strategy shown in Fig. 1c [26–28]. The fast and precise control limit has been studied extensively for Hamiltonian estimation [27, 29, 30], where the signal is encoded unitarily by $\hat{H}$ in the presence of noise jump operators $\{\hat{L}_{k,(n)}\}_{k=K+1}^{K+N}$ in Eq. 1 and the root mean square error in estimating the signal falls either as $1/t$ or $1/\sqrt{t}$ over time t, depending on the geometry of the signal and noise operators. The fast and precise control limit has also been studied previously for Lindblad estimation [31–34], where the signal is

* james.gardner@anu.edu.au
† tuvia.gefen@mail.huji.ac.il

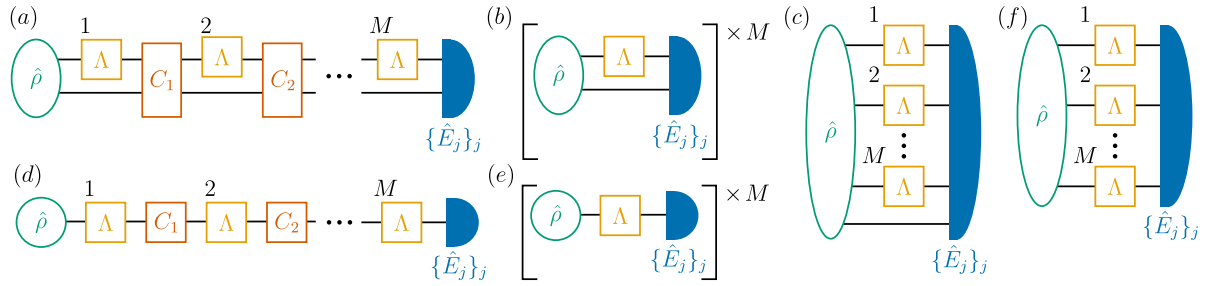

FIG. 1. Block diagrams of different quantum metrological strategies. **(a)** Sequential strategy with fast and precise control where the encoding channel Λ is queried M times on a short timescale t and between each query a control operation C is performed. The final state after a total time of $T = Mt$ is then measured. **(b)** Measure-and-reset strategy where the encoding channel is queried once, a measurement is performed, and then the state is re-initialised and the procedure is repeated a total of M times. **(c)** Parallel strategy where M devices are simultaneously queried and measured. **(d–f)** Unextended (i.e. ancilla-free) cases of the extended (i.e. noiseless ancilla–assisted) cases shown in panels (a–c), respectively.

encoded non-unitarily by signal jump operators and the noise corresponds to the nuisance Hamiltonian $\hat{H}$ and noise jump operators. In this case, the error in estimating the signal always falls as $1/\sqrt{t}$. But, the optimal initial state, control sequence, and measurement scheme for Lindblad estimation remain unknown.

In this paper, we consider the fast and precise control limit of Lindblad estimation. Firstly, in Sec. I, we review the theory of quantum estimation. Then, in Sec. II, we apply this formalism to find the optimal strategy for Lindblad estimation. Finally, we apply our results to Lindblad estimation of a single qubit in Sec. III, multiple qubits in Sec. IV, and a stochastic waveform in Sec. V.

I. REVIEW OF QUANTUM ESTIMATION

We now review the quantum theory of asymptotic unbiased estimation using the Fisher information formalism. For an introduction to this topic, see Refs. [35, 36].

1. *Parameter estimation*

Let us model an experiment as taking an initial quantum state described by a density matrix $\hat{\rho}$ and transforming it to some final state $\hat{\rho}'$ that encodes some real signal θ of interest but also introduces some noise. This transformation is described by a quantum channel Λ, i.e. a completely positive trace-preserving linear map between density matrices, such that $\hat{\rho}' = \Lambda(\hat{\rho})$. The quantum channel Λ has the following Kraus representation:

$$\hat{\rho}' = \sum_j \hat{K}_j \hat{\rho} \hat{K}_j^\dagger, \qquad \sum_j \hat{K}_j^\dagger \hat{K}_j = \hat{I}$$

where $\{\hat{K}_j\}_j$ are the non-unique Kraus operators of Λ and $\hat{I}$ is the identity operator. We want to estimate θ from measuring M independent and identical copies of the final state $\hat{\rho}'$. This measurement is described by a positive operator-valued measure (POVM) with a set of effects $\{\hat{E}_j\}_j$ such that the measurement outcome j occurs with probability $\mathrm{Tr}[\hat{\rho}'\hat{E}_j]$. (The effects satisfy $\sum_j \hat{E}_j = \hat{I}$ but are not necessarily a set of orthogonal projections.) We want to find the protocol that yields an unbiased estimate of θ with the minimum root mean square error (RMSE) $\Delta\theta$. This unextended measure-and-reset strategy is shown in Fig. 1e. We do not yet assume anything about the timescale of Λ.

Let the outcomes from a given measurement of $\hat{\rho}'$ be described by a random variable X with probability distribution $p(X = x|\theta)$. The classical Cramér-Rao bound on the RMSE is then

$$\Delta\theta \geq \frac{1}{\sqrt{M\mathcal{I}_C(\theta)}}$$

where the inequality is asymptotically tight as $M \to \infty$ using the maximum likelihood estimator, the factor of $1/\sqrt{M}$ comes from the Central Limit Theorem, and the classical Fisher information (CFI) about θ is defined as

$$\mathcal{I}_C(\theta) = \sum_j \frac{[\partial_\theta p(X = x_j|\theta)]^2}{p(X = x_j|\theta)} \tag{2}$$

where the sum excludes any j such that $p(X = x_j|\theta) = 0$.

The CFI about θ is defined for a given measurement, but there are many possible measurements of $\hat{\rho}'$ to perform. The quantum Fisher information (QFI) about θ, $\mathcal{I}_Q(\theta)$, is defined as the maximum possible value of the CFI about θ maximised over all possible measurements such that the quantum Cramér-Rao bound on the RMSE is

$$\Delta\theta \geq \frac{1}{\sqrt{M\mathcal{I}_Q(\theta)}}.$$

Here, the QFI is given by $\mathcal{I}_Q(\theta) = \mathrm{Tr}[\hat{\rho}'\hat{S}^2]$ where $\hat{S}$ is the symmetric logarithmic derivative of $\hat{\rho}'$ with respect to θ which satisfies $\partial_\theta\hat{\rho}' = \frac{1}{2}\{\hat{\rho}', \hat{S}\}$ and is given,

in terms of the spectral decomposition of the final state $\hat{\rho}' = \sum_j p_j |\phi_j\rangle\langle\phi_j|$, by

$$\hat{S} = \sum_{j,k} \frac{2}{p_j + p_k} \langle\phi_j|\partial_\theta \hat{\rho}'|\phi_k\rangle |\phi_j\rangle\langle\phi_k|$$

where the sum excludes any j, k such that $p_j + p_k = 0$. The QFI thus equals

$$\mathcal{I}_Q(\theta) = \sum_{j,k} \frac{2}{p_j + p_k} \left|\langle\phi_j|\partial_\theta \hat{\rho}'|\phi_k\rangle\right|^2. \tag{3}$$

An optimal measurement of $\hat{\rho}'$ is to project onto the eigenbasis of $\hat{S}$. Two useful properties are that the Fisher information from independent and identical repetitions sum and that the fractional RMSE is bounded by

$$\frac{\Delta\theta}{\theta} \geq \frac{1}{\sqrt{M\theta^2 \mathcal{I}_Q(\theta)}} \tag{4}$$

such that changing parameter from estimating θ to θ^2 only scales the fractional RMSE by a factor of two since $\theta^2 \mathcal{I}_Q(\theta) = 4\theta^4 \mathcal{I}_Q(\theta^2)$ by the chain rule.

We are often also interested in sensing a weak parameter θ, i.e. in studying the limit of vanishing signal: $\theta \to 0$. If $\lim_{\theta\to 0} \partial_\theta \hat{\rho}' = 0$ holds, then the optimal measurement in this limit is the projection $\hat{\Pi}_0$ onto the eigenbasis of the signal-free final state $\lim_{\theta\to 0} \hat{\rho}'$ such that [14, 37, 38]

$$\lim_{\theta\to 0} \partial_\theta \hat{\rho}' = 0 \implies \mathcal{I}_Q(\theta = 0) = \mathcal{I}_C^{\hat{\Pi}_0}(\theta = 0). \tag{5}$$

This simplifies calculating the QFI in the vanishing signal limit into finding the CFI from the results of measuring $\hat{\Pi}_0$.

2. *Channel estimation*

While finding the optimal measurement is straightforward, finding the optimal initial state $\hat{\rho}$ to maximise the QFI is more complicated. Due to the convexity of the space of density matrices and the QFI, the initial state can be assumed to be pure, i.e. $\hat{\rho} = |\psi\rangle\langle\psi|$ for some $|\psi\rangle$. Given an initial mixed state, there exists an initial pure state with greater than or equal QFI. This means that, without loss of generality, the state $\hat{\rho}$ shown in Fig. 1 is pure and the POVM is a projection-valued measure $\{\hat{\Pi}_j\}_j$. The initial pure state $|\psi\rangle$ may include entanglement of the system with a noiseless ancilla such that we can then perform joint measurements of the final state of the system and the unchanged ancilla. (This is represented by extending the Kraus operators of Λ to $\{\hat{K}_j \otimes \hat{I}\}_j$.) We will call states entangled with a noiseless ancilla "extended states" and ancilla-free states "unextended states". The extended measure-and-reset strategy is shown in Fig. 1b. In comparison to the performance of a given extended pure state, the sensitivity of the mixed state corresponding to its reduced density matrix will often be greatly diminished as the entanglement with the noiseless ancilla is crucial. Finding the optimal extended and unextended states for a particular channel Λ will be the main focus of our work. The optimal extended state always performs as well as the optimal unextended state and, as we will see, sometimes the optimal performance can only be attained using entanglement. Numerically, the optimal extended state and its accompanying QFI and measurement scheme can be found efficiently using a semi-definite program (SDP) [28, 39, 40].

3. *Fast and precise quantum control*

Finally, we can consider improving the QFI by using fast and precise quantum control. Given an infinitesimal channel Λ which is repeated M times with control as shown in Fig. 1a, the upper bound on the total QFI is [26, 27]

$$\mathcal{I}_Q(\theta) \leq 4M \|\hat{\alpha}\| + 4M(M-1)\|\hat{\beta}\|(\|\hat{\beta}\| + 2\sqrt{\|\hat{\alpha}\|}) \tag{6}$$

where $\hat{\alpha} = \dot{\mathbf{K}}^\dagger \dot{\mathbf{K}}$ and $\hat{\beta} = i\dot{\mathbf{K}}^\dagger \mathbf{K}$ given the vector $\mathbf{K} = (\hat{K}_1, \ldots, \hat{K}_r)^{\mathrm{T}}$ from a minimum rank r Kraus representation $\{\hat{K}_j\}_{j=1}^r$ of Λ. (Here, $\dot{\mathbf{K}}$ is the derivative of $\mathbf{K}$ with respect to the parameter of interest θ.) The Kraus representation of Λ chosen above is not unique, however, as any r-by-r unitary matrix u forms an equally valid Kraus representation $u\mathbf{K}$. All minimum rank Kraus representations of Λ can be reached this way. Under this gauge transformation, $\hat{\alpha}$ and $\hat{\beta}$ become

$$\hat{\alpha} = (\dot{\mathbf{K}} - ih\mathbf{K})^\dagger (\dot{\mathbf{K}} - ih\mathbf{K}), \quad \hat{\beta} = i(\dot{\mathbf{K}} - ih\mathbf{K})^\dagger \mathbf{K} \tag{7}$$

where h is any r-by-r Hermitian matrix. In the relevant case that $\hat{\beta} = 0$, then Eq. 6 minimised over all possible Kraus representations becomes

$$\mathcal{I}_Q(\theta) \leq 4M \min_h \left\| (\dot{\mathbf{K}} - ih\mathbf{K})^\dagger (\dot{\mathbf{K}} - ih\mathbf{K}) \right\|. \tag{8}$$

This upper bound is tight assuming access to a noiseless ancilla, i.e. the extended case shown in Fig. 1a rather than the unextended case shown in Fig. 1d. Since fast and precise control is hierarchically the most powerful metrological strategy, Eq. 8 provides the ultimate sensitivity limit assuming that $\hat{\beta} = 0$. Eq. 8 can also be efficiently calculated numerically via an SDP.

II. LINDBLAD ESTIMATION

We now return to our Lindblad estimation problem [41]. Before applying the above quantum estimation formalism to this problem, we need to explain some of the simplifying assumptions that we make. We assume that the signal in Eq. 1 is much weaker than

the noise such that the following condition holds for all $k = K+1, \ldots, K+N$

$$\gamma_1 \ll \frac{\|[\hat{H}, \hat{\rho}]\|}{\|\sum_{j=1}^{K} \mathcal{L}_{j,(s)}(\hat{\rho})\|}, \quad \gamma_1 \ll \frac{\gamma_k \|\mathcal{L}_{k,(n)}(\hat{\rho})\|}{\|\sum_{j=1}^{K} \mathcal{L}_{j,(s)}(\hat{\rho})\|} \tag{9}$$

where $\|\cdot\|$ is the operator norm [42]. Informally, this condition means that the signal decay rate γ_1 is much slower than the natural frequencies of the nuisance Hamiltonian $\hat{H}$ and all of the noise decay rates γ_k [43]. We choose to estimate $\sqrt{\gamma_1}$ instead of γ_1 without loss of generality to avoid a divergence in the QFI in this vanishing signal limit in the noiseless case [1]. The fractional RMSEs with respect to $\sqrt{\gamma_1}$ and γ_1 are the same up to a factor of two by Eq. 4 such that any sensitivity improvement that we demonstrate for $\sqrt{\gamma_1}$ is similar for γ_1. When estimating $\sqrt{\gamma_1}$, we assume that all of the jump operators $\{\hat{L}_j\}_{j=1}^{K+N}$ are known. We also assume for now that we know the nuisance Hamiltonian $\hat{H}$ and noise decay rates $\{\gamma_j\}_{j=K+1}^{K+N}$, but we will show later that we can remove this assumption. The jump operators $\hat{L}_j$ are traceless, i.e. $\mathrm{Tr}[\hat{L}_j] = 0$, without loss of generality since any component of $\hat{L}_j$ along $\hat{I}$ will not affect Eq. 1. We make no assumption about the dimension of the Hilbert space which may be infinite.

We want to find the optimal metrological strategy and ultimate precision limit of this Lindblad estimation problem. We thus study the sequential strategy with fast and precise control since it is hierarchically the most powerful metrological strategy. This means that we can perform arbitrary quantum control operations including the use of noiseless ancillae as shown in Fig. 1a. To understand the most powerful metrological strategies, the result in Eq. 6 implies that we must study the infinitesimal channel Λ. I.e., we need to find the quantum channel Λ that represents the time evolution under the master equation in Eq. 1 after an infinitesimal time t. Given the initial state $\hat{\rho}$, then the final state $\hat{\rho}' = \Lambda(\hat{\rho})$ is

$$\begin{aligned} \hat{\rho}' &= \hat{\rho} + t\left(-i[\hat{H}, \hat{\rho}] + \sum_{j=1}^{K+N} \gamma_j \mathcal{L}_j(\hat{\rho})\right) \\ &= \sum_{j=0}^{K+N} \hat{K}_j \hat{\rho} \hat{K}_j^\dagger \end{aligned} \tag{10}$$

where one possible Kraus representation of Λ is

$$\begin{aligned} \hat{K}_0 &= \hat{I} - t\left(i\hat{H} + \frac{1}{2}\sum_{k=1}^{K+N} \gamma_k \hat{L}_k^\dagger \hat{L}_k\right) \\ \hat{K}_j &= \sqrt{\gamma_j t}\hat{L}_j, \quad j = 1, \ldots, K+N. \end{aligned} \tag{11}$$

In Eq. 10 and henceforth, we drop all $\mathcal{O}(t^2)$ terms provided that the short evolution time t satisfies

$$t \ll \frac{4\|\sum_{j=1}^{K+N} \gamma_j \mathcal{L}_j(\hat{\rho}) - i[\hat{H}, \hat{\rho}]\|}{\|\sum_{j,k=1}^{K+N} \gamma_j \gamma_k \mathcal{L}_j(\mathcal{L}_k(\hat{\rho})) - [\hat{H}, [\hat{H}, \hat{\rho}]] + \hat{A}\|} \tag{12}$$

where the cross term is the following operator

$$\hat{A} = -i \sum_{j=1}^{K+N} \gamma_j \left(\mathcal{L}_j([\hat{H}, \hat{\rho}]) + [\hat{H}, \mathcal{L}_j(\hat{\rho})]\right).$$

Intuitively, the time t must be faster than the shortest timescale $\min[1/\gamma, 1/\omega]$ set by the largest decay rate γ or natural frequency ω of the nuisance Hamiltonian $\hat{H}$. For example, for sensing a bosonic loss given by the annihilation operator $\hat{L}_{1,(s)} = \hat{a}$ with a Fock state $|\hat{n} = n\rangle$ such that $\hat{n}|\hat{n} = n\rangle = n|\hat{n} = n\rangle$ where $\hat{n} = \hat{a}^\dagger \hat{a}$ and $n \geq 1$, then Eq. 12 implies that the time must satisfy $t \ll 4/[\gamma_1(2n-1)]$. We now want to find the optimal initial state, control sequence, and projective measurement for the sequential strategy with fast and precise control.

Let us briefly address the existing literature on Lindblad estimation. Estimating a weak decay rate while knowing the jump operators is different to the problem of estimating the whole Lindbladian with no prior information. In particular, the methods developed previously to estimate the whole Lindbladian [18, 44, 45] are not fine-tuned to isolate the noise-free component of a weak signal and will have vanishing QFI in this limit. While fast and precise control has been previously studied in Ref. [32], the particular assumption made in that work about the jump operators does not hold in general, unlike our results (as we discuss in Sec. II D 1). Similarly, the measure-and-reset limit for correlated noise sensing was examined in Refs. [46, 47] but without the proof of optimality that we now provide.

A. Fast and precise control

Let us first calculate the ultimate sensitivity limit given by Eq. 8 for sensing a stochastic signal or Lindblad estimation with fast and precise control as shown in Fig. 1a. In the vanishing signal limit of $\sqrt{\gamma_1} \to 0$, the vector of Kraus operators in Eq. 11 becomes

$$\mathbf{K} = \begin{pmatrix} \hat{I} - t\left(i\hat{H} + \frac{1}{2}\sum_{k=K+1}^{K+N} \gamma_k \hat{L}_{k,(n)}^\dagger \hat{L}_{k,(n)}\right) \\ 0 \\ \vdots \\ 0 \\ \sqrt{\gamma_{K+1} t}\hat{L}_{K+1,(n)} \\ \vdots \\ \sqrt{\gamma_{K+N} t}\hat{L}_{K+N,(n)} \end{pmatrix} \tag{13}$$

and its derivative with respect to $\sqrt{\gamma_1}$ becomes

$$\dot{\mathbf{K}} = \begin{pmatrix} 0 \\ \sqrt{t}\hat{L}_{1,(s)} \\ \vdots \\ \sqrt{t}\hat{L}_{K,(s)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

We choose the following ansatz for the gauge matrix:

$$h = \begin{pmatrix} 0 & \mathbf{g}^\dagger & \mathbf{0}_{1,N} \\ \mathbf{g} & \mathbf{0}_{K,K} & \mathbf{G} \\ \mathbf{0}_{N,1} & \mathbf{G}^\dagger & \mathbf{0}_{N,N} \end{pmatrix} \tag{14}$$

where $\mathbf{0}_{m,n}$ is the m-by-n zero matrix and $\vec{c}_j \in \mathbb{C}^{N+1}$ for $j = 1, \ldots, K$ such that h is determined by

$$\mathbf{g} = -i\sqrt{t}\begin{pmatrix} (\vec{c}_1)_1 \\ \vdots \\ (\vec{c}_K)_1 \end{pmatrix}, \quad \mathbf{G} = -i\begin{pmatrix} \frac{(\vec{c}_1)_2}{\sqrt{\gamma_{K+1}}} & \cdots & \frac{(\vec{c}_1)_{N+1}}{\sqrt{\gamma_{K+N}}} \\ \vdots & \ddots & \vdots \\ \frac{(\vec{c}_K)_2}{\sqrt{\gamma_{K+1}}} & \cdots & \frac{(\vec{c}_K)_{N+1}}{\sqrt{\gamma_{K+N}}} \end{pmatrix}$$

Using this gauge matrix h, the Kraus operators in Eq. 13 become

$$\dot{\mathbf{K}} - ih\mathbf{K} = \begin{pmatrix} 0 \\ \hat{L}_{1,(s)} - \vec{c}_1^{\mathrm{T}}\mathbf{L} \\ \vdots \\ \hat{L}_{K,(s)} - \vec{c}_K^{\mathrm{T}}\mathbf{L} \\ 0 \\ \vdots \\ 0 \end{pmatrix}\sqrt{t} \tag{15}$$

where $\mathbf{L} = (\hat{I}, \hat{L}_{K+1,(n)}, \ldots, \hat{L}_{K+N,(n)})^{\mathrm{T}}$. This implies that $\hat{\beta} = 0$ in Eq. 7. The upper bound on the QFI with fast and precise control in Eq. 8 is thus:

$$\mathcal{I}_Q(\sqrt{\gamma_1} = 0) \leq 4T \min_{\{\vec{c}_k\}_{k=1}^K} \left\| \sum_{k=1}^K \hat{A}_k^\dagger \hat{A}_k \right\| \tag{16}$$

where $\hat{A}_k = \hat{L}_{k,(s)} - \vec{c}_k^{\mathrm{T}}\mathbf{L}$ and $T = Mt$ is the total time. This upper bound is independent of the nuisance Hamiltonian $\hat{H}$ which agrees with Appendix G of Ref. [34] and will be explained below once we find the optimal strategy.

If $\hat{L}_{k,(s)} \in \mathrm{span}\{\mathbf{L}\}$ for all $k = 1, \ldots, K$, i.e. all of the signal operators lie in the subspace spanned by the noise operators and identity, then the upper bound in Eq. 16 vanishes. Similarly, any component of a signal operator $\hat{L}_{k,(s)}$ within $\mathrm{span}\{\mathbf{L}\}$ does not affect the upper bound, since it can be compensated by appropriately choosing $\vec{c}_k$, and thus may be ignored without loss of generality. We discuss the orthogonality of the signal and noise operators further in Sec. II C.

In the case of a single signal ($K = 1$), Eq. 16 has the following geometric interpretation

$$\begin{aligned} \mathcal{I}_Q(\sqrt{\gamma_1} = 0) &\leq 4T \min_{\vec{c}} \|\hat{L}_{1,(s)} - \vec{c}^{\mathrm{T}}\mathbf{L}\|^2 \\ &= 4T\, d(\hat{L}_{1,(s)}, \mathrm{span}\{\mathbf{L}\})^2 \end{aligned} \tag{17}$$

where we used that $\|X^\dagger X\| = \|X\|^2$ for the operator norm and defined $d(\hat{L}_{1,(s)}, \mathrm{span}\{\mathbf{L}\})$ as the distance induced by the operator norm between the signal operator $\hat{L}_{1,(s)}$ and the subspace of noise operators $\mathrm{span}\{\mathbf{L}\}$. This distance vanishes if and only if $\hat{L}_{1,(s)} \in \mathrm{span}\{\mathbf{L}\}$ such that the QFI is nonzero only if $\hat{L}_{1,(s)} \notin \mathrm{span}\{\mathbf{L}\}$, i.e. some component of the signal operator lies outside of the subspace of noise operators.

We prove that the upper bound in Eq. 16 is tight and determine the optimal sensing strategy:

Theorem 1. Consider sensing the decay rate γ_1 common to K signal operators $\{\hat{L}_{k,(s)}\}_{k=1}^K$ in the presence of N noise operators $\{\hat{L}_{j,(n)}\}_{j=K+1}^{K+N}$ and a nuisance Hamiltonian $\hat{H}$. In the vanishing signal limit of $\gamma_1 \to 0$, the optimal QFI with respect to $\sqrt{\gamma_1}$ is:

$$\mathcal{I}_Q(\sqrt{\gamma_1} = 0) = 4T \max_{|\psi\rangle} \sum_{k=1}^K \langle\psi|\hat{L}_{k,(s)}^\dagger(\hat{I} - \hat{\Pi})\hat{L}_{k,(s)}|\psi\rangle$$

where $\hat{\Pi}$ is the projection onto the noisy subspace $\mathrm{span}\{\mathbf{L}|\psi\rangle\}$ of the images of the state $|\psi\rangle$ under the noise operators $\mathbf{L} = (\hat{I}, \hat{L}_{K+1,(n)}, \ldots, \hat{L}_{K+N,(n)})^{\mathrm{T}}$. This QFI is achieved by a measure-and-reset strategy that continuously prepares the optimal state $|\psi\rangle$ above and projects onto the noisy subspace with $\hat{\Pi}$.

This is the first main result of our work. In Appendix A 1, we show that the expression in Theorem 1 provides an upper bound on the optimal QFI by evaluating the minimisation over $\{\vec{c}_k\}_{k=1}^K$ in Eq. 16. We now show that this upper bound is tight and can be saturated by a measure-and-reset strategy.

B. Measure-and-reset strategy

One possible control operation is to rapidly measure projectively and then re-initialise the state. If we repeatedly perform this control operation, then we reach the measure-and-reset strategy shown in Fig. 1b. (The same is true in the ancilla-free case as shown in Figs. 1d–e.) The measure-and-reset strategy is thus similar to a quantum jump experiment [48].

In general, the optimal measure-and-reset strategy is outperformed by the optimal fast and precise control strategy [26–28]. For example, rapidly projectively measuring a state undergoing Hamiltonian evolution leads to the quantum Zeno effect [49] unlike the optimal fast and precise control strategy which leads to a coherent build-up of signal [30]. For sensing a stochastic signal, however,

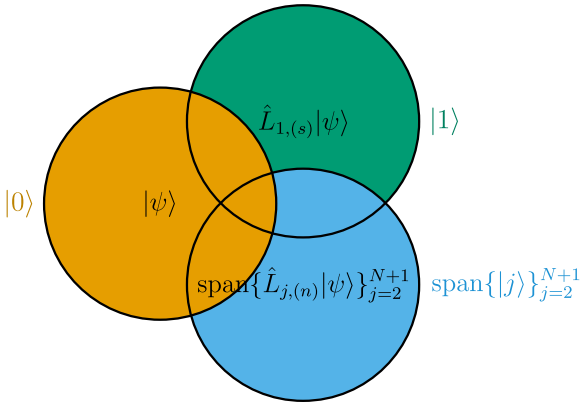

FIG. 2. Visualisation of the Gram-Schmidt process for orthogonalising the images of the initial state $|\psi\rangle$ under N noise operators $\text{span}\{\hat{L}_{j,(n)}|\psi\rangle\}_{j=2}^{N+1}$ and one signal operator $\hat{L}_{1,(s)}|\psi\rangle$. The component of the weak signal within the noisy subspace $\text{span}\{|0\rangle, |j\rangle\}_{j=2}^{N+1}$ is lost leaving only the noise-free component $|1\rangle$.

we will show that the optimal measure-and-reset strategy attains the ultimate sensitivity limit with fast and precise control. Intuitively, this is because the stochastic signal does not build up coherently over time. This also explains why the nuisance Hamiltonian $\hat{H}$ does not affect the optimal measure-and-reset strategy: its effect on the dynamics is frozen out due to the Zeno effect assuming the Eq. 12 holds. This also means that we can drop the assumption that we know what the nuisance Hamiltonian is and can treat it as being zero $\hat{H} = 0$ without loss of generality. We show this formally in Appendix A 2.

Let us find the optimal measure-and-reset strategy and show that it saturates the ultimate sensitivity limit in Theorem 1. We assume that we can rapidly projectively measure $\hat{\rho}'$ in Eq. 10 and then re-initialise the state to $\hat{\rho} = |\psi\rangle\langle\psi|$ if we measure that a jump occurred. We repeat this process $M \gg 1$ times over a long total time $T = Mt$. The total QFI is thus M times the QFI from a single measurement after a time t. Although the probability of a signal jump in a single measurement is proportional to $\gamma_1 t$ and therefore infinitesimal, the probability of a signal jump over the total time is proportional to $\gamma_1 T$ which is finite if we wait for a long enough time T. We now need to find the optimal initial pure state $|\psi\rangle$ and projective measurement to estimate γ_1. By Eq. 10, $\partial_{\sqrt{\gamma_1}}\hat{\rho}' = 0$ holds in the vanishing signal limit of $\sqrt{\gamma_1} \to 0$ such that the optimal measurement is to project onto the eigenbasis of the signal-free final state $\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'$ by Eq. 5. To calculate this eigenbasis, let us first find an orthonormal basis to express the final state.

1. *Gram-Schmidt process*

We perform the Gram-Schmidt process to the images of the initial state $|\psi\rangle$ under the noise and signal jump operators in the following order:

$$|\psi\rangle, \hat{L}_{K+1,(n)}|\psi\rangle, \hat{L}_{K+2,(n)}|\psi\rangle, \ldots, \hat{L}_{K+N,(n)}|\psi\rangle, \quad (18)$$
$$\hat{L}_{1,(s)}|\psi\rangle, \hat{L}_{2,(s)}|\psi\rangle, \ldots, \hat{L}_{K,(s)}|\psi\rangle.$$

That is, we orthogonalise first the initial state, then the images under the noise operators, and finally the images under the signal operators. (We discuss below what happens should the Gram-Schmidt process stop at any point because the images of the intial state have been overfitted, i.e. if they are linearly dependent.) The orthogonalisation is shown for a single signal operator in Fig. 2. This process defines the orthonormal basis $\{|j\rangle\}_{j=0}^{K+N}$ where $|0\rangle = |\psi\rangle$ such that

$$\hat{L}_{K+j,(n)}|0\rangle = c_{K+j,0}|0\rangle + \sum_{k=1}^{j} c_{K+j,K+k}|K+k\rangle$$
$$\hat{L}_{j,(s)}|0\rangle = \sum_{k=1}^{N} c_{j,K+k}|K+k\rangle + \sum_{k=0}^{j} c_{j,k}|k\rangle.$$

We have thus split the space into the noisy subspace $\text{span}\{|0\rangle, |j\rangle\}_{j=K+1}^{K+N} = \text{span}\{\mathbf{L}|\psi\rangle\}$ and the noise-free signal subspace $\text{span}\{|j\rangle\}_{j=1}^{K}$ which can only be reached if a signal jump occurs. Let $\mathbf{v} = (|0\rangle, |K+1\rangle, \ldots, |K+N\rangle)^{\mathrm{T}}$ such that $\mathbf{L}|\psi\rangle = \mathbf{M}\mathbf{v}$ where

$$\mathbf{M} = \begin{bmatrix} \langle 0|\mathbf{L}|0\rangle & \langle K+1|\mathbf{L}|0\rangle & \cdots & \langle K+N|\mathbf{L}|0\rangle \end{bmatrix}.$$

The projection onto the noisy subspace is thus

$$\begin{aligned}\hat{\Pi} &= |0\rangle\langle 0| + \sum_{j=K+1}^{K+N} |j\rangle\langle j| \\ &= \mathbf{v}^{\mathrm{T}}\mathbf{v}^{*} \\ &= \mathbf{L}^{\mathrm{T}}|\psi\rangle\langle\psi|\mathbf{L}^{*}\mathbf{L}^{\mathrm{T}}|\psi\rangle^{-1}\langle\psi|\mathbf{L}^{*} \end{aligned} \quad (19)$$

where we used that $\mathbf{v} = \mathbf{M}^{-1}\mathbf{L}|\psi\rangle$ and $\hat{\Pi}\mathbf{L}|\psi\rangle = \mathbf{L}|\psi\rangle$. Analytically, we thus need to invert the $N+1$ by $N+1$ matrix $\langle\psi|\mathbf{L}^{*}\mathbf{L}^{\mathrm{T}}|\psi\rangle$ to find $\hat{\Pi}$, whereas numerically the Gram-Schmidt process can be efficiently performed to find $\hat{\Pi}$.

Here, we have assumed that the images under the noises in Eq. 18 are linearly independent such that the noisy subspace has dimension $N+1$. If the noisy subspace instead has dimension $N'+1 < N+1$, then we can calculate $\hat{\Pi} = \mathbf{v}^{\mathrm{T}}\mathbf{v}^{*}$ from $\mathbf{L}|\psi\rangle = \mathbf{M}\mathbf{v}$ and the Moore-Penrose pseudoinverse of the $N+1$ by $N'+1$ matrix $\mathbf{M}$. In the case that the noise-free signal subspace $\text{span}\{|j\rangle\}_{j=1}^{K}$ is also overfitted and instead has dimension $K' < K$, then we can similarly calculate the orthonormal basis from the pseudoinverse of the $K+N+1$ by $K'+N'+1$ matrix $\mathbf{M}'$ formed by appending the columns $\langle j|\mathbf{L}'|0\rangle$ for $j = 1, \ldots, K$ to the matrix $\mathbf{M}$ above where

$$\mathbf{L}' = (\hat{I}, \hat{L}_{K+1,(n)}, \ldots, \hat{L}_{K+N,(n)}, \hat{L}_{1,(s)}, \ldots, \hat{L}_{K,(s)})^{\mathrm{T}}.$$

Here, the order of operators is the same as in Eq. 18. In both cases of overfitting the signal and noise, we could

alternatively allow the degenerate elements of $\{|j\rangle\}_{j=0}^{K+N}$ to be zero. Finding an orthogonal basis for the noise-free signal subspace, however, is not necessary to calculate the upper bound in Theorem 1 or the QFI of the measure-and-reset strategy here.

We calculate the QFI of the measure-and-reset strategy with a given state $|\psi\rangle$ using Eq. 5. We determine the eigenbasis of the signal-free final state $\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'$ from Eq. 10 in the orthonormal basis $\{|j\rangle\}_{j=0}^{K+N}$. We show that the QFI given a state $|\psi\rangle$ equals

$$\mathcal{I}_Q(\sqrt{\gamma_1}=0) = 4T\sum_{k=1}^{K}\langle\psi|\hat{L}_{k,(s)}^\dagger(\hat{I}-\hat{\Pi})\hat{L}_{k,(s)}|\psi\rangle. \quad (20)$$

This QFI for the optimal initial state $|\psi\rangle$ thus saturates the ultimate sensitivity limit in Theorem 1. We prove these results in Appendix A 2. We show that the eigenbasis is not exactly $\{|j\rangle\}_{j=0}^{K+N}$ but that the CFI from projecting onto $\{|j\rangle\}_{j=0}^{K+N}$ equals the QFI to leading order in time t. Moreover, it suffices to project onto the noisy subspace with $\hat{\Pi}$ since any nonzero noise is louder than the vanishing signal such that the component of the signal within the noisy subspace is completely lost. It remains to find the optimal initial state $|\psi\rangle$ for a given set of jump operators. We derive the conditions on a given state being optimal in Appendix B. Although we have found the optimal strategy overall, we have not identified the optimal strategy when the initial state, control sequence, or final measurement is non-optimal and fixed.

C. Properties of the optimal strategy

Let us now discuss some properties of the optimal measure-and-reset strategy which saturates the ultimate sensitivity limit in Theorem 1.

1. *Geometric condition on the QFI vanishing*

The necessary and sufficient condition for the QFI to be nonzero in the limit of vanishing signal is the following:

$$\exists j\in\{1,\dots,K\},\quad \hat{L}_{j,(s)}\notin \text{span}\{\mathbf{L}\}. \quad (21)$$

Let us discuss in turn why this condition is necessary and sufficient.

Eq. 21 is a necessary condition since any component of a signal operator parallel to a noise operator or the identity does not contribute to the QFI in Theorem 1. If all of the signal operators lie in span$\{\mathbf{L}\}$, then the QFI is thus zero in the limit of vanishing signal. We previously observed that this is a necessary condition in Eq. 16 and, for the case of a single signal, in Eq. 17.

Eq. 21 is a sufficient condition because if $\hat{L}_{j,(s)}\notin$ span$\{\mathbf{L}\}$ then $\hat{Y}=\hat{L}_{j,(s)}^\dagger(\hat{I}-\hat{\Pi})\hat{L}_{j,(s)}$ is nonzero and there exists a state $|\psi\rangle$ such that $\langle\psi|\hat{Y}|\psi\rangle>0$, e.g. an eigenstate of the positive semidefinite operator $\hat{Y}$ with positive eigenvalue. Theorem 1 is then nonzero because the other terms in the sum over $k=1,\dots,K$ for $k\neq j$ cannot cancel with $\langle\psi|\hat{Y}|\psi\rangle$ since they are nonnegative. We can also show directly that the QFI does not vanish for a particular extended state. Consider the Bell state defined as $|\psi\rangle=\frac{1}{\sqrt{d}}\sum_{j=1}^{d}|e_j\rangle\otimes|e_j\rangle$ for a given basis $\{|e_j\rangle\}_{j=1}^{d}$. (Here, we consider a Hilbert space with finite dimension d.) The Bell state has the property that $\langle\psi|\hat{A}\otimes\hat{I}|\psi\rangle=\frac{1}{d}\text{Tr}[\hat{A}]$ for any $\hat{A}$ such that the QFI in Theorem 1 becomes

$$\mathcal{I}_Q(\sqrt{\gamma_1}=0) = 4T\frac{1}{d}\sum_{j=1}^{K}\text{Tr}[\hat{L}_{j,(s)}^\dagger\hat{L}_{j,(s)}] \quad (22)$$

where we replaced each signal operator $\hat{L}_{j,(s)}$ with its component orthogonal to span$\{\mathbf{L}\}$. If Eq. 21 holds, then at least one of these components is nonzero and the QFI in Eq. 22 is nonzero. The Bell state is not necessarily the optimal initial state, however, as we will see later.

We may assume henceforth without loss of generality that each signal operator is orthogonal to each noise operator, and that the noise operators are pairwise orthogonal. But, the signal operators need not be pairwise orthogonal. In particular, although a nonzero component of a signal operator parallel to a noise operator can affect the dynamics in Eq. 1, we have shown that the optimal strategy and QFI are unaffected such that the component can be ignored.

This phenomenon in noise sensing where the QFI vanishes as the parameter does is called the "Rayleigh curse" by analogy to quantum super-resolution [50]. The noise-induced Rayleigh curse was recently studied in several works, e.g. see Refs. [1, 14, 51, 52]. Here, we present the general criteria for avoiding it.

2. *Comparison to the deterministic case*

The condition in Eq. 21 states when it is possible to overcome the noise limitations in Lindblad estimation. Similar noise limitations were studied extensively in the deterministic (unitary) case. In the deterministic case, we want to sense some parameter θ from measurements of $\hat{\rho}'=\hat{U}\hat{\rho}\hat{U}^\dagger$ where $\hat{U}=\exp(-i\theta\hat{H}T)$ for some Hermitian $\hat{H}$. The QFI in this case is $\mathcal{I}_Q(\theta)=4T^2\text{Var}_{|\psi\rangle}[\hat{H}]$. Here, the QFI is independent of the parameter θ and coherently increases as T^2 compared to the stochastic case which only scales as T. These rates are sometimes called "Heisenberg scaling with time" for T^2 and "Standard Quantum Limit scaling with time" for T, respectively. The fact that Lindblad estimation with fast and precise control cannot achieve Heisenberg scaling with time is well-known [31–34]. For the deterministic case, the optimal strategy with fast and precise control is not to measure-and-reset since we want to coherently build up the signal as T^2, instead the optimal strategy is to perform quantum error correction. This strategy achieves

Heisenberg scaling with time provided that the following necessary and sufficient condition holds: [30]

$$\hat{H} \notin \text{span}\{\hat{I}, \hat{L}_{j,(n)}, \hat{L}^\dagger_{j,(n)}, \hat{L}^\dagger_{j,(n)}\hat{L}_{k,(n)}\}^{K+N}_{j,k=K}.$$

In comparison, in the stochastic case, the condition in Eq. 21 is on the QFI not vanishing and can be met even if, e.g., $\hat{L}_{1,(s)} = \hat{L}^\dagger_{2,(n)}$ or $\hat{L}_{1,(s)} = \hat{L}^\dagger_{2,(n)}\hat{L}_{2,(n)}$. Thus, the condition in the deterministic case to achieve T^2 scaling is not equivalent to the condition in the stochastic case to achieve nonzero QFI scaling as T.

3. *Independence of the QFI from the size of the noise*

We also observe in the vanishing signal limit that the measurement probabilities in Eq. A4 and thus the QFI of the optimal strategy do not depend on the noise decay rates $\{\gamma_j\}^{K+N}_{j=K+1}$ or the size of the noise jump operators $\{\|\hat{L}_{j,(n)}\|\}^{K+N}_{j=K+1}$ provided that they are nonzero. Instead, they only depend on the geometry of the noise operators which determines the noisy subspace projection $\hat{\Pi}$ since any nonzero noise is louder than the vanishing signal. We can thus relax our assumption that we know the noise decay rates and can assume that the noise jump operators have unit norm without loss of generality. For finite signals, however, the QFI will depend on and require knowing the noise decay rates and jump operator norms [1].

4. *It suffices to study the reduced density matrix*

Finally, we see that the QFI also only depends on the reduced density matrix of the initial state, $\hat{\rho}_S = \text{Tr}_A[|\psi\rangle\langle\psi|]$, after tracing out any ancilla. This is because the QFI in Theorem 1 depends only on the expectation values of operators that act on the system. This means that the maximisation can be performed over the reduced density matrix $\hat{\rho}_S$ instead of the full extended state $|\psi\rangle$ [53]. The QFI here is different than if we prepare the system in the mixed state $\hat{\rho}_S$, which is suboptimal, because to derive Theorem 1 we assumed that we can initialise and measure the ancilla too.

D. Examples of the optimal QFI

Let us now study the QFI of the optimal strategy from Theorem 1 for different numbers of noise operators.

1. *Noiseless case*

Consider the noiseless case with K signal jump operators, where the optimal QFI is

$$\mathcal{I}_Q(\sqrt{\gamma_1} = 0) = 4T \max_{|\psi\rangle} \sum_{j=1}^{K} \text{Var}_{|\psi\rangle}[\hat{L}_{j,(s)}]. \tag{23}$$

Here, the variance of an operator $\hat{A}$ is defined as $\text{Var}_{|\psi\rangle}[\hat{A}] = \text{Cov}_{|\psi\rangle}[\hat{A}, \hat{A}]$ where the covariance of two operators $\hat{A}$ and $\hat{B}$ is defined as

$$\text{Cov}_{|\psi\rangle}[\hat{A}, \hat{B}] = \langle\psi|\hat{A}^\dagger\hat{B}|\psi\rangle - \langle\psi|\hat{A}^\dagger|\psi\rangle\langle\psi|\hat{B}|\psi\rangle.$$

The following measure-and-reset strategy achieves the noiseless QFI in Eq. 23: Prepare the optimal initial state $|\psi\rangle$ that maximises the sum of variances above and project onto it since $\hat{\Pi} = |\psi\rangle\langle\psi|$.

Let us consider the noiseless case with a single signal jump operator $\hat{L}_{1,(s)}$. We need to find the state $|\psi\rangle$ that maximises the variance of $\hat{L}_{1,(s)}$ since the QFI is

$$\mathcal{I}_Q(\sqrt{\gamma_1} = 0) = 4T \max_{|\psi\rangle} \text{Var}_{|\psi\rangle}[\hat{L}_{1,(s)}]. \tag{24}$$

If $\hat{L}_{1,(s)}$ is Hermitian, then this process corresponds to a random unitary channel with $\hat{L}_{1,(s)}$ as the generator [1]. The optimal states are then equal superpositions of the eigenvectors of $\hat{L}_{1,(s)}$ with the maximum λ_+ and minimum λ_- eigenvalues, i.e. $|\psi\rangle = \frac{1}{\sqrt{2}}(|\lambda_+\rangle + e^{i\phi}|\lambda_-\rangle)$ for any ϕ, such that the QFI is $\mathcal{I}_Q(\sqrt{\gamma_1} = 0) = T(\lambda_+ - \lambda_-)^2$. Since this is the optimal state, there is no benefit from entangling with an ancilla and thus Fig. 1b and Fig. 1e are equivalent in this case.

If instead $\hat{L}_{1,(s)}$ is non-Hermitian, then finding the optimal state is more difficult. The maximal variance and optimal state can be obtained numerically by using an SDP to compute $\min_{c\in\mathbb{C}}\|\hat{A}^\dagger\hat{A}\|$ where $\hat{A} = \hat{L}_{1,(s)} - c\hat{I}$. The optimal state $|\psi\rangle$ lies in the maximal eigenspace of $\hat{A}^\dagger\hat{A}$ and satisfies $\langle\psi|\hat{A}|\psi\rangle = 0$.

The optimality and QFI of a continuous measure-and-reset strategy in the noiseless case have previously been shown in Ref. [32]. The authors therein also show that if the signal and noise jump operators comprise unique jumps between orthogonal subspaces, then the noise can be ignored. In practice, however, the signal and noise do not satisfy this condition in most applications, including the ones that we will examine later, and the noise leads to a penalty in sensitivity. This noise penalty is crucial to understanding the realistic performance of our measurement devices. We determine this penalty below in the case of a single source of noise and more generally in Theorem 1.

2. *One source of noise*

Let us now consider the case where there is a single noise jump operator $\hat{L}_{K+1,(n)}$ such that the QFI in Eq. 20

becomes

$$\mathcal{I}_Q(\sqrt{\gamma_1}=0) \tag{25}$$
$$= 4T\sum_{j=1}^{K}\left(\mathrm{Var}_{|\psi\rangle}[\hat{L}_{j,(s)}] - \frac{\left|\mathrm{Cov}_{|\psi\rangle}[\hat{L}_{j,(s)},\hat{L}_{K+1,(n)}]\right|^2}{\mathrm{Var}_{|\psi\rangle}[\hat{L}_{K+1,(n)}]}\right)$$

where the optimal QFI in Theorem 1 is obtained by maximising this expression over all possible $|\psi\rangle$. Compared to the noiseless QFI in Eq. 23, the noisy QFI loses the component of each signal within the noisy subspace resulting in a penalty related to the covariance between the signal and noise operators. This lost signal component within the noisy subspace is shown as the intersection of the two corresponding circles in Fig. 2 in the single signal case. This leads to a tradeoff when optimising Eq. 25 against $|\psi\rangle$ between maximising the sum of the variances and minimising the covariance penalty term. We will study some examples later where the noiseless QFI either can or cannot be recovered due to this tradeoff.

3. *More sources of noise*

In general, the expression for the QFI in terms of only the optimal initial state $|\psi\rangle$ and jump operators is given by expanding $\hat{\Pi}$ from Eq. 19 in Eq. 20. This results in an expression that depends only on the variances and covariances of the jump operators with respect to $|\psi\rangle$. For example, for two noise operators $N=2$, we provide the QFI explicitly in Appendix C. For additional noise operators $N>2$, inverting the $N+1$ by $N+1$ matrix $\langle\psi|\mathbf{L}^*\mathbf{L}^{\mathrm{T}}|\psi\rangle$ in Eq. 19 analytically is more complicated but can be done numerically instead.

E. Impact of finite signal and time

To derive Theorem 1, we first assume in Eq. 12 that the interrogation time is infinitesimally short to reach the sequential limit with fast and precise control in Eq. 6 and then assume in Eq. 9 that the signal is vanishingly weak to derive the Kraus operators in Eq. 13. We now address what happens to the QFI when the signal and time are instead finite with a couple of examples.

For example, consider sensing decay $\hat{L}_{1,(s)}=\hat{\sigma}_-$ of a qubit prepared in $|\uparrow\rangle$ such that $\hat{\sigma}_-|\uparrow\rangle=|\downarrow\rangle$. Let us examine a delayed measure-and-reset strategy that waits for a time t before measuring the state given by the exact solution to the master equation in Eq. 1. The total QFI from $M=T/t$ repetitions of this strategy for a total time T is

$$\mathcal{I}_Q(\sqrt{\gamma_1}) = \frac{4\gamma_1 tT}{e^{\gamma_1 t}-1}$$

which depends on both the finite signal γ_1 and time t. For a given finite signal γ_1, it is optimal to measure and reset as fast as possible since the QFI is maximised with value $4T$ in the short time limit of $t\ll 2/\gamma_1$. This agrees with Eq. 24 and is independent of the finite signal γ_1. We study the noiseless case with a finite signal further in Appendix D.

The fact that we first take the short time limit and then the vanishing signal limit matters. For example, consider sensing double decay $\hat{L}_{1,(s)}=\hat{J}_-^2$ of a three-level system prepared in $|\hat{J}_z=1\rangle$ in the presence of single decay noise $\hat{L}_{2,(n)}=\hat{J}_-$ such that $\hat{J}_-|\hat{J}_z=1\rangle=\sqrt{2}|\hat{J}_z=0\rangle$ and $\hat{J}_-^2|\hat{J}_z=1\rangle=2|\hat{J}_z=-1\rangle$. The QFI of a single measurement in the short time limit is

$$\mathcal{I}_Q(\sqrt{\gamma_1}) = 16t - 8t^2\left(4\gamma_1+2\gamma_2+\frac{\gamma_2^2}{\gamma_1}\right)+\mathcal{O}(t^3).$$

If we first take the short time limit, i.e. assume that $t\ll 2\gamma_1/\gamma_2^2$, then this QFI equals $16t$ independent of γ_1. This agrees with Eq. 25. In comparison, if we instead first take the vanishing signal limit, i.e. assume that $t\gg 2\gamma_1/\gamma_2^2$, then the QFI calculated from Eq. 3 vanishes. We study the first case where $t\ll 2\gamma_1/\gamma_2^2$ because it corresponds to finding the optimal metrological strategy with fast and precise control.

F. Summary of applications

In the following Secs. III–V, we analyse a variety of applications of the general results above. We now provide a brief summary of the main results with experimental impacts to guide the reader:

1. For qubit noise sensing, overcoming loss (i.e. amplitude-damping noise) in dephasing estimation with a single qubit requires entanglement either between a probe and an ancilla (see Sec. III D) or between multiple probes (see Sec. III G).

2. We characterise the cases where entanglement with an ancilla provides a metrological gain in single-qubit Lindblad estimation (see Sec. III F and Figs. 3–4).

3. For sensing the correlated dephasing and decay of many qubits, we determine the optimal initial state (see Sec. IV A). We also determine methods to entirely overcome Pauli noise and recover the noiseless QFI for many-qubit Lindblad estimation (see Claim 1 and Examples 1–2).

4. It is possible to arbitrarily exceed the loss-limited sensitivity of stochastic waveform estimation without control, which is relevant to searches for stochastic gravitational waves and ultra-light dark matter (see Claim 2 and Fig. 5).

III. A SINGLE QUBIT

Let us now consider a single qubit with one signal $\hat{L}_{1,(s)}$ and one noise $\hat{L}_{2,(n)}$. We want to find the optimal initial state and QFI for all possible choices of $\hat{L}_{1,(s)}$ and $\hat{L}_{2,(n)}$ to minimise the loss of precision compared to the noiseless QFI due to the noise. Furthermore, we want to determine the hierarchy of metrological strategies shown in Fig. 1: Whether an unextended measure-and-reset strategy (Fig. 1e) can saturate the optimal QFI or whether entanglement with other noisy probes (Fig. 1f) or a noiseless ancilla (Fig. 1b) is necessary. We will answer these questions by directly maximising the QFI in Eq. 25 over the initial state. As shown in Sec. II C, we can assume without loss of generality that $\hat{L}_{1,(s)}$ and $\hat{L}_{2,(n)}$ are orthogonal, traceless, and normalised. This leads to the following result:

Lemma 1. For a single qubit, the only unextended states with nonzero QFI are the eigenstate/s of $\hat{L}_{2,(n)}$. The optimal unextended QFI equals $4T\mathrm{Var}_{|\psi\rangle}[\hat{L}_{1,(s)}]$ maximised over these eigenstates.

This result can be shown geometrically by seeing that if $\mathrm{Var}_{|\psi\rangle}[\hat{L}_{2,(n)}] > 0$, then any unextended state has vanishing QFI since the signal $\hat{L}_{1,(s)}|\psi\rangle$ must be in the span of $|\psi\rangle$ and $\hat{L}_{2,(n)}|\psi\rangle$ since the space is two-dimensional. The optimal unextended QFI then follows from Eq. 25, by observing that if $|\psi\rangle$ is an eigenstate of $\hat{L}_{2,(n)}$ then $\mathrm{Cov}_{|\psi\rangle}[\hat{L}_{1,(s)}, \hat{L}_{2,(n)}] = 0$.

A. Hierarchy of metrological strategies

Let us now summarise the results here for the hierarchy of metrological strategies which we prove below. We denote the optimal noiseless QFI as $\mathcal{I}_Q^{\text{noiseless}}$ and the optimal noisy QFI as $\mathcal{I}_Q^{\text{optimal}}$, both corresponding to an extended sequential strategy with fast and precise control (Fig. 1a) which we have shown in Theorem 1 is attained by an extended measure-and-reset strategy (Fig. 1b). Let us also denote the optimal QFI with an unextended measure-and-reset strategy (Fig. 1e) as $\mathcal{I}_Q^{\text{unextended}}$, where we have not determined whether this attains the QFI of the optimal unextended sequential strategy (Fig. 1d). Finally, let us denote the optimal QFI per qubit with an unextended parallel measure-and-reset strategy (Fig. 1f) as $\mathcal{I}_Q^{\text{parallel}}$. The hierarchy of metrological strategies for a single qubit then depends on whether $\hat{L}_{1,(s)}$ and $\hat{L}_{2,(n)}$ are each Hermitian or non-Hermitian as follows:

1. $\hat{L}_{2,(n)}$ is Hermitian and $\hat{L}_{1,(s)}$ is Hermitian or non-Hermitian (Secs. III B or III C, respectively):

$$\mathcal{I}_Q^{\text{unextended}} = \mathcal{I}_Q^{\text{parallel}} = \mathcal{I}_Q^{\text{optimal}} = \mathcal{I}_Q^{\text{noiseless}}.$$

2. $\hat{L}_{2,(n)}$ is non-Hermitian and $\hat{L}_{1,(s)}$ is Hermitian (Sec. III D):

$$\mathcal{I}_Q^{\text{unextended}} \leq \mathcal{I}_Q^{\text{parallel}} = \mathcal{I}_Q^{\text{optimal}} = \mathcal{I}_Q^{\text{noiseless}}.$$

3. $\hat{L}_{1,(s)}$, $\hat{L}_{2,(n)}$ are non-Hermitian and in-plane (Sec. III E):

$$\mathcal{I}_Q^{\text{unextended}} = \mathcal{I}_Q^{\text{parallel}} = \mathcal{I}_Q^{\text{optimal}} \leq \mathcal{I}_Q^{\text{noiseless}}.$$

4. $\hat{L}_{1,(s)}, \hat{L}_{2,(n)}$ are non-Hermitian and $\hat{L}_{2,(n)}$ out-of-plane (Sec. III F):

$$\mathcal{I}_Q^{\text{unextended}} \leq \mathcal{I}_Q^{\text{parallel}} \leq \mathcal{I}_Q^{\text{optimal}} \leq \mathcal{I}_Q^{\text{noiseless}}.$$

For each of the above inequalities (i.e. $\leq$), we provides examples in the referenced section below for where it becomes a strict inequality (i.e. $<$) or an equality (i.e. $=$).

B. $\hat{L}_{1,(s)}, \hat{L}_{2,(n)}$ are both Hermitian

We first consider the case where $\hat{L}_{1,(s)}$ and $\hat{L}_{2,(n)}$ are Hermitian. A traceless Hermitian operator acting on a single qubit is a real linear combination of the Pauli operators $\hat{\sigma}_x$, $\hat{\sigma}_y$, and $\hat{\sigma}_z$. Without loss of generality, we may assume that $\hat{L}_{1,(s)} = \frac{1}{\sqrt{2}}\hat{\sigma}_z$ and $\hat{L}_{2,(n)} = \frac{1}{\sqrt{2}}(\cos(\theta)\hat{\sigma}_x + \sin(\theta)\hat{\sigma}_y)$. The maximal variance of $\hat{L}_{1,(s)}$ is thus $\frac{1}{2}$ using any pure state on the equator of the Bloch sphere, i.e. $|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + e^{i\phi}|\downarrow\rangle)$ where $\hat{\sigma}_z|\uparrow\rangle = |\uparrow\rangle$ and $\hat{\sigma}_z|\downarrow\rangle = -|\downarrow\rangle$, such that the optimal noiseless QFI in Eq. 23 is $\mathcal{I}_Q(\sqrt{\gamma_1} = 0) = 2T$.

In the noisy case, since $|\psi\rangle$ is an eigenstate of $\hat{L}_{2,(n)}$ for $\phi = \theta$, we recover the optimal noiseless QFI using $|\psi\rangle$ as an input state. A noiseless ancilla qubit is therefore not needed in this case. For reference, the Bell state in Eq. 22 also recovers the optimal noiseless QFI since $\frac{1}{d}\mathrm{Tr}[\hat{L}_{1,(s)}^\dagger \hat{L}_{1,(s)}] = \frac{1}{2}$.

C. $\hat{L}_{1,(s)}$ is non-Hermitian and $\hat{L}_{2,(n)}$ is Hermitian

We now consider the case where $\hat{L}_{1,(s)}$ is non-Hermitian and $\hat{L}_{2,(n)}$ is Hermitian. Without loss of generality, we can assume that $\hat{L}_{1,(s)} = a\hat{\sigma}_+ + b\hat{\sigma}_-$ and $\hat{L}_{2,(n)} = \frac{1}{\sqrt{2}}\hat{\sigma}_z$ where $\hat{\sigma}_\pm = (\hat{\sigma}_x \pm i\hat{\sigma}_y)/2$, $a \neq b^*$, and $|a|^2 + |b|^2 = 1$. The noiseless QFI is $4T\max\{|a|^2, |b|^2\}$ since $\hat{L}_{1,(s)}^\dagger \hat{L}_{1,(s)} = \mathrm{diag}(|b|^2, |a|^2)$. The noiseless QFI is attained, e.g., by $|\downarrow\rangle$ if $|a| \geq |b|$ and $|\uparrow\rangle$ if $|a| < |b|$. The noiseless QFI is recovered in the noisy case with the same unextended states since they are eigenstates of $\hat{L}_{2,(n)}$.

D. $\hat{L}_{1,(s)}$ is Hermitian and $\hat{L}_{2,(n)}$ is non-Hermitian

We now consider the case where $\hat{L}_{1,(s)}$ is Hermitian and $\hat{L}_{2,(n)}$ is non-Hermitian. Without loss of generality, let $\hat{L}_{1,(s)} = \frac{1}{\sqrt{2}}\hat{\sigma}_z$ and $\hat{L}_{2,(n)} = a\hat{\sigma}_+ + b\hat{\sigma}_-$ where $a \neq b^*$ and $|a|^2 + |b|^2 = 1$. The noiseless QFI remains $\mathcal{I}_Q(\sqrt{\gamma_1} = 0) = 2T$. The Bell state in Eq. 22 recovers the optimal noiseless QFI for every a, b since $\frac{1}{2}\mathrm{Tr}[\hat{L}^\dagger_{1,(s)}\hat{L}_{1,(s)}] = \max_{|\psi\rangle} \mathrm{Var}_{|\psi\rangle}[\hat{L}_{1,(s)}]$. Let us determine whether we can recover the optimal noiseless QFI without entanglement. The eigenstates of $\hat{L}_{2,(n)}$ are

$$|\psi\rangle = \frac{\sqrt{|a|}|\uparrow\rangle \pm \sqrt{|b|}e^{\frac{1}{2}i[\arg(b)-\arg(a)]}|\downarrow\rangle}{\sqrt{|a|+|b|}}.$$

By Lemma 1, these are the only unextended states with nonzero QFI equal to

$$\mathcal{I}_Q(\sqrt{\gamma_1} = 0) = 4T\left(1 - \frac{1}{1+2|ab|}\right).$$

There is therefore a gap between the optimal QFI obtained with a Bell state and the optimal unextended QFI obtained with an eigenstate of $\hat{L}_{2,(n)}$. The size of the gap depends on the values of a, b as follows:

1. For $|a| = |b| = \frac{1}{\sqrt{2}}$, the gap vanishes and the unextended QFI is optimal.
2. For any $0 < |ab| < \frac{1}{2}$, the optimal unextended QFI is suboptimal but nonzero.
3. For $a = 0$ or $b = 0$, the optimal unextended QFI vanishes.

The case of $a = 0$ and $b = 1$ such that $\hat{L}_{1,(s)} = \frac{1}{\sqrt{2}}\hat{\sigma}_z$ and $\hat{L}_{2,(n)} = \hat{\sigma}_-$ is of particular practical interest. This case corresponds to estimating a weak dephasing rate (T_2) [19] and to qubit spectroscopy [12, 14] in the presence of strong amplitude-damping noise (T_1). These results imply that entanglement is necessary to achieve nonzero QFI for this sensing task. Hence, T_2 estimation in this regime can be significantly improved by using a Bell state compared to the standard unextended scheme. We remark that while this optimal protocol assumes a noiseless ancilla, we will show later in Sec. III G that this assumption can be removed: The same enhancement is achieved with a parallel strategy that entangles two noisy qubit probes.

E. $\hat{L}_{1,(s)}, \hat{L}_{2,(n)}$ are both non-Hermitian and in-plane

We now consider the case where $\hat{L}_{1,(s)}$ and $\hat{L}_{2,(n)}$ are non-Hermitian. We can further subdivide this case based on whether the real and imaginary parts of $\hat{L}_{1,(s)}$ and

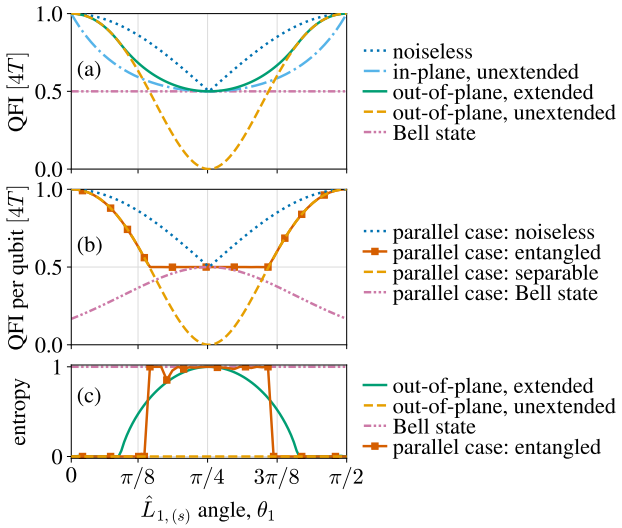

FIG. 3. **(a)** QFI versus θ_1 for the case of $\hat{L}_{1,(s)}, \hat{L}_{2,(n)}$ both non-Hermitian given in Eq. 29 where $\phi_1 = \phi_2 = 0$ and $\theta_2 = 0$ for the $\hat{L}_{2,(n)}$ in-plane case and $\theta_2 = \pi/4$ in the $\hat{L}_{2,(n)}$ out-of-plane case. In the in-plane case, an unextended state is optimal and a gap exists between the noisy and noiseless QFIs. In the out-of-plane case, there is a transition as θ_1 increases below which an unextended state is optimal and above which a gap emerges between the optimal unextended and extended QFIs. **(b)** For the unextended parallel strategy, the optimal state allowing for entanglement discontinuously changes from being separable to highly entangled when the QFI per qubit equals $2T$. **(c)** Entanglement entropy versus θ_1 normalised to the entanglement entropy of a maximally entangled state.

$\hat{L}_{2,(n)}$ all lie in the same plane. Let us first consider the in-plane case where, without loss of generality,

$$\hat{L}_{1,(s)} = \cos(\theta_1)\hat{\sigma}_+ + \sin(\theta_1)e^{i\varphi_1}\hat{\sigma}_-$$
$$\hat{L}_{2,(n)} = \sin(\theta_1)e^{-i\varphi_1}\hat{\sigma}_+ - \cos(\theta_1)\hat{\sigma}_-.$$

The noiseless QFI is thus

$$\mathcal{I}_Q(\sqrt{\gamma_1} = 0) = 4T\max\{\cos^2(\theta_1), \sin^2(\theta_1)\} \tag{26}$$

such that the Bell state, which only achieves $2T$, is suboptimal for $\theta \neq (2k+1)\pi/4$ with $k \in \mathbb{Z}$. In comparison, the optimal QFI in the noisy case is

$$\mathcal{I}_Q(\sqrt{\gamma_1} = 0) = \frac{4T}{\max_\pm[\cos(\theta_1) \pm \sin(\theta_1)]^2} \tag{27}$$

which is attained by preparing either of the two eigenstates of $\hat{L}_{2,(n)}$ denoted as $|\lambda'_j\rangle$ for $j = 1, 2$. This implies that any extended state of the form $|\psi\rangle = a_1|\lambda'_1, e_1\rangle + a_2|\lambda'_2, e_2\rangle$ with $|e_1\rangle$ and $|e_2\rangle$ orthogonal is also optimal. There is therefore no gap between the extended and unextended cases, however, there is a gap between the optimal noisy QFI in Eq. 27 and the optimal noiseless QFI in Eq. 26 for $\theta_1 \neq k\pi/4$ with $k \in \mathbb{Z}$ as shown in Fig. 3a.

This result is obtained by direct optimisation of the QFI over the input state as follows. We parameterise the

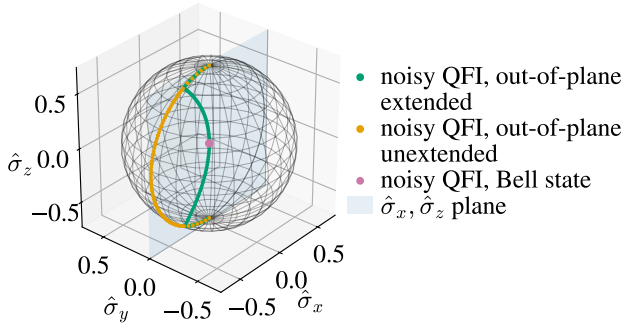

FIG. 4. Bloch ball representation of the reduced density matrix of the optimal state for the case of $\hat{L}_{1,(s)}, \hat{L}_{2,(n)}$ both non-Hermitian and $\hat{L}_{2,(n)}$ out-of-plane. The families of optimal extended states and unextended states are shown for different values of $\theta_1 \in (0, \pi/2)$ where $\theta_2 = \pi/4$ and $\varphi_1 = \varphi_2 = 0$. States on the surface of the Bloch ball are unextended, states inside the Bloch ball are extended, and states at the origin are maximally entangled such as the Bell state. Starting at $\theta_1 = 0$, the optimal state is $|\uparrow\rangle$ which is shown at the North pole of the Bloch ball. The optimal state remains unextended for increasing values of θ_1 and rotates down the $\hat{\sigma}_x, \hat{\sigma}_z$ plane until a transition point determined by $\theta_2 = \pi/4$ is reached after which the optimal state becomes extended and increasingly entangled until it reaches a maximally entangled state when $\theta_1 = \pi/4$. The behaviour is then symmetric about $\theta_1 = \pi/4$.

reduced density matrix of an arbitrary input state as:

$$\hat{\rho}_S = \begin{pmatrix} p & re^{i\phi} \\ re^{-i\phi} & 1-p \end{pmatrix}$$

where $0 \leq r^2 \leq p(1-p)$. The QFI in Eq. 25 for this input state then equals [54]

$$\mathcal{I}_Q(\sqrt{\gamma_1} = 0) = \frac{8T\left[(1-p)p - r^2\right]}{1 + (2p-1)\cos(2\theta_1) - 2r^2[1 - \sin(2\theta_1)\cos(2\phi)]}. \tag{28}$$

Maximising this QFI against the state parameters p, r, and ϕ leads to the optimal QFI in Eq. 27. This optimisation can be done analytically first against r for fixed values of p and ϕ: the QFI is either constant in r or its derivative with respect to r^2 is never zero such that it suffices to calculate the QFI on the boundary at $r = 0$ and $r^2 = p(1-p)$. Taking $r = 0$ and optimising Eq. 28 over p then yields Eq. 27.

F. $\hat{L}_{1,(s)}, \hat{L}_{2,(n)}$ are both non-Hermitian and $\hat{L}_{2,(n)}$ is out-of-plane

We now consider the case where $\hat{L}_{1,(s)}$ and $\hat{L}_{2,(n)}$ are non-Hermitian but the real and imaginary parts of $\hat{L}_{1,(s)}$ and $\hat{L}_{2,(n)}$ do not lie in the same plane. Without loss of generality,

$$\begin{aligned} \hat{L}_{1,(s)} &= \cos(\theta_1)\hat{\sigma}_+ + \sin(\theta_1)e^{i\varphi_1}\hat{\sigma}_- \\ \hat{L}_{2,(n)} &= \cos(\theta_2)[\sin(\theta_1)e^{-i\varphi_1}\hat{\sigma}_+ - \cos(\theta_1)\hat{\sigma}_-] \\ &\quad + \frac{1}{\sqrt{2}}\sin(\theta_2)e^{i\varphi_2}\hat{\sigma}_z. \end{aligned} \tag{29}$$

The noiseless QFI is still given in Eq. 26 and the Bell state still achieves $2T$, but the remaining results from the previous case are different. As in the in-plane case, we want to optimise the QFI against the initial state parameters (p, r, ϕ) for any given values of $(\theta_1, \theta_2, \varphi_1, \varphi_2)$. The expression of the QFI in this case is too verbose to optimise analytically and thus we determine the optimal state numerically instead. As shown in Fig. 3a, a gap in the QFI exists in parts of the $(\theta_1, \theta_2, \varphi_1, \varphi_2)$ parameter space, not just between the noiseless and noisy cases as seen for $\theta_2 = 0$, but also between the unextended and extended cases. This means that entanglement is sometimes necessary to achieve the maximum QFI. The optimal state transitions from the unextended state to the Bell state as shown in Fig. 3c and Fig. 4.

G. Unextended parallel strategy

We now consider the unextended parallel strategy shown in Fig. 1f with $M = 2$. Compared to the previous case shown in Fig. 1b, the noiseless ancilla is replaced by a copy of the system that is coupled independently to the same signal and noise. Experimentally, we have two copies of the same probe device which we can entangle with each other and perform joint measurements of. We model this as $K = 2$ local signal operators and $N = 2$ local noise operators and focus on the unextended states of this two-qubit system. Since we can always prepare the two noisy qubits in a separable state, the optimal QFI per qubit of this unextended parallel strategy (Fig. 1f) is always at least the optimal QFI of the unextended measure-and-reset strategy (Fig. 1e).

We previously required access to a noiseless ancilla to achieve the optimal QFI for the case of Hermitian $\hat{L}_{1,(s)}$ and non-Hermitian $\hat{L}_{2,(n)}$ (Sec. III D). We observed that for $|a| \neq |b|$, there was a gap between the optimal unextended QFI and the optimal QFI attained by the Bell state. We want to compare this now to the parallel strategy. Without loss of generality, the signal operators are $\hat{L}_{j,(s)} = \frac{1}{\sqrt{2}}\hat{\sigma}_z^{(j)}$ for $j = 1, 2$ where $\hat{\sigma}_z^{(j)}$ acts on the jth qubit and the noise operators are $\hat{L}_{j+2,(n)} = a\hat{\sigma}_+^{(j)} + b\hat{\sigma}_-^{(j)}$ for $j = 1, 2$ where $|a|^2 + |b|^2 = 1$ and $|a| \neq |b|$. The optimal noiseless QFI is $4T$ by Eq. 23 which should be normalised to a QFI per qubit of $2T$. By Eq. C1, this noiseless QFI per qubit is recovered by a Bell state since each covariance between a signal and a noise vanishes. This means that a Bell state of two noisy qubit probes is

optimal in this case and access to noiseless ancilla is not required.

Revisiting the practical problem with $a = 0$ and $b = 1$ of sensing a weak dephasing rate (T_2) in the presence of strong amplitude-damping noise (T_1), this result implies that we can overcome the vanishing QFI using a Bell state of two noisy qubit probes without requiring noiseless ancilla. Intuitively, the Bell state of the two probes allows quantum error detection of the amplitude-damping noises to differentiate them from the signals.

We also required access to a noiseless ancilla in Sec. III F where both $\hat{L}_{1,(s)}$ and $\hat{L}_{2,(n)}$ were non-Hermitian but their real and imaginary parts did not lie in the same plane. In the parallel case, without loss of generality, the signal operators are

$$\hat{L}_{j,(s)} = \cos(\theta_1)\hat{\sigma}_+^{(j)} + \sin(\theta_1)e^{i\varphi_1}\hat{\sigma}_-^{(j)}$$

where $j = 1, 2$ and the noise operators are

$$\begin{aligned}\hat{L}_{j+2,(n)} = \cos(\theta_2)[\sin(\theta_1)e^{-i\varphi_1}\hat{\sigma}_+^{(j)} - \cos(\theta_1)\hat{\sigma}_-^{(j)}] \\ + \frac{1}{\sqrt{2}}\sin(\theta_2)e^{i\varphi_2}\hat{\sigma}_z^{(j)}.\end{aligned}$$

In the noiseless case, by Eq. 23 and Eq. 26, the optimal QFI per qubit is

$$\frac{1}{2}\mathcal{I}_Q(\sqrt{\gamma_1} = 0) = 4T\max\{\cos^2(\theta), \sin^2(\theta)\}$$

and the Bell state only achieves a QFI per qubit of $2T$. We numerically optimise the noisy QFI per qubit from Eq. C1 over the initial state allowing for entanglement and compare this to the QFI per qubit of a product state of the eigenstates of $\hat{L}_{3,(n)}$ and $\hat{L}_{4,(n)}$. As shown in Fig. 3b, we observe a discontinuity in the optimal state versus θ_1 as the optimal state is separable until the QFI per qubit equals $2T$ and the optimal state becomes highly entangled. These highly entangled states maintain a QFI per qubit of $2T$ and are not necessarily Bell states or the extended states from the previous case. The optimal QFI of the out-of-plane extended state in Fig. 3a is greater than the QFI per qubit of these entangled states although they outperform the separable states. This demonstrates the expected hierarchy of metrology strategies since the extended sequential strategy in Fig. 1a can simulate the unextended parallel strategy in Fig. 1f which can itself simulate the unextended measure-and-reset strategy in Fig. 1e, respectively.

IV. MANY QUBITS

We now consider some applications for n qubits and discuss first the detection of correlated noise, focusing on correlated dephasing and decay, and then the general Pauli Lindblad estimation problem.

A. Correlated dephasing and decay

With correlated signals and noises, such as correlated dephasing and decay, we can restrict ourselves to the $n+1$ dimensional subspace of symmetric pure states in the unextended case. This is different from the unextended parallel case shown in Fig. 1f which represents uncorrelated local dephasing and decay.

Finding optimal estimation strategies of correlated dephasing and decay is relevant for various platforms and applications. Correlated dephasing appears due to laser noise [17, 55, 56] and in nano-scale nuclear magnetic resonance and spectroscopy problems [12, 14]. Correlated decay appears in various scenarios [57] such as superradiant lasers [58], Dicke-driven phase transitions [59], and chiral spin networks [60]. For each of these applications, we know that the measure-and-reset strategy is optimal (see also the discussion in Refs. [46, 47]). We thus only need to determine the optimal initial state in each case.

For example, consider sensing correlated decay $\hat{L}_{1,(s)} = \hat{J}_-$ in the presence of correlated amplification $\hat{L}_{2,(n)} = \hat{J}_+$ and dephasing $\hat{L}_{3,(n)} = \hat{J}_z$ such that we want to optimise Eq. C1. Here, the collective spin operators are defined by $\hat{J}_j = \frac{1}{2}\sum_{k=1}^n \hat{\sigma}_j^{(k)}$ for $j = x, y, z$ and $\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$. Since $(\hat{J}_-)_{j,j-1} = \sqrt{j(n+1-j)}$, then $(\hat{J}_+\hat{J}_-)_{j,k} = \delta_{jk}(j+1)(n-j)$ and the optimal state is thus the jth eigenstate of $\hat{J}_z$ where $j = (n-1)/2$ and the QFI is $4T(n+1)^2$ if n is odd or $j = n/2, n/2-1$ and the QFI equals $4Tn(n+2)$ if n is even. The QFI grows quadratically in the number of correlated sensors n but only linearly in the total time T.

In comparison, to sense correlated dephasing along $\hat{L}_{1,(s)} = \hat{J}_z$ in the presence of amplitude damping $\hat{L}_{2,(n)} = \hat{J}_-$ and amplification $\hat{L}_{3,(n)} = \hat{J}_+$ as well as dephasing along $\hat{J}_x$ and $\hat{J}_y$, the optimal state for $n > 1$ is $(|\hat{J}_z = n/2\rangle + |\hat{J}_z = -n/2\rangle)/\sqrt{2}$. This state maximises the variance of $\hat{J}_z$ and its images under the noises are orthogonal such that it recovers the noiseless QFI of Tn^2.

In more complicated cases than these two examples, we can instead use an SDP to find the optimal state.

B. Pauli Lindblad estimation

Let us now consider the full 2^n dimensional space of pure states of n qubits. We assume that each jump operator $\hat{L}$ is an n-qubit Pauli operator, i.e. a tensor product of n single qubit Pauli operators including the identity, such that it is Hermitian ($\hat{L}^\dagger = \hat{L}$) and involutory ($\hat{L}^2 = \hat{I}$). We also assume that all of the signal and noise jump operators, $\{\hat{L}_{j,(s)}\}_{j=1}^K$ and $\{\hat{L}_{j,(n)}\}_{j=K+1}^{K+N}$, correspond to different Pauli operators and thus are pairwise orthogonal. There are 4^n-1 traceless orthogonal Pauli operators. The Pauli weight of a given Pauli operator is the number of non-identity Pauli operators it contains such that there are 3^n traceless orthogonal Pauli operators with a

Pauli weight of n.

This Pauli structure simplifies the analysis and leads to the following result about the QFI in the extended case:

Claim 1. For any set of Pauli jump operators, the noiseless QFI can be recovered by the extended measure-and-reset strategy with a Bell state.

This claim can be shown as follows. By Eq. 23, the noiseless QFI is $4KT$. The noisy QFI of a Bell state given in Eq. 22 then recovers this value because $\mathrm{Tr}[\hat{L}^\dagger_{j,(s)}\hat{L}_{j,(s)}] = d$ for each signal Pauli operator. This can be understood similarly to superdense coding as the initial Bell state and its up to $4^n - 1$ images under the different Pauli operators are all pairwise orthogonal.

In the unextended case, however, Claim 1 does not hold. For example, for a single qubit and sensing $\hat{L}_{1,(s)} = \hat{\sigma}_x$ in the presence of $\hat{L}_{2,(n)} = \hat{\sigma}_y$ and $\hat{L}_{3,(n)} = \hat{\sigma}_z$, the unextended QFI vanishes. This is because the unextended initial state and its two images under the noises always span the two-dimensional space since the state cannot be a simultaneous eigenstate of $\hat{\sigma}_y$ and $\hat{\sigma}_z$. While the general sufficient and necessary conditions for attaining the noiseless QFI or for the noisy QFI vanishing are unknown in the unextended case, we now provide a couple of examples where the performance can be evaluated:

Example 1. If the Pauli weight of the signal operator is strictly greater than the maximum Pauli weight of the noise operators, then the noiseless QFI can be recovered by preparing an unextended product state.

We prove this result in Appendix F. For example, if sensing $\hat{L}_{1,(s)} = \hat{\sigma}^{(1)}_x\hat{\sigma}^{(2)}_x$ in the presence of all six local noise operators of two qubits, then preparing an unextended product state $|\uparrow\rangle \otimes |\uparrow\rangle$ recovers the noiseless QFI. Here, we restrict to one signal operator ($K = 1$) while Claim 1 applies for any $K \geq 1$. Alternatively, the noiseless QFI can always be recovered in the commuting case:

Example 2. For any commuting set of Pauli jump operators, the noiseless QFI can be recovered by the unextended measure-and-reset strategy.

This result can be shown as follows. Claim 1 implies that a Bell state recovers the noiseless QFI and Eq. E1 implies that the ancilla is not required for the commuting Hermitian case. (Appendix E considers the case more generally where all of the jump operators are Hermitian and commute.) This means that a uniform superposition $|\psi\rangle = \frac{1}{\sqrt{d}}\sum_{j=1}^{d}|e_j\rangle$ in the simultaneous eigenbasis $\{|e_j\rangle\}_{j=1}^{d}$ recovers the noiseless QFI.

V. STOCHASTIC WAVEFORM ESTIMATION

Let us consider stochastic waveform estimation using a linear quantum device, where we want to estimate the power spectral density of some continuous random variable $y(t)$ in time t by measuring an outgoing bosonic mode from the device. This scenario describes several fundamental physics experiments including searches for quantum gravity, stochastic gravitational waves, and axionic dark matter. In Ref. [1], the problem of stochastic waveform estimation in the vanishing signal limit was shown to be equivalent to the problem studied here of estimating $\sqrt{\gamma_1}$ for $\hat{L}_{1,(s)} = \hat{x}$ which is a random displacement channel along $\hat{p}$. Here, given the annihilation operator $\hat{a}$ of a single harmonic oscillator, the quadrature operators are $\hat{x} = (\hat{a} + \hat{a}^\dagger)/\sqrt{2}$ and $\hat{p} = (-i\hat{a} + i\hat{a}^\dagger)/\sqrt{2}$ such that a generalised quadrature at angle θ is given by $\hat{x}_\theta = \cos(\theta)\hat{x} + \sin(\theta)\hat{p}$. The signal and noise processes in Ref. [1], however, were modelled in separate stages and fast and precise control of the device was not considered. We now assume that fast and precise control is possible on timescales for which the signal remains stochastic.

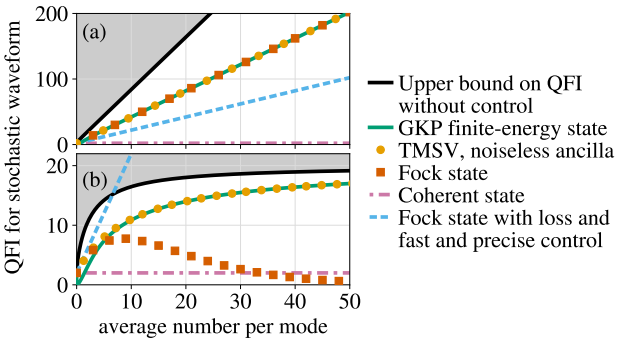

FIG. 5. QFI for stochastic waveform estimation versus the average occupation number per mode of the initial state. **(a)** In the noiseless case without control, the upper bound on the QFI for any initial state is attained by an SMSV state. **(b)** For a loss of 10%, the QFI without control is bounded at 20 and attained by a Gottesman-Kitaev-Preskill (GKP) state [61]. In comparison, the QFI with fast and precise control is unbounded and attained by a Fock state. The noiseless QFI is not recovered, however, as shown in panel (a). We omit the factor of T in the QFI with control. (Figure adapted from Fig. 4 of Ref. [1].)

1. *Noiseless case*

Let us first consider the noiseless case. By Eq. 24, the noiseless QFI is unbounded as $\mathrm{Var}_{|\psi\rangle}[\hat{x}]$ can be made arbitrarily large by preparing, e.g., a single-mode squeezed vacuum (SMSV) state as shown in Fig. 5a. Given a constraint on the average number $\langle\psi|\hat{n}|\psi\rangle = \bar{n}$ where the number operator is $\hat{n} = \hat{a}^\dagger\hat{a}$, then an SMSV state has $\mathrm{Var}_{|\psi\rangle}[\hat{x}] = \bar{n} + \frac{1}{2} + \sqrt{\bar{n}(\bar{n}+1)}$ and is optimal [1]. The noiseless QFI is thus $8\bar{n}T$ in the high energy limit of large $\bar{n} \gg 1$. In comparison, if $\bar{n} \in \mathbb{Z}$, then the Fock state $|\hat{n} = \bar{n}\rangle$ such that $\hat{n}|\hat{n} = \bar{n}\rangle = \bar{n}|\hat{n} = \bar{n}\rangle$ has $\mathrm{Var}_{|\hat{n}=\bar{n}\rangle}[\hat{x}] = \bar{n} + \frac{1}{2}$ and thus a noiseless QFI of $4\bar{n}T$ in the high energy limit as shown in Fig. 5a.

2. Noisy case

Let us now consider the case of a single noise source. We prove the following claim:

Claim 2. The optimal QFI with fast and precise control for sensing a random displacement channel generated by a quadrature $\hat{x}_\theta$ in the presence of loss $\hat{a}$ or gain $\hat{a}^\dagger$ is $2(\bar{n}+1)T$ or $2\bar{n}T$, respectively, given a constraint on the average number per mode of $\langle\psi|\hat{n}|\psi\rangle = \bar{n}$. This QFI is doubled if sensing isotropic random displacements generated by $\hat{x}$ and $\hat{p}$. The QFI is achieved by a measure-and-reset strategy for any initial state $|\psi\rangle$ that satisfies the following necessary and sufficient condition:

$$\langle\psi|\hat{a}|\psi\rangle = \langle\psi|\hat{a}^2|\psi\rangle = 0. \tag{30}$$

Examples of optimal protocols include continuously preparing a Fock state (if $\bar{n} \in \mathbb{Z}$) or a finite superposition of Fock states or a binomial code state (if $\bar{n} \notin \mathbb{Z}$) and performing a number-resolving measurement. Alternatively, preparing a two-mode squeezed vacuum (TMSV) state and performing number-resolving measurements of the system and noiseless ancilla is also optimal.

This claim can be proved as follows. We showed in Sec. II C that only the component of $\hat{L}_{1,(s)}$ orthogonal to $\hat{L}_{2,(n)}$ contributes to the QFI. The component of $\hat{x}_\theta$ orthogonal to $\hat{a}$ is $\hat{a}^\dagger/\sqrt{2}$ up to a complex phase factor $e^{i\theta}$ which does not affect the QFI. The QFI in Eq. 25 for $\hat{L}_{1,(s)} = \hat{a}^\dagger/\sqrt{2}$ and $\hat{L}_{2,(n)} = \hat{a}$ is

$$\mathcal{I}_Q(\sqrt{\gamma_1} = 0) \tag{31}$$
$$= 2T \max_{|\psi\rangle} \left(\bar{n} + 1 - |\langle\psi|\hat{a}|\psi\rangle|^2 - \frac{|\langle\psi|\hat{a}^2|\psi\rangle - \langle\psi|\hat{a}|\psi\rangle^2|^2}{\bar{n} - |\langle\psi|\hat{a}|\psi\rangle|^2} \right).$$

This equals $2(\bar{n}+1)T$ for a fixed $\langle\psi|\hat{n}|\psi\rangle = \bar{n}$ and is attained by a given state $|\psi\rangle$ if and only if it satisfies the condition in Eq. 30. If $\bar{n} \in \mathbb{Z}$, then a Fock state $|\hat{n} = \bar{n}\rangle$ suffices. If $\bar{n} \notin \mathbb{Z}$, then, for example, we can instead prepare the following state

$$|\psi\rangle = \sqrt{1 - \frac{\bar{n}}{n_\star}}|\hat{n} = 0\rangle + \sqrt{\frac{\bar{n}}{n_\star}}|\hat{n} = n_\star\rangle$$

where $n_\star = \max\{3, \lceil\bar{n}\rceil\}$. Any binomial code state [62] $|\psi\rangle = \sum_{n=0}^{\infty} c_n|\hat{n} = qn+r\rangle$ with $q, r \in \mathbb{Z}$ satisfying $q \geq 3$, $0 \leq r < q$, and $\bar{n} = r + q\sum_{n=0}^{\infty} n\,|c_n|^2$ is also optimal. In the extended case, a TMSV state given by

$$|\psi\rangle = \frac{1}{\cosh(r)} \sum_{k=0}^{\infty} \left[-e^{i\phi}\tanh(r)\right]^k |\hat{n} = k\rangle|\hat{n} = k\rangle$$

is optimal and requires continuously performing a number-resolving measurement of the system and ancilla. Similarly, the QFI in the presence of a gain $\hat{a}^\dagger$ is

$$\mathcal{I}_Q(\sqrt{\gamma_1} = 0) \tag{32}$$
$$= 2T \max_{|\psi\rangle} \left(\bar{n} - |\langle\psi|\hat{a}|\psi\rangle|^2 - \frac{|\langle\psi|\hat{a}^2|\psi\rangle - \langle\psi|\hat{a}|\psi\rangle^2|^2}{\bar{n} + 1 - |\langle\psi|\hat{a}|\psi\rangle|^2} \right)$$

which equals $2\bar{n}T$ and Eq. 30 remains the optimality condition such that the same states above are optimal. Finally, for the $K = 2, N = 1$ isotropic case of $\hat{L}_{1,(s)} = \hat{x}$ and $\hat{L}_{2,(s)} = \hat{p}$ in the presence of $\hat{L}_{3,(n)} = \hat{a}$ or $\hat{a}^\dagger$, the total QFI in Eq. 25 sums the QFI from each signal operator. This proves Claim 2.

3. Discussion of Claim 2

With fast and precise control, the QFI in the presence of loss or gain is thus unbounded and goes as $2\bar{n}T$ for $\bar{n} \gg 1$ but does not recover the noiseless QFI of $8\bar{n}T$ as shown in Fig. 5 for a loss. Without fast and precise control, however, the QFI for any state is bounded [1] as shown in Fig. 5b. Fast and precise control can therefore significantly increase the sensitivity of stochastic waveform estimation compared to the case without control.

In comparison to the optimal states above, preparing an SMSV state yields vanishing QFI in the presence of loss or gain, even with fast and precise control. This result has been previously observed in the case without fast and precise control for a fixed nonzero loss in the weak signal limit [1]. With fast and precise control, the QFI in the presence of loss in Eq. 31 or gain in Eq. 32 vanishes since $\langle\psi|\hat{a}|\psi\rangle = 0$ but $|\langle\psi|\hat{a}^2|\psi\rangle|^2 = \bar{n}(\bar{n}+1)$ for the SMSV state.

Other noise sources are also relevant to experiments sensing, e.g., random displacements generated by $\hat{x}$, such as parallel classical noise $\hat{x}$ and perpendicular classical noise $\hat{p}$. The orthogonal component of $\hat{x}$ is zero for parallel classical noise and nonzero for perpendicular classical noise. This implies that the QFI in the presence of parallel classical noise vanishes, and that with fast and precise control the QFI in the presence of perpendicular classical noise recovers the noiseless QFI using an SMSV state since $\mathrm{Cov}_{|\psi\rangle}[\hat{x}, \hat{p}] = 0$. Without fast and precise control, however, the QFI in the presence of perpendicular classical noise vanishes for an SMSV state with fixed energy in the weak signal limit [1]. This again shows how fast and precise control can improve stochastic waveform estimation.

4. Practical considerations

Let us now interrogate our assumption of fast and precise control for stochastic waveform estimation. This requirement is stringent: we must perform projective measurements and re-initialise the state of the device on the timescale of t satisfying Eq. 12 before a second jump of the noise operators occurs. For example, if a single photon is lost from the differential mode of the arms of an optical interferometer, then we need to respond and re-initialise the state before a second photon is lost. This demand is beyond the present and projected capabilities of optical interferometry experiments such as searching

for stochastic gravitational waves with the Laser Interferometric Gravitational-wave Observatory (LIGO) [2–4] or quantum gravity with the Gravity from the Quantum Entanglement of Space-Time (GQuEST) future experiment [7]. Our results in the fast and precise control limit establish the best possible performance of these interferometric devices but do not reflect their attainable near-future performance. We will discuss later possible future work to close this gap and determine the optimal near-future performance of these interferometers under more realistic assumptions.

A more promising application of our present results is to search for ultra-light dark matter such as axions using a superconducting microwave cavity in a static magnetic field [8–10, 63–67]. In the fast and precise control limit, we consider operating the device on a timescale shorter than the noise processes but still longer than the coherence time of the axion such that the signal remains stochastic [68]. The stochastic signal appears isotropically in the $\hat{x}$ and $\hat{p}$ quadratures such that the QFI in the presence of loss $\hat{a}$ is $4(\bar{n}+1)T$ and is attained, e.g., by continuously preparing a Fock state and performing a number-resolving measurement by Claim 2. The ability to prepare highly non-Gaussian states in the microwave domain [10, 69, 70] is encouraging for eventually reaching the fast and precise control limit as technology further improves.

VI. CONCLUSIONS

In this paper, we have investigated the fast and precise control limit of Lindblad estimation using a quantum device. This limit was previously well-understood for Hamiltonian estimation, i.e. sensing a deterministic signal, but not for probing a weak stochastic signal. Several diverse fundamental physics searches look for stochastic signals, from stochastic gravitational waves to axionic dark matter, and so understanding this limit is crucial to establishing the ultimate sensitivity that these devices may reach. We showed that the optimal strategy is to rapidly project the quantum state onto a basis formed from the initial state and orthogonalised images of the initial state under each Lindblad jump operator and then re-initialise the state if a jump is recorded. This is a significant departure from Hamiltonian estimation where the signal should be allowed to coherently build up rather than rapidly measured. We demonstrated this result for sensing a stochastic signal with a single qubit and found a gap between the sensitivity in the noiseless and noisy cases and another gap between the performance with and without entanglement depending on the geometry of the jump operators. We also investigated sensing a stochastic signal with many qubits and Lindblad estimation of Pauli operators. Finally, we discussed the feasibility and benefits of reaching the fast and precise control limit for stochastic waveform estimation.

There are several possible avenues for future work. On a fundamental level, investigating the multi-parameter estimation of different weak decay rates and the $M > 2$ extended parallel strategy shown in Fig. 1c for the qubit case may be interesting. Additionally, it would be useful to know whether the unextended measure-and-reset strategy shown in Fig. 1e is the optimal unextended fast and precise control strategy shown in Fig. 1d for Lindblad estimation as we have only proven this for the extended case.

It would also be interesting to compare our results to other bounds that consider monitoring of the environment [71, 72]. In various cases, the environment that induces the Lindblad equation can be continuously measured along with the probe. It has been previously shown that with monitoring the environment one can outperform the sequential ECQFI bound in deterministic signal sensing [73–75]. It would be useful to know whether environment detection leads to improved precision also in Lindblad estimation. We discuss this briefly in Appendix G but it can be further explored in the future.

Meanwhile, on a practical level, it may be valuable to consider weaker strategies than fast and precise control. For example, in optical interferometers it is difficult to measure and reset the state as fast as the time taken to lose a single photon. However, it may be possible to control the state on longer timescales such that the noise may jump multiple times but the weak signal may only jump once. Similarly, it may be fruitful to investigate the optimal strategy when restricted to weaker controls than projectively measuring and resetting.

ACKNOWLEDGMENTS

We thank the following people for their advice: Richard Allen, Wenjie Gong, Daniel Mark, Lee McCuller, Chris Pattinson, John Preskill, Alex Retzker, Thomas Schuster, and Sisi Zhou. We also thank the Caltech Chen Quantum Group and the ANU CGA Squeezer Group. This research is supported by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (Project No. CE170100004 and CE230100016). J.W.G. and this research are supported by an Australian Government Research Training Program Scholarship and also partially supported by the US NSF grant PHY-2011968. In addition, Y.C. acknowledges the support by the Simons Foundation (Award Number 568762) and the NSF grant PHY-2309231. T.G. acknowledges funding provided by the Institute for Quantum Information and Matter (a US NSF Physics Frontiers Center) and the Quantum Science and Technology Scholarship of the Israel Council for Higher Education. S.A.H. acknowledges support through an Australian Research Council Future Fellowship grant FT210100809. This paper has been assigned LIGO Document No. P2400542.

Appendix A: Proof of Theorem 1

We want to show that the ultimate QFI with fast and precise control is given in Theorem 1 and that it is achieved by a measure-and-reset strategy.

1. Proof that Theorem 1 is an upper bound

We want to show the upper bound on the QFI with fast and precise control given in Theorem 1 by calculating the minimum in Eq. 16. We may use the minimax theorem since the argument is concave in $\hat{\rho}$ and convex in $\{\vec{c}_j\}_{j=1}^K$ [76]. This implies that Eq. 16 becomes

$$\begin{aligned}
&\mathcal{I}_Q(\sqrt{\gamma_1}=0)\\
&\leq 4T\max_{\hat{\rho}}\min_{\{\vec{c}_j\}_{j=1}^K}\mathrm{Tr}[\hat{\rho}\sum_{j=1}^K \hat{A}_j^\dagger\hat{A}_j]\\
&= 4T\max_{|\psi\rangle}\sum_{j=1}^K \min_{\vec{c}_j}\langle\psi|(\hat{L}_{j,(s)}-\vec{c}_j^{\mathrm{T}}\mathbf{L})^\dagger(\hat{L}_{j,(s)}-\vec{c}_j^{\mathrm{T}}\mathbf{L})|\psi\rangle
\end{aligned} \tag{A1}$$

where we replaced the density matrix $\hat{\rho}$ by the extended pure state $|\psi\rangle$ that purifies it such that $\hat{\rho}=\mathrm{Tr}_A[|\psi\rangle\langle\psi|]$.

We now want to show that Eq. A1 is equivalent to Theorem 1. This proof can be done in two different ways. The first way is to differentiate Eq. A1 with respect to $\vec{c}_j^\dagger$, which yields that the minimum is achieved at

$$\vec{c}_j = \langle\psi|\mathbf{L}^*\mathbf{L}^{\mathrm{T}}|\psi\rangle^{-1}\langle\psi|\mathbf{L}^*\hat{L}_{j,(s)}|\psi\rangle. \tag{A2}$$

For example, in the noiseless case this means that $\vec{c}_j = \langle\psi|\hat{L}_{j,(s)}|\psi\rangle$. Substituting this value for $\vec{c}_j$ in Eq. 16 and using Eq. 19 then shows that the QFI is upper bounded as given in Theorem 1. The second way is to rewrite Eq. A1 as follows:

$$\mathcal{I}_Q(\sqrt{\gamma_1}=0)\leq 4T\max_{|\psi\rangle}\sum_{j=1}^K\min_{\vec{c}_j}||(\hat{L}_{j,(s)}-\vec{c}_j^{\mathrm{T}}\mathbf{L})|\psi\rangle||^2$$

such that for every $|\psi\rangle$ and j, we want to find $\min_{\vec{c}_j}||(\hat{L}_{j,(s)}-\vec{c}_j^{\mathrm{T}}\mathbf{L})|\psi\rangle||^2$. By splitting the identity into projections onto orthogonal subspaces with $\hat{I}=\hat{\Pi}+(\hat{I}-\hat{\Pi})$, we can show that

$$\begin{aligned}
&\min_{\vec{c}_j}||(\hat{L}_{j,(s)}-\vec{c}_j^{\mathrm{T}}\mathbf{L})|\psi\rangle||^2\\
&=\min_{\vec{c}_j}||(\hat{\Pi}\hat{L}_{j,(s)}-\vec{c}_j^{\mathrm{T}}\mathbf{L})|\psi\rangle||^2+||(\hat{I}-\hat{\Pi})\hat{L}_{j,(s)}|\psi\rangle||^2\\
&=\langle\psi|\hat{L}_{j,(s)}^\dagger(\hat{I}-\hat{\Pi})\hat{L}_{j,(s)}|\psi\rangle
\end{aligned}$$

where the last equality is obtained by choosing $\vec{c}_j$ such that $\vec{c}_j^{\mathrm{T}}\mathbf{L}|\psi\rangle=\hat{\Pi}\hat{L}_{j,(s)}|\psi\rangle$. This again proves that Theorem 1 is an upper bound.

2. Proof that Theorem 1 can be achieved

We want to show that the QFI for the measure-and-reset strategy achieves the ultimate sensitivity limit in Theorem 1. Let us assume for now that there is no nuisance Hamiltonian, i.e. $\hat{H}=0$, and then we will show later that it has no effect on the optimal strategy. By Eq. 10, the signal-free density matrix is

$$\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}' \approx \hat{\rho}+\sum_{j=K+1}^{K+N}\gamma_j t\mathcal{L}_{j,(n)}(\hat{\rho})$$

such that its matrix elements in the $\{|j\rangle\}_{j=0}^{K+N}$ basis are

$$\begin{aligned}
\langle 0|\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'|0\rangle &= 1-\sum_{j=K+1}^{K+N}\gamma_j t\sum_{k=K+1}^{j}|c_{j,k}|^2\\
\langle j|\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'|k\rangle &= 0\\
\langle j|\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'|0\rangle &= -\frac{1}{2}\sum_{k=K+1}^{K+N}\gamma_k t\langle j|\hat{L}_{k,(n)}^\dagger\hat{L}_{k,(n)}|0\rangle\\
\langle j'|\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'|k'\rangle &= \sum_{l=\max\{j',k'\}}^{K+N}\gamma_l t c_{l,j'}c_{l,k'}^*\\
\langle j'|\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'|k\rangle &= 0\\
\langle j'|\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'|0\rangle &= \sum_{k=K+1}^{K+N}\gamma_k t\Big(c_{k,0}^*\langle j'|\hat{L}_{k,(n)}|0\rangle\\
&\qquad -\frac{\langle j'|\hat{L}_{k,(n)}^\dagger\hat{L}_{k,(n)}|0\rangle}{2}\Big)
\end{aligned} \tag{A3}$$

where $1\leq j,k\leq K$ and $K+1\leq j',k'\leq K+N$ for the free indices above. Let us find the eigenvectors $\{|\lambda_j\rangle\}_{j=0}^{K+N}$ of this matrix since they form the optimal measurement basis by Eq. 5. Firstly, let us try the following ansatz

$$|\lambda_0\rangle=|0\rangle+t\sum_{j=1}^{K+N}a_{0,j}|j\rangle,\quad \lambda_0=\langle 0|\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'|0\rangle$$

such that the eigenvector equation is

$$\left(\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'-\lambda_0\hat{I}\right)|\lambda_0\rangle=\sum_{j=1}^{K+N}\left(\langle j|\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'|0\rangle-a_{0,j}t\right)|j\rangle$$

which is zero if $a_{0,j}=\langle j|\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'|0\rangle/t$ for $j=1,\ldots,K+N$. Secondly, let us try the following ansatz for $j=1,\ldots,K$

$$|\lambda_j\rangle=\sum_{k=1}^{K}a'_{j,k}|k\rangle+t\sum_{k=0}^{K+N}a_{j,k}|k\rangle,\quad \lambda_j=0$$

such that the eigenvector equation is

$$\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'|\lambda_j\rangle=\left(a_{j,0}t+\sum_{k=1}^{K}a'_{j,k}\langle 0|\lim_{\sqrt{\gamma_1}\to 0}\hat{\rho}'|k\rangle\right)|0\rangle$$

which is zero if

$$a_{j,0} = -\sum_{k=1}^{K} a'_{j,k} \frac{\langle 0| \lim_{\sqrt{\gamma_1}\to 0} \hat{\rho}'|k\rangle}{t}.$$

This means that there is a degenerate K-dimensional subspace of eigenvectors with eigenvalue zero for which we may freely choose which to project onto. Without loss of generality, we choose the eigenvectors which are $\mathcal{O}(t)$ perturbations of the $\{|j\rangle\}_{j=1}^{K}$ states such that

$$|\lambda_j\rangle = |j\rangle + t\sum_{k=0}^{K+N} a_{j,k}|k\rangle, \quad \lambda_j = 0$$

where $a_{j,0} = -\langle 0| \lim_{\sqrt{\gamma_1}\to 0} \hat{\rho}'|j\rangle/t$ for $j = 1,\ldots,K$. The remaining N eigenvectors, corresponding to the noisy subspace, are orthogonal to $\{|\lambda_j\rangle\}_{j=0}^{K}$ such that they are of the following form for $j = K+1,\ldots,K+N$

$$|\lambda_j\rangle = \sum_{k=K+1}^{K+N} a'_{j,k}|k\rangle + t\sum_{k=0}^{K+N} a_{j,k}|k\rangle.$$

Unlike the first $K+1$ eigenvectors which are $\mathcal{O}(t)$ perturbations of the initial state and signal eigenbasis $\{|j\rangle\}_{j=1}^{K}$, these last N eigenvectors are not $\mathcal{O}(t)$ perturbations of $\{|j\rangle\}_{j=K+1}^{K+N}$ and are not necessarily degenerate.

We now calculate the probabilities $p_j = \mathrm{Tr}\left[|\lambda_j\rangle\langle\lambda_j|\hat{\rho}'\right]$ of projecting $\hat{\rho}'$ onto this optimal measurement basis. The first $K+1$ probabilities are equal to the corresponding diagonal elements of $\hat{\rho}'$, i.e. $p_j = \langle j|\hat{\rho}'|j\rangle$ for $j = 0,1,\ldots,K$. Since p_0 has a constant term, its contribution to the CFI in Eq. 2 vanishes. Meanwhile, the probabilities for $j = 1,\ldots,K$ are

$$\begin{aligned} p_j &= \gamma_1 t\sum_{k=j}^{K} |c_{k,j}|^2 \\ &= \gamma_1 t\langle\psi|\hat{L}^{\dagger}_{j,(s)}(\hat{I}-\hat{\Pi})\hat{L}_{j,(s)}|\psi\rangle \end{aligned} \tag{A4}$$

such that the sum of their contributions to the CFI in Eq. 2 equals the QFI in Theorem 1 after multiplying by the number of measurements $M = T/t$. Moreover, all of the information about $\sqrt{\gamma_1}$ comes from projecting onto the combined noise-free signal subspace with $\hat{I} - \hat{\Pi} = \sum_{j=1}^{K}|j\rangle\langle j|$, i.e. we do not need to differentiate the different signals since the decay rate is common.

We now need to show that the remaining probabilities p_j for $j = K+1,\ldots,K+N$ from the noisy N-dimensional subspace contribute no information in the limit of $\sqrt{\gamma_1}\to 0$. These probabilities only depend on the constant coefficients $a'_{j,k}$ of $|\lambda_j\rangle$ and thus on the $\langle j'|\hat{\rho}'|k'\rangle$ matrix elements where $K+1 \le j', k' \le K+N$, such that p_j for $j = K+1,\ldots,K+N$ can be shown to equal

$$\gamma_1 t\sum_{m=1}^{K}\left|\sum_{k=K+1}^{K+N} c^*_{m,k}a'_{j,k}\right|^2 + \sum_{m=K+1}^{K+N}\gamma_m t\left|\sum_{k=K+1}^{m} c^*_{m,k}a'_{j,k}\right|^2$$

The contribution to the CFI in Eq. 2 corresponding to p_j vanishes provided that the second term above is nonzero. That is, if $|\sum_{k=K+1}^{m} c^*_{m,k}a'_{j,k}|^2$ is nonzero for at least one $m = K+1,\ldots,K+N$. Let us show that this must be the case since otherwise we will find a contradiction if we try to set this term to zero for all $m = K+1,\ldots,K+N$. Firstly, consider $m = K+1$ such that $|c^*_{K+1,K+1}a'_{j,K+1}|^2 = 0$ if and only if $a'_{j,K+1} = 0$ since we assume that $c_{j,j} \neq 0$ without loss of generality. Then, consider $m = K+2$ such that the term is $|c^*_{K+2,K+2}a'_{j,K+2}|^2$ given $a'_{j,K+1} = 0$, which then demands that $a'_{j,K+2} = 0$ too. This process continues similarly with the $m = K+3$ term and so on up to the $m = K+N$ term with $a'_{j,m} = 0$ required at each stage. Since $|\lambda_j\rangle$ is normalised, however, $a'_{j,m} \neq 0$ for at least one $m = K+1,\ldots,K+N$ such that this process must fail and thus the contribution to the CFI from p_j vanishes.

Projecting onto $\{|j\rangle\}_{j=0}^{K+N}$ instead of $\{|\lambda_j\rangle\}_{j=0}^{K+N}$ yields the same CFI since the only differences between the two bases, i.e. $\mathcal{O}(t)$ perturbations and a rotation of the noisy subspace, do not affect the CFI to leading order. We can now address what happens to the CFI if an unknown nuisance Hamiltonian $\hat{H}$ is also present. The final state in Eq. 10 becomes $\hat{\rho}' = \hat{\rho}'|_{\hat{H}=0} - it[\hat{H},\hat{\rho}]$ such that the probability p_j from projecting onto $|j\rangle$ for $j = 0,\ldots,K+N$ is unchanged since

$$\begin{aligned} p_j &= \mathrm{Tr}\left[|j\rangle\langle j|\hat{\rho}'\right] \\ &= p_j|_{\hat{H}=0} + 2t\,\mathrm{Im}\left[\langle j|\hat{H}|\psi\rangle\langle\psi|j\rangle\right] \\ &= p_j|_{\hat{H}=0}. \end{aligned}$$

The nuisance Hamiltonian only affects the off-diagonal elements of the final density matrix such as $\langle j|\hat{\rho}'|\psi\rangle$ for $j > 0$. Since the probabilities are unchanged, the CFI is too and thus the nuisance Hamiltonian does not affect the optimal strategy. This completes the proof of Theorem 1.

Appendix B: Optimal initial states

We establish here the sufficient and necessary conditions for a state $\hat{\rho}$ to be optimal. Let the solution of the SDP in Eq. 16 be $\vec{c}_{j,\min}$ such that $\hat{A}_{j,\min} = \hat{L}_j - \vec{c}^{\,\mathrm{T}}_{j,\min}\mathbf{L}$ for $1 \le j \le K$. A state $\hat{\rho}$ is then optimal if and only if it satisfies the following conditions:

(I) The support of $\hat{\rho}$ lies within the eigenspace of the maximal eigenvalue of $\sum_{j=1}^{K}\hat{A}^{\dagger}_{j,\min}\hat{A}_{j,\min}$.

(II) For any signal index $1 \le j \le K$ and noise index $K+1 \le k \le K+N$, then $\mathrm{Tr}[\hat{\rho}\hat{A}_{j,\min}] = 0$ and $\mathrm{Tr}[\hat{\rho}\hat{L}^{\dagger}_{k,(n)}\hat{A}_{j,\min}] = 0$ hold.

These conditions can be proven as a special case of the conditions found in Ref. [28], however, we now present a self-contained derivation for our problem. $\hat{\rho}$ is an optimal

initial state if and only if it satisfies:

$$\mathrm{Tr}\left[\hat{\rho}\sum_{j=1}^{K}\hat{A}_{j,\min}^{\dagger}\hat{A}_{j,\min}\right] = \max_{\hat{\varrho}}\min_{\vec{c}}\mathrm{Tr}\left[\hat{\varrho}\sum_{j=1}^{K}\hat{A}_{j}^{\dagger}\hat{A}_{j}\right] \tag{B1}$$

$$= \min_{\vec{c}}\max_{\hat{\varrho}}\mathrm{Tr}\left[\hat{\varrho}\sum_{j=1}^{K}\hat{A}_{j}^{\dagger}\hat{A}_{j}\right] \tag{B2}$$

Condition (I) guarantees that $\hat{\rho}$ and $\vec{c}_{\min}$ saturate Eq. B2. We have shown that $\hat{\rho}$ and $\vec{c}_{\min}$ saturate Eq. B1 if and only if $A_{j,\min}|\psi\rangle$ is orthogonal to $|\psi\rangle$ and $\hat{L}_{k,(n)}|\psi\rangle$ for every $K+1 \le k \le K+N$ where $|\psi\rangle$ is a purification of $\hat{\rho}$. This is exactly Condition (II).

Appendix C: QFI with two sources of noise

In the case where there are K signals and two noises $\hat{L}_{K+1,(n)}$ and $\hat{L}_{K+2,(n)}$, the QFI in Theorem 1 is

$$\mathcal{I}_Q(\sqrt{\gamma_1}=0) = 4T\max_{|\psi\rangle}\sum_{j=1}^{K} f_{|\psi\rangle}(\hat{L}_{j,(s)},\hat{L}_{K+1,(n)},\hat{L}_{K+2,(n)}). \tag{C1}$$

Here, the function $f_{|\psi\rangle}$ is defined as

$$\begin{aligned} &f_{|\psi\rangle}(\hat{L}_{j,(s)},\hat{L}_{K+1,(n)},\hat{L}_{K+2,(n)}) \\ &= \mathrm{Var}_{|\psi\rangle}[\hat{L}_{j,(s)}] - \frac{\left|\mathrm{Cov}_{|\psi\rangle}[\hat{L}_{K+1,(n)},\hat{L}_{j,(s)}]\right|^2}{\mathrm{Var}_{|\psi\rangle}[\hat{L}_{K+1,(n)}]} \\ &\quad - \frac{\left|\mathrm{Cov}_{|\psi\rangle}[\hat{L}_{K+2,(n)},\hat{L}_{j,(s)}] - \zeta\right|^2}{\mathrm{Var}_{|\psi\rangle}[\hat{L}_{K+2,(n)}] - \xi} \end{aligned}$$

where ζ and ξ are given by

$$\zeta = \frac{\mathrm{Cov}_{|\psi\rangle}[\hat{L}_{K+2,(n)},\hat{L}_{K+1,(n)}]\mathrm{Cov}_{|\psi\rangle}[\hat{L}_{K+1,(n)},\hat{L}_{j,(s)}]}{\mathrm{Var}_{|\psi\rangle}[\hat{L}_{K+1,(n)}]}$$

$$\xi = \frac{\left|\mathrm{Cov}_{|\psi\rangle}[\hat{L}_{K+1,(n)},\hat{L}_{K+2,(n)}]\right|^2}{\mathrm{Var}_{|\psi\rangle}[\hat{L}_{K+1,(n)}]}$$

The second term in $f_{|\psi\rangle}$ removes the $\hat{L}_{K+1,(n)}$ component of the signal and is the same as the case of a single noise operator in Eq. 25. The third term in $f_{|\psi\rangle}$ then removes the remaining $\hat{L}_{K+2,(n)}$ component of the signal that is orthogonal to the $\hat{L}_{K+1,(n)}$ component already removed.

Appendix D: Sensing a finite signal with no noise

We now discuss the noiseless case of sensing a finite signal as discussed in Sec. II E. We want to show that the QFI for a finite signal is the same as in the vanishing signal limit, provided that all of the signal jump operators are Hermitian. The proof is analogous to the vanishing signal case. Compared to Eq. 13, the Kraus operators from Eq. 11 are now

$$\mathbf{K} = \begin{pmatrix} \hat{I} - \frac{t}{2}\gamma_1\sum_{k=1}^{K}\hat{L}_{k,(s)}^{\dagger}\hat{L}_{k,(s)} \\ \sqrt{\gamma_1 t}\hat{L}_{1,(s)} \\ \vdots \\ \sqrt{\gamma_1 t}\hat{L}_{K,(s)} \end{pmatrix}$$

$$\dot{\mathbf{K}} = \begin{pmatrix} -t\sqrt{\gamma_1}\sum_{k=1}^{K}\hat{L}_{k,(s)}^{\dagger}\hat{L}_{k,(s)} \\ \hat{L}_{1,(s)}\sqrt{t} \\ \vdots \\ \hat{L}_{K,(s)}\sqrt{t} \end{pmatrix}.$$

We choose the same gauge as the vanishing signal limit in Eq. 14:

$$h = \begin{pmatrix} 0 & \mathbf{g}^{\dagger} \\ \mathbf{g} & \mathbf{0}_{K,K} \end{pmatrix}, \quad \mathbf{g} = -i\sqrt{t}\begin{pmatrix} c_1 \\ \vdots \\ c_K \end{pmatrix}$$

such that $\dot{\mathbf{K}} - ih\mathbf{K}$ now equals Eq. 15 plus higher order corrections in time t as follows:

$$\dot{\mathbf{K}} - ih\mathbf{K} = \begin{pmatrix} 0 \\ \hat{L}_{1,(s)} - c_1\hat{I} \\ \vdots \\ \hat{L}_{K,(s)} - c_K\hat{I} \end{pmatrix}\sqrt{t} + \begin{pmatrix} t\sqrt{\gamma_1}\sum_{k=1}^{K}c_k^*\hat{L}_{k,(s)} - \hat{L}_{k,(s)}^{\dagger}\hat{L}_{k,(s)} \\ c_1\frac{1}{2}t^{3/2}\gamma_1\sum_{k=1}^{K}\hat{L}_{k,(s)}^{\dagger}\hat{L}_{k,(s)} \\ \vdots \\ c_K\frac{1}{2}t^{3/2}\gamma_1\sum_{k=1}^{K}\hat{L}_{k,(s)}^{\dagger}\hat{L}_{k,(s)} \end{pmatrix}.$$

This means that $\hat{\alpha}$ and $\hat{\beta}$ in Eq. 7 are now

$$\hat{\alpha} = t\sum_{k=1}^{K}\hat{A}_k^{\dagger}\hat{A}_k + \mathcal{O}(t^2)$$

$$\hat{\beta} = i\sqrt{\gamma_1}t\sum_{k=1}^{K}\left(c_k\hat{L}_{k,(s)}^{\dagger} - c_k^*\hat{L}_{k,(s)}\right) + \mathcal{O}(t^2)$$

where $\hat{A}_k = \hat{L}_{k,(s)} - c_k\hat{I}$ here. We now apply our assumption that all of the signal jump operators are Hermitian to imply that $\hat{\beta} = \mathcal{O}(t^2)$ since the coefficients $\{c_k\}_{k=1}^{K}$ are then real by Eq. A2. Since $\hat{\alpha} = \mathcal{O}(t)$ and $M = T/t$ with T fixed, the second term in Eq. 6 can then be dropped since it is $\mathcal{O}(T^2\sqrt{t})$ compared to the first term which is $\mathcal{O}(T)$. This establishes that the noiseless QFI for a non-vanishing signal $\sqrt{\gamma_1} > 0$ is upper bounded by the vanishing signal limit in Eq. 23, provided that the signal jump operators are Hermitian.

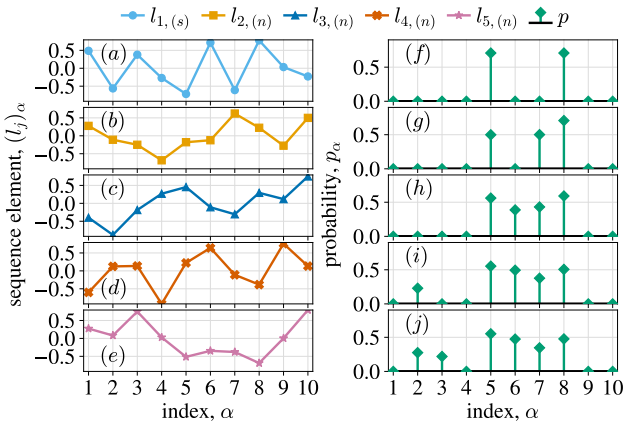

FIG. 6. In the commuting Hermitian case, the jump operators $\hat{L}_j$ become real vectors l_j and the state $|\psi\rangle$ becomes a probability distribution p. **(a)** The signal vector $l_{1,(s)}$ and **(b–d)** noise vectors $l_{j,(n)}$ are randomly sampled from a uniform distribution. **(f)** Optimal probability distribution in the noiseless case. **(g–j)** Optimal distribution given the signal $l_{1,(s)}$ and first 1–4 noises, respectively. For example, panel (i) shows the distribution given the signal $l_{1,(s)}$ and three noises: $l_{2,(n)}$, $l_{3,(n)}$, and $l_{4,(n)}$. The support of the optimal distribution is length $N+2$ given N noise operators.

We now show that the CFI of the same measure-and-reset strategy as the vanishing signal limit, i.e. preparing and projecting onto the optimal state $|\psi\rangle$, attains this upper bound on the QFI. The probability $p = \langle\psi|\hat{\rho}'|\psi\rangle$ of projecting the density matrix $\hat{\rho}'$, expressed in the $\{|j\rangle\}_{j=0}^{K}$ basis, onto the initial state $|\psi\rangle$ is

$$p = 1 - \gamma_1 t \sum_{j=1}^{K}\sum_{k=1}^{j} |c_{jk}|^2 .$$

The CFI in Eq. 2 of p and $1-p$ with respect to $\sqrt{\gamma_1}$ therefore equals the upper bound to leading order in time, once multiplied by the number of measurements $M = T/t$. This is because the probability p is the same in the vanishing signal limit in Eq. A4.

This completes the proof, assuming that the signal jump operators are Hermitian. In the non-Hermitian case, the measure-and-reset CFI above is the same since we did not use the assumption of Hermitiancy in deriving it. We defer to future work showing that the same upper bound holds in the non-Hermitian case such that the optimal QFI is independent of γ_1.

Appendix E: Commuting Hermitian case

Let us discuss the case where the jump operators are all Hermitian and pairwise commute. This implies that a simultaneous eigenbasis exists where $\hat{L}_j = \mathrm{diag}\,(l_j)$ and $l_j \in \mathbb{R}^d$ for each $j = 1,\ldots,K+N$. We can assume without loss of generality that the jump operators are traceless and pairwise orthonormal, i.e. $\sum_{\alpha=1}^{d}(l_j)_\alpha = 0$ and $\sum_{\alpha=1}^{d}(l_j)_\alpha(l_k)_\alpha = \delta_{jk}$ where the $\{l_j\}_{j=1}^{K+N}$ are thus zero-sum and pairwise orthonormal as real vectors. Here, we use Latin indices (e.g. $j,k,l \in \{1,\ldots,K+N\}$) to label a particular vector and Greek indices (e.g. $\alpha,\beta,\gamma \in \{1,\ldots,d\}$) to label the elements of that vector. Let us now consider the pure quantum state $|\psi\rangle = \sum_{\alpha,\beta=1}^{d} c_{\alpha\beta}|e_\alpha\rangle \otimes |e_\beta\rangle$ and its reduced density matrix $\hat{\rho}_S$. The probability for the system to be in $|e_\alpha\rangle$ is $p_\alpha = \sum_{\beta=1}^{d}|c_{\alpha\beta}|^2$ for $\alpha = 1,\ldots,d$. Let us define the expectation value of a vector $l \in \mathbb{R}^d$ with respect to a probability distribution $p \in \mathbb{R}^d$ where $\sum_{\alpha=1}^{d} p_\alpha = 1$ as $\mathbb{E}_p[l] = \sum_{\alpha=1}^{d} p_\alpha l_\alpha$. For any operator $\hat{A} = \mathrm{diag}\,(l)$ where $l \in \mathbb{R}^d$ and $\sum_{\alpha=1}^{d} l_\alpha = 0$, then, since $\langle\psi|\hat{A}\otimes\hat{I}|\psi\rangle = \mathrm{Tr}[\hat{\rho}_S\hat{A}]$, the expectation value is

$$\langle\psi|\hat{A}\otimes\hat{I}|\psi\rangle = \mathbb{E}_p[l]. \tag{E1}$$

The QFI in Theorem 1 only involves expectation values of operators $\{\hat{L}_j\}_{j=1}^{K+N}$ or products of two such operators with respect to $|\psi\rangle$. By Eq. E1, this means that entanglement is not required and the QFI only depends on $\mathbb{E}_p[l_j]$ and $\mathbb{E}_p[l_j l_k]$.

We can now restate the estimation problem for the case of a single source of noise. Given $l_j \in \mathbb{R}^d$ with $\sum_{\alpha=1}^{d}(l_j)_\alpha = 0$ and $\sum_{\alpha=1}^{d}(l_j)_\alpha(l_k)_\alpha = \delta_{jk}$ for $j,k \in \{1,2\}$, then the QFI in Eq. 25 is

$$\begin{aligned}&\mathcal{I}_Q(\sqrt{\gamma_1}=0)\\ &= 4T\max_p\left(\mathrm{Var}_p[l_{1,(s)}] - \frac{|\mathrm{Cov}_p[l_{1,(s)},l_{2,(n)}]|^2}{\mathrm{Var}_p[l_{2,(n)}]}\right).\end{aligned}$$

We want to find the optimal probability distribution $p \in \mathbb{R}^d$ above. Similarly, for the case of N noise sources, we replace all operator (co)variances in Theorem 1 for $K=1$ with their respective vector (co)variances. The uniform distribution $p_\alpha = 1/d$ makes the covariances vanish since the $\{l_j\}_{j=1}^{K+N}$ are pairwise orthogonal and leaves $\mathcal{I}_Q(\sqrt{\gamma_1}=0) = 4T/d$ which is nonzero but not necessarily the optimal distribution, similarly to the Bell state in Eq. 22. Since the QFI is now only expressed in terms of real vectors and probability distributions, the commuting Hermitian case can also be viewed as the classical limit of Lindblad estimation.

Let us now consider the commuting Hermitian case with one signal ($K=1$) and $N = d-2$ noises. Provided that a regularity condition of generalised non-degeneracy given in Appendix E 1 holds, then the optimal QFI is

$$\mathcal{I}_Q(\sqrt{\gamma_1}=0) = 4T\frac{||l_1||_2^4}{||l_1||_1^2} \tag{E2}$$

where $\|l\|_2^2 = \sum_{\alpha=1}^{d}|l_\alpha|^2$ and $\|l\|_1 = \sum_{\alpha=1}^{d}|l_\alpha|$. The optimal distribution is $p_\alpha \propto |l_{1,\alpha}|$ for $\alpha = 1,\ldots,d$, i.e. it mimics the magnitude of the signal l_1 up to normalisation. The regularity condition is unnecessary for $d=3$,

i.e. for one signal and one noise in three dimensions. For $d = 2$, i.e. the noiseless case in two dimensions, the QFI in Eq. E2 recovers the noiseless QFI of $2T||l_1||_2^2$. We prove these results in Appendix E 1.

In the general case with one signal ($K = 1$), N noises, and any dimension d, we can determine the optimal QFI and initial state given a solution $c_{1,\min}, c_{i,\min}$ to the following SDP:

$$l_1 = c_1 \vec{1} + \sum_{i=2}^{N+1} c_i l_i + a_{\min}$$

where $\vec{1} = (1, \ldots, 1)^{\mathrm{T}}$. By Appendix B, the optimal initial state is supported only on indices $\alpha \in I$ that satisfy $|\,(a_{\min})_\alpha\,| = ||a_{\min}||_\infty$ where $\|l\|_\infty = \max_{\alpha=1,\ldots,d} |l_\alpha|$. We can thus restrict ourselves to the m-dimensional subspace spanned by the indices in I and consider the corresponding sub-vectors $l_j^{(I)}$ and $a_{\min}^{(I)}$. The optimal QFI is then:

$$\mathcal{I}_Q(\sqrt{\gamma_1} = 0) = 4T \frac{||l_{1,\perp}^{(I)}||_2^4}{||l_{1,\perp}^{(I)}||_1^2} \tag{E3}$$

where $l_{1,\perp}^{(I)}$ is the projection of $l_1^{(I)}$ onto the subspace orthogonal to $\mathrm{span}\{\vec{1}, l_2^{(I)}, ..., l_{N+1}^{(I)}\}$. The optimal distribution is then $p_\alpha \propto |(l_{1,\perp}^{(I)})_\alpha|$ for $\alpha = 1, \ldots, m$ up to normalisation. We derive this optimal QFI and distribution in Appendix E 1.

We can thus calculate the optimal QFI and initial state from the support of the optimal distribution. It remains to understand this support. We numerically observe the following phenomenon: the optimal distribution for $K = 1$ in the presence of N noises has a support of length $N+2$. (Our code for this result and the rest of our work is available online [77].) This is shown in Fig. 6 for $d = 10$ where we see that each additional noise adds one more dimension to the support of the optimal distribution. Analytically, we know that given N noise operators and a generalised condition of non-degeneracy, the minimal length of the support is $N + 2$ since otherwise the QFI must vanish by the dimension of the noisy subspace. This numerical phenomenon suggests that a length $N+2$ support is also sufficient, but we defer proving this result to future work.

1. Proofs for the commuting Hermitian case

We derive here the above results for the commuting Hermitian case. Let us first prove for $d = 3$ and $N = 1$ that the QFI is given in Eq. E2 and that the optimal distribution is $p_i \propto |l_{1,i}|$. We derive this using the SDP formulation of Eq. 16. The solution of this SDP is given in the common eigenbasis as:

$$l_1 = c_{1,\min}\vec{1} + c_{2,\min} l_2 + a_{\min} \tag{E4}$$

where $a_{\min}$ is the diagonal of $\hat{A}_{\min}$ given $c_{1,\min}$ and $c_{2,\min}$. The optimal QFI is then given by $4T||a_{\min}||_\infty^2$. Eq. E4 implies that $a_{\min} \notin \mathrm{span}\{l_2, \vec{1}\}$ such that $a_{\min}$ is not orthogonal to l_1 and

$$|l_1|^2 = l_1 \cdot a_{\min}. \tag{E5}$$

If $(l_2)_1 = (l_2)_3$ then the optimal QFI is equal to the noiseless one: $4T\,(l_{1,\max} - l_{1,\min})^2$ with the optimal state being $(1/\sqrt{2}, 0, 1/\sqrt{2})^{\mathrm{T}}$.

Let us now split into cases regarding the size of the support of the optimal state which is either length two or three. If the support is length two, then $l_1 = (1, 0, -1)$ and $l_2 = (1, -2, 1)$ up to constants and permutations, and the optimal QFI equals the two-dimensional noiseless QFI. If the support is instead length three, then $a_{\min} = x v_i$ where $v_1 = (1, -1, -1)$, $v_2 = (1, 1, -1)$, or $v_3 = (1, -1, 1)$ and $x = |l_1|^2 / \sum_i |l_{1,i}|$ by Eq. E5 since $\max_i (l_1 \cdot v_i) = \sum_i |l_i|$. This also implies that $(a_{\min})_i = x\ \mathrm{sign}(l_{1,i})$. The optimal QFI is therefore $4Tx^2$ which is Eq. E2.

Let us now derive the optimal probability distribution $(p_i)_{i=1}^3$ for the initial state. By Condition (II) in Appendix B and using $(a_{\min})_i \propto \mathrm{sign}(l_{1,i})$, the necessary and sufficient conditions on the optimal distribution are the following:

$$\sum_{i=1}^{3} p_i\, \mathrm{sign}(l_{1,i}) = 0, \quad \sum_{i=1}^{3} p_i\, \mathrm{sign}(l_{1,i}) l_{2,i} = 0.$$

Since $l_1 \cdot \vec{1} = l_1 \cdot l_2 = 0$, then the distribution $p_i \propto |l_{1,i}|$ is thus optimal. This completes the proof of Eq. E2 for $d = 3$ and $N = 1$.

This derivation can be generalised for any $d = N + 2$. We introduce the following regularity or generalised non-degeneracy condition: The vectors $\{\vec{1}, l_2, \ldots, l_{N+1}\}$ are regular if any k-dimensional sub-vectors of them span $\mathbb{R}^k$. (For any set of indices $J = \{j_1, \ldots, j_k\}$, the sub-vectors of $\{\vec{1}, l_2, \ldots, l_{N+1}\}$ that correspond to these indices span $\mathbb{R}^k$.) This regularity condition does not always hold, e.g. it is not satisfied for commuting Pauli operators, but it does typically hold for uniformly randomly sampled commuting operators.

Given this regularity condition, the optimal QFI is given in Eq. E2 and the optimal state corresponds to a distribution $p_i \propto |l_{1,i}|$. The proof of this is similar to the $d = 3$ case. For example, Eq. E4 becomes:

$$l_1 = c_{1,\min}\vec{1} + \sum_{j=2}^{N+1} c_{j,\min} l_j + a_{\min}. \tag{E6}$$

The regularity condition guarantees that the optimal state is supported by all d eigenstates, which implies that $a_{\min}$ takes the form of $a_{\min} = xv$ where $v \in \{1, -1\}^d$ such that

$$x = \frac{|l_1|^2}{\max_v (l_1 \cdot v)} = \frac{|l_1|^2}{\sum_i |l_{1,i}|} = \frac{||l_1||_2^2}{||l_1||_1}. \tag{E7}$$

The derivation of the optimal input is then analogous to the $d = 3$ case.

We can now also derive the more general optimal QFI expression of Eq. E3. Given a solution of the SDP in Eq. E6, we can restrict ourselves to the support of the optimal state, i.e the eigenspace of the maximal eigenvalue of $\hat{A}^\dagger_{\min}\hat{A}_{\min}$. Eq. E6 then becomes

$$l_1^{(I)} = c_{1,\min}\vec{1} + \sum_{j=2}^{N+1} c_{j,\min} l_j^{(I)} + xv \tag{E8}$$

where $a^{(I)}_{\min} = xv$ and $v \in \{1, -1\}^{|I|}$. We now cannot use Eq. E7 to find x since $l_1^{(I)}$ is not necessarily orthogonal to $\text{span}\{\vec{1}, l_2^{(I)}, \ldots, l_{N+1}^{(I)}\}$. Instead, we can replace $l_1^{(I)}$ with $l_{1,\perp}^{(I)}$ as this will only modify the coefficients $\{c_i\}_{i=1}^{N+1}$ such that Eq. E8 becomes

$$l_{1,\perp}^{(I)} = c_1'\vec{1} + \sum_{j=2}^{N+1} c_j' l_j^{(I)} + xv.$$

By orthogonality, we then obtain that $x = ||l_{1,\perp}^{(I)}||_2^2/||l_{1,\perp}^{(I)}||_1$ which yields the optimal QFI expression. The proof for the optimal distribution is analogous to the previous case.

Appendix F: Proof of Example 1

We want to prove Example 1 concerning the QFI of the unextended measure-and-reset strategy for Pauli Lindblad estimation. We can assume that the signal operator is $\hat{L}_{1,(s)} = \prod_{j=1}^n \hat{\sigma}_x^{(j)}$ without loss of generality. Then, the noiseless QFI of $4T$ is recovered by preparing $|\uparrow\rangle^{\otimes n}$ since it is a maximal variance state of $\hat{L}_{1,(s)}$ and, for any noise operator $\hat{L}_{j,(n)}$ with Pauli weight less than n, there exists at least one qubit for which the state remains $|\uparrow\rangle$. Thus, the image of the initial state under the signal, $|\downarrow\rangle^{\otimes n}$, is orthogonal to the initial state and the image under each noise operator such that the noiseless QFI is achieved.

Appendix G: Impact of monitoring the environment

The Lindblad master equation in Eq. 1 is typically induced by interaction of the probe with an environment. In various cases, the environment can also be measured together with the probe. The QFI for this monitored environment scenario is an upper bound on the QFI for the case that we consider where only the probe is measured. This upper bound was calculated in Ref. [71] where it was shown that

$$\mathcal{I}_Q = 4\,\partial_{\sqrt{\gamma_1}}\partial_{\sqrt{\gamma_2}} \log(|\text{Tr}\,[\hat{\rho}(\gamma_1, \gamma_2, t)]|),$$

where $\hat{\rho}(\gamma_1, \gamma_2, t)$ is the solution of the modified master equation $\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = \mathcal{L}_{\gamma_1,\gamma_2}(\hat{\rho})$ where the modified Lindbladian $\mathcal{L}_{\gamma_1,\gamma_2}(\hat{\rho})$ is defined as

$$\mathcal{L}_{\gamma_1,\gamma_2}(\hat{\rho}) := \sum_{j=1}^{K}\left(\sqrt{\gamma_1\gamma_2}\hat{L}_{j,(s)}\hat{\rho}\hat{L}^\dagger_{j,(s)} - \frac{\gamma_1}{2}\hat{L}^\dagger_{j,(s)}\hat{L}_{j,(s)}\hat{\rho} - \frac{\gamma_2}{2}\hat{\rho}\hat{L}^\dagger_{j,(s)}\hat{L}_{j,(s)}\right) + \sum_{k=K+1}^{K+N}\gamma_k\left(\hat{L}_{k,(n)}\hat{\rho}\hat{L}^\dagger_{k,(n)} - \frac{1}{2}\{\hat{L}^\dagger_{k,(n)}\hat{L}_{k,(n)}, \hat{\rho}\}\right).$$

It was shown in Ref. [71] that in the limit of large t we can bound this QFI with an expression that depends only on the Lindbladian:

$$\mathcal{I}_Q = 4t \max_\lambda \left(\partial_{\sqrt{\gamma_1}}\partial_{\sqrt{\gamma_2}}\text{Re}\,[\lambda]\right)|_{\gamma_1=\gamma_2=\gamma}, \tag{G1}$$

where the maximisation is over the eigenvalues λ of the linear superoperator $\mathcal{L}_{\gamma_1,\gamma_2}$. The maximum is typically obtained with eigenvalues that correspond to steady states of the standard Lindbladian $\mathcal{L}_{\gamma,\gamma}$, i.e. eigenvalues whose real part vanishes when $\gamma_1 = \gamma_2$. While in some cases this bound in Eq. G1 coincides with our bound, it can be considerably larger than ours, due to the extra information retrieved from the environment. Consequently, if we restrict ourselves to only monitoring the probe, then this bound can be quite loose. We now consider some examples to demonstrate this.

Let us first consider some examples where the two bounds are the same. For any single Pauli jump operator, i.e. $\mathcal{L}_{\gamma_1,\gamma_2}(\hat{\rho}) = \sqrt{\gamma_1\gamma_2}\hat{\sigma}_j\hat{\rho}\hat{\sigma}_j - (\frac{\gamma_1}{2} + \frac{\gamma_2}{2})\hat{\rho}$, the bounds are the same. The relevant eigenvalue is $\lambda = \sqrt{\gamma_1\gamma_2} - \frac{\gamma_1}{2} - \frac{\gamma_2}{2}$ such that either bound provides that $\mathcal{I}_Q = 4t$. Similarly, for amplitude damping, the modified Lindbladian is $\mathcal{L}_{\gamma_1,\gamma_2}(\hat{\rho}) = \sqrt{\gamma_1\gamma_2}\hat{\sigma}_-\hat{\rho}\hat{\sigma}_+ - \frac{\gamma_1}{2}\hat{\sigma}_+\hat{\sigma}_-\hat{\rho} - \frac{\gamma_2}{2}\hat{\rho}\hat{\sigma}_+\hat{\sigma}_-$. The relevant eigenvalue is then $\lambda = \sqrt{\gamma_1\gamma_2}$ such that the QFI is $\mathcal{I}_Q = 4t$.

More generally, consider any Hermitian jump operator $\hat{L}$ such that $\mathcal{L}_{\gamma_1,\gamma_2}(\rho) = \sqrt{\gamma_1\gamma_2}\hat{L}\hat{\rho}\hat{L}^\dagger - \frac{\gamma_1}{2}\hat{L}^\dagger\hat{L}\hat{\rho} - \frac{\gamma_2}{2}\hat{\rho}\hat{L}^\dagger\hat{L}$. If $|\lambda\rangle$ is an eigenstate of $\hat{L}$ with eigenvalue λ, then $|\lambda\rangle\langle\lambda|$ is an eigenstate of $\mathcal{L}_{\gamma_1,\gamma_2}(\hat{\rho})$ with eigenvalue $\lambda_* := \lambda^2(\sqrt{\gamma_1\gamma_2} - \frac{\gamma_1}{2} - \frac{\gamma_2}{2})$. Therefore, the bound in Eq. G1 is $\mathcal{I}_Q = 4t\max_\lambda \lambda^2$, which is in general larger than our bound $\mathcal{I}_Q = (\lambda_{\max} - \lambda_{\min})^2 t$.

Similarly, let us consider two commuting Hermitian operators, $\hat{L}_{1,(s)}$ and $\hat{L}_{2,(n)}$. For any mutual eigenstate

$|\lambda_1,\lambda_2\rangle$ of $\hat{L}_{1,(s)}$ and $\hat{L}_{2,(n)}$, $|\lambda_1,\lambda_2\rangle\langle\lambda_1,\lambda_2|$ is an eigenstate of $\mathcal{L}_{\gamma_1,\gamma_2}(\hat{\rho})$ with eigenvalue λ_*, which is the same eigenvalue as the case without $\hat{L}_{2,(n)}$. Hence, this bound is completely unaffected by adding the noise $\hat{L}_{2,(n)}$, unlike our case. The reason for this difference is that the environment allows us to totally differentiate between $\hat{L}_{1,(s)}$ and $\hat{L}_{2,(n)}$ in cases where we cannot do this by measuring only the probe. These results, where we see metrological gain due to also detecting the environment, align with previous results that showed such improvement for Hamiltonian sensing [73–75].

While we have considered a few examples above, it remains to be established in general what the sufficient and necessary conditions are for a gap to exist between the cases with and without monitoring of the environment.

[1] J. W. Gardner, T. Gefen, S. A. Haine, J. J. Hope, J. Preskill, Y. Chen, and L. McCuller, Stochastic waveform estimation at the fundamental quantum limit, PRX Quantum **6**, 030311 (2025).

[2] J. Aasi, B. P. Abbott, R. Abbott, et al., Advanced LIGO, Class. Quantum Grav. **32**, 074001 (2015).

[3] J. D. Romano and Neil. J. Cornish, Detection methods for stochastic gravitational-wave backgrounds: A unified treatment, Living. Rev. Relativ. **20**, 2 (2017).

[4] A. I. Renzini, B. Goncharov, A. C. Jenkins, and P. M. Meyers, Stochastic gravitational-wave backgrounds: Current detection efforts and future prospects, Galaxies **10**, 34 (2022).

[5] E. P. Verlinde and K. M. Zurek, Observational Signatures of Quantum Gravity in Interferometers, Phys. Lett. B **822**, 136663 (2021).

[6] L. McCuller, Single-Photon Signal Sideband Detection for High-Power Michelson Interferometers (2022), arxiv:2211.04016 [hep-ex, physics:physics, physics:quant-ph].

[7] S. M. Vermeulen, T. Cullen, D. Grass, I. A. O. MacMillan, A. J. Ramirez, J. Wack, B. Korzh, V. S. H. Lee, K. M. Zurek, C. Stoughton, and L. McCuller, Photon Counting Interferome- try to Detect Geontropic Space-Time Fluctuations with GQuEST (2024), arxiv:2404.07524 [astro-ph, physics:gr-qc, physics:physics, physics:quant-ph].

[8] L. J. Rosenberg and K. A. Van Bibber, Searches for invisible axions, Phys. Rep. **325**, 1 (2000).

[9] P. W. Graham, I. G. Irastorza, S. K. Lamoreaux, A. Lindner, and K. A. van Bibber, Experimental searches for the axion and axion-like particles, Annu. Rev. Nucl. Part. Sci. **65**, 485 (2015).

[10] A. Agrawal, A. V. Dixit, T. Roy, S. Chakram, K. He, R. K. Naik, D. I. Schuster, and A. Chou, Stimulated Emission of Signal Photons from Dark Matter Waves, Phys. Rev. Lett. **132**, 140801 (2024).

[11] H. Shi and Q. Zhuang, Ultimate precision limit of noise sensing and dark matter search, npj Quantum Inf **9**, 1 (2023).

[12] S. L. Mouradian, N. Glikin, E. Megidish, K.-I. Ellers, and H. Haeffner, Quantum sensing of intermittent stochastic signals, Phys. Rev. A **103**, 032419 (2021).

[13] A. Rotem, T. Gefen, S. Oviedo-Casado, J. Prior, S. Schmitt, Y. Burak, L. McGuiness, F. Jelezko, and A. Retzker, Limits on spectral resolution measurements by quantum probes, Physical review letters **122**, 060503 (2019).

[14] T. Gefen, A. Rotem, and A. Retzker, Overcoming resolution limits with quantum sensing, Nat. Comm. **10**, 4992 (2019).

[15] D. Cohen, T. Gefen, L. Ortiz, and A. Retzker, Achieving the ultimate precision limit with a weakly interacting quantum probe, npj Quantum Information **6**, 83 (2020).

[16] L. M. Norris, G. A. Paz-Silva, and L. Viola, Qubit noise spectroscopy for non-gaussian dephasing environments, Physical review letters **116**, 150503 (2016).

[17] A. L. Shaw, R. Finkelstein, R. B.-S. Tsai, P. Scholl, T. H. Yoon, J. Choi, and M. Endres, Multi-ensemble metrology by programming local rotations with atom movements, Nature Physics **20**, 195 (2024).

[18] D. Stilck França, L. A. Markovich, V. Dobrovitski, A. H. Werner, and J. Borregaard, Efficient and robust estimation of many-qubit hamiltonians, Nature Communications **15**, 311 (2024).

[19] M. J. Arshad, C. Bekker, B. Haylock, K. Skrzypczak, D. White, B. Griffiths, J. Gore, G. W. Morley, P. Salter, J. Smith, et al., Real-time adaptive estimation of decoherence timescales for a single qubit, Physical Review Applied **21**, 024026 (2024).

[20] M. Tsang, Quantum noise spectroscopy as an incoherent imaging problem, Phys. Rev. A **107**, 012611 (2023).

[21] R. Alicki and K. Lendi, *Quantum dynamical semigroups and applications*, Vol. 717 (Springer Science & Business Media, 2007).

[22] H. Carmichael, *An open systems approach to quantum optics*, Vol. 18 (Springer Science & Business Media, 2009).

[23] D. Manzano, A short introduction to the lindblad master equation, Aip advances **10** (2020).

[24] F. Campaioli, J. H. Cole, and H. Hapuarachchi, Quantum master equations: Tips and tricks for quantum optics, quantum computing, and beyond, PRX Quantum **5**, 020202 (2024).

[25] V. V. Albert, Lindbladians with multiple steady states: theory and applications, arXiv preprint arXiv:1802.00010 (2018).

[26] R. Demkowicz-Dobrzański and L. Maccone, Using entanglement against noise in quantum metrology, Phys. Rev. Lett. **113**, 250801 (2014).

[27] P. Sekatski, M. Skotiniotis, J. Kołodyński, and W. Dür, Quantum metrology with full and fast quantum control, Quantum **1**, 27 (2017).

[28] S. Zhou and L. Jiang, Asymptotic theory of quantum channel estimation, PRX Quantum **2**, 010343 (2021).

[29] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Adaptive quantum metrology under general markovian noise, Physical Review X **7**, 041009 (2017).

[30] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Achieving the Heisenberg limit in quantum metrology using quan-

tum error correction, Nat. Commun. **9**, 78 (2018).
[31] K. Wan and R. Lasenby, Bounds on adaptive quantum metrology under markovian noise, Phys. Rev. Research **4**, 033092 (2022).
[32] P. Sekatski and M. Perarnau-Llobet, Optimal nonequilibrium thermometry in markovian environments, Quantum **6**, 869 (2022).
[33] S. Kurdziałek, W. Górecki, F. Albarelli, and R. Demkowicz-Dobrzański, Using adaptiveness and causal superpositions against noise in quantum metrology, Phys. Rev. Letters **131**, 090801 (2023).
[34] S. Zhou, A. G. Manes, and L. Jiang, Achieving metrological limits using ancilla-free quantum error-correcting codes, Phys. Rev. A **109**, 042406 (2024).
[35] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, 2009).
[36] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Quantum metrology with nonclassical states of atomic ensembles, Rev. Mod. Phys. **90**, 035005 (2018).
[37] W. Górecki, A. Riccardi, and L. Maccone, Quantum metrology of noisy spreading channels, Phys. Rev. Lett. **129**, 240503 (2022).
[38] S. A. Haine, Using interaction-based readouts to approach the ultimate limit of detection-noise robustness for quantum-enhanced metrology in collective spin systems, Phys. Rev. A **98**, 030303 (2018).
[39] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, The elusive heisenberg limit in quantum-enhanced metrology, Nature communications **3**, 1063 (2012).
[40] J. Kołodyński and R. Demkowicz-Dobrzański, Efficient tools for quantum metrology with uncorrelated noise, New J. Phys. **15**, 073043 (2013).
[41] We use the terms "Lindblad estimation" and "stochastic signal sensing" interchangeably. If the Born-Markov approximation does not hold, however, then the stochastic signal may not be described by a Lindbladian and a different approach is necessary.
[42] The operator norm $\|\cdot\|$ is induced by the ℓ_2 norm of the kets. For a given operator $\hat{A}$, $\|\hat{A}\|$ is also called the spectral norm and equals the largest singular value of $\hat{A}$.
[43] In the noiseless case, there are no other natural scales to compare to the size of the signal. Instead, we first derive the Kraus operators in Eq. 11 assuming that the short time condition in Eq. 12 holds which implies that $\gamma_1 t \ll 1$. Then, we take the limit of $\gamma_1 t \to 0$ of the Kraus operators in Eq. 11 to reach Eq. 13. This is what we mean by the vanishing signal limit in the noiseless case.
[44] N. Boulant, T. F. Havel, M. A. Pravia, and D. G. Cory, Robust method for estimating the lindblad operators of a dissipative quantum process from measurements of the density operator at multiple time points, Phys. Rev. A **67**, 042322 (2003).
[45] E. Ben Av, Y. Shapira, N. Akerman, and R. Ozeri, Direct reconstruction of the quantum-master-equation dynamics of a trapped-ion qubit, Phys. Rev. A **101**, 062305 (2020).
[46] Y.-X. Wang, J. Bringewatt, A. Seif, A. J. Brady, C. Oh, and A. V. Gorshkov, Exponential entanglement advantage in sensing correlated noise, arXiv preprint arXiv:2410.05878 (2024).
[47] A. J. Brady, Y.-X. Wang, V. V. Albert, A. V. Gorshkov, and Q. Zhuang, Correlated noise estimation with quantum sensor networks, arXiv preprint arXiv:2412.17903 (2024).
[48] W. Nagourney, J. Sandberg, and H. Dehmelt, Shelved optical electron amplifier: Observation of quantum jumps, Phys. Rev. Letters **56**, 2797 (1986).
[49] W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland, Quantum zeno effect, Phys. Rev. A **41**, 2295 (1990).
[50] M. Tsang, R. Nair, and X.-M. Lu, Quantum theory of superresolution for two incoherent optical point sources, Phys. Rev. X **6**, 031033 (2016).
[51] C. Oh, S. Zhou, Y. Wong, and L. Jiang, Quantum limits of superresolution in a noisy environment, Physical Review Letters **126**, 120502 (2021).
[52] H. Shi, A. J. Brady, W. Górecki, L. Maccone, R. Di Candia, and Q. Zhuang, Quantum-enhanced dark matter detection with in-cavity control: mitigating the rayleigh curse, arXiv preprint arXiv:2409.04656 (2024).
[53] S. A. Haine and S. S. Szigeti, Quantum metrology with mixed states: When recovering lost information is better than never losing it, Physical Review A **92**, 032317 (2015).
[54] Here, the complex phase φ_1 in $\hat{L}_{1,(s)}$ and $\hat{L}_{2,(n)}$ has been compensated by setting $\phi \mapsto \phi + \varphi_1/2$ such that we may assume that $\varphi_1 = 0$ without loss of generality.
[55] K. Macieszczak, M. Fraas, and R. Demkowicz-Dobrzański, Bayesian quantum frequency estimation in presence of collective dephasing, New Journal of Physics **16**, 113002 (2014).
[56] S. Direkci, R. Finkelstein, M. Endres, and T. Gefen, Heisenberg-limited bayesian phase estimation with low-depth digital quantum circuits, arXiv preprint arXiv:2407.06006 (2024).
[57] W.-K. Mok, A. Poddar, E. Sierra, C. C. Rusconi, J. Preskill, and A. Asenjo-Garcia, Universal scaling laws for correlated decay of many-body quantum systems, arXiv preprint arXiv:2406.00722 (2024).
[58] J. G. Bohnet, Z. Chen, J. M. Weiner, D. Meiser, M. J. Holland, and J. K. Thompson, A steady-state superradiant laser with less than one intracavity photon, Nature **484**, 78 (2012).
[59] G. Ferioli, A. Glicenstein, I. Ferrier-Barbut, and A. Browaeys, A non-equilibrium superradiant phase transition in free space, Nature Physics **19**, 1345 (2023).
[60] H. Pichler, T. Ramos, A. J. Daley, and P. Zoller, Quantum optics of chiral spin networks, Physical Review A **91**, 042116 (2015).
[61] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, Phys. Rev. A **64**, 012310 (2001).
[62] M. H. Michael, M. Silveri, R. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, New class of quantum error-correcting codes for a bosonic mode, Phys. Rev. X **6**, 031006 (2016).
[63] J. E. Kim and G. Carosi, Axions and the strong CP problem, Rev. Mod. Phys. **82**, 557 (2010).
[64] K. Choi, S. H. Im, and C. S. Shin, Recent progress in the physics of axions and axion-like particles, Annu. Rev. Nucl. Part. Sci. **71**, 225 (2021).
[65] R. Cameron, G. Cantatore, A. Melissinos, G. Ruoso, Y. Semertzidis, H. Halama, D. Lazarus, A. Prodell, F. Nezrick, C. Rizzo, et al., Search for nearly massless, weakly coupled particles by optical techniques, Phys. Rev. D **47**, 3707 (1993).

[66] N. Du, N. Force, R. Khatiwada, E. Lentz, R. Ottens, L. Rosenberg, G. Rybka, G. Carosi, N. Woollett, D. Bowring, et al., Search for invisible axion dark matter with the axion dark matter experiment, Phys. Rev. Lett. **120**, 151301 (2018).
[67] A. V. Dixit, S. Chakram, K. He, A. Agrawal, R. K. Naik, D. I. Schuster, and A. Chou, Searching for Dark Matter with a Superconducting Qubit, Phys. Rev. Lett. **126**, 141302 (2021).
[68] On these short timescales, whether the signal is entirely stochastic or instead needs to be modelled as a combination of a coherent part and an incoherent part is a topic of ongoing research [1, 52].
[69] X. Deng, S. Li, Z.-J. Chen, Z. Ni, Y. Cai, J. Mai, L. Zhang, P. Zheng, H. Yu, C.-L. Zou, et al., Quantum-enhanced metrology with large fock states, Nat. Phys. , 1 (2024).
[70] A. Eickbusch, V. Sivak, A. Z. Ding, S. S. Elder, S. R. Jha, J. Venkatraman, B. Royer, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, Fast universal control of an oscillator with weak dispersive coupling to a qubit, Nature Physics **18**, 1464 (2022).
[71] S. Gammelmark and K. Mølmer, Fisher information and the quantum cramér-rao sensitivity limit of continuous measurements, Physical review letters **112**, 170401 (2014).
[72] D. Yang, S. F. Huelga, and M. B. Plenio, Efficient information retrieval for sensing via continuous measurement, Physical Review X **13**, 031012 (2023).
[73] T. Gefen, D. A. Herrera-Martí, and A. Retzker, Parameter estimation with efficient photodetectors, Physical Review A **93**, 032133 (2016).
[74] M. B. Plenio and S. F. Huelga, Sensing in the presence of an observed environment, Physical Review A **93**, 032123 (2016).
[75] F. Albarelli, M. A. Rossi, D. Tamascelli, and M. G. Genoni, Restoring heisenberg scaling in noisy quantum metrology by monitoring the environment, Quantum **2**, 110 (2018).
[76] J. v. Neumann, Zur theorie der gesellschaftsspiele, Mathematische annalen **100**, 295 (1928).
[77] slipperyslider, J. W. Gardner. *slipperySlider*. 2024. https://git.ligo.org/jameswalter.gardner/slipperyslider.

Chapter 6

Bayesian waveform estimation at the fundamental quantum limit

The clouds methought would open and show riches
Ready to drop upon me, that when I waked,
I cried to dream again.

(Caliban, Act 3 Scene 2)
William Shakespeare, The Tempest

We now present Ref. [52] where we return to deterministic waveform sensing and restrict ourselves to just optimising over measurements, much like in Chapter 3, but now consider the Bayesian regime, leaving the asymptotic Fisher limit behind. The Bayesian regime presents a much higher computational challenge but removes the sometimes troublesome assumption of many measurements to deliver the exact answer for optimal sensing when limited in time and frequency resolution. This opens a different dimension of the landscape of waveform estimation as shown in Fig. 1.1. We came to this problem of optimal measurement historically because of two recent exciting works: (1) a head-to-head comparison of number and quadrature measurements in the Bayesian regime [57] and (2) a proposal to run Grover's algorithm on a quantum computer to search for a signal at an unknown frequency [55]. Here, we reconcile these two works by showing that the optimal measurement to search for an unknown frequency is to project onto approximately the "whitened" waveform basis. This outperforms either number or quadrature measurements and does not necessarily require a quantum computer to implement. Although we do not assume access to different state preparation or quantum control, unlike the previous two

papers, the proposed measurement scheme still is only a far-future possibility and so in the broader arc of the thesis this paper still deserves to be at the more ambitious end.

We now list the important steps and results from this paper. Firstly, our main result in Eq.29 of Sec.II.C.2 is the fundamental Bayesian limit for a Toeplitz family of states if we have no prior information about the signal. This is more general than the frequency estimation of the coherent state shown in Fig.1 which we handle in Sec.V.C. Secondly, much of the behaviour can be understood purely from the geometry/inner product of the singular-particle and coherent states in question which we determine in Eq.40 of Sec.III and Eq.43 of Sec.IV, respectively. In particular, the relation between the classical waveforms and the quantum states. Finally, Sec.V.C.2 and Fig.2 show the numerical results which are independent of the proofs of the rest of the paper. The numerical approach in Appendix H may prove useful for other applications.

Bayesian frequency estimation at the fundamental quantum limit

James W. Gardner [1,2,*] Tuvia Gefen [3] Ethan Payne [2,4] Su Direkci [2] Sander M. Vermeulen [2] Simon A. Haine [5] Joseph J. Hope [5] Lee McCuller [2] and Yanbei Chen [2]

[1] *OzGrav-ANU, Centre for Gravitational Astrophysics, Research Schools of Physics, and of Astronomy and Astrophysics, The Australian National University, Canberra, ACT 2601, Australia*
[2] *Division of Physics, Mathematics and Astronomy, California Institute of Technology, Pasadena, California 91125, USA*
[3] *Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem 91904, Givat Ram, Israel*
[4] *LIGO Laboratory, California Institute of Technology, Pasadena, California 91125, USA*
[5] *Department of Quantum Science and Technology and Department of Fundamental and Theoretical Physics, Research School of Physics, The Australian National University, Canberra, ACT 0200, Australia*

(Dated: July 29, 2025)

Searching for a weak signal at an unknown frequency is a canonical task in experiments probing fundamental physics such as gravitational-wave observatories and ultra-light dark matter haloscopes. These state-of-the-art sensors are limited by quantum noise arising from the fundamental uncertainty about the state of the device. Classically, frequency estimation suffers from a threshold effect in the signal-to-noise ratio such that weak signals are extremely hard to localise in frequency. We show that this phenomenon persists at the fundamental quantum limit but that the classical approach, a quadrature measurement, can nevertheless be beaten by a coherent protocol of projecting onto the "quantum whitened" possible quantum states. Quantum whitening is a covariant measurement, and we examine it analytically in the wide-prior limit and numerically for finite-width priors. Beyond accelerating searches for unknown frequencies, quantum whitening may be used generally to sense the parameter of a unitary encoding given no prior information about the parameter.

Many promising searches for fundamental physics probe for weak signals at unknown frequencies encoded in the displacement of a bosonic mode. For example, searches using ground-based gravitational-wave observatories for the post-merger remnant of binary neutron star mergers could test our theories of exotic matter at supranuclear densities [1–3] and searches for axions using table-top haloscopes could directly detect dark matter and resolve the strong $\mathcal{CP}$ problem [4–7]. However, the fundamental Bayesian quantum limits of the sensitivity of these searches are not well understood. Here, a Bayesian analysis is necessary because we assume that only a single event is seen and that the prior on the possible frequencies of the signal is wide. For example, the range of possible axion masses spans five orders of magnitude [8] at the ultra-light end of the vast 90 orders of magnitude of possible masses for dark matter more generally [9]. Understanding how to achieve the fundamental quantum limit of these devices using unconventional measurement protocols may lead to improved sensitivity and thus accelerate these searches for fundamental physics.

Let us now define a canonical frequency estimation problem. We want to sense the following sinusoidal signal

$$h_\omega(t) = A\cos(\omega t + \phi) \tag{1}$$

where the amplitude A (with units of square root Hertz) and phase ϕ are known, but the frequency ω is unknown. We assume a flat prior $\pi(\omega)$ on the possible frequencies within the interval $(\omega_0 - \Delta\omega/2, \omega_0 + \Delta\omega/2)$ with centre frequency ω_0 and a wide prior width $\Delta\omega$. Suppose that the sinusoidal signal interacts with a linear device from time $t = 0$ to $t = T$, where T is the integration time. If the bosonic bath of incoming modes to the sensor is in the multi-mode vacuum state $|0\rangle$, then the coherent quantum state of the bath of outgoing modes is as follows

$$|h_\omega\rangle = e^{i\int_0^T \mathrm{d}t\, h_\omega(t)\hat{x}(t)}|0\rangle, \tag{2}$$

where the amplitude quadrature of the bath is $\hat{x}(t) = [\hat{a}(t) + \hat{a}^\dagger(t)]/\sqrt{2}$ given the bath annihilation and creation operators $\hat{a}(t)$ and $\hat{a}^\dagger(t)$, respectively, which obey the commutator $[\hat{a}(t), \hat{a}^\dagger(t')] = \delta(t - t')$. Here, we set $\hbar = 1$ and have absorbed the convolution with the linear susceptibility of the sensor into the signal [10, 11]. This coherent state consists of the vacuum displaced along the phase quadrature of the bath $\hat{p}(t) = [-i\hat{a}(t) + i\hat{a}^\dagger(t)]/\sqrt{2}$ by the signal as shown in Fig. 1. We assume here that the direction of the displacement is known and fixed, because we want to model ground-based gravitational-wave observatories consisting of tuned resonant optical cavities and a Michelson interferometer held at dark port such that the signal only appears in the phase quadrature [12].

We want to determine the optimal measurement of the final state $|h_\omega\rangle$ in Eq. 2 to best constrain the posterior distribution on the possible signal frequencies. Ref. [1] recently showed an advantage in frequency estimation by performing a frequency-domain number-resolving measurement $\hat{n}(\Omega)$ over a time-domain quadrature measurement $\hat{p}(t)$. In this paper, we go beyond this head-to-head comparison and find the fundamental quantum limit of frequency estimation and how to achieve it. First, in

* james.gardner@anu.edu.au

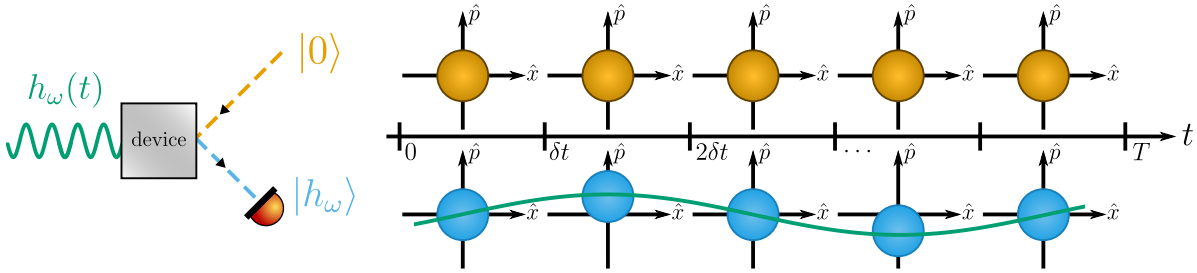

FIG. 1. A sinusoidal signal $h_\omega(t)$ in time t displaces a bath of vacuum modes $|0\rangle$ to an oscillating coherent state $|h_\omega\rangle$ via a linear quantum device. The ball-and-stick diagrams on the right idealise the Wigner function of the bath discretised in time in Appendix E 2. The signal is windowed to the bounded interval $(0, T)$ which is discretised into $M = T/\delta t$ time-bins of width δt. While here we choose $M = 5$ for the purposes of illustration, we take the continuum limit of $\delta t \to 0$ with T fixed in the text.

Sec. I, we review the techniques of optimal quantum sensing. Second, in Sec. II, we calculate the general fundamental limit and optimal measurement. Third, in Sec. III and Sec. IV, we review the geometry of the single-particle and coherent states, respectively, to establish a waveform basis of quantum states. Finally, we apply the results to different cases of single-particle and coherent states in Sec. V and draw conclusions in Sec. VI.

I. BACKGROUND

Let us review quantum parameter estimation, with Bayesian and Fisher information approaches, and classical frequency estimation.

A. Review of Bayesian quantum estimation

Let us consider Bayesian quantum estimation of a single generic parameter θ (see, e.g., Refs. [13–18]). We suppose that we have a prior distribution $\pi(\theta)$ on the parameter θ that is encoded in a quantum state $\hat{\rho}(\theta)$. We then perform some measurement described by a positive operator-valued measure (POVM) with a set of effects $\{\hat{E}_x\}_x$ such that $E_x \geq 0$ and $\int \mathrm{d}x\ E_x = 1$. We obtain the measurement outcome x given θ with likelihood $L(x|\theta) = \mathrm{Tr}[\hat{\rho}(\theta)\hat{E}_x]$. Marginalising over θ, then the total probability of obtaining x is given by the evidence $p(x) = \int \mathrm{d}\theta\ \pi(\theta)L(x|\theta)$. By Bayes's rule, the posterior on θ given x is then given by

$$p(\theta|x) = \frac{\pi(\theta)L(x|\theta)}{p(x)}. \tag{3}$$

From the posterior given x, we form some estimator $\check{\theta}_x$ of θ, which can be biased. The mean-square error (MSE) of this estimator at a given θ is

$$\mathrm{MSE}(\theta) = \int \mathrm{d}x\ L(x|\theta)(\check{\theta}_x - \theta)^2. \tag{4}$$

We wish to minimise the Bayesian MSE (BMSE) defined as the MSE averaged against the prior or, equivalently by Eq. 3, the posterior variance averaged against the evidence:

$$\mathrm{BMSE} = \int \mathrm{d}\theta\ \pi(\theta)\ \mathrm{MSE}(\theta) = \int \mathrm{d}x\ p(x)\ V_{\mathrm{post}}(x) \tag{5}$$

where the variance of the posterior given x is

$$V_{\mathrm{post}}(x) = \int \mathrm{d}\theta\ p(\theta|x)(\check{\theta}_x - \theta)^2. \tag{6}$$

Here, the optimal estimator $\check{\theta}_x$ to minimise the BMSE for a given measurement scheme is the mean of the posterior distribution. For a given quantum state $\hat{\rho}(\theta)$, the minimum BMSE (MBMSE) over all possible measurements is given as follows [13]

$$\Delta^2\theta = \Delta^2\theta^{(0)} - \mathrm{Tr}[\hat{\bar{\rho}}\hat{L}^2] \tag{7}$$

where $\Delta^2\theta^{(0)} = \int \mathrm{d}\theta\ \pi(\theta)\,\theta^2$ is the variance of the prior $\pi(\theta)$. The optimal measurement basis $|x\rangle$ and estimator $\check{\theta}_x$ is captured by the Bayesian symmetric logarithmic derivative (BSLD) defined as

$$\hat{L} = \int \mathrm{d}x\ \check{\theta}_x|x\rangle\langle x| \tag{8}$$

which is the solution to the following Lyapunov equation

$$\hat{\bar{\rho}}' = \frac{1}{2}\{\hat{\bar{\rho}}, \hat{L}\} \tag{9}$$

where $\{\cdot,\cdot\}$ is the anticommutator and the mixed state is

$$\hat{\bar{\rho}} = \int \mathrm{d}\theta\ \pi(\theta)\hat{\rho}(\theta), \quad \hat{\bar{\rho}}' = \int \mathrm{d}\theta\ \pi(\theta)\hat{\rho}(\theta)\,\theta. \tag{10}$$

Here, $\hat{\bar{\rho}}$ captures the classical uncertainty in the final state $\hat{\rho}(\theta)$ given the prior, and $\hat{\bar{\rho}}'$ which we call the "Bayesian derivative" acts informally like the average derivative of $\hat{\bar{\rho}}$ with respect to the parameter θ. Indeed, in the case of a Gaussian prior then $\hat{\bar{\rho}}' = \Delta^2\theta^{(0)} \int \mathrm{d}\theta\ \pi(\theta)\partial_\theta\hat{\rho}(\theta)$ [13]. For a given prior $\pi(\theta)$, the Lyapunov equation in Eq. 9 can then be solved given the spectral decomposition of $\hat{\bar{\rho}} = \sum_j \lambda_j|\phi_j\rangle\langle\phi_j|$ as

$$\hat{L} = \sum_{j,k} \frac{2}{\lambda_j + \lambda_k}\langle\phi_j|\hat{\bar{\rho}}'|\phi_k\rangle|\phi_j\rangle\langle\phi_k| \tag{11}$$

Method	Bayesian	Fisher
Figure-of-merit	BMSE (Eq. 5)	MSE (Eq. 4)
Fundamental limit	MBMSE (Eq. 7)	QCRB (Eq. 15)
Optimal measurement	BSLD (Eq. 11)	SLD (Eq. 17)
Prior	Any prior	No prior
Estimator	Biased	Unbiased
Valid	Everywhere	Asymptotically
Computation	Harder	Easier

TABLE I. Comparison of Bayesian and Fisher approaches to quantum parameter estimation.

where the sum excludes j,k such that $\lambda_j + \lambda_k = 0$. We can then calculate the MBMSE in Eq. 7 as

$$\Delta^2\theta = \Delta^2\theta^{(0)} - \sum_{j,k} \frac{4\lambda_j}{(\lambda_j+\lambda_k)^2} \left|\langle\phi_j|\hat{\tilde{\rho}}'|\phi_k\rangle\right|^2. \tag{12}$$

We assume in Eqs. 7–10 that the prior is centred such that $\int \mathrm{d}\theta\ \pi(\theta)\theta = 0$. For example, for sensing a frequency within $\omega \in (\omega_0 - \Delta\omega/2, \omega_0 + \Delta\omega/2)$ we formally estimate $\omega - \omega_0$ instead of ω. However, we will ignore this technicality throughout the rest of the paper and just refer to estimating the frequency ω.

B. Review of Fisher information

Let us discuss the asymptotic limit of Bayesian estimation where we measure $M \gg 1$ independent and identical copies of the quantum state. In this frequentist regime, the prior width is negligibly narrow and we instead study local estimation of infinitesimal deviations about a given θ. Formally, this connection between the Bayesian and frequentist approaches is given by the Bernstein-von Mises theorem.

We now review the Fisher information approach for estimating a single parameter θ. In our motivating problem, although this parameter determines a classical waveform in Eq. 1 which is encoded in a quantum state in Eq. 2, this should be distinguished from true waveform estimation where there are a continuum of parameters to estimate (see, e.g., Refs. [19–21]). The residual MSE in Eq. 4 at θ for a given measurement obeys the classical Cramér-Rao bound (CCRB)

$$\mathrm{MSE}(\theta) \geq [M\mathcal{I}_C(\theta)]^{-1}, \tag{13}$$

where the classical Fisher information (CFI) is given by

$$\mathcal{I}_C(\theta) = \sum_x \frac{[\partial_\theta L(x|\theta)]^2}{L(x|\theta)}. \tag{14}$$

The CFIs from independent and identical observations sum, which contributes the factor of M in Eq. 13. Lifting from a given measurement to the quantum state, the quantum Cramér-Rao bound (QCRB) is

$$\mathrm{MSE}(\theta) \geq [M\mathcal{I}_Q(\theta)]^{-1}, \tag{15}$$

where the maximum CFI over all possible measurements is the quantum Fisher information (QFI) which is given similarly to Eq. 12 by

$$\mathcal{I}_Q(\theta) = \sum_{j,k} \frac{4\lambda_j}{(\lambda_j+\lambda_k)^2} \left|\langle\phi_j|\partial_\theta\hat{\rho}(\theta)|\phi_k\rangle\right|^2 \tag{16}$$

where the spectral decomposition of the state is $\hat{\rho}(\theta) = \sum_j \lambda_j|\phi_j\rangle\langle\phi_j|$ and the sum excludes j,k such that $\lambda_j + \lambda_k = 0$. Similarly to how Eq. 7 relates the MBMSE to the BSLD, the QFI is $\mathcal{I}_Q(\theta) = \mathrm{Tr}[\hat{\rho}\hat{S}^2]$ in terms of $\hat{S}$, the symmetric logarithmic derivative (SLD). The SLD solves the Lyapunov equation $\partial_\theta\hat{\rho}(\theta) = \frac{1}{2}\{\hat{\rho}(\theta), \hat{S}\}$ similarly to Eq. 9. The solution is similar to Eq. 11 and is given by

$$\hat{S} = \sum_{j,k} \frac{2}{\lambda_j+\lambda_k} \langle\phi_j|\partial_\theta\hat{\rho}(\theta)|\phi_k\rangle|\phi_j\rangle\langle\phi_k| \tag{17}$$

where we exclude $\lambda_j + \lambda_k = 0$ as usual. The expectation value $\mathrm{Tr}[\hat{\rho}\hat{S}^2]$ thus recovers the QFI in Eq. 16. The equality between the maximum CFI and the QFI only holds in the single-parameter case, which we focus on, as the multi-parameter case requires handling non-commutativity [21, 22].

We compare the Bayesian and Fisher methods in Table I. Although the Fisher information describes the asymptotic scaling of the MSE, questions remain about how fast a given optimal measurement reaches the asymptotic limit such that we need to maximise the domain of unambiguous estimation along with the point estimate sensitivity [23]. I.e., two different measurements can both be optimal asymptotically (their CFIs attain the QFI) but the MSE of one can converge to the quantum limit much faster than the other. This behaviour is already accounted for implicitly in the Bayesian method, however, such that we need not alter our approach.

The QFI is also related to the infinitesimal expansion of the fidelity between quantum states as follows

$$\begin{aligned}\mathcal{F}(\theta) &= \mathrm{Tr}\left[\sqrt{\sqrt{\hat{\rho}(\theta)}\hat{\rho}(\theta+\delta\theta)\sqrt{\hat{\rho}(\theta)}}\right]^2 \\ &= 1 - \frac{1}{4}\delta\theta^2\mathcal{I}_Q(\theta) + \mathcal{O}(\delta\theta^3).\end{aligned} \tag{18}$$

Here, we assume that the rank of $\hat{\rho}(\theta)$ is constant in θ, e.g. if the state is pure $\hat{\rho}(\theta) = |h_\theta\rangle\langle h_\theta|$ such that the fidelity is $\mathcal{F}(\theta) = |\langle h_\theta|h_{\theta+\delta\theta}\rangle|^2$.

For the Gaussian case, the Fisher approach simplifies considerably. The CFI in Eq. 14 for a multi-variate normal distribution of n variables with n-by-1 parameter-dependent mean vector $\vec{\mu}_n(\theta)$ and n-by-n parameter-independent covariance matrix Σ_n is equal to

$$\mathcal{I}_C(\theta) = \partial_\theta\vec{\mu}_n(\theta)^{\mathrm{T}}\Sigma_n^{-1}\partial_\theta\vec{\mu}_n(\theta). \tag{19}$$

Similarly, the QFI in Eq. 16 for an n-mode Gaussian state with $2n$-by-1 parameter-dependent mean vector $\vec{\mu}_{2n}(\theta)$, since there are two quadratures for each mode, and $2n$-by-$2n$ parameter-independent covariance matrix Σ_{2n} is

$$\mathcal{I}_Q(\theta) = \partial_\theta\vec{\mu}_{2n}(\theta)^{\mathrm{T}}\Sigma_{2n}^{-1}\partial_\theta\vec{\mu}_{2n}(\theta) \tag{20}$$

which is achieved by measuring the quadrature along $\Sigma^{-1}\partial_\theta \vec{\mu}_{2n}(\theta)$ [24]. The Gaussian shift QFI is thus saturated by a Gaussian measurement.

We can also incorporate a priori information about θ from a prior distribution $\pi(\theta)$, thus treating the parameter as a random variable. In the case that the MSE and QFI are independent of the parameter, then van Trees's Bayesian QCRB states that [25–27]

$$\text{MSE} \geq [M(\mathcal{I}_Q + \mathcal{I}_{\text{prior}})]^{-1} \tag{21}$$

where the Fisher information of the prior is

$$\mathcal{I}_{\text{prior}} = \int \mathrm{d}\theta \, \frac{\pi'(\theta)^2}{\pi(\theta)}. \tag{22}$$

Here, the integral should be restricted to the support of the prior. For example, if the prior is a Gaussian with variance σ^2, then the prior Fisher information is $\mathcal{I}_{\text{prior}} = 1/\sigma^2$. This bound is not tight except asymptotically, however, and the full Bayesian approach is required for true incorporation of the prior.

C. Review of classical frequency estimation

Let us briefly discuss the classical problem of estimating the frequency of a weak sinusoidal signal hidden in white noise. This is the scenario for our motivating problem in Eq. 2 if we perform a time-domain quadrature measurement $\hat{p}(t)$, since the quantum shot noise of the vacuum is white and the expectation values follow the signal as $\langle \hat{p}(t) \rangle = h(t)$. The noise here comes entirely from the quantum state entering the linear device shown in Fig. 1 and the classical signal is deterministic, i.e. $h(t)$ is not a continuous random variable (in contrast to Ref. [28, 29]). It is well-known for this classical problem that the CCRB in Eq. 13 is not a tight bound below some signal-to-noise ratio (SNR) threshold [30–32]. Here, the SNR is $A\sqrt{2T}$ in amplitude units where the factor of two comes from the quantum noise of the vacuum state of a harmonic oscillator which equals $\text{Var}[\hat{p}] = \frac{1}{2}$ given $[\hat{x}, \hat{p}] = i$. Below the SNR threshold, we need to instead use a different classical bound such as the BMSE. For example, later in Fig. 3, we will see that the BMSE for the time-domain quadrature measurement $\hat{p}(t)$ does not attain the CCRB except for above $A\sqrt{2T} \sim 10$. Instead of the BMSE, we could use the Barankin bound but this requires an unbiasedness assumption which does not hold as generally as the Bayesian approach [33].

The harmonic frequencies appearing in the Fourier series on the interval $(0, T)$ are $\omega_n = 2\pi n/T$ for $n \in \mathbb{Z}$. One might then expect the classical frequency resolution to scale as $1/T$ in amplitude units, but the CCRB actually decreases as $T^{-3/2}$ [31, 34, 35]. This can be shown from Eq. 19. The CFI acquires a factor of T^2 from the square of the derivative $-At\sin(\omega t + \phi)$ of the signal in Eq. 1 and obtains the third factor of T from the integral over time.

We now want to investigate what happens for the case of quantum frequency estimation above and below the SNR threshold, and whether other measurement schemes than quadrature measurement can improve the sensitivity.

II. QUANTUM WHITENING

Before we return to our motivating problem of frequency estimation of the coherent state in Eq. 2, we will first solve a more general problem. We will consider general families of pure states $\{|\psi_\theta\rangle\}_{\theta\in\mathbb{R}}$, where θ is the parameter to be estimated. We will focus on the case where this family of states has a Toeplitz symmetry, i.e. the inner product $\langle\psi_\theta|\psi_{\theta'}\rangle$ is a function of only $\theta - \theta'$. This will be useful to our motivating problem, since the geometry of the coherent states has a Toeplitz symmetry in a particular limit which means that we want to understand Toeplitz families of states. Thus, in this section, we will first review the covariant measurement of an additive unitary. Second, we will show that a Toeplitz family of states is always generated by an additive unitary, and we will interpret the covariant measurement of a continuous Toeplitz family as projecting on the "quantum whitened" possible states. Finally, we will calculate the fundamental quantum limit, the MBMSE, for a Toeplitz family when we have no prior information about the parameter.

A. Additive unitary case

Let us first review the case where the parameter is encoded by an additive unitary. Let the final state be $|\psi_\theta\rangle = e^{-i\hat{H}\theta}|\psi_0\rangle$ for some Hamiltonian $\hat{H}$ and initial state $|\psi_0\rangle$. The unitary $\hat{U}(\theta) = e^{-i\hat{H}\theta}$ is additive since $\hat{U}(\theta)\hat{U}(\theta') = \hat{U}(\theta + \theta')$ and $\hat{U}^\dagger(\theta) = \hat{U}(-\theta)$. The Hamiltonian $\hat{H}$ is therefore the generator of the unitary group $\{\hat{U}(\theta)\}_{\theta\in\mathbb{R}}$, which is the relevant symmetry transformation of this problem. The inner product between quantum states is thus $\langle\psi_\theta|\psi_{\theta'}\rangle = \langle\psi_0|e^{i\hat{H}(\theta-\theta')}|\psi_0\rangle$. This inner product is Toeplitz, i.e. it is only a function of $\theta - \theta'$. As such, we refer to $\{|\psi_\theta\rangle\}_{\theta\in\mathbb{R}}$ as a Toeplitz family of states.

Since we are ultimately interested in frequency estimation of the coherent states, we consider the case where the Hilbert space is infinite-dimensional and the Hamiltonian $\hat{H}$ has a continuous spectrum. (We discuss the finite and discrete cases in Appendix A.) Let the spectral decomposition be $\hat{H} = \int_{-\infty}^{\infty} \mathrm{d}k \; k|k\rangle\langle k|$ for some continuous basis satisfying $\langle k|l\rangle = \delta(k-l)\chi_S(k)$ where $\chi_S(k)$ is the indicator function for some set $S \subset \mathbb{R}$ to be determined. The eigenbasis of the Hamiltonian is thus $\{|k\rangle\}_{k\in S}$. We will show later that S is the support of the spectral measure, i.e. the Fourier transform of the Toeplitz inner product between quantum states above. Since $\langle k|k\rangle = 0$ for $k \notin S$

such that $|k\rangle = 0$, then we may write the Hamiltonian as

$$\hat{H} = \int_S \mathrm{d}k\ k|k\rangle\langle k|. \tag{23}$$

In terms of the spectral decomposition of $\hat{H}$, the unitary is $\hat{U}(\theta) = \int_S \mathrm{d}k\ e^{-ik\theta}|k\rangle\langle k|$. This implies that the mapping, $\theta \mapsto \hat{U}(\theta)$ is injective, i.e. for every $\theta_1 \neq \theta_2$ in $\mathbb{R}$, we have that $\hat{U}(\theta_1) \neq \hat{U}(\theta_2)$. The initial state is $|\psi_0\rangle = \int_S \mathrm{d}k\ \psi_0(k)|k\rangle$ for some complex wavefunction $\psi_0(k)$ supported on S. The initial state's probability distribution in k is thus $|\psi_0(k)|^2$, and the final state is $|\psi_\theta\rangle = \int_S \mathrm{d}k\ e^{-ik\theta}\psi_0(k)|k\rangle$. The inner product between final states, $\langle\psi_\theta|\psi_{\theta'}\rangle = \int_S \mathrm{d}k\ e^{ik(\theta-\theta')}\,|\psi_0(k)|^2$, equals the value at $\theta - \theta'$ of the Fourier transform of the initial state's probability distribution in k.

1. *Covariant measurement*

What is the optimal measurement of the final state to estimate the parameter θ? Since the inner product is not necessarily a delta function, we cannot estimate the parameter perfectly but some measurement bases may still outperform others. If the figure-of-merit is covariant with respect to the symmetry group $\{\hat{U}(\theta)\}_{\theta\in\mathbb{R}}$, i.e. if the prior $\pi(\theta)$ is flat, then it is known that the optimal measurement is also covariant with respect to this symmetry [36–38]. We discuss covariant measurements more in Appendix A 2, and we will return later to the assumption that the prior is flat.

We thus want to find the covariant states, $\{|\theta\rangle\}_{\theta\in\mathbb{R}}$, that satisfy $\hat{U}(\theta)|\theta'\rangle = |\theta + \theta'\rangle$. Let the states be $|\theta\rangle = \int_S \mathrm{d}k\ c(\theta,k)|k\rangle$ for some $c(\theta,k)$ such that covariance implies that $c(\theta,k) = e^{-ik\theta}c(0,k)$. The inner product between states is then $\langle\theta|\theta'\rangle = \int_S \mathrm{d}k\ \ e^{i(\theta-\theta')k}\,|c(0,k)|^2$ which equals $\delta(\theta-\theta')$ such that the states are orthogonal if $S = \mathbb{R}$ and $|c(0,k)| = 1/\sqrt{2\pi}$. The covariant states are thus unique up to the complex phase of $c(0,k)$ which is a gauge freedom described by the unitary $\hat{V} = \int_S \mathrm{d}k\ e^{i\vartheta_k}|k\rangle\langle k|$ for a given set of a phases $\{\vartheta_k\}_{k\in S}$. Since this gauge unitary commutes with the group action, $[\hat{U}(\theta), \hat{V}] = 0$, it can be ignored without loss of generality. Therefore, we define the covariant states as a Fourier transform of the eigenbasis of the Hamiltonian as follows

$$|\theta\rangle = \frac{1}{\sqrt{2\pi}}\int_S \mathrm{d}k\ e^{-i\theta k}|k\rangle \tag{24}$$

where the inverse Fourier transformation is $|k\rangle = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}\theta\ e^{i\theta k}|\theta\rangle$. The inner product between covariant states is thus

$$\langle\theta|\theta'\rangle = \frac{1}{2\pi}\int_S \mathrm{d}k\ e^{i(\theta-\theta')k}. \tag{25}$$

It is optimal to measure the covariant states provided that the prior is flat, which we will revisit shortly. If $S = \mathbb{R}$ such that the covariant states are orthogonal ($\langle\theta|\theta'\rangle = \delta(\theta-\theta')$), then we project onto them by measuring the covariant operator, $\hat{W} = \int_{-\infty}^{\infty}\mathrm{d}\theta\ \theta|\theta\rangle\langle\theta|$. This is the canonical conjugate of the Hamiltonian, $[\hat{W}, \hat{H}] = i$, such that $\hat{U}^\dagger(\theta)\hat{W}\hat{U}(\theta) = \hat{W} + \theta$. This is similar to the canonical quadratures which obey $[\hat{x}, \hat{p}] = i$ and $e^{i\theta\hat{p}}\hat{x}e^{-i\theta\hat{p}} = \hat{x} + \theta$ and whose eigenstates are related by the Fourier transform.

If instead $S \subsetneq \mathbb{R}$ such that the covariant states are not orthogonal, then they form a covariant non-projective POVM on the subspace spanned by the eigenbasis of the Hamiltonian. The POVM is $\{\hat{E}_\theta\}_{\theta\in(-\pi,\pi]}$ where the effects are $\hat{E}_\theta = |\theta\rangle\langle\theta|$. This POVM is normalised although the covariant states may not be. In this case, we can still form the covariant operator, $\hat{W} = \int_{-\infty}^{\infty}\mathrm{d}\theta\ \theta|\theta\rangle\langle\theta|$, but now the covariant states are not its eigenbasis since they are not orthogonal. Ref. [13] showed that replacing the non-projective POVM $\{\hat{E}_\theta\}_{\theta\in(-\pi,\pi]}$ with projection onto the orthogonal eigenbasis of $\hat{W}$ cannot increase the BMSE. Since we already know that the covariant POVM is optimal, the projective measurement onto the eigenbasis of $\hat{W}$ is also optimal (its BMSE equals the MBMSE). This means that $\hat{W}$ equals the BSLD $\hat{L}$ in Eq. 8. Conversely, to show that a given POVM is optimal in general, it suffices to form the operator like $\hat{W}$ (which can also be interpreted as projection onto its eigenbasis) and show that it equals the BSLD.

From the initial family of non-orthogonal quantum states $\{|\psi_\theta\rangle\}_{\theta\in\mathbb{R}}$ related by an additive unitary, we have defined two continuous families of states: the stationary states $\{|k\rangle\}_k$ and covariant states $\{|\theta\rangle\}_{\theta\in\mathbb{R}}$. These span the same subspace of possible quantum states spanned by $\{|\psi_\theta\rangle\}_{\theta\in\mathbb{R}}$. We may restrict ourselves to this subspace without loss of generality, which can significantly reduce the dimensionality of the Hilbert space as we will see later.

B. Toeplitz family of states

We now consider a general Toeplitz family of states $\{|\psi_\theta\rangle\}_{\theta\in\mathbb{R}}$ and want to show that it is generated by an additive unitary. Since the geometry of the quantum states is Toeplitz, there exists some positive semi-definite function $G : \mathbb{R} \to \mathbb{C}$, which is called the symbol of the Toeplitz family, such that $\langle\psi_\theta|\psi_{\theta'}\rangle = G(\theta-\theta')$ and $G(0) = 1$. Since G is not necessarily a delta function, the states $\{|\psi_\theta\rangle\}_{\theta\in\mathbb{R}}$ are not orthogonal and so we first want to find a basis to express them in.

Let us perform a Fourier transform and then normalise to define the following Fourier states

$$|k\rangle = \frac{1}{2\pi\sqrt{g(k)}}\int_{-\infty}^{\infty}\mathrm{d}\theta\ e^{ik\theta}|\psi_\theta\rangle \tag{26}$$

Here, we define the spectral measure of the symbol

$G(\theta) = \int_{-\infty}^{\infty} \mathrm{d}k\ e^{ik\theta} g(k)$ as

$$g(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\theta\ e^{-ik\theta} G(\theta). \tag{27}$$

If $g(k) = 0$ for a given k, then we take $|k\rangle = 0$. Let $S = \{k \in \mathbb{R} : g(k) > 0\}$ be the support of g. The inner product between Fourier states in Eq. 26 is thus

$$\langle k|l\rangle = \delta(k-l)\chi_S(k). \tag{28}$$

On the support of g, the Fourier states are orthonormal, $\langle k|l\rangle = \delta(k-l)$ for $k \in S$ and $l \in S$, and otherwise the inner product vanishes, $\langle k|l\rangle = 0$ if $k \notin S$ or $l \notin S$. The inverse transformation between Fourier states and the original family is then

$$|\psi_\theta\rangle = \int_S \mathrm{d}k\ e^{-ik\theta} \sqrt{g(k)}|k\rangle \tag{29}$$

The function $g(k)$ is a valid probability distribution by Bochner's theorem because it is normalised ($\int_S \mathrm{d}k\ g(k) = G(0) = 1$) and non-negative since it is the Fourier transform of a positive semi-definite function. Similarly, the Gram matrix of inner products is Hermitian and positive semi-definite such that its eigenvalues $g(k)$ are non-negative. While this construction could be applied to finite-dimensional systems such as qubits or qudits, we focus on the infinite-dimensional case. We also assume here that θ ranges over all of $\mathbb{R}$. If instead the Hilbert space is finite dimensional or θ is continuous within a bounded interval because G is periodic, then S is discrete. In this latter case, we need to use the Fourier series instead and then normalise as we show in Appendix A.

We can now show that, for any Toeplitz family of states, there exists an additive unitary that generates it. (We already showed that the additive unitary case is Toeplitz.) For a given Toeplitz symbol $G(\theta)$, let the Hamiltonian be defined as $\hat{H} = \int_S \mathrm{d}k\ k|k\rangle\langle k|$ with eigenbasis $\{|k\rangle\}_{k\in S}$ from Eq. 26. Let the initial state be

$$|\psi_0\rangle = \int_S \mathrm{d}k\ \sqrt{g(k)}|k\rangle \tag{30}$$

such that its probability distribution in k is the spectral measure $g(k)$. The wavefunction in the previous section is thus $\psi_0(k) = \sqrt{g(k)}$. Then, the final state $|\psi_\theta\rangle = \hat{U}(\theta)|\psi_0\rangle$ agrees with Eq. 29 and the Gram matrix equals the symbol $G(\theta-\theta') = \langle\psi_\theta|\psi_{\theta'}\rangle$ as required. The additive unitary and Toeplitz family cases are thus equivalent.

If the prior is flat, then the covariant measurement is thus optimal by the result in the additive unitary case. The inner product between the original states $\{|\psi_\theta\rangle\}_{\theta\in\mathbb{R}}$ and covariant states $\{|\theta\rangle\}_{\theta\in\mathbb{R}}$ defined in Eq. 24 is

$$\langle\theta|\psi_{\theta'}\rangle = \frac{1}{\sqrt{2\pi}} \int_S \mathrm{d}k\ \sqrt{g(k)} e^{i(\theta-\theta')k}. \tag{31}$$

The likelihood of measuring θ from quantum whitening ($\hat{E}_\theta = |\theta\rangle\langle\theta|$) for a given parameter θ' and state $\hat{\rho}(\theta') = |\psi_{\theta'}\rangle\langle\psi_{\theta'}|$ is $L(\theta|\theta') = \mathrm{Tr}[\hat{\rho}(\theta')\hat{E}_\theta] = |\langle\theta|\psi_{\theta'}\rangle|^2$ which equals

$$L(\theta|\theta') = \frac{1}{2\pi} \int_S \mathrm{d}k \int_S \mathrm{d}l\ \sqrt{g(k)g(l)} e^{i(\theta-\theta')(k-l)}. \tag{32}$$

For the Toeplitz case, and in particular the coherent states that we consider later, we call the covariant states $\{|\theta\rangle\}_{\theta\in\mathbb{R}}$ the "quantum whitened states" and the covariant operator $\hat{W}$ the "quantum whitening measurement operator". This is in analogy to the classical whitening process, where a continuous random variable with a stationary two-point correlation function is mapped to a white noise random variable by a similar procedure of Fourier transforming, dividing by the square root of the eigenvalues, and then inverse Fourier transforming.

Similarly to the previous section, from the initial non-orthogonal Toeplitz family $\{|\psi_\theta\rangle\}_{\theta\in\mathbb{R}}$, we have defined the families of the Fourier states (stationary states) $\{|k\rangle\}_{k\in\mathbb{R}}$ and quantum whitened states (covariant states) $\{|\theta\rangle\}_{\theta\in\mathbb{R}}$. The Fourier states are orthogonal on S and the quantum whitened states are orthogonal if $S = \mathbb{R}$.

C. Fundamental quantum limit

We now calculate the MBMSE analytically in the general case of a Toeplitz family with the uninformative and uniform prior.

Let us start with just the Toeplitz assumption and assuming nothing about the prior. For a given Toeplitz family of states $\{|\psi_\theta\rangle\}_{\theta\in\mathbb{R}}$ and prior $\pi(\theta)$, the mixed state and Bayesian derivative in Eq. 10 in the Fourier basis in Eq. 26 are

$$\begin{aligned}
\hat{\bar{\rho}} &= \int_S \mathrm{d}k \int_S \mathrm{d}l\ \tilde{\pi}(k-l)\sqrt{g(k)g(l)}|k\rangle\langle l|, \\
\hat{\bar{\rho}}' &= \int_S \mathrm{d}k \int_S \mathrm{d}l\ i\tilde{\pi}'(k-l)\sqrt{g(k)g(l)}|k\rangle\langle l|
\end{aligned} \tag{33}$$

where we use $\hat{\rho}(\theta) = |\psi_\theta\rangle\langle\psi_\theta|$ and Eq. 29. The Fourier transform of the prior is $\tilde{\pi}(k) = \int_{-\infty}^{\infty} \mathrm{d}\theta\ e^{-ik\theta}\pi(\theta)$ such that $i\tilde{\pi}'(k) = \int_{-\infty}^{\infty} \mathrm{d}\theta\ \theta e^{-ik\theta}\pi(\theta)$ and the inverse Fourier transform is $\pi(\theta) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}k\ e^{ik\theta}\tilde{\pi}(k)$. We show this in Appendix B. The next step in finding the MBMSE is to calculate the BSLD in Eq. 11 from the spectral decomposition of the mixed state $\hat{\bar{\rho}}$, however, the eigenbasis is not known for a general prior. We will show below that the mixed state is diagonal in the Fourier basis for a flat prior. Otherwise, we can diagonalise the mixed state numerically to calculate the BSLD as we do in Appendix. K. Similarly, a covariant measurement is only necessarily optimal if the prior on θ is flat such that the figure-of-merit is covariant. For arbitrary priors, it may not be optimal to project onto the quantum whitened states.

1. Uninformative, uniform prior on $\mathbb{R}$

Let us now additionally assume that the prior is flat such that we can apply the covariant measurement result. This needs to be done with care since we assumed that θ and k were continuous on $\mathbb{R}$. This prior is the limit of a sequence of wider and wider normalised priors such as rectangular or Gaussian functions. In principle, we would want to calculate the MBMSE for each finite prior and then take the limit of the sequence of MBMSEs as the prior width goes to infinity. However, as discussed above, we cannot diagonalise the mixed state $\hat{\tilde{\rho}}$ in Eq. 33 for an arbitrary prior analytically. We conjecture that the MBMSE and BMSE of quantum whitening for any such sequence will converge to the value that we obtain below. We show in Appendix D 5 that this limit works in the case of amplitude estimation. Numerically, we show later in Sec. V C that this also works for frequency estimation.

Here, we instead calculate the MBMSE directly for a prior that is normalised and uniform on $\mathbb{R}$. These two properties are all that the following proof needs to assume, i.e. that $\partial_\theta \pi(\theta) = 0$ and $\int_{-\infty}^{\infty} \mathrm{d}\theta\ \pi(\theta) = 1$. The uniformity means that $\tilde{\pi}(k-l) \propto \delta(k-l)$ such that the mixed state in Eq. 33 is diagonal in the Fourier basis. The normalisation of the prior is necessary, else the non-unit norm of the prior will appear in the MBMSE as we discuss in Appendix D 5.

We need to be careful assigning an exact value to the prior. We do so nevertheless so that we can write the following equations explicitly, but we emphasise that our conclusions only require that the prior be normalised and flat on $\mathbb{R}$. Suppose that $\pi(\theta) \equiv \varepsilon$ for some infinitesimal $0 < \varepsilon \ll 1$ such that $\int_{-\infty}^{\infty} \mathrm{d}\theta\ \pi(\theta) = 1$ and $\tilde{\pi}(k-l) = \varepsilon 2\pi\delta(k-l)$. Informally, we choose that $\varepsilon = 1/[2\pi\delta(0)]$ in the following equations such that $\tilde{\pi}(k-l) = \delta(k-l)/\delta(0)$ but this division by $\delta(0)$ needs to be handled with the upmost care. The only property that we use is that $\delta(0)/\delta(0) = 1$ such that the flat prior and evidence will cancel in Bayes's rule in Eq. 3 below.

2. BMSE of quantum whitening

The prior being flat implies that the covariant measurement, i.e. quantum whitening, is optimal. The MBMSE is thus the BMSE of quantum whitening $\hat{W}$. Let us find the posterior $p(\theta'|\theta)$ on the parameter θ' for a given measurement result θ. Given the likelihood $L(\theta|\theta')$ in Eq. 32, the evidence is $p(\theta) \equiv 1/[2\pi\delta(0)]$ which is normalised like the prior ($\int_{-\infty}^{\infty} \mathrm{d}\theta\ p(\theta) = 1$). The posterior thus equals the likelihood ($p(\theta'|\theta) = L(\theta|\theta')$) by Bayes's rule in Eq. 3. The posterior is normalised ($\int_{-\infty}^{\infty} \mathrm{d}\theta'\ p(\theta'|\theta) = 1$) and centred on the measurement result ($\check{\theta}_\theta = \theta$). The MBMSE, which equals the BMSE in Eq. 5 for $\hat{W}$, is thus

$$\Delta^2\theta = \mathcal{V} := \frac{1}{4}\int_S \mathrm{d}k\ \frac{g'(k)^2}{g(k)}. \tag{34}$$

Here, the MSE in Eq. 4 and variance of the posterior in Eq. 6 are both constant and equal to the MBMSE

$$\mathrm{MSE}(\theta') = V_{\mathrm{post}}(\theta) \equiv \mathcal{V} \tag{35}$$

and the prior and evidence are normalised. We prove this result in Appendix B 1.

The value of $\mathcal{V}$ in Eq. 34 has a familiar functional form. It equals a quarter of the Fisher information in Eq. 22 were we to treat $g(k)$ in Eq. 27 as a prior on k for some hypothetical dual estimation problem. We emphasise that this hypothetical dual problem of estimating k and its classical Fisher information $4\mathcal{V}$ are not the same as our actual problem of estimating θ and the quantum Fisher information which we discuss below. In particular, we wish to minimise the MBMSE with respect to θ and thus the Fisher information of $g(k)$. Informally, since θ and k are related by a Fourier transform, this scenario is similar to the classical time-frequency uncertainty principle known as the Gabor limit in signal processing. This intuition will become clearer when we discuss the Gaussian case in Sec. V A.

3. Comparison to the QFI

Let us now calculate the actual QFI for this estimation problem, which we distinguish from the intuition about the functional form of the MBMSE provided above. The value of $\mathcal{V}$ in Eq. 34 for the MBMSE is in contrast with the QFI which will not be reached except for at high SNR. Using the additive unitary representation with the Hamiltonian $\hat{H}$ in Eq. 23 and initial state $|\psi_0\rangle$ in Eq. 30, the QFI with respect to θ is $\mathcal{I}_Q(\theta) = 4\mathrm{Var}_{|\psi_0\rangle}[\hat{H}]$ which equals

$$\mathcal{I}_Q(\theta) = 4\left(\int_S \mathrm{d}k\ k^2 g(k) - \left[\int_S \mathrm{d}k\ k g(k)\right]^2\right) \tag{36}$$

which is four times the variance of $g(k)$. In contrast to the QCRB, the MBMSE in Eq. 34 in principle depends on more than just the second moment of $g(k)$, although we will show that it is equal in the Gaussian case below. In particular, we will show cases in Sec. V where the QFI is finite and the MBMSE diverges.

The value for the QFI in Eq. 36 can also be shown without using the additive unitary representation. Let us Taylor expand the Toeplitz symbol for a small parameter difference $\delta\theta$ as

$$G(\delta\theta) \approx 1 + \delta\theta G'(0) + \frac{1}{2}\delta\theta^2 G''(0) \tag{37}$$

where $\mathrm{Re}\,[G'(0)] = 0$ holds such that the states are normalised. The derivatives of the symbol at zero are related to the moments of the spectral measure as $G'(0) = i\int_S \mathrm{d}k\ k g(k)$ and $G''(0) = -\int_S \mathrm{d}k\ k^2 g(k)$. The QFI from the fidelity in Eq. 18 thus equals the value in Eq. 36.

4. Optimality of quantum whitening

We can also prove that quantum whitening is optimal without using the result that covariant measurement is optimal. We include this alternative proof here for completeness and provide further details in Appendix B 2.

For the uninformative, uniform prior ($\tilde{\pi}(k-l) = \delta(k-l)/\delta(0)$), the mixed state in Eq. 33 is normalised ($\mathrm{Tr}[\hat{\tilde{\rho}}] = 1$) and diagonal in the Fourier basis since

$$\hat{\tilde{\rho}} = [1/\delta(0)] \int_S \mathrm{d}k \; g(k)|k\rangle\langle k|. \tag{38}$$

This was the point of choosing the flat prior: we now know that the eigenbasis of the mixed state is the Fourier basis and thus can calculate the BSLD. The matrix coefficients of the Bayesian derivative in Eq. 33 are

$$\langle k|\hat{\tilde{\rho}}'|l\rangle = i[\delta'(k-l)/\delta(0)]\sqrt{g(k)g(l)} \tag{39}$$

such that the coefficients of the BSLD in Eq. 11 are

$$\langle k|\hat{L}|l\rangle = i\delta'(k-l)\frac{2\sqrt{g(k)g(l)}}{g(k)+g(l)}\chi_S(k)\chi_S(l). \tag{40}$$

The distributions $\delta'(k-l)f_1(k,l)$ and $\delta'(k-l)f_2(k,l)$ are equal if the functions f_1 and f_2 and their partial derivatives are equal at $k = l$, i.e. if the following conditions are satisfied

$$\begin{aligned} f_1(k,k) &= f_2(k,k), \\ \partial_x f_1(x+l,l)|_{x=0} &= \partial_x f_2(x+l,l)|_{x=0}, \\ \partial_x f_1(k,k-x)|_{x=0} &= \partial_x f_2(k,k-x)|_{x=0}. \end{aligned} \tag{41}$$

This can be shown using integration by parts against an arbitrary test function. Eq. 40 thus implies that here

$$\langle k|\hat{L}|l\rangle = i\delta'(k-l)\chi_S(k)\chi_S(l). \tag{42}$$

In comparison, the matrix coefficients of the quantum whitening measurement operator are

$$\langle k|\hat{W}|l\rangle = i\delta'(k-l)\chi_S(k)\chi_S(l) \tag{43}$$

which equal those of the BSLD $\hat{L}$ in Eq. 42. This can be verified by showing that $\hat{W}$ satisfies the Lyapunov equation in Eq. 9. In particular, the matrix coefficients of the anticommutator are

$$\langle k|\frac{1}{2}\{\hat{\tilde{\rho}}, \hat{W}\}|l\rangle = \frac{1}{2}i[\delta'(k-l)/\delta(0)][g(k)+g(l)] \tag{44}$$

which equal the matrix coefficients of the Bayesian derivative in Eq. 39 since the conditions in Eq. 41 hold. This means that it is optimal to measure the quantum whitened states $|\theta'\rangle$, either via the non-projective POVM if $S \subsetneq \mathbb{R}$ or by projection if $S = \mathbb{R}$. Either way, the optimal estimator for θ is simply the measurement result θ'.

We have now determined the fundamental quantum limit and optimal measurement for an arbitrary Toeplitz family given no prior information about the parameter. We will now apply these results to various families of single-particle and coherent states. However, first we must establish the inner product structure, i.e. the geometry, of single-particle and coherent states.

III. GEOMETRY OF SINGLE-PARTICLE STATES

Before we return to our problem of determining the fundamental quantum limit on frequency estimation of the coherent state in Eq. 2, it will be informative to first study the simpler problem of estimating the parameter governing the waveform of a single-particle state, i.e. the multi-mode Fock state with a single particle in a given temporal mode. This single-particle state will provide a good introduction to the techniques that we use later for the coherent state. Although the inner product structure of the multi-mode single-particle and coherent states is well-known, we review it here for completeness as we will build upon it later in our Bayesian analysis.

Let us provide a basic physical model for the single-particle state. It may be imagined as the result of coupling an optical field to a two-level atomic transition with the field initially in vacuum and the qubit excited. Suppose that after some energy-conserving unitary dynamics the final state of the qubit is the ground state, then the final state of the optical field will be a single-particle state of exactly one photon. The temporal mode of the state is determined by the dynamics and the unknown signal.

We will demonstrate the following key idea: that we may restrict ourselves to the "waveform subspace" spanned by the possible quantum states which significantly reduces the dimension of the Hilbert space as will be shown. To find a basis for this waveform subspace, we need to first understand how the geometry of the quantum states depends on the corresponding geometry of the classical waveforms, i.e. temporal modes. We will show that these geometries are the same for single-particle states but different for coherent states which is what makes our motivating problem about coherent states interesting.

A. General single-particle state case

Let us generalise the classical waveforms beyond Eq. 1 to now allow for any family parameterised by θ of $L^2(\mathbb{R})$ functions, i.e. square-integrable complex waveforms:

$$\left\{ h_\theta \,\middle|\, \theta \in \mathbb{R},\, h_\theta : \mathbb{R} \to \mathbb{C},\, \int_{-\infty}^{\infty} \mathrm{d}t \;\; |h_\theta(t)|^2 < \infty \right\} \tag{45}$$

We want to estimate the classical waveform–governing parameter θ from measurements of the quantum state that encodes it. There are many possible ways for a classical waveform to be encoded in a quantum state, but we

focus on the single-particle and coherent state cases. For a given classical waveform and encoding scheme, we may refer to the corresponding quantum state as a "quantum waveform".

Let the single-particle quantum state corresponding to a given classical waveform be defined as $|x_\theta\rangle = \hat{b}_\theta^\dagger|0\rangle$ where the temporal mode annihilation operator is defined as

$$\hat{b}_\theta = \int_{-\infty}^{\infty} \mathrm{d}t\ x_\theta^*(t)\hat{a}(t) \tag{46}$$

and we decompose the classical waveform as $h_\theta(t) = -i\sqrt{2}\alpha_\theta x_\theta(t)$ where $\alpha_\theta \in \mathbb{C}$ and $\int_{-\infty}^{\infty} \mathrm{d}t\ |x_\theta(t)|^2 = 1$. Here, the normalisation α_θ is discarded since we assume that the state only has one particle but later this will be the complex amplitude of the corresponding coherent state. Thus, without loss of generality, let us assume that $\alpha_\theta \equiv 1$ for the remainder of this section. The inner product between single-particle states is then related to the commutator between different temporal modes as

$$\langle x_\theta|x_{\theta'}\rangle = [\hat{b}_\theta, \hat{b}_{\theta'}^\dagger] = \int_{-\infty}^{\infty} \mathrm{d}t\ x_\theta^*(t)x_{\theta'}(t) \tag{47}$$

which is the $L^2(\mathbb{R})$ inner product of the corresponding classical waveforms which are normalised.

The classical and quantum geometries are thus the same since their inner products are the same: There is no added structure or quantum noise from measuring the single-particle states compared to the classical waveforms. This means that quantum whitening the single-particle states is the same as classically whitening their corresponding classical waveforms, i.e. temporal modes, as we show in Appendix C. We will see that these geometries are not the same in the coherent state case.

B. QFI of single-particle states

To determine the QFI, suppose that $\theta' = \theta + \delta\theta$ with $\delta\theta$ small such that we can Taylor expand the normalised waveform as

$$x_{\theta+\delta\theta}(t) \approx x_\theta(t) + \delta\theta\partial_\theta x_\theta(t) + \frac{1}{2}\delta\theta^2\partial_\theta^2 x_\theta(t) \tag{48}$$

where we keep only terms $\mathcal{O}(\delta\theta^2)$ henceforth. The inner product in Eq. 47 is thus

$$\begin{aligned}\langle x_\theta|x_{\theta+\delta\theta}\rangle \approx 1 &+ \delta\theta \int_{-\infty}^{\infty} \mathrm{d}t\ x_\theta^*(t)\partial_\theta x_\theta(t) \\ &+ \frac{1}{2}\delta\theta^2 \int_{-\infty}^{\infty} \mathrm{d}t\ x_\theta^*(t)\partial_\theta^2 x_\theta(t).\end{aligned} \tag{49}$$

Since the waveforms $x_\theta(t)$ are normalised, by computing the norm of both sides of Eq. 48, we can show that from the linear term in $\delta\theta$ that

$$\int_{-\infty}^{\infty} \mathrm{d}t\ \mathrm{Re}\left[x_\theta(t)\partial_\theta x_\theta^*(t)\right] = 0 \tag{50}$$

and from the quadratic term in $\delta\theta^2$ that

$$\int_{-\infty}^{\infty} \mathrm{d}t\ \mathrm{Re}\left[x_\theta(t)\partial_\theta^2 x_\theta^*(t)\right] = -\int_{-\infty}^{\infty} \mathrm{d}t\ |\partial_\theta x_\theta(t)|^2 . \tag{51}$$

This means that the infinitesimal fidelity between single-particle states is

$$\begin{aligned}\mathcal{F}(\theta) &= |\langle x_\theta|x_{\theta+\delta\theta}\rangle|^2 \\ &\approx 1 - \delta\theta^2 \int_{-\infty}^{\infty} \mathrm{d}t\ |\partial_\theta x_\theta(t)|^2 \\ &\quad + \delta\theta^2 \left(\int_{-\infty}^{\infty} \mathrm{d}t\ \mathrm{Im}\left[x_\theta^*(t)\partial_\theta x_\theta(t)\right]\right)^2\end{aligned} \tag{52}$$

such that by Eq. 18 the QFI is

$$\begin{aligned}\mathcal{I}_Q(\theta) &= 4\int_{-\infty}^{\infty} \mathrm{d}t\ |\partial_\theta x_\theta(t)|^2 \\ &\quad - 4\left(\int_{-\infty}^{\infty} \mathrm{d}t\ \mathrm{Im}\left[x_\theta^*(t)\partial_\theta x_\theta(t)\right]\right)^2 .\end{aligned} \tag{53}$$

This agrees with the QFI in Eq. 36 if the family of waveforms is Toeplitz. We will compare this below to the QFI of a coherent state with an average particle number of one.

IV. GEOMETRY OF COHERENT STATES

We now finally return to our problem of estimating the frequency ω from the coherent state in Eq. 2. We address the toy problem of estimating the displacement of a single harmonic oscillator in Appendix D. Here, we need to find the optimal measurement across many displaced harmonic oscillators in time. The total Hilbert space is too large to simulate numerically, e.g., the dimension is d^M if we discretise the total time T into $M = T/\delta t$ oscillators in time-bins of width δt and truncate each oscillator to dimension d in the Fock basis, see Appendix E 2. However, the waveform subspace $\mathrm{span}\{|h_\omega\rangle\}_{\omega\in\mathbb{R}}$ spanned by the possible quantum states can be much smaller as we will see later. Intuitively, this is because the inner product between neighbouring coherent states in the family $\{|h_\omega\rangle\}_{\omega\in\mathbb{R}}$ may not decay quickly.

We want to find a basis for the waveform subspace and then calculate the MBMSE in Eq. 7 from computing $\hat{\bar{\rho}}$ and $\hat{\bar{\rho}}'$ in that basis. To find a basis, we need to first calculate the Gram matrix $G_{\omega,\omega'} = \langle h_\omega|h_{\omega'}\rangle$ of inner products between different coherent states to understand the structure of the waveform subspace. We will calculate the Gram matrix for coherent states generally and then consider the case of frequency estimation and other examples summarised in Table II.

A. General coherent state case

Let us now generalise the classical waveforms beyond Eq. 1 and coherent states beyond Eq. 2 to now any fam-

ily of classical waveforms in Eq. 45 and coherent states defined as follows

$$|h_\theta\rangle = e^{\frac{i}{\sqrt{2}}\int_{-\infty}^{\infty}\mathrm{d}t\,\left[h_\theta(t)\hat{a}^\dagger(t)+h_\theta^*(t)\hat{a}(t)\right]}|0\rangle. \tag{54}$$

Here, we choose a convention such that the expectation values of the quadratures follow the classical waveform as

$$\begin{aligned}\langle h_\theta|\hat{x}(t)|h_\theta\rangle &= -\mathrm{Im}\left[h_\theta(t)\right],\\ \langle h_\theta|\hat{p}(t)|h_\theta\rangle &= \mathrm{Re}\left[h_\theta(t)\right].\end{aligned} \tag{55}$$

We now want to relate the Hilbert space structure of the classical waveforms to that of the coherent quantum states like we did above for the single-particle states. The inner product between the coherent states is

$$\langle h_\theta|h_{\theta'}\rangle = e^{-\frac{1}{4}\int_{-\infty}^{\infty}\mathrm{d}t\,\left[|h_\theta(t)|^2+|h_{\theta'}(t)|^2-2h_\theta^*(t)h_{\theta'}(t)\right]}, \tag{56}$$

where the first two terms in the exponent are the $L^2(\mathbb{R})$ norms squared of the classical waveforms h_θ and $h_{\theta'}$, and the third term is twice the $L^2(\mathbb{R})$ inner product between them. We prove this result in Appendix E. This inner product is not necessarily Toeplitz but we will see examples where it will be.

Even if the classical waveforms are orthogonal, the inner product between the corresponding quantum states is nonzero for finite SNR because of the quantum noise of the coherent state, unlike the single-particle states. If the $L^2(\mathbb{R})$ inner product $\int_{-\infty}^{\infty}\mathrm{d}t\,h_\theta^*(t)h_{\theta'}(t)$ is real, e.g. if the classical waveforms are real, then the inner product between quantum states in Eq. 56 becomes

$$\langle h_\theta|h_{\theta'}\rangle = e^{-\frac{1}{4}\int_{-\infty}^{\infty}\mathrm{d}t\,|h_\theta(t)-h_{\theta'}(t)|^2} \tag{57}$$

where the exponent is the $L^2(\mathbb{R})$ distance between the classical waveforms h_θ and $h_{\theta'}$.

The $L^2(\mathbb{R})$ norm squared of the classical waveform h_θ which appears in the exponent of Eq. 56 is equal to twice the average number of particles in the corresponding coherent state since

$$\bar{N}_\theta = \langle h_\theta|\hat{N}|h_\theta\rangle = \frac{1}{2}\int_{-\infty}^{\infty}\mathrm{d}t\;|h_\theta(t)|^2 \tag{58}$$

where $\hat{N} = \int_{-\infty}^{\infty}\mathrm{d}t\,\hat{a}^\dagger(t)\hat{a}(t)$ is the total number operator and the units of $h_\theta(t)$ are square root Hertz. This is implied by the coherent state being an eigenstate of the annihilation operator $\hat{a}(t)|h_\theta\rangle = \frac{i}{\sqrt{2}}h_\theta(t)|h_\theta\rangle$.

B. QFI of coherent states

Let us consider how distinguishable parameter values are when they are infinitesimally far apart. Suppose that $\theta' = \theta+\delta\theta$ with $\delta\theta$ small such that we can Taylor expand the signal as

$$h_{\theta'}(t) \approx h_\theta(t) + \delta\theta\partial_\theta h_\theta(t) + \frac{1}{2}\delta\theta^2\partial_\theta^2 h_\theta(t) \tag{59}$$

and we drop $\mathcal{O}(\delta\theta^3)$ terms henceforth. The inner product in Eq. 56 is then

$$\langle h_\theta|h_{\theta+\delta\theta}\rangle \approx e^{-\frac{1}{4}\delta\theta^2\int_{-\infty}^{\infty}\mathrm{d}t\,|\partial_\theta h_\theta(t)|^2+i\varphi} \tag{60}$$

where the phase is

$$\varphi = \frac{1}{2}\delta\theta\int_{-\infty}^{\infty}\mathrm{d}t\;\mathrm{Im}\left[h_\theta^*(t)\left(\partial_\theta h_\theta(t)+\frac{1}{2}\delta\theta\partial_\theta^2 h_\theta(t)\right)\right] \tag{61}$$

which vanishes if the $L^2(\mathbb{R})$ inner product between waveforms $\int_{-\infty}^{\infty}\mathrm{d}t\;h_\theta^*(t)h_{\theta+\delta\theta}(t)$ is real. The fidelity between these states is thus

$$\mathcal{F}(\theta) = |\langle h_\theta|h_{\theta+\delta\theta}\rangle|^2 \approx e^{-\frac{1}{2}\delta\theta^2\int_{-\infty}^{\infty}\mathrm{d}t\,|\partial_\theta h_\theta(t)|^2} \tag{62}$$

such that by Eq. 18 the QFI is

$$\mathcal{I}_Q(\theta) = 2\int_{-\infty}^{\infty}\mathrm{d}t\;|\partial_\theta h_\theta(t)|^2 \tag{63}$$

which is the $L^2(\mathbb{R})$ norm squared of the partial derivative of the classical waveform $\partial_\theta h_\theta$. This value for the QFI can also be shown from the Gaussian formula in Eq. 20. It can furthermore be recovered from the QFI in Eq. 36 if the family is Toeplitz.

The CFI for a time-varying quadrature measurement $\hat{x}_{\vartheta(t)}(t) = \cos[\vartheta(t)]\hat{x}(t) + \sin[\vartheta(t)]\hat{p}(t)$ saturates the QFI, where $\vartheta(t) = \arg[\partial_\theta h_\theta(t)] + \frac{\pi}{2}$ such that $\partial_\theta\langle h_\theta|\hat{x}_{\vartheta(t)}(t)|h_\theta\rangle = |\partial_\theta h_\theta(t)|$ by Eq. 55. (Here, $\hat{x}_{\vartheta(t)}(t)$ should be distinguished from the normalised temporal mode $x_\theta(t)$ in Eq. 46.) This can be shown, e.g., by discretising the state in time following Appendix E 2, calculating the CFI in Eq. 19, and then taking the continuum limit. This Fisher information approach, however, is only tight asymptotically and so we need to calculate the MBMSE instead to understand the fundamental quantum limit for finite SNR.

1. *Comparison to the QFI of single-particle states*

Suppose that we decompose the classical waveform as $h_\theta(t) = -i\sqrt{2}\alpha_\theta x_\theta(t)$ into its norm $\alpha_\theta = \sqrt{\bar{N}_\theta}$ and $L^2(\mathbb{R})$-normalised waveform $x_\theta(t)$. Here, we assume that $\alpha_\theta > 0$ without loss of generality, since any overall complex phase of $h_\theta(t)$ can be absorbed into $x_\theta(t)$. Then, the QFI in Eq. 63 decomposes as follows

$$\mathcal{I}_Q(\theta) = 4(\partial_\theta\alpha_\theta)^2 + 4\alpha_\theta^2\int_{-\infty}^{\infty}\mathrm{d}t\;|\partial_\theta x_\theta(t)|^2 \tag{64}$$

where the first term comes from changes in the norm of the classical waveform (average number of particles) and the second term comes from changes in the normalised waveform (temporal mode function).

The QFI for a coherent state with an average particle number of one ($\alpha_\theta \equiv 1$) is thus

$$\mathcal{I}_Q(\theta) = 4\int_{-\infty}^{\infty}\mathrm{d}t\;|\partial_\theta x_\theta(t)|^2. \tag{65}$$

Let us compare this $\bar{N}_\theta = 1$ coherent state $|h_\theta\rangle = e^{\hat{b}_\theta^\dagger - 1/2}|0\rangle$ from Eq. E1 to the single-particle state $|x_\theta\rangle = \hat{b}_\theta^\dagger|0\rangle$ with the same temporal mode operator $\hat{b}_\theta$ in Eq. 46 given by the same normalised waveform $x_\theta(t)$. The coherent state QFI in Eq. 65 is an upper bound on the single-particle state QFI in Eq. 53. The difference in QFI is the second term in Eq. 53 which equals

$$\mathcal{I}_Q^{|h_\theta\rangle}(\theta) - \mathcal{I}_Q^{|x_\theta\rangle}(\theta) = 4\left(\int_{-\infty}^{\infty} \mathrm{d}t \;\; |x_\theta(t)|^2 \, \partial_\theta \arg[x_\theta(t)]\right)^2. \tag{66}$$

This difference vanishes, e.g., (1) if $x_\theta(t)$ is real or (2) if $x_\theta(t)$ is complex but the parameter does not affect its phase such that $\partial_\theta \arg[x_\theta(t)] = 0$. This is because non-vacuum coherent states are sensitive to phase changes but Fock states are not since Fock states are rotationally symmetric in the phase plane.

For example, suppose that we want to sense a rotation of phase space by $\hat{U}(\theta) = e^{-i\theta\hat{N}}$. A given coherent state $|h_0\rangle$ with $h_0(t) = -i\sqrt{2}\alpha_0 x_0(t)$ simply rotates around the origin in the phase plane under this unitary as $|h_\theta\rangle = |h_0 e^{-i\theta}\rangle$ such that the QFI is $\mathcal{I}_Q(\theta) = 4\alpha_0^2$. In particle, if the average particle number is one, then the coherent state QFI is 4. In comparison, a given single-particle state $|x_0\rangle$ is invariant up to a global phase as $\hat{U}(\theta)|x_0\rangle = e^{-i\theta}|x_0\rangle$ such that the QFI vanishes.

C. Frequency estimation of a windowed sinusoid

Let us examine this inner product structure for the case of the sinusoidal signal. The general state in Eq. 54 recovers the sinusoidal state in Eq. 2 by windowing the signal in Eq. 1 to the time interval $(0, T)$ as

$$h_\omega(t) = A\cos(\omega t + \phi)\chi_{(0,T)}(t) \tag{67}$$

where $\chi_{(0,T)}(t)$ is the indicator function on $(0, T)$. This choice of window function and time interval is only a convention, as we discuss in Appendices F–G. The average number of particles in Eq. 58 is thus

$$\bar{N}_\omega = \frac{A^2 T}{4} + \frac{A^2}{8\omega}\left(\sin[2(\omega T + \phi)] - \sin[2\phi]\right) \tag{68}$$

which equals $\bar{N}_\omega \approx A^2 T/4$ in the relevant many-cycles limit of $T \gg 1/\omega$. In this regime, the average number of particles thus equals an eighth of the SNR squared, since the SNR is $A\sqrt{2T}$. Meanwhile, the $L^2(\mathbb{R})$ inner product between classical waveforms is

$$\int_{-\infty}^{\infty} \mathrm{d}t \; h_\omega^*(t) h_{\omega'}(t) = \frac{1}{2}A^2 T \,\mathrm{sinc}[(\omega - \omega')T] + \frac{A^2 \sin[(\omega + \omega')T + 2\phi] - A^2\sin(2\phi)}{2(\omega + \omega')} \tag{69}$$

where $\mathrm{sinc}(0) = 1$ and $\mathrm{sinc}(x) = \sin(x)/x$ for $x \neq 0$. This equals $\frac{1}{2}A^2 T\,\mathrm{sinc}[(\omega - \omega')T]$ in the many-cycles limit of $T \gg 1/(\omega + \omega')$. The inner product between quantum states in Eq. 56, or Eq. 57 since the waveform is real, is thus

$$\langle h_\omega | h_{\omega'}\rangle = e^{-\frac{1}{4}A^2 T g(\omega,\omega')} \tag{70}$$

where the exponent is defined as

$$\begin{aligned} g(\omega,\omega') = 1 &- \mathrm{sinc}[(\omega - \omega')T] \\ &+ \frac{1}{4\omega T}\left(\sin[2(\omega T + \phi)] - \sin[2\phi]\right) \\ &+ \frac{1}{4\omega' T}\left(\sin[2(\omega' T + \phi)] - \sin[2\phi]\right) \\ &- \frac{\sin[(\omega + \omega')T + 2\phi] - \sin(2\phi)}{(\omega + \omega')T} \end{aligned} \tag{71}$$

such that in the many-cycles limit of $T \gg 1/\min(\omega, \omega')$ the Gram matrix of inner products becomes

$$\langle h_\omega | h_{\omega'}\rangle \approx e^{-\frac{1}{4}A^2 T\left(1 - \mathrm{sinc}[(\omega - \omega')T]\right)} \tag{72}$$

In this regime, the Gram matrix is a Toeplitz matrix since it is purely a function of $\omega - \omega'$, whereas in general the Gram matrix in Eq. 70 is not Toeplitz. The many-cycles limit is relevant because we consider long integration times and want to know how the sensing performance scales with time.

The sinc function in the exponent of the inner product of states in Eq. 72 comes from the asymptotic value of the $L^2(\mathbb{R})$ inner product of classical waveforms $\int_{-\infty}^{\infty} \mathrm{d}t \; h_\omega^*(t) h_{\omega'}(t)$ in the many-cycles limit. This sinc function vanishes for frequencies that are far apart $(\omega - \omega')T \gg 1$ or separated by a harmonic frequency of $\omega - \omega' = n\pi/T$ for $(0, T)$ in Eq. 72. When the sinc function vanishes such that the classical waveforms are orthogonal, however, the inner product between the waveform states is nonzero and equals $\langle h_\omega | h_{\omega'}\rangle \approx e^{-A^2 T/4}$ for finite SNR because of the quantum noise of the coherent state.

The QFI in Eq. 63 for frequency estimation of the signal defined on $(0, T)$ in Eq. 67 is

$$\mathcal{I}_Q(\omega) \approx \frac{1}{3}A^2 T^3 \tag{73}$$

where we take the many-cycles limit of $\omega T \gg 1$. The QCRB in Eq. 15 in amplitude units is thus

$$\mathrm{MSE}(\omega)^{-1/2} \approx \sqrt{3}A^{-1}T^{-3/2} \tag{74}$$

which scales as $1/T^{3/2}$ instead of the naïve harmonic frequency spacing of $1/T$, as discussed in Sec. I C. To reach a given frequency resolution $\delta\omega$, this means that we require only a time T proportional to $1/\delta\omega^{2/3}$ rather than $1/\delta\omega$ (cf. Ref. [35]). Heuristically, this scaling can be understood as the number of states that are approximately orthogonal. Suppose that we require that $|\langle h_\omega | h_{\omega+\delta\omega}\rangle| \leq \varepsilon$ for some $\varepsilon > 0$ such that the probability ε^2 of measuring $|h_{\omega+\delta\omega}\rangle$ if the signal is actually $|h_\omega\rangle$ is small. For two

Parameter	Signal	Gram matrix	Toeplitz	Circulant	QFI
General case, θ	Eq. 45: family of square-integrable functions	Eq. 56	—	—	Eq. 63
Frequency, ω	Eq. 67: sinusoidal	Eq. 72*	✓*	✗	Eq. 73*
	Eq. G1: complex exponential	Eq. G3	✓	✗	Eq. G6
Amplitude, A	Eq. H1: general scale factor family	Eq. H2	✓	✗	Eq. H3
	Eq. 67: sinusoidal	Eq. H5*	✓	✗	Eq. H6*
	Eq. G1: complex exponential	Eq. H7	✓	✗	Eq. H8
Phase, ϕ	Eq. 67: sinusoidal	Eq. J1*	✓*	✓*	Eq. J2*
	Eq. G1: complex exponential	Eq. J3	✓	✓	Eq. J4

TABLE II. Summary of the properties and QFI of the Gram matrix $G_{\theta,\theta'} = \langle h_\theta | h_{\theta'} \rangle$ of inner products between coherent states for estimating different parameters from different signals. The Gram matrix is Toeplitz if $G_{\theta,\theta'} = G_{\theta-\theta'}$ and circulant if $G_{\theta,\theta'} = G_{\theta-\theta'+nP}$ for $n \in \mathbb{Z}$ and some finite period P. We reference the equations for the signal defined on the time interval $(0, T)$ but the $(-T/2, T/2)$ case is similar. *This property or equation holds only in the many-cycles limit of $T \gg 1/\min(\omega, \omega')$.

close frequencies satisfying $\delta\omega\, T \ll 1$, the inner product in Eq. 60 and Eq. 72 is

$$\langle h_\omega | h_{\omega+\delta\omega} \rangle \approx e^{-\frac{1}{24}\delta\omega^2 A^2 T^3} \tag{75}$$

such that the frequencies need to be at least as far apart as

$$\delta\omega \geq 2A^{-1}T^{-3/2}\sqrt{6\log(1/\varepsilon)}. \tag{76}$$

This means that the maximum number of approximately orthogonal waveform states $N = \Delta\omega/\delta\omega$ in a frequency interval of width $\Delta\omega$ is

$$N = \frac{\Delta\omega A T^{3/2}}{2\sqrt{6\log(1/\varepsilon)}}. \tag{77}$$

Suppose that ε is small enough that we can well-determine which of these N waveform states the measured state is closest in frequency to. Then, the remaining error in our estimate of ω simply comes from the resolution of the frequency bins of width $\delta\omega$. If the prior is flat, then we may use that the standard deviation of the continuous uniform distribution is $(b-a)/(2\sqrt{3})$ on the interval (a, b) to conclude that the error in estimating the frequency given Eq. 76 is

$$\sigma_\omega = A^{-1}T^{-3/2}\sqrt{2\log(1/\varepsilon)} \tag{78}$$

which has the same scaling with $A^{-1}T^{-3/2}$ as the QCRB in Eq. 74. This provides an intuitive explanation of the scaling of the QCRB. We assumed above that $\delta\omega\, T \ll 1$ such that we require that the SNR $A\sqrt{2T}$ satisfies

$$A\sqrt{T} \gg 2\sqrt{6\log(1/\varepsilon)}. \tag{79}$$

This illustrates that the QFI is only valid for high SNRs since ε is small. At low SNR, the QCRB cannot be achieved by any measurement which motivates our use of Bayesian techniques.

We discuss frequency estimation in the case of a complex exponential signal instead of the real sinusoidal one above in Appendix G, including for different window functions than the rectangular one above. We also discuss instead estimating the amplitude with the frequency and phase known in Appendix H and estimating the phase with the amplitude and frequency known in Appendix J. We summarise our results in Table II including whether the Gram matrix is Toeplitz and circulant (i.e. periodic).

V. APPLICATIONS

Let us now examine some examples of the MBMSE in Eq. 34 for different Toeplitz families of states $\{|\psi_\theta\rangle\}_{\theta\in\mathbb{R}}$ with different symbols $G(\theta)$ and the uninformative, uniform prior on the parameter θ. We will then present our numerical results for frequency estimation of a sinusoidal signal in a coherent states where the Gram matrix is only approximately Toeplitz and the prior is uniform on a bounded interval.

A. Gaussian case

Suppose that the symbol is $G(\theta) = e^{-\frac{1}{2}\varsigma^2\theta^2}$ which is proportional to a Gaussian distribution with variance $1/\varsigma^2$. Then, the spectral measure $g(k) = e^{-k^2/(2\varsigma^2)}/(\varsigma\sqrt{2\pi})$ is a Gaussian probability distribution with variance ς^2. The likelihood in Eq. 31 is $L(\theta|\theta') = \frac{2\varsigma}{\sqrt{2\pi}}e^{-2\varsigma^2(\theta-\theta')^2}$ which is Gaussian with mean θ' and variance $1/(4\varsigma^2)$. The MBMSE in Eq. 34 is thus $\mathcal{V} = 1/(4\varsigma^2)$ from the MSE. This can also be shown from the dual estimation problem since the Fisher information of $g(k)$ with respect to k is $1/\varsigma^2$. The MBMSE decreases as ς increases since $G(\theta)$ becomes narrower such that neighbouring parameters can be more easily distinguished. Thus we want to minimise the Fisher information of $g(k)$. This value of $\mathcal{V}$ also equals the QCRB in Eq. 36 (cf. Eq. H3) because the QCRB is tight for a finite number of measurements in the Gaussian case.

For example, the Gram matrix for amplitude estima-

tion of a coherent state in Eq. H2 is Gaussian with $\varsigma^2 = \bar{N}_1$. The MBMSE in Eq. H4 is $1/(4\bar{N}_1)$ in the limit of wider and wider priors. This result agrees with the value of $\mathcal{V}$ for the MBMSE from Eq. 34 above.

For another example, consider estimating the centre of a single-particle state corresponding to a Gaussian pulse. Suppose that the normalised classical waveform in Eq. 46 is

$$x_\theta(t) = \frac{1}{(2\pi)^{1/4}\sqrt{\sigma}} e^{-(t-\theta)^2/(4\sigma^2)} \tag{80}$$

The Gram matrix in Eq. 47 is thus $G(\theta) = e^{-\theta^2/(8\sigma^2)}$ and so the QCRB and MBMSE are both σ^2. In comparison, as we show in Appendix I, the coherent state case driven by $h_\theta(t) = -i\sqrt{2}\alpha x_\theta(t)$ has the same QCRB of σ^2 if the average number of particles is one ($\alpha = 1$) but has divergent MBMSE since the Gram matrix is non-Gaussian and asymptotically nonzero.

1. *Constant case*

Suppose that $G(\theta) \equiv 1$ such that the state is independent of the parameter and the posterior equals the prior since no information is gained from any measurement. Then, $g(k) = \delta(k)$ and the MBMSE in Eq. 34 diverges as it equals the prior variance. Similarly, the QFI in Eq. 36 vanishes such that the QCRB diverges. This is the wide limit $\varsigma \to 0$ of the Gaussian case.

2. *Singular case*

Let us now consider the opposite case. Suppose that $G(\theta) = \delta(\theta)/\delta(0)$ such that the states are all orthogonal, e.g. the position eigenstates, then $g(k) \equiv 1/[2\pi\delta(0)]$ which is the uninformative, uniform probability distribution on $\mathbb{R}$ similarly to the prior. The MBMSE in Eq. 34 vanishes as the parameter can be exactly determined from a single measurement projecting onto the states themselves. Similarly, the QFI in Eq. 36 diverges such that the QCRB vanishes. This is the narrow limit $\varsigma \to \infty$ of the Gaussian case.

B. Frequency estimation of single-particle states

Suppose that the symbol is $G(\theta) = \mathrm{sinc}(a\theta)$ for some $a > 0$. We give an example of this below for frequency estimation of a single-particle state. The spectral measure is then $g(k) = \frac{1}{2a}\chi_{(-a,a)}(k)$ which is the uniform probability distribution on the bounded interval $(-a, a)$. This rectangular function is discontinuous at $k = \pm a$ such that the MBMSE in Eq. 34 diverges. The proof of the posterior variance in Eq. 34 given in Appendix B 1 assumes that g is smooth which it is not here. Nevertheless, we can independently verify that the MBMSE diverges since Eq. 31 implies that the posterior of quantum whitening is $p(\theta|\theta') = (a/\pi)\,\mathrm{sinc}^2[a(\theta' - \theta)]$ which has unbounded variance. This is because the sinc function $\mathrm{sinc}(\theta)$ falls off as $1/\theta$, which is too slow for the integral in the posterior variance in Eq. 6 to converge. Since the symbol $G(\theta)$ is not constant, however, we do learn something from the quantum whitening measurement such that the posterior is better than the prior in a sense, but their variances are both undefined. This highlights that the BMSE can be a poor figure-of-merit for the uninformative, uniform prior on $\mathbb{R}$. In contrast, the QFI in Eq. 36 is $\mathcal{I}_Q(\theta) = 4a^2/3$ such that the QCRB in Eq. 15 is finite.

A more robust metric would be to compute the limit of the ratio between the posterior and prior variances for a sequence of, e.g., uniform priors on wider and wider bounded intervals that converge to the uninformative, uniform prior on $\mathbb{R}$. Analytically, we do not know the MBMSE for finite prior width. Numerically, however, we see that the variance of the uniform prior on $\theta \in (-\Delta/2, \Delta/2)$ scales as Δ^2 and the quantum whitening posterior variance scales as Δ as expected. This means that the ratio between the posterior and prior variances scales as $1/\Delta$ such that the relative gain improves for wider and wider priors. We proved above that quantum whitening is optimal in the limit of wide priors such that this relative gain proportional to $1/\Delta$ is the best that we can achieve.

For example, suppose that we want to estimate the frequency of a single-particle state corresponding to the normalised classical waveform $x_\omega(t) = -i\frac{1}{\sqrt{T}}e^{i(\omega t+\phi)}\chi_{(-T/2,T/2)}(t)$ in Eq. 46 (cf. Eq. G4). The Gram matrix is thus $G(\omega) = \mathrm{sinc}(\omega T/2)$ such that $g(k) = \frac{1}{T}\chi_{(-T/2,T/2)}(k)$ and the QFI is $\mathcal{I}_Q(\omega) = T^2/3$ which is finite unlike the MBMSE. For the real sinusoid case, the same result holds for $x_\omega(t) = \sqrt{\frac{2}{T}}\cos(\omega t + \phi)\chi_{(-T/2,T/2)}(t)$ (cf. Eq. F1) in the many-cycles limit since then the Gram matrix is the same. We address the coherent state case of frequency estimation below.

C. Frequency estimation of coherent states

Let us now finally return to the problem of frequency estimation of the coherent states. For the sinusoidal signal defined on $(0, T)$ in Eq. 67, the exact Gram matrix in Eq. 70 is not Toeplitz but the Gram matrix in Eq. 72 in the many-cycles limit is Toeplitz. If we wanted to avoid the many-cycles limit, we could instead study the complex exponential signal defined on $(-T/2, T/2)$ in Eq. G4 whose exact Gram matrix in Eq. G5 is Toeplitz and has the same functional form as Eq. 72.

Suppose that the symbol is $G(\omega) = e^{-\frac{1}{4}A^2T[1-\mathrm{sinc}(\omega T)]}$ whose Fourier transform, the spectral measure $g(k)$, we cannot compute analytically. However, as shown in Fig. 2, we observe numerically that $g(k)$ is discontinuous at the same points $k = \pm T$ as the Fourier transform $\frac{1}{2T}\chi_{(-T,T)}(k)$ of $\mathrm{sinc}(\omega T)$ such that the MBMSE

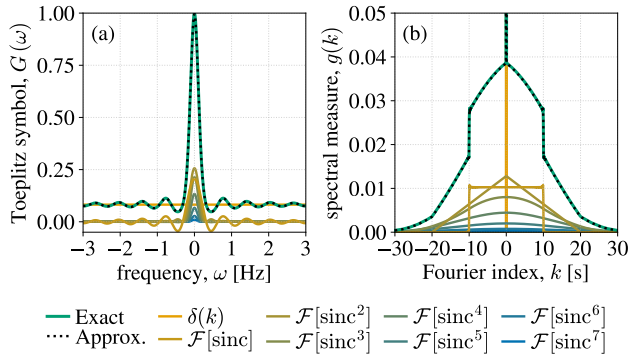

FIG. 2. Exact and approximate (truncated at 7th order) values for the Taylor expansion of the (a) symbol $G(\omega)$ and (b) spectral measure $g(k)$ for the frequency estimation Gram matrix in Eq. 72 with $A = 1$ Hz$^{-0.5}$ and $T = 10$ s such that the SNR is $A\sqrt{2T} \approx 4.5$. A divergence at DC and discontinuities at $\pm T$ are introduced in $g(k)$ from the Fourier transform of the constant and sinc parts of $G(\omega)$, respectively.

in Eq. 34 again diverges. This is in comparison to the QCRB in Eq. 74 which is finite but unattainable. There is also a singularity at DC due to the $e^{-\frac{1}{4}A^2T}$ asymptotic value of $G(\omega)$ for large ω.

1. *Low-SNR expansion*

These divergences can be explained analytically in the low-SNR limit $A^2T \ll 1$ since then the symbol is

$$G(\omega) \approx e^{-\frac{1}{4}A^2T}\left[1 + \frac{A^2T}{4}\mathrm{sinc}(\omega T) + \frac{A^4T^2}{32}\mathrm{sinc}^2(\omega T)\right] \tag{81}$$

where we drop terms of order $\mathcal{O}(A^6T^3/384)$. Each term here corresponds to one additional particle: the first term to zero particles, the second to one particle, the third to two particles etc. The spectral measure is then

$$g(k) \approx e^{-\frac{1}{4}A^2T}\left[\delta(k) + \frac{A^2}{8}\chi_{(-T,T)}(k) + \frac{A^4}{128}\max(0, 2T - |k|)\right]. \tag{82}$$

Here, the nth term in the symbol $G(\omega)$ is proportional to $\mathrm{sinc}^n(\omega T)$ and thus goes as $1/\omega^n$ and its Fourier transform, i.e. the corresponding term in the spectral measure $g(k)$, is a piecewise degree-$(n-1)$ polynomial with support $(-nT, nT)$ and discontinuities in the nth derivative if $n \geq 1$. For example, as shown above, the $n = 0$ term goes to the Dirac delta function at $k = 0$, the $n = 1$ term goes to a rectangular function on $(-T, T)$, the $n = 2$ term goes to a triangular function on $(-2T, 2T)$, the $n = 3$ term goes to a piecewise quadratic function on $(-3T, 3T)$, etc. This expansion is only valid in the low SNR–limit but we see numerically the contributions to $g(k)$ from these different terms and the appearance of the discontinuities at $k = \pm T$ and divergence at DC for finite SNR too as shown in Fig. 2.

2. *Limiting behaviour for wider and wider priors*

Similarly to the discussion in the sinc case above, this means that the BMSE is a poor figure-of-merit for frequency estimation with the uninformative, uniform prior on $\mathbb{R}$ and that we want to instead study the ratio between the MBMSE, i.e. the minimum average posterior variance, and the prior variance for a sequence of wider and wider priors. Unlike the sinc case above, however, the symbol $G(\omega)$ here asymptotes to a nonzero value $e^{-\frac{1}{4}A^2T}$ for large ω. This means that the posterior variance scales as $\Delta\omega^2$ like the variance of the uniform prior on $(-\Delta\omega/2, \Delta\omega/2)$ for large $\Delta\omega$. We give another example in Eq. I1 where the Gram matrix has a nonzero constant component and thus the posterior variance diverges quadratically.

We observe numerically that the gap between the prior and posterior variances closes, albeit quite slowly, for increasing $\Delta\omega$ and fixed SNR. This is expected as the effect on the variance from the central peak due to the other terms in the expansion of $G(\omega)$ is slowly suppressed for wider and wider priors, even though it may still contain a significant probability mass for a given finite prior. For example, for SNR 4.5, the numerical ratio of posterior to prior variance is approximately 0.06 for a prior spanning two orders of magnitude and only rises to 0.08 for a prior spanning four orders of magnitude. This means that the decrease in variance of the posterior compared to the prior is significant for wide but finite priors, even if we expect it to diminish in the limit of an infinitely wide prior.

The comparison above is between the prior and posterior distribution from quantum whitening. More precisely, we approximate the posterior distribution with the likelihood below Eq. 31 which only equals the posterior in the limit of an infinitely wide prior. Is quantum whitening optimal for finite priors? We have only shown that quantum whitening is optimal for an infinite, flat prior. We calculate the MBMSE numerically from Eq. 12 using the Toeplitz symbol above and methods in Sec. K. The results for the MBMSE are similar to the posterior variance but not identical. The small residual difference between the MBMSE and posterior variance could be because (1) we use a finite numerical resolution, (2) we use the likelihood instead of the true posterior, and/or (3) quantum whitening is close to but not fully optimal for a wide but finite prior. We defer better understanding the finite prior case to future work, including what happens with non-uniform priors.

For the coherent state case, the quantum whitened states in Eq. 24 are not necessarily coherent states themselves since they are particular superpositions of coherent

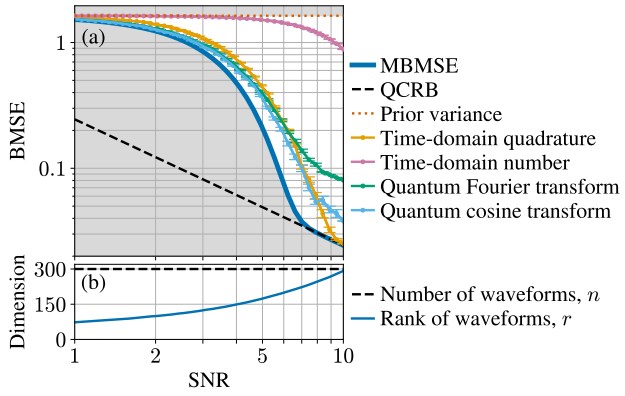

FIG. 3. (a) BMSE in amplitude units for different measurement schemes versus the SNR in amplitude units $A\sqrt{2T}$ for different signal amplitudes A and fixed total time T compared to the MBMSE, QCRB, and prior variance for the parameters given in Table III. The shaded grey region is inaccessible. (b) Rank of the Gram matrix versus the SNR compared to the number of waveforms simulated numerically.

Parameter	Value
Known phase, ϕ	0
Total time, T	10 s
Harmonic frequency spacing, $2\pi/T$	$2\pi \times 0.1$ Hz
Prior width, $\Delta\omega$	$2\pi \times 0.9$ Hz
Centre frequency, ω_0	$2\pi \times 0.5$ Hz
Centre number of cycles, $\omega_0 T/(2\pi)$	0.5
Minimum frequency, $\omega_{\min} := \omega_0 - \Delta\omega/2$	$2\pi \times 0.05$ Hz
Minimum number of cycles, $\omega_{\min} T/(2\pi)$	5
Maximum frequency, $\omega_{\max} := \omega_0 + \Delta\omega/2$	$2\pi \times 0.95$ Hz
Maximum number of cycles, $\omega_{\max} T/(2\pi)$	9.5
Number of oscillators, M	101
Time resolution, δt	0.1 s
Nyquist frequency, $\pi/\delta t$	$2\pi \times 5$ Hz
Number of waveforms, n	300
Frequency resolution, $\Delta\omega/n$	$2\pi \times 3$ mHz
Phase difference per time bin, $\Delta\omega\delta t$	0.18π

TABLE III. Parameters for the BMSE versus SNR results shown in Fig. 3. The first group of parameters are physically meaningful in the continuum limit whereas the second group of parameters, below the grey line, only arise from discretising the time domain numerically as done in Appendix E 2.

states. The coherent states form an overcomplete basis such that any state in the subspace is some superposition of them. This is in contrast to the single-particle state case where the quantum whitened states are themselves single-particle states with the classically whitened waveforms/temporal modes. The quantum whitened coherent states and the coherent states with classically whitened waveforms are thus not necessarily the same.

3. *Non-Toeplitz case*

Let us now examine what happens if we do not assume that the Gram matrix is Toeplitz by not taking the many-cycles limit. Using the numerical approach in Appendix K, we can study our motivating example of frequency estimation of the sinusoidal signal in Eq. 67 with known amplitude and phase and the exact non-Toeplitz Gram matrix in Eq. 70. Let us fix the parameters given in Table III such as the total time T and prior width $\Delta\omega$ and vary the known amplitude A to reach different values of the SNR $A\sqrt{2T}$. As the SNR increases, the MBMSE exhibits an SNR threshold as it goes from the prior variance to the QCRB as shown in Fig. 3a. For intermediate values of the SNR, the MBMSE is not achieved by quadrature measurement in time, which is the classical measurement discussed in Sec. I C. The MBMSE is also not achieved by performing a number-resolving measurement in time, e.g. illuminating a single-photon detector in the case of an optical device. Additionally, it is not saturated by first performing a discrete quantum Fourier or cosine transform to the bath discretised in the time-domain as in Appendix E 2 and then performing a number-resolving measurement to each of the resulting modes. This is the filter-based photon-counting strategy recently compared in Ref. [1] to the time-domain quadrature measurement, although the advantage that we see here between these two strategies is smaller than that seen in Ref. [1] likely due to the simplifications that we have made here to the signal and filter models and in assuming that all other parameters are known. This quantum Fourier transform and number-resolving measurement should be distinguished from the Fourier transform in Eq. 26. The optimal measurement that achieves the MBMSE here is not quantum whitening, because we have not taken the limit of wider and wider priors, but is to project onto analogous superpositions of the coherent waveform states. The rank r of the Gram matrix remains stable under increasing the frequency resolution as discussed in Appendix K but increases with the SNR for a fixed number of waveforms n as shown in Fig. 3b. This means that the Fourier transform in Eq. 26 is an efficient representation of the waveform subspace.

VI. OUTLOOK

In this paper, we have determined the fundamental Bayesian quantum limit of frequency estimation of a bosonic system. This Bayesian perspective yields an understanding of phenomena, such as the SNR threshold effect, which are hidden from the traditional Fisher information approach. The key is to understand the Hilbert space geometry of the multi-mode coherent states, and how it relates to their corresponding classical waveforms. For example, the coherent states all overlap for finite SNR, even if their classical waveforms are orthogonal. In the many-cycles limit, such that the inner product becomes Toeplitz, and if we have no prior information

about the frequency, we prove that quantum whitening the states is optimal. This means that we project onto particular superpositions of the coherent states. This result applies beyond frequency estimation to any Toeplitz family of states. A divergence in the Bayesian limit arises, however, in the case of frequency estimation with an infinitely wide prior. We handle this numerically by using finite-width priors and taking the singular value decomposition of the waveform basis, which works even when the inner product is not Toeplitz. We observe that, although the SNR threshold appears just as it does classically, the classical limit can nevertheless be surpassed by a measurement similar to quantum whitening. These results may accelerate future searches for fundamental physics, such as gravitational-wave probes of neutron-star physics and the direct detection of ultra-light dark matter, as well as improve emerging quantum technologies using nonstandard, nonclassical techniques.

Many avenues of future work are promising such as determining how to experimentally realise quantum whitening, studying other parameterised families of waveforms such as Lorentzians, and considering covariant measurements for estimating the phase. We are also curious as to what happens if we marginalise over the amplitude and phase in the case of frequency estimation rather than treating them as known. Ref. [1] shows that the head-to-head comparison between quadrature and number-resolving measurements changes dramatically whether or not we marginalise. We may want to simultaneously estimate the amplitude and phase rather than marginalise over them. However, tight general Bayesian multiparameter bounds are unknown. Additionally, the Bayesian limit is not known if we optimise over the intervening quantum control and preparation of the initial quantum state, potentially allowing for entanglement with noiseless ancillae (such as in Ref. [39]). Much therefore remains to be understood about Bayesian waveform estimation, especially compared to the Fisher information regime.

ACKNOWLEDGMENTS

We thank the following people for their advice provided during this research: Katerina Chatziioannou, Rana Adhikari and the Caltech quantum control group, the Caltech Chen quantum group, and the ANU CGA squeezer group. This research is supported by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (Project No. CE170100004 and CE230100016). J.W.G. and this research are supported by an Australian Government Research Training Program Scholarship and also partially supported by the US NSF grant PHY-2011968. In addition, Y.C. acknowledges the support by the Simons Foundation (Award Number 568762) and the NSF grant PHY-2309231. T.G. acknowledges funding provided by the Quantum Science and Technology early-career fellowship of the Israel Council for Higher Education. S.A.H. acknowledges support through an Australian Research Council Future Fellowship grant FT210100809.

DATA AVAILABILITY

The code and data that supports this work is openly available online [40].

Appendix A: Circulant family and flat prior

Suppose that we have a circulant family of non-orthogonal states $\{|\psi_\theta\rangle\}_{\theta\in\mathbb{R}}$ such that $\langle\psi_\theta|\psi_{\theta'}\rangle = G(\theta-\theta')$ for some periodic function G and the prior is uniform on the finite period. Let the period of the symbol G be 2π without loss of generality such that the prior is $\pi(\theta) = [1/(2\pi)]\chi_{(-\pi,\pi]}(\theta)$ and we may restrict the family to $\{|\psi_\theta\rangle\}_{\theta\in(-\pi,\pi]}$.

Similarly to the Toeplitz case in Sec. II B, we can define the Fourier $\{|\varphi_k\rangle\}_{k\in\mathbb{Z}}$ and whitened $\{|\phi_\theta\rangle\}_{\theta\in(-\pi,\pi]}$ families of states. Many of these formulae are directly analogous to those in the main text, but we provide them in full here regardless for completeness. Instead of the Fourier transform in Eq. 26 which has $\theta\in\mathbb{R}$ and thus $k\in\mathbb{R}$, we now calculate the Fourier series for $\theta\in(-\pi,\pi]$ and $k\in\mathbb{Z}$ as

$$|\varphi_k\rangle = \frac{1}{2\pi\sqrt{g_k}}\int_{-\pi}^{\pi} \mathrm{d}\theta\; e^{ik\theta}|\psi_\theta\rangle \tag{A1}$$

By analogy to the continuous case in Eq. 27, we define the discrete spectral measure as

$$g_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{d}\theta\; e^{-ik\theta}G(\theta). \tag{A2}$$

Here, we take $|\varphi_k\rangle = 0$ if $g_k = 0$ for a given k. Let $S = \{k\in\mathbb{Z} : g_k > 0\}$ be the support of g. Similarly to Eq. 28, the inner product between Fourier states is thus

$$\langle\varphi_k|\varphi_l\rangle = \delta_{kl}\chi_S(k). \tag{A3}$$

Instead of Eq. 29, the inverse transformation between states is now

$$|\psi_\theta\rangle = \sum_{k\in S} e^{-ik\theta}\sqrt{g_k}|\varphi_k\rangle \tag{A4}$$

and the symbol is $G(\theta) = \sum_{k\in S} e^{ik\theta}g_k$ such that $|\psi_\theta\rangle$ and $G(\theta)$ both have period 2π. Similarly to the continuous case, the discrete spectral measure g_k is a valid probability distribution since it is normalised ($\sum_{k\in S} g_k = 1$) and non-negative due to the positive semi-definiteness of the inner product.

Similarly to the Toeplitz case, for any circulant family of states, there exists an additive unitary and initial state that generates it. The additive unitary $\hat{U}(\theta) = e^{-i\hat{H}\theta}$

for this circulant family is generated by the Hamiltonian $\hat{H} = \sum_{k\in S} k|\varphi_k\rangle\langle\varphi_k|$ with the discrete eigenbasis $\{|\varphi_k\rangle\}_{k\in S}$. If the support of g is partial, $S \subsetneq \mathbb{Z}$, then, e.g., the eigenbasis could only be for $k \geq 0$ if $S = \mathbb{N}_0$ or be finite dimensional if S is finite (we consider both of these cases below). If the initial state is $|\psi_0\rangle = \sum_{k\in S}\sqrt{g_k}|\varphi_k\rangle$, then the final state $|\psi_\theta\rangle = \hat{U}(\theta)|\psi_0\rangle$ agrees with Eq. A4 and the Gram matrix recovers the periodic symbol $G(\theta-\theta') = \langle\psi_\theta|\psi_{\theta'}\rangle$. The QFI is thus

$$\mathcal{I}_Q(\theta) = 4\left(\sum_{k\in S} k^2 g_k - \left[\sum_{k\in S} k g_k\right]^2\right) \tag{A5}$$

which is four times the variance of g_k and similar to Eq. 36.

Instead of Eq. 24, the quantum whitened states are now defined for $\theta \in (-\pi, \pi]$ as

$$|\phi_\theta\rangle = \frac{1}{\sqrt{2\pi}}\sum_{k\in S} e^{-i\theta k}|\varphi_k\rangle \tag{A6}$$

such that the inverse transformation is $|\varphi_k\rangle = \frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} \mathrm{d}\theta\, e^{i\theta k}|\phi_\theta\rangle$. Although these states are not necessarily normalised, the effects $\hat{E}_\theta = |\phi_\theta\rangle\langle\phi_\theta|$ of the quantum whitening POVM $\{\hat{E}_\theta\}_{\theta\in(-\pi,\pi]}$ are normalised since $\int_{-\pi}^{\pi}\mathrm{d}\theta\ \hat{E}_\theta = \hat{\Pi}_S$, where $\hat{\Pi}_S$ is the projector onto the subspace spanned by the eigenbasis of the Hamiltonian $\{|\varphi_k\rangle\}_{k\in S}$. The quantum whitened states are also covariant $\hat{U}(\theta)|\phi_{\theta'}\rangle = |\phi_{\theta+\theta'}\rangle$ as before. Furthermore, if $S = \mathbb{Z}$, then these are the unique covariant and orthogonal states, up to a gauge unitary that can be ignored. The proof of this result is similar to the $S = \mathbb{R}$ case given in Sec. II A 1 just with integrals replaced with sums.

Instead of Eq. 25, the inner product between quantum whitened states is

$$\langle\phi_\theta|\phi_{\theta'}\rangle = \frac{1}{2\pi}\sum_{k\in S} e^{i(\theta-\theta')k} \tag{A7}$$

If $S = \mathbb{Z}$, then the quantum whitened states are orthogonal ($\langle\phi_\theta|\phi_{\theta'}\rangle = \delta(\theta-\theta')$) and we may construct a projective measurement operator describing quantum whitening, $\hat{W} = \int_{-\pi}^{\pi}\mathrm{d}\theta\ \theta|\phi_\theta\rangle\langle\phi_\theta|$. (We can form the same operator in the non-projective case below, but its eigenbasis may be different.) For example, if estimating the rotation of a particle constrained to a one-dimensional ring, then $S = \mathbb{Z}$ since the angular momentum eigenstates are indexed by $\mathbb{Z}$ and quantum whitening is a projective measurement.

If $S \subsetneq \mathbb{Z}$, however, then the quantum whitened states are not orthogonal and form a non-projective POVM. For example, consider the phase estimation of a coherent state in Appendix J. There, $S = \mathbb{N}_0$ since the Fock states are indexed by $\mathbb{N}_0$ such that the whitened states are not orthogonal ($\langle\phi_\theta|\phi_{\theta'}\rangle = \frac{1}{2\pi}[1-e^{i(\theta-\theta')}]^{-1}$). The normalisation of the POVM is then $\int_{-\pi}^{\pi}\mathrm{d}\theta\ \hat{E}_\theta = \hat{\Pi}_S \neq 1$.

For another example where $S \subsetneq \mathbb{Z}$, the relevant covariant measurement of the rotation of a qudit is well known [22]. Let the Hamiltonian be $\hat{H} = \hat{J}_z$ where $\{|\varphi_k\rangle\}_{k\in S}$ are the $\hat{J}_z$ eigenstates for finite $S = \{-j,\ldots,j\}$, the spin is $j = (d-1)/2$, and the dimension is d. The final state is thus $|\psi_\theta\rangle = \sum_{k=-j}^{j} e^{-ik\theta}\sqrt{g_k}|\varphi_k\rangle$ for some g_k determined by the initial state. We assume that the dimension d is odd and the spin j is integer such that the symbol $G(\theta) = \sum_{k=-j}^{j} e^{ik\theta}g_k$ is 2π periodic. (If the dimension d is instead even and the spin j is half-integer, then the system is only 4π periodic and we need to appropriately modify the formalism.) The covariant states $|\phi_\theta\rangle = \frac{1}{\sqrt{2\pi}}\sum_{k=-j}^{j} e^{-i\theta k}|\varphi_k\rangle$ are sometimes called the "quantum Fourier transform" states. Their inner product in Eq. A7 is the following Dirichlet kernel

$$\langle\phi_\theta|\phi_{\theta'}\rangle = \frac{\sin[(\theta-\theta')(j+\frac{1}{2})]}{2\pi\sin(\frac{\theta-\theta'}{2})}.$$

This converges to the Dirac comb $\sum_{n\in\mathbb{Z}}\delta(\theta-\theta'+2\pi n)$ in the limit of $j\to\infty$ corresponding to $S = \mathbb{Z}$. Although the continuous set $\{|\phi_\theta\rangle\}_{\theta\in(-\pi,\pi]}$ is not orthogonal, there is a d-dimensional orthogonal subset given by θ' satisfying $\theta-\theta' = n\pi/(j+\frac{1}{2})$ for $n = -j,\ldots,j$ and any starting starting value of θ. Up to this continuous degree-of-freedom of choosing θ and the normalisation of the covariant states, this d-dimensional orthogonal subset may be called the "quantum Fourier transform" basis. We saw a similar distinction between a continuum of non-orthogonal vectors and a discrete orthogonal subset of them in Sec. I C between the classical sinusoids indexed by a continuum of possible frequencies versus by just the harmonic frequencies.

For any $S \subset \mathbb{Z}$, the inner product between the original and quantum whitened states is similar to Eq. 31 as

$$\langle\phi_\theta|\psi_{\theta'}\rangle = \frac{1}{\sqrt{2\pi}}\sum_{k\in S}\sqrt{g_k}e^{i(\theta-\theta')k}. \tag{A8}$$

The likelihood of measuring θ for a given θ' with quantum whitening is $L(\theta|\theta') = |\langle\phi_\theta|\psi_{\theta'}\rangle|^2$, regardless of whether it is a projective measurement. If $S = \mathbb{Z}$, then Eq. A8 is the function on $(-\pi,\pi]$ whose Fourier series is $\sqrt{g_k/(2\pi)}$ evaluated at $\theta-\theta'$ mod $(-\pi,\pi]$.

The covariant measurement operator $\hat{W}$ is off-diagonal in the eigenbasis of the Hamiltonian with the following matrix coefficients

$$\langle\varphi_k|\hat{W}|\varphi_l\rangle = i\frac{(-1)^{k-l}}{k-l}\chi_S(k)\chi_S(l). \tag{A9}$$

Unlike the $k\in\mathbb{R}$ case, however, the commutator between the quantum whitening operator and Hamiltonian is not canonical, $[\hat{W},\hat{H}] \neq i$, as the commutator is instead

$$[\hat{W},\hat{H}] = i\sum_{k,l\in S;\, k\neq l}|\varphi_k\rangle\langle\varphi_l|(-1)^{k-l}. \tag{A10}$$

This highlights some of the common issues with defining a Hermitian phase operator.

The continuous and discrete k cases have different delta functions since $\int_{-\infty}^{\infty} \mathrm{d}\theta\ e^{ik\theta} = 2\pi\delta(k)$ for the continuous case and $\int_{-\pi}^{\pi} \mathrm{d}\theta\ e^{ik\theta} = 2\pi\delta_{k0}$ for the discrete case. One difference between the cases is then that $\int_{-\infty}^{\infty} \mathrm{d}\theta\ \theta e^{ik\theta} = -i2\pi\delta'(k)$ for the continuous case, but $\int_{-\pi}^{\pi} \mathrm{d}\theta\ \theta e^{ik\theta} = i2\pi(-1)^k/k$ for $k \neq 0$ and zero otherwise for the discrete case. For example, this leads to the difference in commutator above and also the difference in $\hat{W}$ Fourier coefficients between Eq. 43 for the continuous k case and Eq. A9 for the discrete k case.

Although much of the structure of the circulant case with $\theta \in (-\pi, \pi]$ and $k \in \mathbb{Z}$ is similar to the Toeplitz case with $\theta, k \in \mathbb{R}$, two key results are different. First, the quantum whitening operator and Hamiltonian are not canonical conjugates in the full support case. Second, quantum whitening is not necessarily optimal with respect to the BMSE but is optimal for covariant figures-of-merit, which we will show below.

1. Non-covariant figure-of-merit

To show that quantum whitening is not optimal, we need to show that $\hat{W}$ does not solve the Lyapunov equation for the BSLD $\hat{L}$ in Eq. 9. The BMSE in Eq. 5 is not covariant in this case since the MSE is not periodic. This means that although quantum whitening is covariant it need not be optimal here, unlike the $k \in \mathbb{R}$ case, since the covariant measurement result discussed in Sec. II A 1 requires the figure-of-merit to be covariant. We will discuss this further in Appendix A 2.

For a general prior, instead of Eq. 33, the mixed state and Bayesian derivative in Eq. 10 for the circulant case are

$$\hat{\tilde{\rho}} = \sum_{k\in S}\sum_{l\in S} \tilde{\pi}_{k-l}\sqrt{g_k g_l}|\varphi_k\rangle\langle\varphi_l|, \tag{A11}$$
$$\hat{\tilde{\rho}}' = \sum_{k\in S}\sum_{l\in S} \tilde{\varpi}_{k-l}\sqrt{g_k g_l}|\varphi_k\rangle\langle\varphi_l|$$

where the Fourier series of the prior is $\tilde{\pi}_k = \int_{-\pi}^{\pi} \mathrm{d}\theta\ e^{-ik\theta}\pi(\theta)$, the inverse transformation is $\pi(\theta) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} e^{ik\theta}\tilde{\pi}_k$, and we define $\tilde{\varpi}_{k-l} := \int_{-\pi}^{\pi} \mathrm{d}\theta\ \pi(\theta)e^{-i\theta(k-l)}\theta$.

For the flat prior $\pi(\theta) \equiv 1/(2\pi)$, we have that $\tilde{\pi}_{k-l} = \delta_{kl}$ and $\tilde{\varpi}_{k-l} = i(-1)^{k-l}/(k-l)$ for $k \neq l$ and $\tilde{\varpi}_0 = 0$. Similarly to Eq. 38, the mixed state is thus diagonal in the Fourier basis (the eigenbasis of the Hamiltonian) since

$$\hat{\tilde{\rho}} = \sum_{k\in S} g_k|\varphi_k\rangle\langle\varphi_k|. \tag{A12}$$

And, instead of Eq. 39, the Bayesian derivative is now

$$\hat{\tilde{\rho}}' = i\sum_{k,l\in S;\, k\neq l} \frac{(-1)^{k-l}}{k-l}\sqrt{g_k g_l}|\varphi_k\rangle\langle\varphi_l|. \tag{A13}$$

Instead of Eq. 40 or Eq. 42, the BSLD in Eq. 11 for $k \neq l$ is thus

$$\langle\varphi_k|\hat{L}|\varphi_l\rangle = i\frac{(-1)^{k-l}}{k-l}\frac{2\sqrt{g_k g_l}}{g_k+g_l}\chi_S(k)\chi_S(l) \tag{A14}$$

where the diagonal coefficients of $\hat{L}$ vanish like $\hat{\tilde{\rho}}'$. Unlike the $k \in \mathbb{R}$ case, however, we do not have an analogy of the conditions in Eq. 41 for simplifying expressions involving a function times the "derivative" of a delta function, where in the discrete k case we mean functions times $\tilde{\varpi}$. This means that the coefficients of $\hat{W}$ in Eq. A9 and those of $\hat{L}$ in Eq. A14 are not necessarily equal and quantum whitening is not optimal for the BMSE. This result does not indicate how far the covariant measurement may be from optimal, however, such that it may be close to optimal in practice.

2. Covariant figure-of-merit

The issue above is not with the quantum whitening measurement, which is covariant, but with the cost function, which is not. The periodic symmetry of a circulant family of states $\{|\psi_\theta\rangle\}_{\theta\in(-\pi,\pi]}$ means that we should switch from using the MSE in Eq. 4 to some periodic cost function.

In the periodic case, the additive unitary is generated by the Hamiltonian $\hat{H} = \sum_{k=-\infty}^{\infty} k|\varphi_k\rangle\langle\varphi_k|$. The relevant symmetry group is $\{\hat{U}(\theta)\}_{\theta\in(-\pi,\pi]}$ where $\hat{U}(\theta) = e^{-i\hat{H}\theta}$. This unitary is additive such that $\hat{U}(\theta_1)\hat{U}(\theta_2) = \hat{U}(\theta_1+\theta_2 \bmod (-\pi,\pi])$. Previously, we considered the BMSE in Eq. 5 which is

$$\mathrm{BMSE} = \int \mathrm{d}\theta \int \mathrm{d}x\ \pi(\theta)L(x|\theta)(\check{\theta}_x - \theta)^2. \tag{A15}$$

This is one example of the more general figure-of-merit $\bar{C}$ for a given cost function $C(\theta, \check{\theta}_x)$

$$\bar{C} = \int \mathrm{d}\theta \int \mathrm{d}x\ \pi(\theta)L(x|\theta)C(\theta, \check{\theta}_x). \tag{A16}$$

The BMSE uses the quadratic cost function, $C(\theta, \check{\theta}_x) = (\check{\theta}_x - \theta)^2$, which is not covariant, but other cost functions can be. An estimation problem is covariant with respect to the symmetry group $\{\hat{U}(\theta)\}_{\theta\in(-\pi,\pi]}$ if the following two conditions hold:

1. The cost function, $C(\theta_1, \theta_2)$, is group invariant:

$$C(\theta_1, \theta_2) = C(\theta_1 + x \bmod (-\pi,\pi], \theta_2 + x \bmod (-\pi,\pi]). \tag{A17}$$

2. The prior distribution is group invariant, $\pi(\theta + x \bmod (-\pi,\pi]) = \pi(\theta)$.

An estimation problem that satisfies these conditions is covariant. It is then well-established that the optimal

measurement is also covariant with respect to this group action even if it is a non-projective POVM [22, 37, 41, 42].

In our case, for the problem to be covariant, the prior must be uniform, $\pi(\theta) = 1/(2\pi)$. Regarding the cost function, it was shown that any $C(\theta_1,\theta_2) = \sum_{k=-\infty}^{\infty} a_k e^{ik(\theta_1-\theta_2)}$ with $a_k \leq 0$ is a valid covariant cost function [22, 37]. For example, the Holevo variance, $C(\theta_1,\theta_2) = 4\sin^2(\frac{\theta_1-\theta_2}{2})$, is commonly used since it retrieves the standard variance in the limit of $|\theta_1-\theta_2| \ll 1$. Therefore, for any circulant family and periodic cost function, covariant measurements are optimal given a flat prior. For example, this means that quantum whitening is optimal for the phase estimation of the coherent states in Appendix J.

Appendix B: Proof of the fundamental quantum limit in the Toeplitz case

Here, we prove the results in Sec. II C for the fundamental quantum limit.

Let us start by showing that Eq. 33 holds for the mixed state and Bayesian derivative in the Toeplitz case with general prior. Using Eqs. 10 and 29, $\hat{\rho}(\theta) = |\psi_\theta\rangle\langle\psi_\theta|$, and the definition of $\tilde{\pi}$, we can show that

$$\begin{aligned}
\hat{\bar{\rho}} &= \int_{-\infty}^{\infty} \mathrm{d}\theta\ \pi(\theta)\hat{\rho}(\theta) \qquad \text{(B1)}\\
&= \int_{-\infty}^{\infty} \mathrm{d}\theta\ \pi(\theta) \int_S \mathrm{d}k\ e^{-ik\theta}\sqrt{g(k)}|k\rangle \int_S \mathrm{d}l\ e^{il\theta}\sqrt{g(l)}\langle l|\\
&= \int_S \mathrm{d}k \int_S \mathrm{d}l\ \sqrt{g(k)g(l)}|k\rangle\langle l| \int_{-\infty}^{\infty} \mathrm{d}\theta\ \pi(\theta)e^{-i(k-l)\theta}\\
&= \int_S \mathrm{d}k \int_S \mathrm{d}l\ \tilde{\pi}(k-l)\sqrt{g(k)g(l)}|k\rangle\langle l|.
\end{aligned}$$

This shows Eq. 33. The proof for $\hat{\bar{\rho}}'$ is similar but uses the expression for $i\tilde{\pi}'$.

1. BMSE of quantum whitening

Here, we show that Eq. 34 holds for the MBMSE given the uninformative, uniform prior. We also show the associated results in Sec. II C 2. The covariant measurement is optimal such that we need to calculate the BMSE of quantum whitening. First, the likelihood in Eq. 32 implies that the evidence $p(\theta) = \langle\theta|\hat{\bar{\rho}}|\theta\rangle$ is

$$\begin{aligned}
&p(\theta) \qquad \text{(B2)}\\
&= \int_{-\infty}^{\infty} \mathrm{d}\theta'\ \pi(\theta')L(\theta|\theta')\\
&= \frac{1}{(2\pi)^2\delta(0)} \int_{-\infty}^{\infty} \mathrm{d}\theta' \int_S \mathrm{d}k \int_S \mathrm{d}l\ \sqrt{g(k)g(l)}e^{i(\theta-\theta')(k-l)}\\
&= \frac{1}{2\pi\delta(0)} \int_S \mathrm{d}k \int_S \mathrm{d}l\ \sqrt{g(k)g(l)}\delta(k-l)\\
&= \frac{1}{2\pi\delta(0)} \int_S \mathrm{d}k\ g(k)\\
&= \frac{1}{2\pi\delta(0)}
\end{aligned}$$

which is the same as the prior. Bayes's rule in Eq. 3 then implies that the posterior equals the likelihood, i.e. $p(\theta'|\theta) = L(\theta|\theta')$. Second, the optimal estimator is given by the mean of the posterior $\check{\theta}_\theta = 2\pi\delta(0)\langle\theta|\hat{\bar{\rho}}'|\theta\rangle$ which is

$$\begin{aligned}
&\check{\theta}_\theta \qquad \text{(B3)}\\
&= \int_{-\infty}^{\infty} \mathrm{d}\theta'\ \theta' p(\theta'|\theta)\\
&= \int_{-\infty}^{\infty} \mathrm{d}\theta'\ \theta' \frac{1}{2\pi} \int_S \mathrm{d}k \int_S \mathrm{d}l\ \sqrt{g(k)g(l)}e^{i(\theta-\theta')(k-l)}\\
&= \frac{1}{2\pi} \int_S \mathrm{d}k \int_S \mathrm{d}l\ \sqrt{g(k)g(l)}e^{i\theta(k-l)} \int_{-\infty}^{\infty} \mathrm{d}\theta'\ \theta' e^{-i\theta'(k-l)}\\
&= \frac{1}{2\pi} \int_S \mathrm{d}k \int_S \mathrm{d}l\ \sqrt{g(k)g(l)}e^{i\theta(k-l)}[-i2\pi\delta'(l-k)]\\
&= i\int_S \mathrm{d}l \int_S \mathrm{d}x\ \sqrt{g(x+l)g(l)}e^{i\theta x}\delta'(x)\\
&= -i\int_S \mathrm{d}l\ \partial_x\left[\sqrt{g(x+l)g(l)}e^{i\theta x}\right]_{x=0}\\
&= -i\int_S \mathrm{d}l\ \left[i\theta g(l) + \frac{1}{2}g'(l)\right]\\
&= \theta.
\end{aligned}$$

Here, we assume that g is differentiable and goes to zero on the boundary of S such that $\int_S \mathrm{d}l\ g'(l) = 0$. Finally, we calculate the MSE as a function of the parameter θ' similarly as

$$\begin{aligned}
&\mathrm{MSE}(\theta') \qquad \text{(B4)}\\
&= \int_{-\infty}^{\infty} \mathrm{d}\theta\ L(\theta|\theta')(\check{\theta}_\theta - \theta')^2\\
&= \int_{-\infty}^{\infty} \mathrm{d}\theta\ \frac{1}{2\pi} \int_S \mathrm{d}k \int_S \mathrm{d}l\ \sqrt{g(k)g(l)}e^{i(\theta-\theta')(k-l)}(\theta-\theta')^2\\
&= \frac{1}{2\pi} \int_S \mathrm{d}k \int_S \mathrm{d}l\ \sqrt{g(k)g(l)} \int_{-\infty}^{\infty} \mathrm{d}\theta\ e^{i(\theta-\theta')(k-l)}(\theta-\theta')^2\\
&= \frac{1}{2\pi} \int_S \mathrm{d}k \int_S \mathrm{d}l\ \sqrt{g(k)g(l)}[-2\pi\delta''(k-l)]\\
&= -\int_S \mathrm{d}l \int_S \mathrm{d}x\ \sqrt{g(x+l)g(l)}\delta''(x)\\
&= -\int_S \mathrm{d}l\ \partial_x^2\left[\sqrt{g(x+l)g(l)}\right]_{x=0}\\
&= -\int_S \mathrm{d}l\ \left(\frac{1}{2}g''(l) - \frac{g'(l)^2}{4g(l)}\right)\\
&= \frac{1}{4}\int_S \mathrm{d}l\ \frac{g'(l)^2}{g(l)}.
\end{aligned}$$

Here, we now also assume that g is twice differentiable and its derivative goes to zero on the boundary of S such that $\int_S \mathrm{d}l\ g''(l) = 0$. These assumptions break for

some of the applications that we study which is why the MBMSE diverges there. Since the MSE is independent of the parameter, it is equal to the BMSE and MBMSE which proves Eq. 34.

2. Optimality of quantum whitening

Here, we provide further details for the independent proof of the optimality of quantum whitening in Sec. II C 4. The mixed state $\hat{\tilde{\rho}}$ and its Bayesian derivative $\hat{\tilde{\rho}}'$ in Eqs. 38 and 39 come directly from applying the result for a general prior in Eq. 33 to the flat prior $(\tilde{\pi}(k-l) = \delta(k-l)/\delta(0))$. By the inner product in Eq. 28, the matrix coefficients of the mixed state in Eq. 38 are $\langle k|\hat{\tilde{\rho}}|l\rangle = [1/\delta(0)]g(k)\delta(k-l)$. The BSLD $\hat{L}$ in Eq. 40 then comes directly from Eq. 11 since the mixed state is diagonal in the Fourier basis, taking care to avoid indices where the eigenvalues are zero.

The conditions in Eq. 41 determine when two functions multiplied by the derivative of the Dirac delta function are equal in the weak sense. They can be shown by integrating by parts against a smooth test function $\xi(k,l)$ as follows

$$\begin{aligned}
&\int_{-\infty}^{\infty} \mathrm{d}k \int_{-\infty}^{\infty} \mathrm{d}l\ \delta'(k-l) f_1(k,l)\xi(k,l) \qquad \text{(B5)}\\
&= \int_{-\infty}^{\infty} \mathrm{d}l \int_{-\infty}^{\infty} \mathrm{d}x\ \delta'(x) f_1(x+l,l)\xi(x+l,l)\\
&= -\int_{-\infty}^{\infty} \mathrm{d}l\ \partial_x \left[f_1(x+l,l)\xi(x+l,l)\right]_{x=0}\\
&= -\int_{-\infty}^{\infty} \mathrm{d}l\ \left[\partial_x f_1(x+l,l)|_{x=0}\xi(l,l)\right.\\
&\qquad \left. + f_1(l,l)\partial_x \xi(x+l,l)|_{x=0}\right]
\end{aligned}$$

such that this equals $\int_{-\infty}^{\infty}\mathrm{d}k \int_{-\infty}^{\infty}\mathrm{d}l\ \delta'(k-l)f_2(k,l)\xi(k,l)$ for arbitrary $\xi(k,l)$ if the first two conditions in Eq. 41 hold. The third condition on the partial derivative with respect to the second slot can be shown similarly by substituting $l = k - x$ instead of $k = x + l$. Direct computation of the partial derivatives of the coefficients of the BSLD in Eq. 40 then shows that the BSLD is given by Eq. 42.

We want to show that the quantum whitening operator $\hat{W}$ equals the BSLD and satisfies the Lyapunov equation. We can always form the covariant operator, $\hat{W} = \int_{-\infty}^{\infty}\mathrm{d}\theta\ \theta|\theta\rangle\langle\theta|$, regardless of S and whether the covariant states are orthogonal as discussed in Sec. II A 1. First, let us show that the coefficients of $\hat{W}$ are given by Eq. 43. The inner product between the covariant and Fourier states is $\langle k|\theta\rangle = \frac{1}{\sqrt{2\pi}}e^{-i\theta k}\chi_S(k)$ by Eq. 24. The coefficients of $\hat{W}$ are thus

$$\begin{aligned}
\langle k|\hat{W}|l\rangle &= \langle k|\int_{-\infty}^{\infty}\mathrm{d}\theta\ \theta|\theta\rangle\langle\theta|l\rangle \qquad \text{(B6)}\\
&= \int_{-\infty}^{\infty}\mathrm{d}\theta\ \theta\frac{1}{2\pi}e^{-i\theta(k-l)}\chi_S(k)\chi_S(l)\\
&= \frac{1}{2\pi}\chi_S(k)\chi_S(l)\int_{-\infty}^{\infty}\mathrm{d}\theta\ \theta e^{-i\theta(k-l)}\\
&= i\delta'(k-l)\chi_S(k)\chi_S(l).
\end{aligned}$$

This shows Eq. 43. Second, let us show that $\hat{W}$ satisfies the Lyapunov equation by proving that Eq. 44 holds for the anticommutator $\frac{1}{2}\{\hat{\tilde{\rho}},\hat{W}\}$. We start by computing the coefficients for the product $\hat{\tilde{\rho}}\hat{W}$ using Eqs. 28 and 38 as

$$\begin{aligned}
&\langle k|\hat{\tilde{\rho}}\hat{W}|l\rangle \qquad \text{(B7)}\\
&= \langle k|[1/\delta(0)]\int_S \mathrm{d}j\ g(j)|j\rangle\langle j|\int_{-\infty}^{\infty}\mathrm{d}\theta\ \theta|\theta\rangle\langle\theta|l\rangle\\
&= [1/\delta(0)]g(k)\chi_S(k)\langle k|\int_{-\infty}^{\infty}\mathrm{d}\theta\ \theta|\theta\rangle\langle\theta|l\rangle\\
&= [1/\delta(0)]g(k)\chi_S(k)\langle k|\hat{W}|l\rangle\\
&= i[\delta'(k-l)/\delta(0)]g(k)\chi_S(k)\chi_S(l).
\end{aligned}$$

The conjugate product $\hat{W}\hat{\tilde{\rho}}$ is thus

$$\langle k|\hat{W}\hat{\tilde{\rho}}|l\rangle = i[\delta'(k-l)/\delta(0)]g(l)\chi_S(k)\chi_S(l) \qquad \text{(B8)}$$

such that the coefficients of the anticommutator $\frac{1}{2}\{\hat{\tilde{\rho}},\hat{W}\}$ are

$$\langle k|\frac{1}{2}\{\hat{\tilde{\rho}},\hat{W}\}|l\rangle = \frac{1}{2}i[\delta'(k-l)/\delta(0)][g(k)+g(l)]\chi_S(k)\chi_S(l) \qquad \text{(B9)}$$

which proves Eq. 44 since the characteristic functions $\chi_S(k)\chi_S(l)$ are redundant. The Lyapunov equation in Eq. 9 is thus satisfied by $\hat{W}$ since the anticommutator $\frac{1}{2}\{\hat{\tilde{\rho}},\hat{W}\}$ equals the Bayesian derivative $\hat{\tilde{\rho}}'$ in Eq. 39 by Eq. 41.

Appendix C: Quantum whitening of single-particle states

In the single-particle state case, quantum whitening is the same as classically whitening the corresponding classical waveforms. We show this now. Suppose that the common classical and quantum geometry in Eq. 47 is Toeplitz such that there exists some function $G : \mathbb{R} \to \mathbb{C}$ such that

$$\langle x_\theta|x_{\theta'}\rangle = \int_{-\infty}^{\infty}\mathrm{d}t\ x_\theta^*(t)x_{\theta'}(t) = G(\theta-\theta') \qquad \text{(C1)}$$

where G is positive semi-definite and $G(0) = 1$. For the single-particle state case, the Fourier states in Eq. 26 are $|k\rangle = \hat{c}_k^\dagger|0\rangle$ where the Fourier annihilation operator is defined as $\hat{c}_k = \int_{-\infty}^{\infty}\mathrm{d}t\ y_k^*(t)\hat{a}(t)$ and the Fourier waveform is

$$y_k(t) = \frac{1}{2\pi\sqrt{g(k)}}\int_{-\infty}^{\infty}\mathrm{d}\theta\ e^{ik\theta}x_\theta(t) \qquad \text{(C2)}$$

such that the inverse transformation is $x_\theta(t) = \int_{-\infty}^{\infty} \mathrm{d}k\ e^{-ik\theta}\sqrt{g(k)}y_k(t)$. Since the Fourier states are orthogonal, we have that

$$\langle k|l\rangle = [\hat{c}_k, \hat{c}_l^\dagger] = \int_{-\infty}^{\infty} \mathrm{d}t\ y_k^*(t)y_l(t) = \delta(k-l). \quad \text{(C3)}$$

The Fourier states are not normalised single-particle states, however, since their norm is $\int_{-\infty}^{\infty}\mathrm{d}t\ |y_k(t)|^2 = \delta(0)$ whereas we require that $\int_{-\infty}^{\infty}\mathrm{d}t\ |x_\theta(t)|^2 = 1$. Informally, we could consider the normalised single-particle state corresponding to $y_k(t)/\sqrt{\delta(0)}$ but this division needs to be treated with the upmost mathematical care. The fact that we cannot have a continuous set of quantum state or classical waveforms that are orthogonal and have unit norm is a property of the continuum and thus is expected here.

Similarly, for the single-particle state case, the quantum whitened states in Eq. 24 are $|\theta\rangle = \hat{d}_\theta^\dagger|0\rangle$ where the whitened annihilation operator is defined as $\hat{d}_\theta = \int_{-\infty}^{\infty}\mathrm{d}t\ z_\theta^*(t)\hat{a}(t)$ and the classically whitened waveform is

$$z_\theta(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}k\ e^{-i\theta k}y_k(t) \quad \text{(C4)}$$

such that the inverse transformation is $y_k(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}\theta\ e^{i\theta k}z_\theta(t)$. The whitened state are also orthogonal such that

$$\langle\theta|\theta'\rangle = [\hat{d}_\theta, \hat{d}_{\theta'}^\dagger] = \int_{-\infty}^{\infty}\mathrm{d}t\ z_\theta^*(t)z_{\theta'}(t) = \delta(\theta-\theta'). \quad \text{(C5)}$$

The classical inner product between the original and whitened waveforms is the same as the quantum inner product in Eq. 31 between the original and whitened states such that

$$\int_{-\infty}^{\infty}\mathrm{d}t\ z_\theta^*(t)x_{\theta'}(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}k\ \sqrt{g(k)}e^{i(\theta-\theta')k}. \quad \text{(C6)}$$

The quantum whitened states are also not single-photon states in the sense of Eq. 46 since they are not normalised since $\int_{-\infty}^{\infty}\mathrm{d}t\ |z_\theta(t)|^2 = \delta(0)$. We have shown that classical and quantum whitening are the same for the single-particle state case.

From the original classical waveforms $\{x_\theta : \mathbb{R} \to \mathbb{C}\}_{\theta\in\mathbb{R}}$, we have here defined two continuous families of orthogonal waveforms: the Fourier waveforms $\{y_k : \mathbb{R} \to \mathbb{C}\}_{k\in\mathbb{R}}$ and classically whitened waveforms $\{z_\theta : \mathbb{R} \to \mathbb{C}\}_{\theta\in\mathbb{R}}$. Just as we defined two bases for the subspace of possible quantum states, these are bases for the $L^2(\mathbb{R})$ subspace of possible classical waveforms.

Appendix D: Bayesian estimation of displacement

Let us consider the Bayesian problem of estimating the displacement μ of a single harmonic oscillator along a known direction. Without loss of generality, the final state is a coherent state with coherent amplitude $\alpha = i\mu/\sqrt{2}$ such that $\langle\hat{x}\rangle = 0$ and $\langle\hat{p}\rangle = \mu$ where $[\hat{x}, \hat{p}] = i$. We now calculate the MBMSE, the BMSEs for quadrature and number-resolving measurements, the difference between amplitude and power esitmation, and the wide-prior limit.

1. MBMSE

In the case of a Gaussian prior, $\pi(\mu) = e^{-\mu^2/(2\sigma^2)}/(\sigma\sqrt{2\pi})$ with variance σ^2, then the MBMSE in Eq. 7 becomes [13]

$$\Delta^2\mu = \sigma^2 - \sigma^4\mathcal{I}_Q^{\hat{\varrho}}(\mu) \quad \text{(D1)}$$

where $\mathcal{I}_Q^{\hat{\varrho}}(\mu)$ is the QFI with respect to μ of the displaced mixed state $\hat{\varrho} = e^{i\mu\hat{x}}\hat{\rho}e^{-i\mu\hat{x}}$. This should be distinguished from the QFI of $\hat{\rho}(\mu)$ for a given μ which is $\mathcal{I}_Q^{\hat{\rho}(\mu)}(\mu) = 2$. The mixed state $\hat{\rho}$ in Eq. 10 is a Gaussian state with zero mean and covariance matrix $\Sigma = \mathrm{diag}\left(\frac{1}{2}, \frac{1}{2}+\sigma^2\right)$ such that the displaced mixed state $\hat{\varrho}$ is a Gaussian state with mean vector $\vec{\mu} = (0,\mu)^\mathrm{T}$ and the same covariance matrix Σ. Thus, the QFI in Eq. 20 is $\mathcal{I}_Q^{\hat{\varrho}}(\mu) = \frac{2}{1+2\sigma^2}$, which is independent of μ as expected, and the MBMSE in Eq. D1 is

$$\Delta^2\mu = \frac{\sigma^2}{1+2\sigma^2}. \quad \text{(D2)}$$

In the wide-prior limit of $\sigma^2 \gg \frac{1}{2}$, the MBMSE is $\Delta^2\mu \approx \frac{1}{2}$ whereas the MBMSE is $\Delta^2\mu \approx \sigma^2$ in the narrow-prior limit of $\sigma^2 \ll \frac{1}{2}$.

2. Quadrature measurement

Suppose that we perform a quadrature measurement $\hat{p}$ on $\hat{\rho}(\mu)$ and obtain a measurement result $\mathfrak{p}$ with a likelihood of $L(\mathfrak{p}|\mu) \sim \mathcal{N}(\mu, \frac{1}{\sqrt{2}})$. Then, the evidence is also Gaussian $p(\mathfrak{p}) \sim \mathcal{N}(0,\varsigma)$ with zero mean and variance $\varsigma^2 = \frac{1}{2}+\sigma^2$, and the posterior is Gaussian $p(\mu|\mathfrak{p}) \sim \mathcal{N}(\tilde{\mu}_\mathfrak{p}, \sqrt{V_\mathrm{post}})$ with mean $\tilde{\mu}_\mathfrak{p} = \frac{2\sigma^2}{1+2\sigma^2}\mathfrak{p}$ and variance $V_\mathrm{post} = \frac{\sigma^2}{1+2\sigma^2}$ from Eq. 6 which is independent of $\mathfrak{p}$. The BMSE in Eq. 5 of quadrature measurement is then just the posterior variance V_post which agrees with marginalising $\mathrm{MSE}(\mu) = \frac{\mu^2+2\sigma^4}{(2\sigma^2+1)^2}$ over the prior $\pi(\mu) \sim \mathcal{N}(0,\sigma)$. This value equals the MBMSE in Eq. D2 meaning that quadrature measurement is optimal. This measurement is biased because of the prior since, for a given μ and σ, we obtain $\mathfrak{p}$ with likelihood $L(\mathfrak{p}|\mu)$ and estimate $\check{\mu}_\mathfrak{p}$ whose expected value is $\check{\mu}_\mu = \frac{2\mu\sigma^2}{1+2\sigma^2} \neq \mu$. In the wide-prior limit of $\sigma^2 \gg \frac{1}{2}$, the posterior distribution is simply the likelihood, $p(\mu|\mathfrak{p}) = L(\mathfrak{p}|\mu)$, and the bias vanishes since $\check{\mu}_\mu = \mu$.

3. Number-resolving measurement

In comparison, suppose that we perform a number-resolving measurement $\hat{n}$ and obtain a measurement result $\mathfrak{n}$ with a Poissonian likelihood of $L(\mathfrak{n}|\mu) = e^{-\frac{\mu^2}{2}}\frac{\mu^{2\mathfrak{n}}}{2^{\mathfrak{n}}\mathfrak{n}!}$ such that the posterior distribution is

$$p(\mu|\mathfrak{n}) = \frac{\mu^{2\mathfrak{n}}\left(\sigma^2+1\right)^{\mathfrak{n}+\frac{1}{2}}}{2^{\mathfrak{n}+\frac{1}{2}}\Gamma\left(\mathfrak{n}+\frac{1}{2}\right)\sigma^{2\mathfrak{n}+1}}e^{-\frac{\mu^2\left(\sigma^2+1\right)}{2\sigma^2}} \tag{D3}$$

which has zero mean independent of the measurement result $\mathfrak{n}$. This means that the optimal estimator $\tilde{\mu}_{\mathfrak{n}} = 0$ contains no additional information about the displacement μ and thus the BMSE equals the prior variance σ^2. This is because the likelihood is even in μ, i.e. $L(\mathfrak{n}|\mu) = L(\mathfrak{n}|-\mu)$, such that we cannot distinguish between μ and $-\mu$ with a number-resolving measurement $\hat{n}$.

4. Power and magnitude estimation

To handle this issue with the sign of the displacement μ, suppose that we instead wish to estimate the power $u = \mu^2$ or magnitude $v = |\mu|$ of the displacement. The Gaussian prior on μ implies that the power follows a scaled chi-squared distribution, $\pi(u) = \frac{1}{\sigma\sqrt{2\pi u}}e^{-u/(2\sigma^2)}$ with expected value $\mathrm{E}[u] = \sigma^2$ and variance $\mathrm{Var}[u] = 2\sigma^4$, and the magnitude follows a half-Gaussian distribution, $\pi(v) = \frac{2}{\sigma\sqrt{2\pi}}e^{-v^2/(2\sigma^2)}$ with expected value $\mathrm{E}[v] = \sigma\sqrt{2/\pi}$ and variance $\mathrm{Var}[v] = \sigma^2(1-2/\pi)$. We estimate the deviation of each parameter from the expected value such that the prior is centred. These prior distributions are non-Gaussian such that Eq. D1 does not apply and we must instead solve Eq. 7 which we have been unable to do analytically. However, a numerical analysis suggests that quadrature measurement $\hat{p}$ does not attain the MBMSE for estimating the power μ^2 or magnitude $|\mu|$ despite attaining the MBMSE with respect to μ. We also see that a number-resolving measurement $\hat{n}$ attains the MBMSE for estimating the power μ^2 but not the magnitude $|\mu|$ where a different measurement from either $\hat{p}$ or $\hat{n}$ is required. This is different to the Fisher limit where a change of variables from θ to θ' only introduces a chain rule factor of $(\frac{\mathrm{d}\theta}{\mathrm{d}\theta'})^2$ in the CFI and QFI and does not change the attainability of bounds. Furthermore, we observe that quadrature measurement does not attain the MBMSE with respect to μ in the case of a flat prior rather than a Gaussian one such that a different measurement from $\hat{p}$ or $\hat{n}$ is again required. We defer understanding these different measurements and the behaviour of the BMSE and MBMSE under changes of variables more generally to future work.

5. Wide-prior limit

Returning to the case of estimating μ, we now demonstrate that the wide-prior limit of $\sigma^2 \gg \frac{1}{2}$ mentioned briefly above needs to be handled carefully. This limit represents the uninformative, uniform prior on $\mathbb{R}$ that is nevertheless normalised, i.e. $\partial_\mu\pi(\mu) = 0$ and $\int_{-\infty}^{\infty}\mathrm{d}\mu\,\pi(\mu) = 1$. For the quadrature measurement likelihood of $L(\mathfrak{p}|\mu) \sim \mathcal{N}(\mu, \frac{1}{\sqrt{2}})$, this is the Jeffreys prior for the mean μ [43]. In the direct calculation of the BMSE for quadrature measurement above, or for quantum whitening in Sec. II C 1, we find that the posterior is normalised and can have finite variance in some cases despite the prior. This value of the posterior variance and MBMSE can be calculated by just assuming that the prior is constant on $\mathbb{R}$ and normalised, e.g. $\pi(\mu) \equiv \varepsilon$ for some $\varepsilon > 0$. Since the Fourier transform of the prior is $\tilde{\pi}(k) = e^{-\frac{1}{2}\sigma^2k^2}$ for finite σ which limits to $\tilde{\pi}(k) = \delta(k)/\delta(0) = \delta_{k0}$, then we could assign the prior the infinitesimal value of $\varepsilon = 1/[2\pi\delta(0)]$ but we emphasise that the exact value does not matter, just that the prior is constant and normalised.

Alternatively, we could use an improper, i.e. unnormalised, prior such as $\pi(\mu) \equiv 1$. Then, we find that the posterior is still normalised and has constant variance equal to the MSE, independent of the measurement result. The issue is then that the BMSE in Eq. 5 is proportional to the prior's normalisation which diverges. Similarly, the MSE can be finite but the evidence is improper. This divergence is artificial, however, since it simply comes from how we have choose to represent a total lack of knowledge about the parameter μ as an uninformative prior. As such, we prefer the first method where the prior is chosen to be constant on $\mathbb{R}$ and normalised from the start.

We have shown above that the BMSE of quadrature measurement achieves the MBMSE of $\frac{1}{2}$ in the wide-prior limit and that the quadrature measurement is unbiased. If we naïvely calculated the MBMSE from Eq. 7 instead, however, then we could have ran into an issue if we did not Taylor expand to sufficiently high order. Here, we assume that we have already independently shown that the quadrature measurement is optimal and wish to verify the MBMSE. For finite σ, the BSLD in Eq. 8 is $\hat{L} = \frac{2\sigma^2}{1+2\sigma^2}\hat{p}$ with $\mathrm{Tr}[\hat{\rho}\hat{L}^2] = \frac{2\sigma^4}{1+2\sigma^2}$ such that the MBMSE in Eq. 7 agrees with Eq. D2. If we Taylor expand these expressions to first order in large $\sigma^2 \gg \frac{1}{2}$, then we get that $\hat{L} \approx [1-1/(2\sigma^2)]\hat{p}$ such that $\mathrm{Tr}[\hat{\rho}\hat{L}^2] \approx \sigma^2 - \frac{1}{2} - 1/(2\sigma^2) \approx \sigma^2 - \frac{1}{2}$ and the MBMSE in Eq. 7 is $\Delta^2\mu \approx \frac{1}{2}$ which agrees with Eq. D2. If we instead naïvely only Taylor expand the expressions to zeroth order in large σ, then get that $\hat{L} \approx \hat{p}$ such that $\mathrm{Tr}[\hat{\rho}\hat{L}^2] \approx \sigma^2 + \frac{1}{2} \approx \sigma^2$ and the MBMSE in Eq. 7 is zero. The MBMSE similarly vanishes if we expand $\mathrm{Tr}[\hat{\rho}\hat{L}^2] = \frac{2\sigma^4}{1+2\sigma^2} \approx \sigma^2$ to zeroth order. These results are erroneous, however, since they come from mishandling the indeterminate form $\infty-\infty$, i.e. we have not expanded

the expressions to sufficient order and expanded them too early to safely cancel the prior variance of $\Delta^2\mu^{(0)} = \sigma^2$ in Eq. 7. For this toy problem, therefore, we have shown that either we need to calculate the BMSE directly from Eq. 5 or that we need to take a sufficiently high-order expansion in the prior variance to calculate the MBMSE from Eq. 7.

These results are relevant to our general calculation in Sec. II C 1 where we cannot calculate the MBMSE for finite prior variance and then take the wide-prior limit like we did above for the toy problem. Instead, we work exclusively in the wide-prior limit and directly calculate the BMSE for the optimal measurement in Eq. 5 rather than naïvely apply Eq. 7. We have motivated this here since the former is always well-defined but the latter has the indeterminate form $\infty-\infty$ unless the limit is properly taken.

Appendix E: Inner product of coherent states

We have two independent ways of showing Eq. 56 for the inner product between multi-mode coherent states. The first way is to extract the temporal mode which is excited for each coherent state. The second way is to discretise the time domain like we might do numerically. We provide both proofs here for completeness.

1. Extracting the temporal mode

Let us start with the direct calculation without discretisation. We use the annihilation operator $\hat{b}_\theta$ from Eq. 46 to express the coherent state in Eq. 54 as follows

$$\begin{aligned} |h_\theta\rangle &= e^{\alpha_\theta \hat{b}_\theta^\dagger - \alpha_\theta^* \hat{b}_\theta}|0\rangle \\ &= e^{-\frac{1}{2}|\alpha_\theta|^2} e^{\alpha_\theta \hat{b}_\theta^\dagger} e^{-\alpha_\theta^* \hat{b}_\theta}|0\rangle \\ &= e^{-\frac{1}{2}|\alpha_\theta|^2} e^{\alpha_\theta \hat{b}_\theta^\dagger}|0\rangle. \end{aligned} \tag{E1}$$

The average number of particles in this coherent state is

$$\begin{aligned} \bar{N}_\theta &= \langle h_\theta|\hat{N}|h_\theta\rangle \\ &= \int_{-\infty}^{\infty} \mathrm{d}t\ \langle h_\theta|\hat{a}^\dagger(t)\hat{a}(t)|h_\theta\rangle \\ &= e^{-|\alpha_\theta|^2} \int_{-\infty}^{\infty} \mathrm{d}t\ \langle 0|e^{\alpha_\theta^* \hat{b}_\theta} \hat{a}^\dagger(t)\hat{a}(t) e^{\alpha_\theta \hat{b}_\theta^\dagger}|0\rangle \\ &= e^{-|\alpha_\theta|^2}\ |\alpha_\theta|^2\ \langle 0|e^{\alpha_\theta^* \hat{b}_\theta} e^{\alpha_\theta \hat{b}_\theta^\dagger}|0\rangle \\ &= |\alpha_\theta|^2 \\ &= \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{d}t\ |h_\theta(t)|^2 \end{aligned} \tag{E2}$$

where we used that $e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A},\hat{B}]$ for central $[\hat{A},\hat{B}]$. This result is expected since $\bar{N}_\theta = |\alpha_\theta|^2$ is the average number of particles for a single-mode coherent state with amplitude α_θ. Similarly, we can show that $\hat{a}(t)|h_\theta\rangle = \frac{i}{\sqrt{2}} h_\theta(t)|h_\theta\rangle$ and that Eq. 55 holds for the expectation values of the quadratures. The inner product between two coherent states is thus

$$\begin{aligned} \langle h_\theta|h_{\theta'}\rangle &= e^{-\frac{1}{2}\left(|\alpha_\theta|^2+|\alpha_{\theta'}|^2\right)} \langle 0|e^{\alpha_\theta^* \hat{b}_\theta} e^{\alpha_{\theta'} \hat{b}_{\theta'}^\dagger}|0\rangle \\ &= e^{-\frac{1}{2}\left(|\alpha_\theta|^2+|\alpha_{\theta'}|^2\right)+\alpha_\theta^*\alpha_{\theta'}[\hat{b}_\theta,\hat{b}_{\theta'}^\dagger]} \langle 0|e^{\alpha_{\theta'} \hat{b}_{\theta'}^\dagger} e^{\alpha_\theta^* \hat{b}_\theta}|0\rangle \\ &= e^{-\frac{1}{2}\left(|\alpha_\theta|^2+|\alpha_{\theta'}|^2\right)+\alpha_\theta^*\alpha_{\theta'} \int_{-\infty}^{\infty} \mathrm{d}t\ f_\theta^*(t) f_{\theta'}(t)} \\ &= e^{-\frac{1}{2}\left[\bar{N}_\theta+\bar{N}_{\theta'}-\int_{-\infty}^{\infty} \mathrm{d}t\ h_\theta^*(t)h_{\theta'}(t)\right]} \end{aligned} \tag{E3}$$

where we used that $e^{\hat{A}}e^{\hat{B}} = e^{[\hat{A},\hat{B}]}e^{\hat{B}}e^{\hat{A}}$ for central $[\hat{A},\hat{B}]$. This completes the proof of Eq. 56.

We evaluated the diagonal matrix coefficients of the number operator above in the overcomplete coherent state basis for the subspace. Meanwhile, the off-diagonal matrix coefficients of the number operator are

$$\langle h_\theta|\hat{N}|h_{\theta'}\rangle = \frac{1}{2}\langle h_\theta|h_{\theta'}\rangle \int_{-\infty}^{\infty} \mathrm{d}t\ h_\theta^*(t)h_{\theta'}(t) \tag{E4}$$

which is half of the product of the quantum and classical inner products.

2. Discretising the time domain

Let us now instead discretise the bath of outgoing modes by coarse-graining them according to a short sampling time δt such that the time domain $\mathbb{R}$ is broken into time bins $(j\delta t, (j+t)\delta t)$ indexed by $j \in \mathbb{Z}$. Let the annihilation operator of the jth time bin be

$$\hat{c}_j = \frac{1}{\sqrt{\delta t}} \int_{j\delta t}^{(j+1)\delta t} \mathrm{d}t\ \hat{a}(t) \tag{E5}$$

where $[\hat{c}_j, \hat{a}^\dagger(t)] = \frac{1}{\sqrt{\delta t}}\chi_{(j\delta t,(j+1)\delta t)}(t)$ and $[\hat{c}_j, \hat{c}_k^\dagger] = \delta_{jk}$ for $j,k \in \mathbb{Z}$. (These should be distinguished from the annihilation operator of the Fourier modes $\hat{c}_k$ defined above Eq. C2.) We now coarse-grain the outgoing state $|h_\theta\rangle$ in Eq. 54 as follows

$$\begin{aligned} |h_\theta\rangle &= \bigotimes_{j=-\infty}^{\infty} e^{\frac{i}{\sqrt{2}}\int_{j\delta t}^{(j+1)\delta t} \mathrm{d}t\ \left[h_\theta(t)\hat{a}^\dagger(t)+h_\theta^*(t)\hat{a}(t)\right]}|0\rangle \\ &\approx \bigotimes_{j=-\infty}^{\infty} e^{\alpha_j \hat{c}_j^\dagger - \alpha_j^* \hat{c}_j}|0\rangle \\ &= \bigotimes_{j=-\infty}^{\infty} e^{-\frac{1}{2}|\alpha_j|^2} e^{\alpha_j \hat{c}_j^\dagger}|0\rangle. \end{aligned} \tag{E6}$$

where $\alpha_j := \frac{i}{\sqrt{2}}\sqrt{\delta t}\, h_\theta(j\delta t)$ is the coherent amplitude of the jth time bin and is dimensionless since $h_\theta(t)$ has units of square root Hertz. Here, we have discretised the signal $h_\theta(t)$ with its values at the left endpoints $h_\theta(j\delta t)$ which is a good approximation assuming that the sampling rate is fast compared to the signal.

The coherent state is an eigenstate of the annihilation operator which can shown as follows

$$\begin{aligned}\hat{a}(t)|h_\theta\rangle &\approx \hat{a}(t)\bigotimes_{j=-\infty}^{\infty} e^{-\frac{1}{2}|\alpha_j|^2}e^{\alpha_j\hat{c}_j^\dagger}|0\rangle \\ &= \left[\frac{1}{\sqrt{\delta t}}\sum_{j=-\infty}^{\infty}\alpha_j\chi_{(j\delta t,(j+1)\delta t)}(t)\right]|h_\theta\rangle\end{aligned} \tag{E7}$$

where we again used that $e^{\hat{A}}\hat{B}e^{-\hat{A}} = \hat{B} + [\hat{A},\hat{B}]$ for central $[\hat{A},\hat{B}]$. The average number of particles is thus

$$\begin{aligned}\bar{N}_\theta &= \langle h_\theta|\hat{N}|h_\theta\rangle \\ &= \int_{-\infty}^{\infty}\mathrm{d}t\ \langle h_\theta|\hat{a}^\dagger(t)\hat{a}(t)|h_\theta\rangle \\ &\approx \int_{-\infty}^{\infty}\mathrm{d}t\ \frac{1}{\delta t}\sum_{j=-\infty}^{\infty}|\alpha_j|^2\chi_{(j\delta t,(j+1)\delta t)}(t) \\ &= \sum_{j=-\infty}^{\infty}|\alpha_j|^2 \\ &\xrightarrow[\delta t\to 0]{} \frac{1}{2}\int_{-\infty}^{\infty}\mathrm{d}t\ |h_\theta(t)|^2\,.\end{aligned} \tag{E8}$$

The inner product between coherent states is thus

$$\begin{aligned}\langle h_\theta|h_{\theta'}\rangle &\approx \prod_{j=-\infty}^{\infty}\langle\alpha_j(\theta)|\alpha_j(\theta')\rangle \\ &= e^{-\frac{1}{2}\sum_{j=-\infty}^{\infty}[|\alpha_j(\theta)|^2+|\alpha_j(\theta')|^2-2\alpha_j(\theta)^*\alpha_j(\theta')]} \\ &\xrightarrow[\delta t\to 0]{} e^{-\frac{1}{2}\left[\bar{N}_\theta+\bar{N}_{\theta'}-\int_{-\infty}^{\infty}\mathrm{d}t\ h_\theta^*(t)h_{\theta'}(t)\right]}\end{aligned} \tag{E9}$$

such that we obtain Eq. 56 in the continuum limit.

The mean values of the quadratures of the jth time bin are $\langle\hat{x}_j\rangle = \sqrt{2}\,\mathrm{Re}\,[\alpha_j]$ and $\langle\hat{p}_j\rangle = \sqrt{2}\,\mathrm{Im}\,[\alpha_j]$ where $\hat{x}_j = (\hat{c}_j + \hat{c}_j^\dagger)/\sqrt{2}$ and $\hat{p}_j = (-i\hat{c}_j + i\hat{c}_j^\dagger)/\sqrt{2}$ such that $[\hat{x}_j,\hat{p}_k] = i\delta_{jk}$. This shows Eq. 55 in the continuum limit. For example, for the sinusoidal signal in Eq. 67, $\langle\hat{p}_j\rangle = A\sqrt{\delta t}\,\cos(\omega j\delta t+\phi)$ oscillates in time as shown in Fig. 1 such that the individual SNR of the jth oscillator is $A\sqrt{2\delta t}$ since $\mathrm{Var}[\hat{p}_j] = \frac{1}{2}$ and the integrated SNR from the $M = T/\delta t$ oscillators within $(0,T)$ is $A\sqrt{2T}$.

Let us briefly comment on discretising the time domain in the case of frequency estimation. Similarly to how the long integration time T sets the separation between the harmonic frequencies $\omega_{n+1}-\omega_n = 2\pi/T$, the sampling frequency $f_s = 1/\delta t$ sets the Nyquist frequency $f_N = f_s/2 = 1/(2\delta t)$ as the highest possible frequency to estimate without aliasing. This is also the meaning of assuming above that the sampling rate is fast compared to the signal. To calculate the inner product $\langle h_\omega|h_{\omega'}\rangle$ between two different frequencies ω and ω' we require that $\max(\omega,\omega') \ll 1/(2\delta t)$ to avoid aliasing. For example, if we want to estimate a frequency ω within the interval $(\omega_0-\Delta\omega/2,\omega_0+\Delta\omega/2)$, then we require that $\omega_0+\Delta\omega/2 < 2\pi f_N = \pi/\delta t$. Since we also want to avoid negative frequencies, then we also require that $\omega_0-\Delta\omega/2 > 0$ such that we have that $\Delta\omega\delta t < \pi$. This condition is a constraint on the maximum phase difference accumulated over a single time bin between two different frequencies. Provided that this condition is satisfied, then we are safe to discretise the outgoing bath in time. For example, in Fig. 3, we have that $\Delta\omega\delta t = 0.18\pi$.

Appendix F: Shifted window for sinusoid frequency estimation

The fact that we windowed the sinusoidal signal in Eq. 67 to the particular time interval $(0,T)$ with a square pulse is only a choice of convention. We could instead use a Gaussian window of width T and the results should be similar as we show later for the complex exponential case. We can also shift the time interval, for example, to $(-T/2,T/2)$ such that the signal becomes

$$h_\omega(t) = A\cos(\omega t+\phi)\chi_{(-T/2,T/2)}(t), \tag{F1}$$

then the inner product between states becomes

$$\langle h_\omega|h_{\omega'}\rangle \approx e^{-\frac{1}{4}A^2T(1-\mathrm{sinc}[(\omega-\omega')T/2])} \tag{F2}$$

where we again assume the many-cycles limit. Whereas the sinc function in the exponent of Eq. 72 vanished for harmonic frequency spacing of $\omega-\omega' = n\pi/T$ for $(0,T)$, here the frequency spacing is $\omega-\omega' = 2n\pi/T$ for $(-T/2,T/2)$.

Although the Gram matrices $G_{\omega,\omega'} = \langle h_\omega|h_{\omega'}\rangle$ in Eq. 72 and Eq. F2 for $(0,T)$ and $(-T/2,T/2)$, respectively, are not identical, they have a similar functional form. In particular, they are generally functions of ω and ω' but become Toeplitz in the many-cycles limit.

The QFI for frequency estimation of the signal defined on $(-T/2,T/2)$ in Eq. F1 is

$$\mathcal{I}_Q(\omega) \approx \frac{1}{12}A^2T^3 \tag{F3}$$

where we take the many-cycles limit of $\omega T \gg 1$. This simply differs by a factor of four from the QFI defined on $(0,T)$ in Eq. 73.

Appendix G: Frequency estimation of a complex exponential

Let us now consider the following example of a complex waveform which corresponds to a rotating coherent state

$$h_\omega(t) = -iAe^{i(\omega t+\phi)}\chi_{(0,T)}(t) \tag{G1}$$

such that the average number in Eq. 58 is $\bar{N}_\omega = A^2T/2$ and the quadrature expectation values in Eq. 55 are

$$\langle h_\theta|\hat{x}(t)|h_\theta\rangle = A\cos(\omega t+\phi)\chi_{(0,T)}(t), \tag{G2}$$

$$\langle h_\theta|\hat{p}(t)|h_\theta\rangle = A\sin(\omega t+\phi)\chi_{(0,T)}(t).$$

The $L^2(\mathbb{R})$ inner product between classical waveforms is $\int_{-\infty}^{\infty}\mathrm{d}t\ h_\omega^*(t)h_{\omega'}(t) = A^2\delta_T(\omega'-\omega)$ where $\delta_T(x) = (e^{ixT}-1)/(ix)$ is a finite time approximation to the Dirac delta function obeying $\delta_T(0)=T$ and $\delta_T(x)=\int_0^T \mathrm{d}t\ e^{ixt}$. The inner product between states in Eq. 56 is thus

$$\langle h_\omega|h_{\omega'}\rangle = e^{-\frac{1}{2}A^2\left[T-\delta_T(\omega'-\omega)\right]}. \tag{G3}$$

This Gram matrix is Hermitian but not real. In general, the Gram matrix is always Hermitian and positive semi-definite such that its eigenvalues are real and non-negative. If we shift the time interval to $(-T/2,T/2)$ instead of $(0,T)$, then the signal becomes

$$h_\omega(t) = -iAe^{i(\omega t+\phi)}\chi_{(-T/2,T/2)}(t) \tag{G4}$$

such that the inner product is instead

$$\langle h_\omega|h_{\omega'}\rangle = e^{-\frac{1}{2}A^2T(1-\mathrm{sinc}\left[(\omega-\omega')T/2\right])}. \tag{G5}$$

This is the same as the sinusoidal case in Eq. F2 except for an overall factor of two in the exponent. This is likely because the complex exponential is composed of sinusoidal real and imaginary parts. Unlike the sinusoidal case, however, the Gram matrix is Toeplitz here even outside of the many-cycles limit.

The QFI in Eq. 63 for frequency estimation of the complex exponential signal defined on $(0,T)$ in Eq. G3 is

$$\mathcal{I}_Q(\omega) = \frac{2}{3}A^2T^3 \tag{G6}$$

and for the signal defined on $(-T/2,T/2)$ in Eq. G4 is

$$\mathcal{I}_Q(\omega) = \frac{1}{6}A^2T^3. \tag{G7}$$

These are twice the sinusoidal QFI in Eq. 73 and Eq. F3, respectively, because of the sinusoidal real and imaginary parts here.

1. Alternate window functions

We assumed a rectangular window function above but the results are similar with other window functions. Let us show this now for Gaussian and double-exponential window functions. Although we do not do this for the later amplitude and phase estimation cases, the results should again be similar to the results there with the rectangular window.

a. *Frequency estimation of a sine-Gaussian*

Consider the following complex sine-Gaussian signal

$$h_\omega(t) = Ae^{-(t-t_0)^2/(2\sigma^2)+i\omega(t-t_0)+i\phi} \tag{G8}$$

where A is the amplitude, σ is the standard deviation of the Gaussian in units of time, t_0 is the centre time, and ϕ is the phase. We assume that these four parameters are known and we just want to estimate the frequency ω. Suppose that this signal drives a coherent state such that the average number in Eq. 58 is $\bar{N}_\omega = \sqrt{\pi}A^2\sigma/2$. The inner product between coherent states in Eq. 57 is then

$$\langle h_\omega|h_{\omega'}\rangle = e^{-\frac{1}{2}\sqrt{\pi}A^2\sigma\left(1-e^{-\frac{1}{4}\sigma^2(\omega-\omega')^2}\right)} \tag{G9}$$

which is asymptotically $e^{-\bar{N}_\omega}$ for large frequency differences. The QFI in Eq. 63 is thus

$$\mathcal{I}_Q(\omega) = \sqrt{\pi}A^2\sigma^3 = 2\bar{N}_\omega\sigma^2. \tag{G10}$$

This has the same scaling as Eqs. G6–G7 given $\sigma\sim T$.

b. *Frequency estimation of a Lorentzian*

Let us now consider a double-exponential window function:

$$h_\omega(t) = Ae^{-\gamma|t-t_0|+i\omega(t-t_0)+i\phi} \tag{G11}$$

where γ is the decay rate which is known. The Fourier transform of this signal is a Lorentzian since

$$H_\omega(\Omega) = \int_{-\infty}^{\infty}\mathrm{d}t\ e^{-i\Omega t}h_\omega(t) = \frac{2A\gamma e^{-i\Omega t_0+i\phi}}{\gamma^2+(\Omega-\omega)^2}. \tag{G12}$$

This is similar to how fact that the Laplace and Cauchy distributions are related by the Fourier transform.

Similarly, the average number in Eq. 58 of the coherent state is $\bar{N}_\omega = A^2/(2\gamma)$. And, the inner product in Eq. 57 is

$$\langle h_\omega|h_{\omega'}\rangle = \exp\left(-\frac{A^2(\omega-\omega')^2}{2\gamma\left[4\gamma^2+(\omega-\omega')^2\right]}\right) \tag{G13}$$

The QFI in Eq. 63 is also

$$\mathcal{I}_Q(\omega) = \frac{A^2}{\gamma^3} = \frac{2\bar{N}_\omega}{\gamma^2} \tag{G14}$$

which has the same scaling again given $\gamma\sim 1/T$.

In comparison, the inner product between single-particle states in Eq. 47 for $x_\omega(t) = \sqrt{\gamma}h_\omega(t)/A$ is

$$\langle x_\omega|x_{\omega'}\rangle = \frac{4\gamma^2}{4\gamma^2+(\omega-\omega')^2} \tag{G15}$$

such that the QFI in Eq. 53 is $\mathcal{I}_Q(\omega) = \frac{2}{\gamma^2}$. This is the same as the QFI above for the coherent state with an average of one particle ($\bar{N}_\omega = 1$). Meanwhile, the likelihood is $L(\theta|\theta') = \frac{2\gamma^3}{\pi[\gamma^2+(\theta-\theta')^2]^2}$ such that the MBMSE in Eq. 34 for the single-particle case is $\mathcal{V} = \gamma^2$ from the MSE. Alternatively, from the dual estimation problem, the prior Fisher information of $g(k) = \gamma e^{-2\gamma|k|}$ with respect to k is $\mathcal{I}_{\text{prior}} = 4\gamma^2$ and the MBMSE is the quarter of this value. Whereas, the MBMSE for the coherent state case diverges since the coherent state inner product is asymptotically nonzero.

Appendix H: Amplitude estimation

We can also consider instead estimating the amplitude A for fixed frequency ω and phase ϕ. In general, suppose that the classical waveform has the following form for a real scale parameter A to estimate and some square-integrable complex function $g(t)$ defined on $\mathbb{R}$

$$h_A(t) = Ag(t) \tag{H1}$$

such that the inner product in Eq. 57 is thus

$$\langle h_A|h_{A'}\rangle = e^{-\frac{1}{2}\bar{N}_1(A-A')^2} \tag{H2}$$

which is Toeplitz and an unnormalised Gaussian with variance $1/\bar{N}_1$ where $\bar{N}_1 = \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{d}t\ |g(t)|^2$. The average number in Eq. 58 is $\bar{N}_A = A^2\bar{N}_1$. The QFI in Eq. 63 is

$$\mathcal{I}_Q(A) = 4\bar{N}_1. \tag{H3}$$

In the temporal mode associated with $g(t)$, the state is a coherent state with amplitude $\alpha_A = A\sqrt{\bar{N}_1}e^{i\varphi}$ for some φ as shown in Appendix E. Without loss of generality, let $\varphi = \pi/2$ such that $\alpha_A = i\mu/\sqrt{2}$ where $\mu = A\sqrt{2\bar{N}_1}$. The MBMSE in Eq. D2 with respect to μ implies that the MBMSE with respect to A is

$$\Delta^2 A = \frac{\sigma^2}{1+4\bar{N}_1\sigma^2} \tag{H4}$$

given a Gaussian prior on A with variance σ^2. The MBMSE is attained by measuring the phase quadrature $\hat{p}$ of the temporal mode associated with $g(t)$. If $4\bar{N}_1\sigma^2 \gg 1$, then the MBMSE is $\Delta^2 A = 1/(4\bar{N}_1)$ which corresponds either to the wide-prior limit of $\sigma^2 \gg 1/\bar{N}_1$ for fixed $\bar{N}_1$ or the high-SNR limit of $\bar{N}_1 \gg 1/\sigma^2$ for fixed σ such that the MBMSE achieves the QCRB. Conversely, if $4\bar{N}_1\sigma^2 \ll 1$, then the MBMSE remains at the prior variance $\Delta^2 A = \sigma^2$ as the information from the prior dominates the information from the measurement. This corresponds to either the narrow-prior limit of $\sigma^2 \ll 1/\bar{N}_1$ for fixed $\bar{N}_1$ or the low-SNR limit of $\bar{N}_1 \ll 1/\sigma^2$ for fixed σ.

For example, in the sinusoidal case, the exact Gram matrix in Eq. H2 is different for the signal defined on $(0,T)$ in Eq. 67 and $(-T/2,T/2)$ in Eq. F1 since the average numbers are different. In the many-cycles limit of $T \gg 1/\omega$, the matrix is equal for both cases to the following

$$\langle h_A|h_{A'}\rangle \approx e^{-\frac{1}{8}(A-A')^2T} \tag{H5}$$

The average number in Eq. 58 is $\bar{N}_A \approx A^2T/4$ with $\bar{N}_1 \approx T/4$ in this limit such that the QFI in Eq. H3 is

$$\mathcal{I}_Q(A) \approx T. \tag{H6}$$

In the complex exponential case, with the signal defined on either $(0,T)$ in Eq. G1 or $(-T/2,T/2)$ in Eq. G4, the Gram matrix in Eq. H2 is

$$\langle h_A|h_{A'}\rangle = e^{-\frac{1}{4}(A-A')^2T} \tag{H7}$$

since the average number is $\bar{N}_A = A^2T/2$ with $\bar{N}_1 = T/2$ such that the QFI is

$$\mathcal{I}_Q(A) = 2T \tag{H8}$$

where the usual factor of two in comparison to the sinusoid case appears.

Appendix I: Estimating the centre of a Gaussian pulse

The following is a different example which might have been expected to yield a Gaussian Gram matrix similarly to amplitude estimation in Appendix H above. Instead, it behaves much more like frequency estimation of a complex sine-Gaussian in Eq. G9 or in the wide-prior limit in the main text.

Suppose that we are estimating the centre of a Gaussian pulse that drives a coherent state. Let $x_\theta(t)$ be given in Eq. 80 and $h_\theta(t) = -i\sqrt{2}\alpha x_\theta(t)$. The average number is $\bar{N}_\theta = \alpha^2$ and the Gram matrix is

$$G(\theta) = \exp\left(-\alpha^2\left[1-\exp\left(-\frac{\theta^2}{8\sigma^2}\right)\right]\right) \tag{I1}$$

which is asymptotically $e^{-\alpha^2}$. This constant term leads to a divergence in $g(k)$ at DC and thus quadratic divergence of the posterior variance similarly to the frequency estimation case, except that the discontinuities from the first-order term of $G(\theta)$ seen in the frequency estimation case do not appear here since the variance of the Gaussian is finite unlike the sinc function. The MBMSE thus diverges. In comparison, the QCRB is $\mathcal{I}_Q(\theta)^{-1} = \sigma^2/\alpha^2$ which is finite. If the average number of particles is one such that $\alpha = 1$, then the QCRB is σ^2 which agrees with the value for the single-particle case. This is because the second term in Eq. 53 vanishes since $x_\theta(t)$ is real.

Appendix J: Phase estimation

Let us now estimate the phase ϕ for fixed amplitude A and frequency ω. In the sinusoidal case, the exact Gram matrix is not Toeplitz and is different for the signal defined on $(0,T)$ in Eq. 67 and $(-T/2,T/2)$ in Eq. F1, but in the many-cycles limit of $T \gg 1/\omega$ the matrix is Toeplitz and equal for both cases to the following

$$\langle h_\phi|h_{\phi'}\rangle \approx e^{-\frac{1}{2}A^2T\sin^2[(\phi-\phi')/2]}. \tag{J1}$$

The average number is $\bar{N}_\phi \approx A^2T/4$ and the QFI is the following in this limit

$$\mathcal{I}_Q(\phi) \approx A^2T. \tag{J2}$$

In the complex exponential case, with the signal defined on either $(0,T)$ in Eq. G1 or $(-T/2,T/2)$ in Eq. G4 but with $\phi \to -\phi$ in either case, the Gram matrix is

$$\langle h_\phi|h_{\phi'}\rangle = e^{-\frac{1}{2}A^2T\left(1-e^{i(\phi-\phi')}\right)} \tag{J3}$$

where the average number is $\bar{N}_\phi = A^2T/2$ and the QFI is

$$\mathcal{I}_Q(\phi) = 2A^2T. \tag{J4}$$

without assuming the many-cycles limit. Similarly to Eq. G3, this Gram matrix is Hermitian and positive semi-definite but not real. Unlike for frequency estimation, the Gram matrices in Eq. J1 and Eq. J3 are not only Toeplitz but also circulant since they are 2π-periodic in $\phi - \phi'$. Matrices which are circulant are necessarily Toeplitz.

We took $\phi \to -\phi$ in the complex exponential case above in Eq. J3 so that the corresponding spectral measure in Eq. A2 would consist of non-negative frequencies, $S = \mathbb{N}_0$. Given a quarter of the SNR squared $\frac{1}{2}A^2T$, $g_k = e^{-\frac{1}{2}A^2T}\frac{1}{k!}(\frac{1}{2}A^2T)^k$ is Poissonian if $k \geq 0$ and $g_k = 0$ if $k < 0$. This $\phi \to -\phi$ case corresponds to $\hat{U}(\phi) = e^{-i\phi\hat{N}}$ such that $\hat{H} = \hat{N}$ instead of $\hat{U}(\phi) = e^{i\phi\hat{N}}$ and $\hat{H} = -\hat{N}$.

Regarding the BMSE, the periodic symmetry of this circulant family of states $\{|h_\phi\rangle\}_{\phi\in(-\pi,\pi]}$ means that we should switch from using the MSE in Eq. 4 to a periodic cost function. We discuss this in Appendix A 2.

Appendix K: Numerical calculation of the MBMSE

In general, the Gram matrix may not be Toeplitz, e.g. in Eq. 70, such that the analytic results in the main text do not apply. Instead, we can calculate the MBMSE numerically as follows. We focus on frequency estimation here but the procedure is general. Suppose that we want to estimate a frequency within $(\omega_0 - \Delta\omega/2, \omega_0 + \Delta\omega/2)$ and use a finite numerical frequency resolution $\delta\omega$ such that there are $n = \Delta\omega/\delta\omega$ waveforms or frequency bins. For example, the Gram matrix G_{nn} is n-by-n dimensional where we make the dimension explicit here for clarity. We find a basis for the waveform subspace by taking the spectral decomposition of the Gram matrix $G_{nn} = U_{nn}S_{nn}U_{nn}^\dagger$ where U_{nn} is unitary and S_{nn} is diagonal. Since G is Hermitian and positive-semidefinite, its eigenvalues (the entries of S_{nn}) are real and non-negative. Let A_{Dn} be the matrix whose n columns are the waveform states $\{|h_\omega\rangle\}_{\omega\in\mathbb{R}}$ each of which is dimension $D = d^M$ such that the Gram matrix is $G_{nn} = A_{nD}^\dagger A_{Dn}$. Here, we discretise the bath in time into M oscillators with truncated dimension d as in Appendix E 2. The goal of working within the waveform subspace is to avoid representing these D-dimensional states directly since the Hilbert space is too large and grows exponentially. We can calculate the Gram matrix for the coherent states numerically by using the analytic result in Eq. 56 without needing to explicitly represent the states in the large D-dimensional space. The spectral decomposition of G_{nn} is the singular value decomposition (SVD) of A_{Dn}. This decomposition is more numerically stable than, e.g., directly performing the Gram-Schmidt process to A_{Dn}.

We may be concerned that improving the numerical frequency resolution $\delta\omega$ will dramatically increase the number of waveforms $n = \Delta\omega/\delta\omega$. However, we observe numerically that the eigenvalues of the Gram matrix quickly decay. For a given numerical threshold, below which the eigenvalues are set to zero, the rank $r \ll n$ of S_{nn} remains stable even as n increases. We may then use the truncated SVD to approximate the Gram matrix as $G_{nn} \approx U_{nr}S_{rr}U_{rn}^\dagger$ by keeping only the eigenvalues above the threshold. Here, U_{nr} is now only semi-unitary since $U_{rn}^\dagger U_{nr} = I_r$ but $U_{nr}U_{rn}^\dagger = P_{nn}$ where P_{nn} is the Hermitian projection onto the column space of U_{nr}. This decomposition is the optimal low-rank approximation to G_{nn} in the sense of minimising the Frobenius norm by the Eckart-Young theorem. The rank r remains finite because as the frequency resolution shrinks the neighbouring frequencies become less orthogonal for finite SNR, i.e. $\langle h_\omega|h_{\omega+\delta\omega}\rangle \to 1$ as $\delta\omega \to 0$ in Eq. 75.

Given the truncated SVD, we can now find a low-rank representation of the waveform states. Let $B_{Dr} = A_{Dn}U_{nr}S_{rr}^{-1/2}$ such that the inverse transformation is $A_{Dn} = B_{Dr}S_{rr}^{1/2}U_{rn}^\dagger$. The r columns of B_{Dr} are orthogonal since $B_{rD}^\dagger B_{Dr} \approx I_r$ given that $A_{nD}^\dagger A_{Dn} \approx U_{nr}S_{rr}U_{rn}^\dagger$. Let the kth column of B_{Dr} be $|\varphi_k\rangle$ for $1 \leq k \leq r$. The transformation between the waveform states $|h_j\rangle$ and these states $|\varphi_k\rangle$ for $1 \leq j \leq n$ and $1 \leq k \leq r$ is

$$|h_j\rangle = \sum_{l=1}^{r}(U_{nr})_{jl}^*(S_{rr})_l^{1/2}|\varphi_l\rangle \tag{K1}$$

$$|\varphi_k\rangle = (S_{rr})_k^{-1/2}\sum_{m=1}^{n}(U_{nr})_{mk}|h_m\rangle.$$

Here, the normalisation of $|\varphi_k\rangle$ includes $(S_{rr})_k^{-1/2}$ which is well-defined since we truncated the SVD to only include positive eigenvalues above the numerical noise floor. We can numerically calculate the matrix coefficients of the mixed state and its Bayesian derivative in Eq. 10 in this basis and thus find the MBMSE in Eq. 12. The BSLD in Eq. 11 is an r-by-r matrix L_{rr} such that we can determine the coefficients of the optimal measurement basis $\{|x\rangle\}_x$ in Eq. 8 in the basis. We can then represent a given $|x\rangle$ as a superposition of the original waveform states, i.e. $|x\rangle = \sum_{j=1}^{n} c_{jx}|h_j\rangle$.

In the Toeplitz case for an infinite continuous family of states, we showed that the Gram matrix is diagonalised by the quantum Fourier transform such that the eigenvalues are $g(k)$ and the eigenstates are the Fourier states $|k\rangle$ and, provided that the prior is uninformative and uniform, that the optimal measurement basis is the quantum whitened superposition of the original waveform states. For finite Toeplitz matrices, the situation is more complicated and requires Szegő's theorem to understand the asymptotic relation between the eigenvalue distribution and the Fourier transform [44]. This is because Toeplitz matrices are only in general asymptotically diagonalised by the Fourier transform unlike finite circulant matrices.

We have considered only families of pure states throughout this work, but most of the techniques will

also work for families of mixed states. In particular, we may want to study the effect of marginalising over unknown parameters rather than assume that they are known. For example, we could estimate the frequency ω of a sinusoidal signal encoded in a coherent state with unknown amplitude A and phase ϕ, where we uniformly marginalise the amplitude over $A \in (0, A_0)$ and phase over $\phi \in (-\pi, \pi)$. In general, suppose that the pure state is $|h_{\theta,\vec{\phi}}\rangle$ for a given parameter θ that we wish to estimate and a vector of unknown parameters $\vec{\phi}$. Our knowledge of the device is thus captured by the classical mixture $\hat{\rho}(\theta) = \int \mathrm{d}\vec{\phi}\ p(\vec{\phi})|h_{\theta,\vec{\phi}}\rangle\langle h_{\theta,\vec{\phi}}|$ and we want to study the family of states $\{\hat{\rho}(\theta)\}_{\theta\in\mathbb{R}}$. These density matrices are bounded operators on the waveform subspace spanned by the possible pure states $|h_{\theta,\vec{\phi}}\rangle$. This means that we can first apply the SVD approach to $\{|h_{\theta,\vec{\phi}}\rangle\}_{\theta,\vec{\phi}}$ to find a basis for the (hopefully low-dimensional) waveform subspace and then represent $\hat{\rho}(\theta)$ in this basis numerically. The procedure to calculate the MBMSE or QFI is then the same as before.

[1] E. Payne et al. (in preparation), Prospects for post-merger remnant signal inference with photon-counting interferometry (2025).

[2] P. D. Lasky, Gravitational wave astronomy, in *Multimessenger Astronomy in Practice*, 2514-3433 (IOP Publishing, 2021) pp. 9–1 to 9–72.

[3] N. Sarin and P. D. Lasky, Multimessenger astronomy with a khz-band gravitational-wave observatory, Publ. Astron. Soc. Aust. **39**, e007 (2022).

[4] J. E. Kim and G. Carosi, Axions and the strong CP problem, Rev. Mod. Phys. **82**, 557 (2010).

[5] K. Choi, S. H. Im, and C. S. Shin, Recent progress in the physics of axions and axion-like particles, Annu. Rev. Nucl. Part. Sci. **71**, 225 (2021).

[6] L. J. Rosenberg and K. A. Van Bibber, Searches for invisible axions, Phys. Rep. **325**, 1 (2000).

[7] P. W. Graham, I. G. Irastorza, S. K. Lamoreaux, A. Lindner, and K. A. van Bibber, Experimental searches for the axion and axion-like particles, Annu. Rev. Nucl. Part. Sci. **65**, 485 (2015).

[8] R. D. Peccei, The strong cp problem and axions, in *Axions: Theory, Cosmology, and Experimental Searches*, edited by M. Kuster, G. Raffelt, and B. Beltrán (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008) pp. 3–17.

[9] M. Battaglieri, A. Belloni, A. Chou, P. Cushman, B. Echenard, R. Essig, J. Estrada, J. L. Feng, B. Flaugher, P. J. Fox, et al., Us cosmic visions: new ideas in dark matter 2017: community report, arXiv preprint arXiv:1707.04591 (2017).

[10] R. Kubo, The fluctuation-dissipation theorem, Rep. Prog. Phys. **29**, 255 (1966).

[11] A. Buonanno and Y. Chen, Signal recycled laser-interferometer gravitational-wave detectors as optical springs, Phys. Rev. D **65**, 042001 (2002).

[12] C. Bond, D. Brown, A. Freise, and K. A. Strain, Interferometer techniques for gravitational-wave detection, Living Rev. Relativ. **19**, 1 (2016).

[13] K. Macieszczak, M. Fraas, and R. Demkowicz-Dobrzański, Bayesian quantum frequency estimation in presence of collective dephasing, New Journal of Physics **16**, 113002 (2014).

[14] M. Jarzyna and R. Demkowicz-Dobrzański, True precision limits in quantum metrology, New Journal of Physics **17**, 013010 (2015).

[15] J. Rubio and J. Dunningham, Quantum metrology in the presence of limited data, New Journal of Physics **21**, 043037 (2019).

[16] R. Kaubruegger, D. V. Vasilyev, M. Schulte, K. Hammerer, and P. Zoller, Quantum variational optimization of ramsey interferometry and atomic clocks, Physical review X **11**, 041045 (2021).

[17] C. D. Marciniak, T. Feldker, I. Pogorelov, R. Kaubruegger, D. V. Vasilyev, R. van Bijnen, P. Schindler, P. Zoller, R. Blatt, and T. Monz, Optimal metrology with programmable quantum sensors, Nature **603**, 604 (2022).

[18] S. Direkci, R. Finkelstein, M. Endres, and T. Gefen, Heisenberg-limited bayesian phase estimation with low-depth digital quantum circuits, arXiv preprint arXiv:2407.06006 (2024).

[19] M. Tsang, H. M. Wiseman, and C. M. Caves, Fundamental Quantum Limit to Waveform Estimation, Phys. Rev. Lett. **106**, 090401 (2011).

[20] H. Miao, R. X. Adhikari, Y. Ma, B. Pang, and Y. Chen, Towards the fundamental quantum limit of linear measurements of classical signals, Phys. Rev. Lett. **119**, 050801 (2017).

[21] J. W. Gardner, T. Gefen, S. A. Haine, J. J. Hope, and Y. Chen, Achieving the fundamental quantum limit of linear waveform estimation, Phys. Rev. Lett. **132**, 130801 (2024).

[22] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (Springer Science & Business Media, 2011).

[23] D. V. Vasilyev, A. Shankar, R. Kaubruegger, and P. Zoller, Optimal multiparameter metrology: The quantum compass solution, arXiv preprint arXiv:2404.14194 (2024).

[24] A. Monras, Phase space formalism for quantum estimation of Gaussian states (2013), arxiv:1303.3682 [quant-ph].

[25] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part IV* (John Wiley & Sons, Incorporated, 2002).

[26] R. D. Gill and B. Y. Levit, Applications of the van Trees inequality: a Bayesian Cramér-Rao bound, Bernoulli , 59 (1995).

[27] M. G. Genoni, M. G. Paris, G. Adesso, H. Nha, P. L. Knight, and M. Kim, Optimal estimation of joint parameters in phase space, Physical Review A—Atomic, Molecular, and Optical Physics **87**, 012107 (2013).

[28] J. W. Gardner, T. Gefen, S. A. Haine, J. J. Hope, J. Preskill, Y. Chen, and L. McCuller, Stochastic waveform estimation at the fundamental quantum limit, PRX Quantum **6**, 030311 (2025).

[29] J. W. Gardner, S. A. Haine, J. J. Hope, Y. Chen, and T. Gefen, Lindblad estimation with fast and precise quantum control, arXiv preprint arXiv:2501.03364 (2025).
[30] D. Rife and R. Boorstyn, Single tone parameter estimation from discrete-time observations, IEEE Trans. Inf. Theory **20**, 591 (1974).
[31] A. Steinhardt and C. Bretherton, Thresholds in frequency estimation, in *ICASSP 85 IEEE Int. Conf. Acoust. Speech Signal Process.*, Vol. 10 (1985) pp. 1273–1276.
[32] B. James, B. D. Anderson, and R. C. Williamson, Characterization of threshold for single tone maximum likelihood frequency estimation, IEEE transactions on signal processing **43**, 817 (2002).
[33] L. Knockaert, The Barankin bound and threshold behavior in frequency estimation, IEEE Trans. Signal Process. **45**, 2398 (1997).
[34] S. Schmitt, T. Gefen, F. M. Stürner, T. Unden, G. Wolff, C. Müller, J. Scheuer, B. Naydenov, M. Markham, S. Pezzagna, et al., Submillihertz magnetic spectroscopy performed with a nanoscale quantum sensor, Science **356**, 832 (2017).
[35] S. Schmitt, T. Gefen, D. Louzon, C. Osterkamp, N. Staudenmaier, J. Lang, M. Markham, A. Retzker, L. P. McGuinness, and F. Jelezko, Optimal frequency measurements with quantum probes, npj Quantum Information **7**, 55 (2021).
[36] A. Holevo, Estimation of shift parameters of a quantum state, Reports on Mathematical Physics **13**, 379 (1978).
[37] G. Chiribella, G. D'ariano, and M. F. Sacchi, Optimal estimation of group transformations using entanglement, Physical Review A—Atomic, Molecular, and Optical Physics **72**, 042338 (2005).
[38] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, 2009).
[39] S. Direkci et al. (in preparation), Extending the dynamic range in atomic clocks with weak measurements (2025).
[40] J. W. Gardner. *boldBat*. 2025. `https://git.ligo.org/jameswalter.gardner/boldbat`.
[41] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, Quantum limits in optical interferometry, Progress in Optics **60**, 345 (2015).
[42] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Reference frames, superselection rules, and quantum information, Reviews of Modern Physics **79**, 555 (2007).
[43] H. Jeffreys, An invariant form for the prior probability in estimation problems, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **186**, 453 (1946).
[44] A. Böttcher and S. M. Grudsky, *Toeplitz matrices, asymptotic linear algebra and functional analysis*, Vol. 67 (Springer, 2000).

Chapter 7

Conclusions and outlook

But release me from my bands
With the help of your good hands:
Gentle breath of yours my sails
Must fill, or else my project fails,
Which was to please.

Prospero, (Epilogue)
William Shakespeare, The Tempest

We now draw the conclusions of this thesis and suggest the outlook for future work in the landscape of waveform estimation.

7.1 Conclusions

This thesis presents the case that modern quantum information science offers new directions to accelerate searches for fundamental physics such as gravitational waves, dark matter, and quantum gravity. These searches are necessary to answer questions across cosmology, particle physics, nuclear physics, and astrophysics. Instead of designing our sensors while informed by outdated classical ways of thinking, we show that quantum-enhanced devices with radically non-classical modes of operation can enable dramatic sensitivity improvements. This arc of increasingly non-classical designs, however, meets a tradeoff of ever more complicated experimental proposals. We see this dual progression of gain and complexity in the different papers presented herein as we optimise the relevant quantum information theoretic quantities first over just the final measurement scheme,

then also the initial quantum state, and finally the intervening quantum control as well. In particular, we show that non-classical protocols are required to optimally search for nonstationary, stochastic, or wide-prior waveform signals. The sensitivity gains made possible by our proposals are measured in orders of magnitude in the right parameter regimes. The key conclusion of this thesis is not the worth of any individual proposal, however, but is to adopt this way of thinking of our sensors as manipulators of quantum information and to always optimise them as such. This may be the only way to enable the searches for fundamental physics to advance our deepest understanding of nature over the next century.

7.2 Outlook

Let us first make an analogy to the progress of audioband squeezing for gravitational-wave detectors. Squeezing for LIGO was first proposed 44 years ago [58]. What will the quantum metrology of gravitational-wave detection be in four decades time? The argument of this thesis is that it should look dramatically different to how we think about the detectors today, but it may take that long to realise our proposals. Similarly, searches for dark matter and quantum gravity should also be reimagined to achieve the fundamental quantum limits of sensitivity.

The ideas presented in this thesis address only a small region of the landscape of waveform estimation shown in Fig. 1.1 which, furthermore, omits the third axis that extends from single-parameter estimation through multi-parameter estimation to many-parameter estimation. The broader avenues of future work, therefore, are naturally laid out as the remaining unknown regions of this landscape, besides the individual outlook within each paper above. Firstly, sensing Non-Markovian stochastic signals applies to, e.g., dark matter searches where the axion's coherence time is similar to the experiment's own. Understanding the transition from approximately Markovian stochastic signals to deterministic ones depending on the ratio of the coherence times is particularly interesting. Secondly, developing methods for optimisation assuming quantum local asymptotic normality should replace the neighbourhood-blind Fisher approach. Finally, no fundamental

tight bounds exist in the Bayesian regime for problems with multi-parameter estimation or quantum control, despite the importance of understanding both of these for experimental design. Addressing this gap in the literature would be most useful.

As a final word, consider that we are only now reaching the quantum noise–limited regime of our sensors in the physical sciences. We may not yet imagine what other applications lie across the life and earth sciences when they too inevitably and inexorably reach the quantum regime. May we be ready to aid them when it happens.

Bibliography

[1] H.-Y. Chen, M. Fishbach, and D. E. Holz. 2018. Nature, 562(7728):545–547.

[2] C. Messenger, K. Takami, S. Gossan, L. Rezzolla, and B. S. Sathyaprakash. 2014. Physical Review X, 4(4):041004.

[3] E. Pian. 2021. Front. Astron. Space Sci., 7.

[4] N. Sarin and P. D. Lasky. 2022. Publ. Astron. Soc. Aust., 39:e007.

[5] D. Martynov, H. Miao, H. Yang, F. H. Vivanco, E. Thrane, R. Smith, P. Lasky, W. E. East, R. Adhikari, A. Bauswein, et al. 2019. Phys. Rev. D, 99(10):102004.

[6] P. D. Lasky. In Multimessenger Astronomy in Practice, 2514-3433, pages 9–1 to 9–72. IOP Publishing, 2021.

[7] M. Shibata, K. Kyutoku, T. Yamamoto, and K. Taniguchi. 2009. Physical Review D—Particles, Fields, Gravitation, and Cosmology, 79(4):044030.

[8] J. Aasi, B. P. Abbott, R. Abbott, T. Abbott, M. R. Abernathy, K. Ackley, C. Adams, T. Adams, P. Addesso, and et al. 2015. Class. Quantum Grav., 32:074001.

[9] C. D. Ott. 2009. Classical and Quantum Gravity, 26(6):063001.

[10] G. Agazie, A. Anumarlapudi, A. M. Archibald, Z. Arzoumanian, P. T. Baker, B. Bécsy, L. Blecha, A. Brazier, P. R. Brook, S. Burke-Spolaor, et al. 2023. The Astrophysical Journal Letters, 951(1):L8.

[11] V. De Luca, G. Franciolini, and A. Riotto. 2021. Physical Review Letters, 126(4):041303.

[12] K. Saikawa. 2017. Universe, 3(2):40.

[13] A. Arvanitaki and A. A. Geraci. 2013. Physical review letters, 110(7):071105.

[14] J. G. Rosa and T. W. Kephart. 2018. Physical review letters, 120(23):231102.

[15] J. E. Kim and G. Carosi. 2010. Rev. Mod. Phys., 82(1):557.

[16] K. Choi, S. H. Im, and C. S. Shin. 2021. Annu. Rev. Nucl. Part. Sci., 71:225–252.

[17] L. J. Rosenberg and K. A. Van Bibber. 2000. Phys. Rep., 325(1):1–39.

[18] P. W. Graham, I. G. Irastorza, S. K. Lamoreaux, A. Lindner, and K. A. van Bibber. 2015. Annu. Rev. Nucl. Part. Sci., 65:485–514.

[19] R. Cameron, G. Cantatore, A. Melissinos, G. Ruoso, Y. Semertzidis, H. Halama, D. Lazarus, A. Prodell, F. Nezrick, C. Rizzo, et al. 1993. Physical Review D, 47(9):3707.

[20] N. Du, N. Force, R. Khatiwada, E. Lentz, R. Ottens, L. Rosenberg, G. Rybka, G. Carosi, N. Woollett, D. Bowring, et al. 2018. Phys. Rev. Lett., 120(15):151301.

[21] A. V. Dixit, S. Chakram, K. He, A. Agrawal, R. K. Naik, D. I. Schuster, and A. Chou. 2021. Phys. Rev. Lett., 126(14):141302.

[22] A. Agrawal, A. V. Dixit, T. Roy, S. Chakram, K. He, R. K. Naik, D. I. Schuster, and A. Chou. 2024. Phys. Rev. Lett., 132(14):140801.

[23] H. Shi and Q. Zhuang. 2023. npj Quantum Inf, 9(1):1–10.

[24] D. G. Boulware and S. Deser. 1975. Annals of Physics, 89(1):193–240.

[25] S. Weinberg. The Quantum Theory of Fields. Cambridge University Press, 1995.

[26] J. F. Donoghue. 1994. Physical Review D, 50(6):3874.

[27] C. P. Burgess. 2004. Living Reviews in Relativity, 7:1–56.

[28] J. F. Donoghue. 1995. Advanced school on effective theories: Almunecar, Granada, Spain, 26:217–240.

[29] S. Weinberg. Gravitation and cosmology: principles and applications of the general theory of relativity. John Wiley & Sons, 2013.

[30] E. P. Verlinde and K. M. Zurek. 2021. Phys. Lett. B, 822:136663.

[31] D. Li, V. S. H. Lee, Y. Chen, and K. M. Zurek. 2023. Phys. Rev. D, 107(2):024002.

[32] S. M. Vermeulen, T. Cullen, D. Grass, I. A. O. MacMillan, A. J. Ramirez, J. Wack, B. Ko- rzh, V. S. H. Lee, K. M. Zurek, C. Stoughton, and L. McCuller, Photon Counting Interferome- try to Detect Geontropic Space-Time Fluctuations with GQuEST (2024).

[33] L. McCuller, Single-Photon Signal Sideband Detection for High-Power Michelson Interferometers (2022).

[34] S. M. Vermeulen, L. Aiello, A. Ejlli, W. L. Griffiths, A. L. James, K. L. Dooley, and H. Grote. 2021. Class. Quantum Grav., 38(8):085008.

[35] A. Chou, H. Glass, H. R. Gustafson, C. Hogan, B. L. Kamai, O. Kwon, R. Lanza, L. McCuller, S. S. Meyer, J. Richardson, et al. 2017. Class. Quantum Grav., 34(6):065005.

[36] D. Carney, P. C. Stamp, and J. M. Taylor. 2019. Classical and Quantum Gravity, 36(3):034001.

[37] L. McCuller, C. Whittle, D. Ganapathy, K. Komori, M. Tse, A. Fernandez-Galiana, L. Barsotti, P. Fritschel, M. MacInnis, F. Matichard, et al. 2020. Phys. Rev. Lett., 124(17):171102.

[38] M. Tsang, H. M. Wiseman, and C. M. Caves. 2011. Phys. Rev. Lett., 106:090401.

[39] H. Miao, R. X. Adhikari, Y. Ma, B. Pang, and Y. Chen. 2017. Phys. Rev. Lett., 119:050801.

[40] C. L. Latune, B. M. Escher, R. L. de Matos Filho, and L. Davidovich. 2013. Phys. Rev. A, 88(4):042112.

[41] S. Zhou, M. Zhang, J. Preskill, and L. Jiang. 2018. Nature communications, 9(1):78.

[42] K. Macieszczak, M. Fraas, and R. Demkowicz-Dobrzański. 2014. New Journal of Physics, 16(11):113002.

[43] S. L. Braunstein and C. M. Caves. 1994. Phys. Rev. Lett., 72(22):3439–3443.

[44] H. M. Wiseman and G. J. Milburn. Quantum Measurement and Control. Cambridge University Press, Cambridge, 2009.

[45] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein. 2018. Rev. Mod. Phys., 90(3):035005.

[46] S. A. Haine. 2018. Phys. Rev. A, 98(3):030303.

[47] M. Tsang, R. Nair, and X.-M. Lu. 2016. Phys. Rev. X, 6:031033.

[48] J. W. Gardner, L. Sun, S. Borhanian, P. D. Lasky, E. Thrane, D. E. McClelland, and B. J. J. Slagmolen. 2023. Phys. Rev. D, 108:123026.

[49] J. W. Gardner, T. Gefen, S. A. Haine, J. J. Hope, and Y. Chen. 2024. Phys. Rev. Lett., 132:130801.

[50] J. W. Gardner, T. Gefen, S. A. Haine, J. J. Hope, J. Preskill, Y. Chen, and L. McCuller. 2025. PRX Quantum, 6:030311.

[51] J. W. Gardner, S. A. Haine, J. J. Hope, Y. Chen, and T. Gefen. 2025. arXiv:2501.03364 [quant-ph].

[52] J. W. Gardner, T. Gefen, E. Payne, S. Direkci, S. M. Vermeulen, S. A. Haine, J. J. Hope, L. McCuller, and Y. Chen. 2025. arXiv:2507.02811 [quant-ph].

[53] Y. Nishino, J. W. Gardner, Y. Chen, and K. Somiya. 2025. arXiv:2505.01013.

[54] Y. Nishino, J. W. Gardner, Y. Chen, and K. Somiya. 2025. arXiv:2504.18111.

[55] R. R. Allen, F. Machado, I. L. Chuang, H.-Y. Huang, and S. Choi. 2025.

[56] D. V. Vasilyev, A. Shankar, R. Kaubruegger, and P. Zoller. 2024.

[57] E. Payne et al. (in preparation). 2025.

[58] C. M. Caves. 1981. Phys. Rev. D, 23:1693–1708.