On the Simulation of Stochastic Processes and Noise Driven Synchronisation Models

Timothy McLennan-Smith

March 2019

A thesis submitted for the degree of Doctor of Philosophy
of the Australian National University

Declaration of Authorship

I, Timothy Andrew McLennan-Smith, declare that this thesis titled, 'On the Simulation of Stochastic Processes and Noise Driven Synchronisation Models' and the work presented in it are my own. I confirm that:

- This work was done wholly while in candidature for a research degree at the Australian Nation University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at the Australian Nation University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where the research is conducted jointly with other persons, the nature and extent of my contribution to the research has been stated the following page titled 'Declaration of Collaboration'.
- I have acknowledged all main sources of help.

Signed: 

Date: 30/11/2018

Declaration of Collaboration

We confirm that the work in the chapters titled 'Two Network Kuramoto-Sakaguchi Model Under Tempered Stable Lévy Noise' and 'Modelling Military Watch Staff Using Stochastic Network Synchronisation' are a result of a research collaboration between the Australian National University and the Defense Science and Technology Group. In particular:

- The work in these chapters forms part of the results of two separate collaborative research projects between the Australian National University and the Defence Science and Technology Group.
- Timothy McLennan-Smith's role in these research projects was in the development and the selection of the numerical methods used to simulate and analyse the stochastic versions of the discussed models, the writing of reports corresponding to this numerical research and in the assistance of writing and editing of the two papers as listed in Section 1.1 resulting from these research projects.
- The contents of these chapters consists of the two papers in full, presented in Chapter 3 and Chapter 7 respectively, with a preceding introduction. In these chapters, the papers have been restructured and in parts rewritten from their original versions to suit the style of this thesis.
- These two papers have been submitted to top-tiered international peer-reviewed journals.
- Timothy McLennan-Smith has acknowledged all collaborating authors involved in this research.

Signed: (Mathew Zupanic)

Date: 30/11/2018

Signed: (A. Kalloniatis)

Date: 30/11/2018

Acknowledgements

First and foremost, I would like to thank my supervisor, Dale Roberts, without whom none of this research would haven been possible. Dale's support was invaluable throughout the course of my PhD. He was always enthusiastic in our discussions about this research, motivating me to explore a vast range of ideas and approaches to problems that I would have likely not have been exposed to under the guidance of another. Dale was kind enough to introduce me to Alex Kalloniatis from the DSTG, with whom I had the pleasure of collaborating with. I would like to thank Alex for the many fruitful discussions we have had in our work together and for pushing forward the ideas that lead to our joint research and the two papers that came about as a result. I would also like to thank Matthew Zuparic from the DSTG who collaborated with us in our first paper.

I would like to thank the staff and my fellow PhD students at RSFAS, whose friendly and welcoming nature has always made working at the CBE an enjoyable and productive experience.

Finally I would like to thank my family and the many friends I have made throughout the course of my PhD whose support and guidance throughout the years has seen me through the highs and lows of my research.

Abstract

We investigate the impact of noise with heavy tails on sychronisation within coupled systems of oscillators, the most classic case being called the 'Kuramoto model'. The initial part of this work focuses on methods of simulation of heavy-tailed stochastic processes that are then implemented in the various synchronisation models discussed throughout the thesis. In particular, we explore the accept-rejection sampling method for stable and tempered stable Lévy processes and the Wiener-Hopf Monte-Carlo algorithm for the simulation of the so-called β-family of Lévy processes. The first application of these heavy-tailed stochastic processes in an extension of an adversarial 'Blue vs Red' Kuramoto model from the deterministic and Gaussian cases to the cases where the models are driven by stable and tempered stable Lévy noise. We apply techniques from Bayesian optimisation and the so-called 0-1 test for chaos to numerically investigate the impact of the stochastic noise under various parameterisations of the driving stochastic noise. This 'Blue vs Red' model is then extended into a 'combat-coordination model' through the development of an adversarial attrition model with spatial dependency driven by tempered stable Lévy noise. We then consider extending our previously considered simulation techniques to the case of simulating conditioned stochastic processes, specifically processes conditioned to stay positive and conditioned on the terminal value. These simulation techniques for conditioned stochastic processes are then applied in the context of a conditioned Kuramoto model with a focus on the impact of the underlying network structure. We then advance an approach to mathematically represent networked human decision makers by adapting a Kuramoto model under heave-tailed noise and applying data from a recent study of military headquarters staff. We finally investigate pairs trading and portfolio optimisation under a general stochastic synchronisation model.

Contents

List of Figures

List of Tables

List of Algorithms

Notation

Var	The variance operator
$\mathbb{E}$	The expected value operator
MSE	The mean square error
$\mathcal{N}(\mu, \sigma^2)$	The normal distribution with mean μ and variance σ^2
$\mathrm{U}(a, b)$	The uniform distribution on $[a, b]$
$\mathrm{Exp}(\lambda)$	The exponential distribution with rate λ
$\mathbb{R}$	The real line
$\overset{iid}{\sim}$	Independent and identically distributed
$\xrightarrow{d}$	Convergence in distribution
$x \vee y$	$\max(x, y)$
$x \wedge y$	$\min(x, y)$
$\mathcal{F}(\mathbb{R})$	Borel σ-algebra on $\mathbb{R}$
$\propto$	Proportional to
W_t	Brownian motion / Wiener process
L_t	Lévy process
$\boldsymbol{I}_n$	The identity matrix of size n
DPE	**D**ynamic **P**rogramming **E**quation
MC	**M**onte **C**arlo
ODE	**O**rdinary **D**ifferential **E**quation
OU	**O**rnstein-**U**hlenbeck
PDE	**P**artial **D**ifferential **E**quation
PDF	**P**robability **D**ensity **F**unction
SDE	**S**tochastic **D**ifferential **E**quation
SLLN	**S**trong **L**aw of **L**arge **N**umbers
WHMC	**W**iener **H**opf **M**onte-**C**arlo

Chapter 1

Introduction

In this thesis we investigate the impact of heavy-tailed stochastic noise on dynamical systems, specifically in the context of synchronisation modelling. Synchronisation is a phenomenon that occurs in many natural, social and technological systems. A common approach to the problem of developing a model of synchronisation, consists of modelling each member of the population as a phase oscillator. That is, we model the dynamics of the phase oscillators as

$$\dot{\beta}_i = F(\beta_i) + \sigma \sum_{j \in \mathcal{V}(i)} H(\beta_i, \beta_j),$$

for some functions $F : \mathbb{R} \to \mathbb{R}$ and $H : \mathbb{R} \to \mathbb{R}$ representing the internal interactions and the linear coupling dynamics of the phase oscillators respectively. One of the central models studied in the literature is the Kuramoto model [1] which takes $F(\beta_i) = \omega_i$ and $H(\beta_i, \beta_j) = \sin(\beta_i - \beta_j)$. The Kuramoto model is simple enough to be mathematically tractable yet complex enough to display a large variety of synchronisation patterns and can be applied in a large variety of situations. One of the underlying features of these models is an underlying network which models the connections between the processes or elements of the system. The topologies of these networks play a key role in determining the sychronisation of the system. Very broadly, a network can be described as a system that can be mathematically represented through a graph, denoted as $G = (V, E)$, whose nodes/vertices V represent the processes or elements of the system and whose connecting links/edges E represent the connections between elements.

Using the Kuramoto model on a network G as the core model, the major aim of this thesis is to perturb the system by a stochastic time-varying white-noise ξ_t at each node of the network and then analyse the resulting dynamics of this system.

In particular, we seek to understand how shocks or jumps in the stochastic component of our model influences the dynamics. For this task, Brownian motion is ill suited due to its continuity property. To this end, we turn a slightly broader class of stochastic processes known as Lévy processes which allow for much more flexibility than a Brownian motion as these processes can allow for jumps. As with Brownian motion, a Lévy process, denoted as L_t in this thesis, is a stochastic process with stationary and independent increments however these processes have several important properties that are of great interest and use in a prospective mathematical model. Noticeably, Lévy processes are the simplest examples of random motion whose samples paths are right-continuous with a countable, either infinite or finite, number of random jump discontinuities occurring at random times with a finite time interval. These processes have several important properties that are of great interest from both a mathematical and modelling perspective, and their application in modelling is one of focal aspects of this thesis.

This work is structured as follows. In Chapter 2 we review some definitions and mathematical objects used throughout the thesis. Our discussion in this chapter briefly reviews the necessary graph theory to describe the underlying networks of our sychronisation models, the Kuramoto model, and stochastic processes where we overview both theory and numerical simulation. Our discussion of the simulation of stochastic processes begins with the simulation of diffusion processes and then proceeds to discuss the formulation of Lévy processes and a few of the special cases of Lévy processes that we shall use in our sychronisation models, specifically stable processes, tempered stable processes, and the β-family of Lévy processes. The rest of this chapter is spent discussing three methods of simulating Lévy processes that will be used in all the subsequent chapters of this thesis. These methods are: the Euler method [2], the acceptance-rejection method for the simulation stable and tempered stable processes [3, 4], and the simulation method proposed by Kuznetsov et al. in [5] known as the Wiener-Hopf Monte Carlo (WHMC) algorithm.

In Chapter 3 we present our first sychronisation model that consists of two interacting populations of phase oscillators labelled 'Blue' and 'Red'. This work was produced as a result of a collaboration with the DSTG[1] who originally proposed the model in the Gaussian case in the work of Holder, Zuparic and Kalloniatis [6]. The novelty of this chapter is the consideration of a 'Red vs. Blue' model in the presence of heavy-tailed noise. In particular, this work models competitive dynamics between the two populations, where each population seeks both internal phase synchronisation and a phase advantage with respect to the other population, subject to exogenous heavy-tailed stochastic shocks. We study the system from an analytic and numerical point of view to

[1]DSTG: Department of Defence, Science and Technology Group. Special thanks to Dr Alexander Kalloniatis who proposed this research project.

understand how the phase lag values and the shape of the noise distribution can lead to steady or noisy behaviour utilising the numerical methods of Bayesian optimisation [7, 8] and the 0-1 test as originally proposed by Gottwald in the works [9, 10]. Our analytic study shows that the bulk behaviour can be described by dynamics in the presence of tilted ratchet potentials manifesting metastability in regimes that would otherwise have been expected to be noisy. The proceeding numerical study of the full nonlinear system shows order and noisy behaviour in regions consistent with our analytic estimates.

A broader goal of our work with the DSTG is the development of a combat-coordination model that can be suitably used to model complex warfighting between adversaries [11]. The scope of Chapter 4 is to propose combat models that can be used as a framework for this research. Specifically, we look to extend the Blue vs Red model considered previously in Chapter 3 to include spatial dynamics using the so-called 'swarming' framework as proposed in the recent work of O'Keeffe et al. in [12]. The general form of the model combines the position and the phase of the oscillators into the following dynamical system

$$\begin{aligned}\dot{\boldsymbol{x}}_i &= \boldsymbol{v}_i + \frac{1}{N}\sum_{j=1}^{N}\left[\boldsymbol{I}_{\text{att}}(\boldsymbol{x}_j - \boldsymbol{x}_i)F(\theta_j - \theta_i) - \boldsymbol{I}_{\text{rep}}\left(\boldsymbol{x}_j - \boldsymbol{x}_i\right)\right],\\ \dot{\theta}_i &= \omega_I + \frac{K}{N}\sum_{j=1}^{N} H_{\text{att}}(\theta_j - \theta_I)G(\boldsymbol{x}_j - \boldsymbol{x}_i),\end{aligned}$$

where $\boldsymbol{x}_i$ is the position of the i'th swarmalator, θ_i, ω_i and $\boldsymbol{v}_i$ are its phase, natural frequency, and self-propulsion velocity respectively. $\boldsymbol{I}_{\text{att}}$ and $\boldsymbol{I}_{\text{rep}}$ are the spatial attractions and repulsions respectively, F measures the influence of phase similarity on spatial attraction, H_{att} is the phase interaction, and G is the influence of spatial proximity on the phase attraction. This new Blue vs Red swarming model is then further extended to take into account attrition of the oscillators where we take inspiration from Lanchester's laws [13]. In this extension, we propose and investigate two different attrition models that can describe direct combat between the two networks.

In considering these models, a question that arose was: "If an oscillator jumps out of synchronisation, to what extent does the network structure play a role in the re-synchronisation of the system?". To focus on these complicated dynamics, which are rare events, we develop new numerical methods to tackle this problem through the use of conditioned Lévy processes. The work in Chapter 5 provides an investigation of the simulation of conditioned diffusion processes and conditioned Lévy processes. Our discussion begins with the diffusion bridge processes, that is conditioning the one-dimensional process X_t on its terminal value $X_T = u$ for some $u \in \mathbb{R}$, consists of a variety of different found in the literature with particular focus upon the method of

Bladt and Sørensen in [14]. Using the WHMC algorithm from Kuznetsov et al. [5] as our foundation, we then propose an extension to conditioned Lévy processes. In addition to the simulation of Lévy bridges, we also consider methods for the simulation of Lévy processes conditioned to stay positive and Lévy processes conditioned on their extrema. Although work has been done upon the theory of conditioned Lévy processes [15–17], discussion on the practical simulation of these conditioned processes is somewhat lacking in the literature. This work on the simulation of conditioned stochastic processes is then applied to the proposed model in Chapter 6 which is a standard synchronised Kuramoto system in both the Gaussian and Lévy noise cases given by the system

$$\mathrm{d}\theta_i(t) = \left(\omega_i - \sigma_G \sum_j^N A_{ij} \sin(\theta_i - \theta_j)\right) \mathrm{d}t + \mathrm{d}{L_t}^{(i)}, \qquad i = \{1, \ldots, N\},$$
$$\theta_n(T_c) = \theta_c \text{ for some } |\theta_c| \geq \pi/2, \quad \theta_1(0) = \theta_2(0) = \cdots = \theta_N(0) = 0,$$

where θ_i is the phase angle for node i, ω_i is the intrinsic frequency of the given node, σ_G is the global coupling constant, L_i is a Lévy process, n is selected from the set of nodes $\{1, \ldots, N\}$, and $T_c > 0$ is the (time) point of conditioning. The simulation of these conditioned processes utilises the results of the simulation of bridge processes discussed in Chapter 5. We then focus our analysis upon how the underlying network structure affects the dynamics of the system before and after the point of conditioning.

As part of the broader aim to study the 'Perception-Action cycle loop' of individual agents in the context of a synchronisation model, as was previously investigated in the adversarial context in Chapters 3 and 4, the work in Chapter 7 advances an approach to mathematically represent networked human decision makers by adapting the Kuramoto model of networked phase oscillators and applying data from a recent study of military headquarters staff. The motivation for this work was originally proposed by the DSTG and is based upon the data studied by Kalloniatis in [18], although the aforementioned study was outside of the context of the synchronisation of decision making. The implementation of stochastic noise in this model captures the idea that agents may randomly 'jump' ahead in their decision state in time-stressed conditions where intuition becomes significant. We observe a high probability that synchronisation may be lost as a consequence of the high information demand of some agents in the crisis scenario; this provides face-validation of the model. We examine the causes of this loss of synchronisation and compare to qualitative approaches in the literature to validate the model. To conclusion this chapter, we discuss the model's utility as a future decision support tool.

As a final model, we present an application of our work outside of the context of an

'OODA loop' synchronisation model. The model that we investigate in Chapter 8 employs a synchronisation model in context of pairs trading. Although the pairs trading model has been studied in the literature [19–21], our goal in this chapter is to extend the pairs scenario to the multiple asset setting in which the modelled financial system has an underlying topology which can be described by a graph and the dynamics of the drifts of the assets are driven by the Kuramoto equation given by the following system

$$\frac{dY_t^k}{Y_t^k} = \alpha_t^k dt + \sum_{i=1}^{n} \sigma_{ki} dW_t^i,$$

where the nonlinear co-integration factor that incorporates the underlying market structure is given by

$$\alpha_t^k = a_0^k + \sum_{i=1}^{n} A_{ki} F\left(\log Y_t^i - \log Y_t^k\right),$$

where $\mathbf{W} = \left(W_t^1, \ldots, W_t^n\right)_{0 \leq t \leq T}$ is a vector of independent Brownian motions, F is some known function $F : \mathbb{R} \to \mathbb{R}$ and $\boldsymbol{A}$ is a weighted adjacency matrix that encodes the network topology. Once the model is established, we then derive the optimal portfolio based upon this model.

Finally Chapter 9 of the thesis presents a brief discussion of the results established in the previous chapters and examines various possible extensions and future works that may come about as a result of this thesis.

To summarise the aim of this thesis, in addition to discussing currently used methods in the literature to simulate stochastic processes and reviewing previous work with the Kuramoto model, the central goal of this work is to present novel research in the following areas:

- The development of numerical methods to simulate Lévy processes, conditioned Lévy processes and various path functionals of these processes.

- The application of Lévy noise in the so-called adversarial and crisis scenario synchronisation models.

- The development and analysis of a variety of synchronisation models utilising Lévy processes and conditioned Lévy processes.

1.1 Publication, presentation, and collaboration details

The work done in Chapter 3 and Chapter 7 of this thesis is a product of the result of two separate research agreements with the Australian Defence Science and Technology Group. This research has been submitted for publication in [22] and [23] and provided in detail below:

- Kalloniatis, Alexander C., McLennan-Smith, Timothy A., Roberts, Dale O. and Zuparic, Mathew L. *Two-network Kuramoto-Sakaguchi model under tempered stable Lévy noise.* Physical Review E 99, no. 1 (2019): 012205.
- Kalloniatis, Alexander C., McLennan-Smith, Timothy A., and Roberts, Dale O. *Modelling military watch staff using stochastic network synchronisation*, Submitted to the Journal of the Operational Research Society.

The research on conditioned Lévy processed from Section 5.1 and attrition models in the swarming setting from Chapter 4 has been presented in the following talks:

- Fourth IMS-FPS (Institute for Mathematical Statistics - Finance, Probability and Statistics) Workshop, Sydney. *Wiener-Hopf Monte Carlo method for Lévy processes conditioned to stay positive*, July 2014
- Modelling Complex Warfighting Strategic Research Initiative Symposium 2, Adelaide. *A ruin adversarial model under tempered stable noise*, May 2018

Chapter 2

Preliminaries

We begin the main body of this thesis with some necessary preliminary details that will go into the various mathematical aspects and definitions used throughout the synchronisation models examined in the later chapters. This chapter gives a brief overview of the following major aspects: in Section 2.1 we review some graph theory which is used to describe the topology of the underlying networks in our models. In Section 2.2 we provide a more detailed discussion of the central model of this thesis, the Kuramoto model. Section 2.3 discusses the simulation of diffusion processes. In Section 2.4 we present a discussion of the mathematical description of our stochastic noise with a focus upon the definition of a Lévy process. Finally in Sections 2.5 and 2.6 the simulation of Lévy processes is discussed through the Euler method, an acceptance-rejection method, and the recent Wiener-Hopf Monte Carlo method.

2.1 Networks and graphs

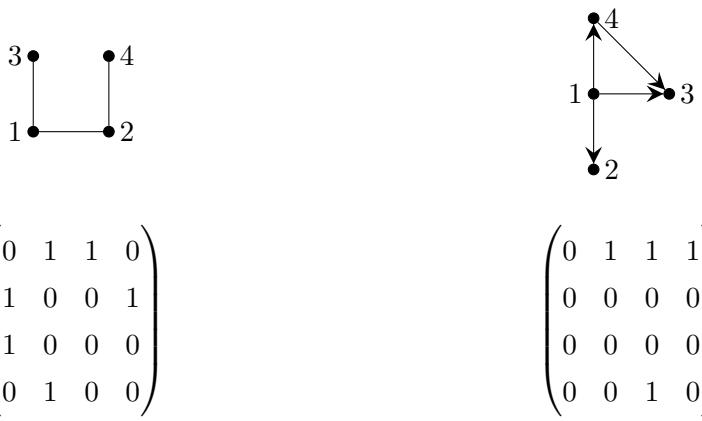
2.1.1 An overview of networks and graphs

A network is a commonly used device to represent the interaction between agents in a complex system; see for example [24, 25]. In order to mathematically describe a network, we utilise the structure known as a graph to represent its topology. We define a graph as the ordered pair $G = (V, E)$ where the nodes/vertices denoted as V represent the processes or elements of the system and the connecting links/edges denoted as E represent the connections between elements. The edge (i, j) joins the nodes i and j and these nodes are said to be adjacent. Adjacent nodes are also referred to as neighbours. The order of a graph, $|V| = N$, is the total number of vertices in the graph N and the size of a graph $|E|$ is the total number of edges. One of the most useful tools used to

describe a graph is the $N \times N$ *adjacency matrix* $\mathbf{A}$ where are the elements of this matrix are given by

$$A_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

There are two distinct classes of graphs that we consider in this thesis, that of directed and undirected graphs. An *undirected graph* is a graph in which edges have no orientation, that is the edge (i,j) is identical to the edge (j,i). In the case of an undirected graph, the adjacency matrix $\mathbf{A}$ is symmetric. A *directed graph* or *digraph* is a graph in which edges have orientations. In this case, the set of vertices E are now ordered pairs. These connections are depicted by arrows where an arrow (i,j) is considered to be directed from i to j and j is called the head and i is called the tail of the arrow. For a directed graph, the adjacency matrix $\mathbf{A}$ is not symmetric. An example of undirected and directed graphs and their respective adjacency matrices are given on Figure 2.1.



$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

FIGURE 2.1: Graphs and adjacency matrices for an undirected graph (left) and a directed graph (right) with $|V| = 4$.

The *degree* of a node i, denoted as d_i, is the number of edges that connect to it. In the case where an edge that connects a vertex to itself, it is counted twice. In terms of the adjacency matrix we can define the degree

$$d_i = \sum_j A_{ij},$$

where A_{ij} is a element of the adjacency matrix $\boldsymbol{A}$. This definition is only valid for undirected graphs. For a directed graph, the degree of a given node i is defined in terms of the sum of edges arriving at the node and departing at the node. That is, for a directed graph, we define the degree by

$$d_i = d_{in,i} + d_{out,i},$$

where $d_{in,i} = \sum_j A_{ji}$ and $d_{out,i} = \sum_j A_{ij}$. A useful matrix representation is that of the *degree matrix* $\mathbf{D}$ which is an $N \times N$ diagonal matrix that which contains all the information about the degrees of the network in matrix form and is defined as

$$D_{ij} = \begin{cases} d_i & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

In addition to the adjacency and degree matrices, we define the $N \times N$ *Laplacian matrix* as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where the elements of $\mathbf{L}$ are given as

$$L_{ij} = \begin{cases} d_i & \text{if } i = j, \\ -1 & \text{if } i \neq j \text{ and node } i \text{ is adjacent to node } j, \\ 0 & \text{otherwise.} \end{cases}$$

As we shall discuss in the proceeding section, the Laplacian matrix provides a useful tool that can be used to find many useful features of a graph.

We refer to the connected component, or simply just the *component*, of an undirected graph as the subgraph in which any two vertices are connected to each other by a path. A connected graph, which contains a path between every pair of nodes, has just one connected component and it consists of the entire graph. If a graph has more than one connected component, then the graph is said to be *disconnected*. More precisely, a graph is said to be *k-connected* if there does not exist a set of $k-1$ edges whose removal disconnects the graph.

2.1.2 Quantifying graph features

In this section we present a variety of different statistics that will be used to measure particular properties of a given graph $G = (V, E)$ with $|V| = N$, adjacency matrix $\mathbf{A}$, degree matrix $\mathbf{D}$, and Laplacian matrix $\mathbf{L}$. For more details, see [24, 25].

The *Fiedler value*, or the algebraic connectivity, of the graph G is given by the second smallest eigenvalue of the Laplacian matrix $\mathbf{L}$ and is denoted as λ_2 [26]. The relative magnitude of the Fielder value provides a measure of how well connected the graph is. A well known result is that a connected graph will always have a non-zero Fiedler value and furthermore that the Fielder value is bounded below by the $\frac{4}{ND}$ where D is the diameter of the graph and N is the number of vertices [27]. As an example of the relationship between the overall connectivity of the graph and the Fiedler value, on Figure 2.2 we plot two 1-connected graphs with low and high connectivity respectively.

FIGURE 2.2: Two sampled graphs from the set of 1-connected graphs with $|V| = 6$. Left: a 1-connected graph with low algebraic connectivity ($\lambda_2 = 0.081$). Right: a 1-connected graph with high algebraic connectivity ($\lambda_2 = 0.485$).

A natural measure of distance in graphs is to count the number of edges that make up the shortest connecting path between two nodes i and j. This distance is referred to as the *shortest-path length* l_{ij}. For two nodes in two disconnected graphs, we define $l_{ij} = \infty$. The diameter of the graph G is than defined through the shortest path length as follows

$$D = \max_{i,j} l_{ij},$$

that is the diameter provides a measure of the maximum shortest-path lengths between any two nodes in the network.

Another measure of the size of the graph is the *average shortest-path length* which is defined as the average value of l_{ij} over all possible pairs of vertices in the network

$$\langle l \rangle = \frac{1}{N(N-1)} \sum_{i,j} l_{ij},$$

where by definition $\langle l \rangle \leq D$.

The *closeness centrality* is a measure of centrality in a network that is calculated as the average distance of a node to all others through the shortest-path length as

$$g_i = \frac{1}{\sum_{j \neq i} l_{ij}},$$

thus the more central a node is in terms of the relative magnitude of its closeness centrality, the closer it is to all other nodes.

The final statistic that we shall discuss is the characterisation of clustering in a graph. The property of clustering in a graph implies that, given the three nodes i, j, k, where there an edge between the nodes i and j and an edge between the nodes i and k, then there is a high probability that there will be a vertex between the nodes j and k. In an undirected graph, the clustering at a given node i can be quantified in terms of the

clustering coefficient [24] which is defined by

$$C_i = \frac{e_i}{d_i(d_i-1)/2},$$

where d_i is the degree of the node i and e_i is the number of edges among the neighbours of the node i which can be calculated in terms of the adjacency matrix $\mathbf{A}$ as

$$e_i = \frac{1}{2}\sum_{jk} A_{ij}A_{jk}A_{ki}.$$

The *average clustering coefficient* of the network is than given by the average clustering coefficient over all nodes as

$$\langle C \rangle = \frac{1}{N}\sum_i C_i.$$

2.1.3 Network models

We conclude our overview on graphs with two well known network models that feature in this thesis, namely the Erdös-Rényi model [28] and the k-ary tree model [29].

2.1.3.1 The Erdös-Rényi model

The Erdös-Rényi model [24, 28, 30] is used to generate random graphs. The idea behind the original form of the model is that nodes are connected at random with some probability p. There are two variants of the model that are used, the first variant denoted as $G(N, E)$ fixes the number of edges E and joins the N nodes at random amongst the E edges. In the second variant, denoted as $G(N, p)$, each edge is included in the graph with probability p independent from every other edge. In the second variant of these two models, the probability of constructing a graph $G(N, E)$ with a $G(N, p)$ model is given by

$$P(G(N, E)) = p^E(1-p)^{\frac{1}{2}N(N-1)-E}.$$

In fact, the two models are statistically equivalent in the limit as $N \to \infty$ with $pN(N-1)/2 = E$ [24]. In the $G(N, p)$ variant of these models, the distribution of the degree of any particular node [31] follows the binomial distribution as

$$P(d_i = k) = \begin{pmatrix} N-1 \\ k \end{pmatrix} p^k(1-p)^{N-1-k}. \tag{2.1}$$

In the limit as $N \to \infty$ and holding Np constant, the distribution in Eq. (2.1) can be approximated by the Poisson distribution [32] as

$$P(d_i = k) = e^{-Np}\frac{(Np)^k}{K!}.$$

Some of the key properties of the $G(N, p)$ model are:

- $Np < 1$: $G(N, p)$ has no connected components of size larger than $O(\log(n))$ almost surely.
- $Np = 1$: $G(n, p)$ has a largest component whose size is of $O(n^{2/3})$ almost surely.
- $p < \frac{(1-\varepsilon)\log N}{N}$: $G(N, p)$ will be disconnected almost surely.
- $p > \frac{(1-\varepsilon)\log N}{N}$: $G(N, p)$ will be connected almost surely.

2.1.3.2 The k-ary tree

Unlike the Erdös-Rényi model which generates a random graph, the k-ary tree model generates a graph with a distinctive structure. This model is a type of tree model [33, 34]. A tree model is an undirected graph in which any two vertices are connected by exactly one path. Simply put, a tree is a undirected, connected, and acyclic (that is there no cycles) graph. From these properties, a tree with N nodes has $N - 1$ edges. Conversely, a connected graph with N nodes and $N - 1$ edges is a tree. The k-ary tree model is an rooted tree model in which each node has no more than k children. By rooted tree, we refer to a tree model in which one node is designated as the root of the tree. We refer to the parent node as the node connected to it on the path to the root and the child of a node v as the node of which v is the parent. Formally, we can define the k-ary tree, denoted as T_k, as the set of nodes that consists of a root, R, and exactly k distinct k-ary trees. That is, the remaining nodes are partitioned into $k \geq 0$ subsets, $T_0, T_1, \ldots, T_{k-1}$, each of which is an k-ary tree such that $T_k = \{R, T_0, T_1, \ldots, T_{k-1}\}$. We shall specifically look at the case of a perfect k-ary tree which is a full k-ary tree in which all leaf nodes are at the same depth. That is every node is either a 0-node or a k-node and all leaves are on the same level. For a perfect k-ary tree with some height h, the total number of nodes N is given by

$$N = \frac{k^{h+1} - 1}{k - 1}.$$

2.2 Synchronisation models

There are many different dynamical processes that can be constructed upon a network. The dynamical processes of interest in this thesis will be those that can effectively model the property of sychronisation amongst the elements of the system. These so-called synchronisation models model each member of the population within the given network as a phase oscillator. In the most general setting, a phase oscillator can be represented by the dynamical system

$$\dot{\beta} = f(\beta),$$

for some function $f : \mathbb{R} \to \mathbb{R}$ and where the phase β has a periodic orbit given by some $\gamma \in \mathbb{R}$. Extending our scope to that of a system of linearly coupled phase oscillators upon a network described by the $G = (V, E)$, the most general form of the dynamical system can then be written as

$$\dot{\beta}_i = F(\beta_i) + \sigma \sum_{j \in \mathcal{V}(i)} H(\beta_i, \beta_j),$$

for some functions $F : \mathbb{R} \to \mathbb{R}$ and $H : \mathbb{R}^{\nvDash} \to \mathbb{R}$ representing the internal interactions and the linear coupling dynamics of the phase oscillators respectively and σ denoting the strength of the linear coupling between the phase oscillators.

2.2.1 The Kuramoto model

One of the most widely studied synchronisation models in the literature is the Kuramoto model and its variants. Many biological, social and economic examples of complex systems contain both cooperative and competitive phenomena in tension that lead to both ordered and disordered behaviours. The Kuramoto model (e.g., see [35–39]) and its extension to multi-networks [40–45] has been shown to offer a viable approach to modelling competitive dynamics between populations in these contexts. Richer behaviour in the dynamics of the Kuramoto was observed when various forms of Gaussian white or coloured noise perturbation [36, 46–51] were introduced and more recently under heavier-tailed noise [52, 53] coming from the class of Lévy processes. This class includes fractional diffusion [54, 55] and other closely related noise distributions. In its most popular form, the Kuramoto model governs the dynamics of coupled phase oscillators according to the following dynamical system

$$\dot{\beta}_i = \omega_i + \frac{K}{N} \sum_{j \in \mathcal{V}(i)} \sin\left(\beta_j - \beta_i\right), \tag{2.2}$$

where ω_i is the natural frequency of oscillator i, $\mathcal{V}(i)$ is the set of vertices connected to i and K is the coupling strength of the oscillators. Thus in the Kuramoto model, each oscillator tries to run independently at its own frequency, while the coupling tends to synchronise it to all the others. When the coupling K is weak, the oscillators evolve incoherently whereas beyond a certain threshold of the coupling, collective synchronisation of the system emerges.

2.2.2 Measures of synchronisation and the mean-field Kuramoto model

In the Kuramoto model, as given by Eq. (2.2), synchronisation can be measured through the *order parameter* $R \in \mathbb{C}$ given by

$$R = re^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j}, \tag{2.3}$$

where $r \in [0, 1]$ represents the phase-coherence of the population of oscillators and ψ indicates the average phase. In the limit as $N \to \infty$, if the oscillators are out of synchronisation, the amplitude of the order parameter approaches 0 that is $r \to 0$. In the case that $0 < r < 1$, a subset of the oscillators are synchronised and in the case where $r \approx 1$, synchronisation is present within all of the oscillators of the system.

An alternative version of the Kuramoto model, the so-called mean-field Kuramoto model, can be obtained by combining the dynamical system of the Kuramoto model and the order parameter in Eq. (2.2) and Eq. (2.3) respectively. This is achieved by multiplying Eq. (2.3) by $e^{-i\phi}$ and considering only the imaginary part we get

$$r\sin(\psi - \beta_i) = \frac{1}{N}\sum_{j=1}^{N} \sin(\beta_j - \beta_i),$$

which can be substituted back into Eq. (2.2) to give

$$\dot{\beta}_i = \omega_i + Kr\sin(\beta_i - \psi).$$

This formulation of the Kuramoto model highlights the mean-field behaviour of the model, that is each oscillator is coupled to the common average phase ψ with coupling strength given by Kr. When $K \to 0$, the dynamics of the system can be approximated by $\beta_i(t) \approx \omega_i t + \beta_i(0)$, that is the oscillators rotate at frequencies given by their own natural frequencies. In the limit of strong coupling, that is as $K \to \infty$, the oscillators become synchronised to their average phase, $\beta_i \approx \phi$ and hence $r \to 1$. For intermediate

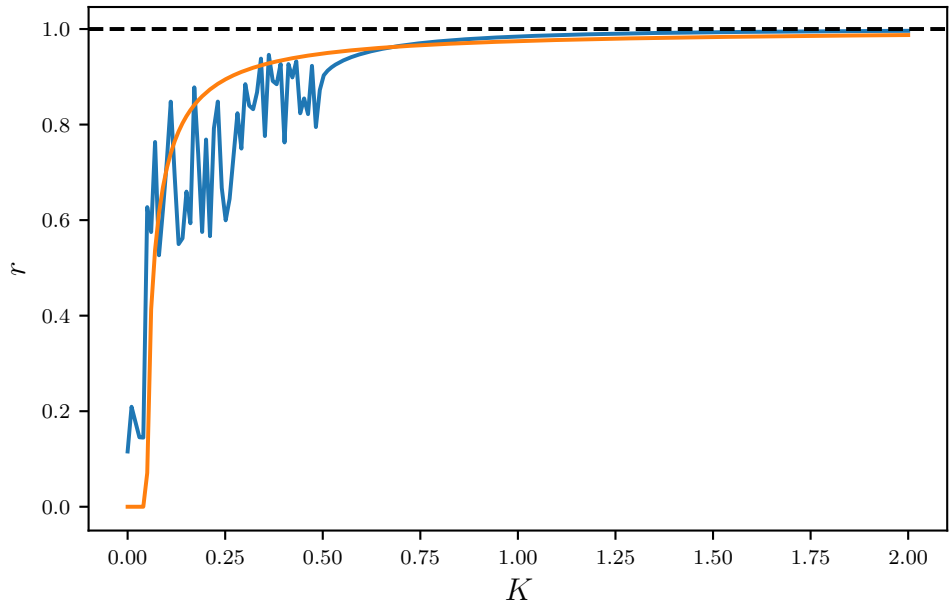


FIGURE 2.3: The dependency of sychronisation on coupling strength using the mean-field Kuramoto model with a Lorentzian frequency distribution as in Eq. (2.4) with $\gamma = 0.5$. The order parameter in the thermodynamic limit is given in orange and the simulation with $N = 50$ and $T = 1000$ is given in blue.

couplings, $K_c < K < \infty$, part of the oscillators are phase-locked $\dot{\beta}_i = 0$ and part are moving out of synchronisation with the locked oscillators.

Taking the Lorentzian frequency distribution for the distribution of the natural frequencies as

$$g(\omega) = \frac{\gamma/\pi}{\gamma^2 + \omega^2}, \tag{2.4}$$

where ω is uniformly distributed as $\omega \sim \mathrm{U}[\pi, \pi]$. It can be shown [56] that in the limit as $N, T \to \infty$, the order parameter behaves as

$$r = \begin{cases} 0 & K < \frac{2\gamma}{N} \\ \sqrt{1 - \frac{2\gamma}{NK}} & K \geq \frac{2\gamma}{N} \end{cases}$$

Setting $\gamma = 0.5$ and approximating the thermodynamic limit with $N = 50$ and $T = 1000$, we present the mean-field result for the Kuramoto model in Figure 2.3.

2.3 Diffusion processes

Most of the models considered in this thesis can be formulated through the stochastic differential equation [57]

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{\mu}(t, \boldsymbol{X}_t)\,\mathrm{d}t + \boldsymbol{\sigma}(t, \boldsymbol{X}_t)\,\mathrm{d}\boldsymbol{L}_t, \tag{2.5}$$

with some initial condition $\boldsymbol{X}(0)$ and some Lévy process $\mathbf{L}(t) = (L_1(t), L_2(t), \dots L_d(t))$, where L_i is independent of L_j for $i \neq j$. In this form, the two deterministic functions $\boldsymbol{\mu} : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\boldsymbol{\sigma} : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are processes of dimension d and $d \times d$ respectively. These two processes are commonly refereed to as the drift and diffusion coefficients. Though we shall continue this chapter without explicit mentioning so, the drift coefficient $\boldsymbol{\mu}$ and the diffusion coefficient $\boldsymbol{\sigma}$ obey certain regularity conditions. In particular they are measurable with respect to the filtration $\{F_t\}$ generated by W_t and are in the functional spaces $L^1([0,T]\times\mathbb{R}^d)$ and $L^2([0,T]\times\mathbb{R}^d)$ respectively. Furthermore, in order to guarantee that a solution exists and is unique, we make the following two assumptions:

- **Global Lipschitz**: For all $x, y \in \mathbb{R}^d$ and $t \in [0,T]$, there exists a constant $K < +\infty$ such that:

$$|\boldsymbol{\mu}(t,x) - \boldsymbol{\mu}(t,y)| + |\boldsymbol{\sigma}(t,x) - \boldsymbol{\sigma}(t,y)| \leq K\,|x-y|\,.$$

- **Linear growth**: For all $x \in \mathbb{R}^d$ and $t \in [0,T]$, there exists a constant $C < +\infty$ such that:

$$|\boldsymbol{\mu}(t,x)| + |\boldsymbol{\sigma}(t,x)| \leq C\,(1 + |x|)\,.$$

The driving noise of the SDE in Eq. (2.5) can be either a Lévy process or a Brownian motion. In this section, we look at the case of a diffusion process which arises when the driving noise in Eq. (2.5) is a Brownian motion $\mathbf{W}(t) = (W_1(t), W_2(t), \dots W_d(t))$, where W_i is independent of W_j for $i \neq j$. (For a more detailed discussion on Brownian motion see e.g., [58].)

Unless otherwise stated, we shall continue with the one-dimensional diffusion case of $d = 1$ in Eq. (2.5). In scalar form the SDE is denoted as $dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$. Some examples of well known diffusion processes are provided on Table 2.1.

	$\mu(t, X_t)$	$\sigma^2(t, X_t)$	
Brownian motion	0	σ^2	
Standard Brownian motion	0	1	
Geometric Brownian motion	μX_t	σX_t	
Brownian motion with drift	μ	σ^2	
Ornstein-Uhlenbeck process	$-\beta(X_t - \alpha)$	σ^2	$\beta > 0$
Cox-Ingersoll-Ross process	$-\beta(X_t - \alpha)$	$\sigma^2 X_t$	$\alpha > 0, \beta > 0$

TABLE 2.1: A list of some well known one-dimensional diffusion processes.

2.3.1 Simulating diffusion processes

The simulation of an unconditioned diffusion processes follows a relatively simple procedure. Given a one-dimensional diffusion processes X_t whose dynamics are driven by the SDE in Eq. (2.5) with $d = 1$, the Euler-Maruyama method for the simulation of X_t with some initial condition $X_0 = x$ on $t \in [0, t]$ is given by the approximation $\hat{X}$ where $\hat{X}$ is defined as follows:

- Partition the time interval $[0, T]$ into N equal subintervals $0 = \tau_0 < \tau_1 < \cdots < \tau_N = T$ where the width of each interval is given by $\Delta t = T/n$.
- Set the initial value $\hat{X}_0 = x$.
- Incrementally solve for $\hat{X}_n$, for $1 \leq n \leq N$ by

$$\hat{X}_{n+1} = \hat{X}_n + \mu(\tau_n, \hat{X}_n)\Delta t + \sigma(\tau_n, \hat{X}_n)\Delta W_n,$$

  where

$$\Delta W_n = W_{\tau_{n+1}} - W_{\tau_n}$$

  and ΔW_n are independent and identically distributed normal random variables with zero mean and variance Δt, that is $\Delta W_n \overset{iid}{\sim} \mathcal{N}(0, \Delta t)$.

If we are interested in evaluating $\mathbb{E}(f(X_t))$ where $f : \mathbb{R} \to \mathbb{R}$ is some known function and X_t is the solution to Eq. (2.5) at a time t, a common method for evaluating such a quantity is the Monte Carlo method. The Monte Carlo method works by simulating n replications $x_1, x_2, \ldots, x_n$ of the random variable X_t, and then approximating $\mathbb{E}(f(X_t))$ through the mean of the observations $f(x_i)$, that is

$$\mathbb{E}(f(X_t)) \approx \frac{1}{n}\sum_{i=1}^{n} f(x_i) =: \bar{f}_t.$$

If a finiteness condition holds, that is $\mathbb{E}|f(X_t)| < \infty$, then the central limit theorem guarantees that

$$\bar{f}_t \xrightarrow{d} \mathcal{N}\left(\mathbb{E}(f(X_t)), \frac{1}{n}\mathrm{Var}\left(f(X_t)\right)\right).$$

We can estimate the standard error of the simulation as

$$\mathrm{SE}(f_t) = \frac{s}{\sqrt{n}},$$

where s is the sample variance of the Monte Carlo replications

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(f(x_i) - \bar{f}_t\right)^2.$$

Although the convergence rate $\frac{1}{\sqrt{n}}$ is slow, the attraction of the method is that it is independent of the smoothness of the function $f(\cdot)$.

To finalise this discussion on the simulation of a standard diffusion process, the following algorithm summaries the Monte Carlo procedure for obtaining an estimate of $\mathbb{E}(f(X_t))$:

Algorithm 1 Monte Carlo simulation of $\mathbb{E}(f(X_t))$

$\Delta t \leftarrow T/N$
for $i = 1, \ldots, n$ **do**
    $X(0) \leftarrow X_0$
    for $j = 0, \ldots, N-1$ **do**
        sample W $\sim \mathcal{N}(0, \Delta t)$
        $X(j+1) \leftarrow X(j) + \mu(\Delta t \cdot j, X(j)) \cdot \Delta t + \sigma(\Delta t \cdot j, X(j)) \cdot W$
    $F(i) \leftarrow f(X(n))$
return mean(F)

2.3.2 The generator of a diffusion process

We now look at the characterisation of the dynamics of the process in an infinitesimal time interval through the infinitesimal generator. The infinitesimal generator of a d-dimensional diffusion process is defined by the operator $\mathcal{L}$ which is given by

$$\mathcal{L}f(\boldsymbol{x}) = \lim_{t \downarrow 0} \frac{\mathbb{E}[f(\boldsymbol{X}_t)] - f(\boldsymbol{x})}{t}, \tag{2.6}$$

where f is some suitable function $f : \mathbb{R}^n \to \mathbb{R}$. The set of all functions f for which the above limit exists is referred to as the domain of the operator and denoted by $\mathcal{D}_{\mathcal{L}}$, where $\mathcal{D}_{\mathcal{L}}$ contains all twice-differentiable functions with compact support. In the context of a diffusion process, it can be shown (for a full proof see e.g, [59]) that the differential operator can be expressed as

$$\mathcal{L} = \sum_{j=1}^{d} \boldsymbol{\mu}_j \frac{\partial}{\partial x_j} + \frac{1}{2}\sum_{i,j=1}^{d} \left(\boldsymbol{\sigma}\boldsymbol{\sigma}^T\right)_{i,j} \frac{\partial^2}{\partial x_i \partial x_j}. \tag{2.7}$$

In the one dimensional case, the generator is defined as $\mathcal{L} = \mu\partial_x + \frac{1}{2}\sigma^2\partial^2_{xx}$.

As an example, for the case of an uncorrelated d-dimensional Brownian motion with fixed drift $\boldsymbol{\mu}$ and diffusion $\boldsymbol{\sigma} = \boldsymbol{I}_n$, where the SDE given by the following equation

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{\mu}\mathrm{d}t + \mathrm{d}\boldsymbol{W}_t,$$

the generator of $\boldsymbol{X}_t$ is then given by

$$\mathcal{L}f = \sum \boldsymbol{\mu}_i \frac{\partial f}{\partial x_i} + \frac{1}{2}\sum \frac{\partial^2 f}{\partial x_i^2}.$$

2.3.3 Transition densities of diffusion processes

The dynamics of diffusion processes can be described in terms of their transition probabilities. Thanks to the Markov property of our diffusion process, we can describe the solutions of the SDE Eq. (2.5) in terms of the transition densities:

$$p(T, y|t, x) := \mathbb{P}\left(X_T \in [y, y + \mathrm{d}y)|X_t = x\right). \tag{2.8}$$

The backward evolution of the transition density $p(T, y|t, x)$ is given by the Kolmogorov backward equation which tells us about the distribution of the process at time $t < T$, given that it is in some future state at time T. This is defined by

$$\frac{\partial p}{\partial t} + \mu(t, x)\frac{\partial p}{\partial x} - \frac{1}{2}\sigma^2(t, x)\frac{\partial^2 p}{\partial x^2} = 0,$$

for an initial condition $p(T, y|T, x) = \delta(x-y)$ where δ is the Dirac measure. Alternatively we can express this in terms of the infinitesimal generator $\mathcal{F}$ from Eq. (2.7) as

$$\frac{\partial p}{\partial t} + \mathcal{F}p = 0.$$

The forward evolution of the transition density $p(T, y|t, x)$ is given by the Kolmogorov forward equation (or Fokker-Planck equation) which tells us about the distribution of the state of the process at a future time $T > t$ from a starting state x. This is defined by

$$\frac{\partial p}{\partial T} + \frac{\partial}{\partial y}\left\{\mu(T, y)p\right\} - \frac{1}{2}\left\{\sigma^2(T, y)p\right\} = 0,$$

for an initial condition $p(t, y|t, x) = \delta(x-y)$ and can alternatively be expressed in terms of $\mathcal{L}^*$, the adjoint operator of $\mathcal{L}$, by

$$\frac{\partial p}{\partial t} = \mathcal{F}^* p.$$

For a more detailed discussion on the simulation of diffusion processes and stochastic differential equations in general see e.g., [59–61].

2.4 Lévy processes

In our discussion on Lévy processes [17, 57, 62], we shall work on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Here we define Ω as

$$\Omega = D\left([0,\infty], \mathbb{R}^d \cup \{\delta\}\right),$$

which is the space of right-continuous functions with left limits with the isolated point δ which shall serve as the cemetery state of the process. The Borel sigma-field of Ω is denoted by $\mathcal{F}$ and we consider the probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ such that under $\mathbb{P}$, the process L_t has independent homogeneous increments and $\mathbb{P}(L_0 = 0) = 1$.

2.4.1 An overview of Lévy processes

Lévy processes are a class of stochastic processes that generalise the idea of a Brownian motion to include jumps in their paths and heavy tails in their transition distributions. They have a rich and deep theory and play an important role in financial mathematics, statistics, and insurance risk theory (e.g., [17, 57, 62–65]). From a mathematical perspective, Lévy processes are interesting as they can be viewed as the scaling limit of random walks. These processes provide a number of challenges as they are on the edge of mathematical tractability due to their characterisation obtained by their characteristic function (i.e., the Fourier transform of their transition densities).

A stochastic process $\{L_t : t \geq 0\}$ on $\mathbb{R}^d$ is said to be a Lévy process if

- $L_0 = 0$ a.s
- (Independence of increments) For any $0 \leq t_1 < t_2 < \cdots < t_n < \infty$, $L_{t_2} - L_{t_1}, L_{t_3} - L_{t_2}, \ldots, L_{t_n} - L_{t_{n-1}}$ are independent.
- (Stationary increments) For every $s, t \geq 0$, the distribution of the process $L_{s+t} - L_s$ does not depend on s.
- For any $\epsilon > 0$ and $t \geq 0$ it holds that $\lim_{h \to 0} \mathbb{P}(|L_{t+h} - L_t| > \epsilon) = 0$.
- There is a $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}[\Omega_0 = 1]$ such that for every $\omega \in \Omega_0$, $L_t(\omega)$ is right-continuous in $t \geq 0$ and has left limits in $t \geq 0$.

The process L_t can be characterised by the triplet (a, σ^2, Π), where $a \in \mathbb{R}$ is the linear drift component, $\sigma \geq 0$ is the volatility of the Gaussian component, and $\Pi(\mathrm{dx})$ is the Lévy measure satisfying $\Pi(\{0\}) = 0$ and $\int_{\mathbb{R}}(x^2 \wedge 1)\,\Pi(\mathrm{dx}) < \infty$. In general the distribution of a Lévy process is described by its characteristic function as explicit expressions for densities are often unobtainable. The characteristic function of a Lévy processes is given by the Lévy-Khintchine formula. This formula states that the characteristic exponent $\Psi(\theta)$ of a Lévy process L_t is given by $\mathbb{E}\left[e^{iuL_t}\right] = e^{-t\Psi(u)}$ for $u \in \mathbb{R}$ where

$$\Psi(u) := iua + \frac{1}{2}\sigma^2u^2 + \int_{\mathbb{R}}\left(1 - e^{iux} + iux\mathbf{1}_{\{|x|<1\}}\right)\Pi(\mathrm{dx}).$$

Some well known examples of the characteristic exponents of Lévy processes can be observed in the case of a linear Brownian motion $\mathcal{N}(a, \sigma^2)$

$$\Psi(u) = \frac{\sigma^2u^2}{2} - iua,$$

the Poisson process $Pois(\lambda)$

$$\Psi(u) = \lambda(1 - e^{iu}),$$

and the Cauchy process $C(\mu, \theta)$

$$\Psi(u) = \theta|u| - iu\mu.$$

A useful tool for understanding the nature of Lévy processes through their paths is that of the Lévy Itô decomposition. This states that, given a Lévy triplet (a, σ^2, Π), there exist three independent Lévy processes, in the same probability space $\mathbb{P}$, denoted as $L_t^{(1)}, L_t^{(2)}, L_t^{(3)}$ where

- $L_t^{(1)} = \sigma W_t - at$ is a Wiener process with drift μ and diffusion σ^2 and corresponds to the continuous part of the measure,
- $L_t^{(2)} = \sum_{i=1}^{N_t}\xi_i$ is a compound Poisson process where $\{N_t : t \geq 0\}$ is a Poisson process with rate $\Pi(\mathbb{R}\backslash(-1,1))$ and $\{\xi_i : i \geq 1\}$ are iid with distribution $\Pi(dx)/\Pi(\mathbb{R}\backslash(-1,1))$ concentrated on $x : |x| \geq 1$, and
- $L_t^{(3)}$ whose characteristic exponent is given by $\Psi(u) = \int_{0<|x|<1}\left(1 - e^{iux} + iux\right)\Pi(\mathrm{dx})$ and it can be shown that it corresponds to the superposition of a countable number of independent compound Poisson processes with different arrival rates and an additional linear drift.

Thus the process defined by $L_t = L_t^{(1)} + L_t^{(2)} + L_t^{(3)}$ is then a Lévy process with triplet (a, σ^2, Π).

2.4.2 Families of Lévy processes

In this section we provide the details of a few of the special cases of Lévy processes that we shall use in our sychronisation models, specifically stable processes, tempered stable processes and the β-family of Lévy processes.

2.4.2.1 The stable and tempered stable processes

A stochastic process L_t is said to be a stable process [66, 67] if it is a Lévy process which satisfies the scaling property, that is under the law $\mathbb{P}$, for every $c > 0$, the process $cL_{tc^{-\alpha}}$ has the same law as L. Here the stability parameter α is bounded as $\alpha \in (0, 2]$, and the case $\alpha = 2$ corresponds to Brownian motion. The distributions of the stable and tempered stable processes can be expressed through their characteristic functions as follows: Let $L_t^{(S)}, t \geq 0$ be a spectrally positive stable process, then

$$\begin{aligned}\mathbb{E}\left[e^{iuL_t^{(S)}}\right] &= \exp\left[ta\Gamma(-\alpha)\cos\left(\frac{\pi\alpha}{2}\right)|u|^\alpha\left(1 - i\tan\frac{\pi\alpha}{2}\mathrm{sgn}(u)\right)\right], \\ &= \begin{cases} \exp\left[t\int_{\mathbb{R}_+}(e^{iux}-1)\frac{a}{x^{\alpha+1}}\mathrm{d}x\right], & \text{if } \alpha \in (0,1), \\ \exp\left[t\int_{\mathbb{R}_+}(e^{iux}-1-iux)\frac{a}{x^{\alpha+1}}\mathrm{d}x\right], & \text{if } \alpha \in (1,2), \end{cases}\end{aligned} \tag{2.9}$$

for some $a > 0$ and stability index $\alpha \in (0, 1) \cup (1, 2)$. Note that in the case of $\alpha = 2$, the stable process corresponds to that of a Brownian motion. We exclude the Cauchy process ($\alpha = 1$) and write $S(\alpha, a)$ for the probability distribution of S_1. The process $(S_t)_{t\geq 0}$ is the scaling limit of a random walk with power-law jumps: $c^{-1/\alpha}(L_1 + \cdots + L|ct|) \to S_t$ in distribution as the time scale $c \to \infty$ when all the random variables L_i satisfy $\mathbb{P}(L_i > x) = ax^{-\alpha}$ for large $x > 0$. The totally positively skewed stable process $(S_t)_{t\geq 0}$ takes values in $\mathbb{R}_+$ if $\alpha \in (0, 1)$ and in $\mathbb{R}$ if $\alpha \in (1, 2)$. The moments of the stable process are given by $\mathbb{E}[(L_t^\theta)^m] = \infty$ for all $m > \alpha$, which implies that for $\alpha < 1$ even the expectation of L_t^θ diverges. The density $f_{S(\alpha,a)}(x)$ of the stable distribution $S(\alpha, a)$ exists however it cannot be obtained as a closed-form solution and only by numerically inverting the characteristic function. More generally, a (two-sided) stable process $(L_t^\theta)_{t\geq 0}$ of index α can be obtained from two independent positively skewed stable processes $(S_t)_{t\geq 0}$ and $(\tilde{S}_t)_{t\geq 0}$ as $L_t^\theta := (1 + \theta)S_t - (1 - \theta)\tilde{S}_t$ with $-1 \leq \theta \leq 1$.

An extension of the stable processes can be seen in the class of processes known as the tempered stable processes [68, 69]. We define the one-sided tempered stable processes in the following manner. Let $L_t^{(TS)}, t \geq 0$ be a centered and spectrally positive tempered

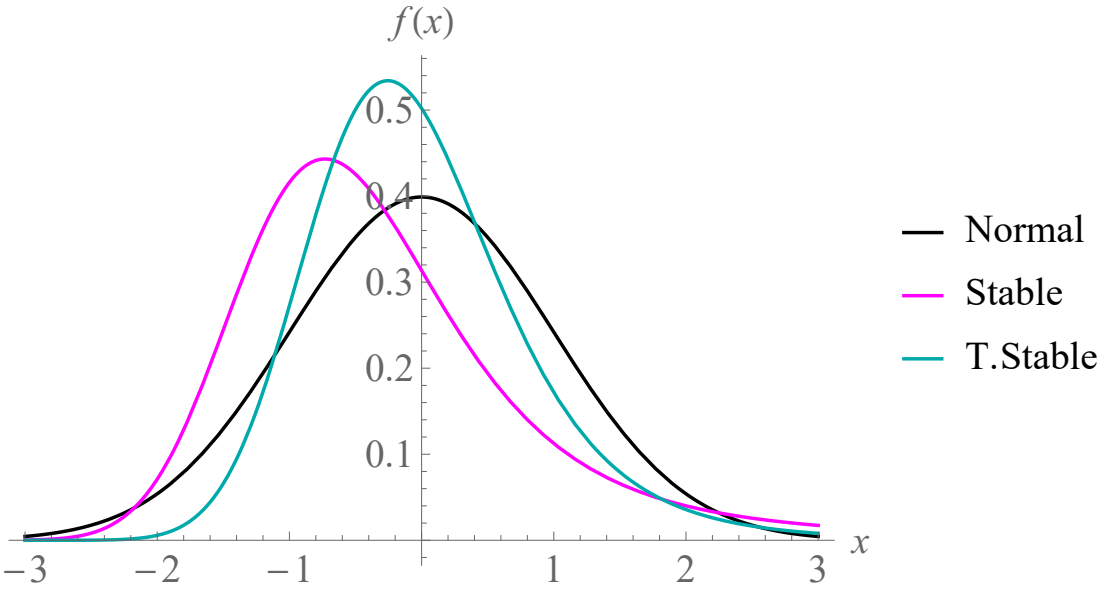


FIGURE 2.4: Comparison of Gaussian, one-sided stable and one-sided tempered stable densities.

stable process, then

$$\begin{aligned}\mathbb{E}\left[e^{iuL_t^{(TS)}}\right] &= \exp\left[ta\Gamma(-\alpha)\left((\lambda-iu)^\alpha-\lambda^\alpha+iu\alpha\lambda^{\alpha-1}\right)\right],\\ &= \exp\left[t\int_{\mathbb{R}_+}(e^{iux}-1-iux)a\frac{e^{-\lambda x}}{x^{\alpha+1}}\mathrm{d}x\right]\end{aligned} \tag{2.10}$$

where $\lambda \geq 0$ is the tempering parameter. Again, a (two-sided) tempered stable process $(L_t^{(TS)})_{t\geq 0}^{\theta}$ of index α can be obtained from two independent positively skewed tempered stable processes $(L_t^{(TS)})_{t\geq 0}^{\theta}$ and $(\tilde{L}_t^{(TS)})_{t\geq 0}$ as $(L_t^{(TS)})_{t\geq 0} := (1+\theta)L_t^{(TS)} - (1-\theta)\tilde{L}_t^{(TS)}$ with $-1 \leq \theta \leq 1$. Thus the centred noise is specified by four parameters $\alpha, \sigma, \theta, \lambda$, generalising from the one parameter of mean-zero Gaussian noise, σ.

On Figure 2.4 we provide a comparison of the densities for the stable and tempered stable processes and the Gaussian distribution for $\sigma = 1$, $\alpha = 1.5$ and $\lambda = 1$ and where we parameterise the scaling parameter as

$$a := -\frac{1}{2}\sigma^2\frac{1}{\Gamma(-\alpha)\cos(\pi\frac{\alpha}{2})}$$

so that $\Psi_t^{(S)}(u) \to \exp(-\frac{1}{2}\sigma^2u^2)$ as $\alpha \to 2$ to recover the Gaussian distribution in the limit.

2.4.2.2 The β-family of Levy processes

In this final section we shall briefly review the so-called β-family of Lévy processes. The β-family of Lévy processes, originally introduced by Kuznetsov [5, 70, 71], is a flexible class of Lévy processes thanks to a large parameter set and its similarity to many of the more popular processes such as the tempered stable processes as mentioned in the previous section. One of the most useful proprieties of the β-family of Lévy processes is that there exists semi-explicit expressions for the laws of the extrema of these processes. The existence of these laws allows for a method of simulating paths of Wiener-Hopf Monte Carlo method which we shall discuss in a following section. Some applications of the β-family can be seen in financial models in [70, 72].

The β-family of Lévy processes is defined as a 10-parameter Lévy process which has the characteristic exponent

$$\begin{aligned}\Psi(u) =& iau + \frac{1}{2}\sigma^2 u^2 + \frac{c_1}{\beta_1}\left\{B(\alpha_1, 1-\lambda_1) - B(\alpha_1 - \frac{iu}{\beta_1}, 1-\lambda_1)\right\} \\ &+ \frac{c_2}{\beta_2}\left\{B(\alpha_2, 1-\lambda_2) - B(\alpha_2 + \frac{iu}{\beta_2}, 1-\lambda_2)\right\}, \end{aligned} \tag{2.11}$$

where $a, \sigma \in \mathbb{R}, c_1, c_2, \alpha_1, \alpha_2, \beta_1, \beta_2 > 0$, $\lambda_1, \lambda_2 \in (0,3)\setminus\{1,2\}$ and $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ is the Beta function and $\Gamma(x)$ is the Gamma function. The large number of parameters in the β-family allows for great flexibility in the choice of paths, specifically we can obtain paths of either bounded or unbounded variation through at least of the parameter choices $\sigma \neq 0, \lambda_1 \in (2,3)$ or $\lambda_2 \in (2,3)$ for unbounded variation or all of the parameter choices $\sigma = 0, \lambda_1 \in (0,2)$ or $\lambda_2 \in (0,2)$ holding for bounding variation. Additionally we can obtain either finite or infinite jump activity in the jump component through the parameter choice $\lambda_1, \lambda_2 \in (1,3)$ for infinite jump activity, and $\lambda_1, \lambda_2 \notin (1,3)$ for finite jump activity. Further details of the β-family are given in Appendix B.1.

2.5 Simulation schemes for Lévy processes

To conclude our preliminary discussion, we discuss three methods of simulating Lévy processes that will be used in all the subsequent chapters of this thesis, namely the the Euler method [2], the acceptance-rejection method for the simulation stable and tempered stable processes [3, 4], and the WHMC algorithm [5, 70, 71].

2.5.1 The Euler scheme for Lévy processes

Similarly to the case of a diffusion process as discussed in Section 2.3.1, there exists an Euler scheme for the simulation of Lévy driven SDEs (see for e.g. [2]). We begin by considering the SDE

$$\mathrm{d}X_t = f(X_t)dL_t. \tag{2.12}$$

with some initial condition X_0 and where $f : \mathbb{R} \to \mathbb{R}$ is some deterministic function and L_t is a Lévy process with characteristic triplet (μ, σ, ν) and initial condition $L_0 = 0$. The Euler method is then to discretise Eq. (2.12) in time where the discretisation step on the time interval $[0, T]$ given by $\Delta t = T/N$. The approximate solution $\hat{X}$ with initial condition $\hat{X}_0 = X_0$ is incrementally given by

$$\hat{X}_{(n+1)\Delta t} = \hat{X}_{n\Delta t} + f(\hat{X}_{n\Delta t})\Delta L,$$

where $\Delta L = L_{(n+1)\Delta t} - L_{n\Delta t}$. Since we assume that the increments ΔL are stationary and independent, we can then take replicated samples $\Delta L = L_{\Delta t}$ for each increment of the algorithm. The Euler scheme for simulating Lévy SDEs is summarised in Algorithm 2.

Algorithm 2 Euler simulation of Lévy driven SDEs

1: $\Delta t \leftarrow T/N$
2: **for** $n = 1, \ldots, N$ **do**
3: $\quad \Delta L \sim \mathbf{Law}(L_{\Delta t})$
4: $\quad X(n+1) = X(n) + f(X(n)) \cdot \Delta L$
5: **return** X

2.5.2 Poisson and Gaussian approximations of Lévy processes

Recall from the Lévy-Ito decomposition of a Lévy process from Section 2.4.1, that we can decompose L_t into the sum of a Wiener process, a linear drift and a pure jump process. In a similar sense to this decomposition we can decompose $L_t = L(t)$ into the following three independent Lévy processes

$$L(t) = L^{(1)}(t) + L^{\epsilon}(t) + L_{\epsilon}(t),$$

where $L^{(1)}(t)$ is a Wiener process, $L^{\epsilon}(t)$ is a compound Poisson process with a drift and the distribution of jumps proportional to $\nu^{\epsilon} = \nu_{\{|x|>\epsilon\}}$ and the process $L_{\epsilon}(t)$ has a zero

mean with a fixed Lévy measure $\nu_\epsilon = \nu_{\{|x|\leq\epsilon\}}$. The last term L_ϵ can be interpreted as the small jump component of $L(t)$. When ϵ is small, we can take the approximation

$$L(t) \approx L^{(1)}(t) + L^\epsilon(t),$$

and thus simulate the L_t as if it were a jump-diffusion process. However this is not feasible in the case the intensity of the small jumps is high. Fortunately we can take an approximation result provided in Theorem 2.1 from Asmussen and Rosinski in [73] which allows us to approximate, under the condition that $v(\epsilon)/\epsilon^2 \to \infty$, the small jumps X_ϵ by an independent Brownian motion.

Theorem 2.1 *[Asmussen and Rosinski.]*
Let L_t be a Lévy process with characteristics $(a, 0, \nu)$. denote $v(\epsilon) = \int_{-\epsilon}^{\epsilon} x^2\nu(\mathrm{d}x)$. If $v(\epsilon)/\epsilon^2 \to \infty$ as $\epsilon \downarrow 0$, then

$$\frac{L(t) - L^\epsilon(t)}{\sqrt{v(\epsilon)}} \xrightarrow{d} W_t \qquad \text{as } \epsilon \downarrow 0,$$

for and an independent Brownian motion W_t.

Thus using the Gaussian approximation of L_ϵ we can write

$$L(t) \approx L^{(1)}(t) + L^\epsilon(t) + v(\epsilon)W(t),$$

and effectively simulate $L(t)$ as if it were a jump-diffusion process.

Stable and tempered stable processes allow for Gaussian approximations of small jumps as is demonstrated in [3] however this is not the case for a gamma process. For more information on the Gaussian approximation of small jumps see for e.g. [73, 74].

2.5.3 Simulation through characteristic inversion

In the simulation of Lévy processes, many methods require some form of sampling from the distribution of the process under consideration. However, in general, we often don't have any information about the process aside from the characteristic function. In this part we discuss a procedure for sampling from the distribution of a Lévy process through an inversion of the characteristic function.

Using the so-called Lévy inversion theorem, which states that if $\phi(\cdot)$ is a characteristic function of the distribution function F for some scalar random variable X, then we can

express the distribution function of the process in terms of its characteristic function as

$$F(x) = F(0) + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1 - e^{-iux}}{iu}\phi(u)\mathrm{d}u,$$

where $F(0)$ is some constant, thus for every interval (a, b) we have

$$\int_a^b \mathrm{d}F(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-iua} - e^{-iub}}{iu}\phi(u)\mathrm{d}u,$$

which then allows us to express the density of the process as

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iux}\phi(u)\mathrm{d}u.$$

The constant term in the expression for the distribution function can be eliminated thanks to the following result from Gil-Pelaez in [75] as

$$F(x) = \frac{1}{2} + \frac{1}{2\pi}\int_{0}^{\infty} \frac{e^{iux}\phi(u) - e^{-iux}\phi(u)}{iu}\mathrm{d}u.$$

Practical implementation of the inversion theorem is done through the use of the so-called fast Fourier transform (FFT) algorithm developed by Cooley and Tukey in [76] which allows for the computation of the discrete Fourier transform, DFT, and its inverse the, IDFT, in $n\log(n)$ operations. For some sequence of complex numbers x_k, $k = [1, 2, \ldots, N]$ the DFT transforms x_k into the sequence X_k, where the transformed sequence is given by

$$X_k = \sum_{n=1}^{N} x_n e^{-\frac{2\pi ikn}{N}}.$$

In order to apply the FFT to the characteristic function, we first approximate the integral through Simpson's rule as follows:

$$f(x) \approx \frac{1}{\pi}\sum_{n=1}^{N} e^{-i2\pi u_n x}\phi(u_n), \Delta u,$$

where $u_n = \Delta u \cdot n$ for some Δu. Discretising the x range as $x_j = x_L + \Delta_x(k-1)$ for $k = 1, \ldots, N$ where $[x_L, x_U]$ is the support of x and $\Delta_x = \frac{x_U - x_L}{N-1}$, we write

$$f(x_k) = \mathrm{Re}\left[\frac{1}{2}e^{-i\Delta u\cdot(n-1)x_L}\sum_{n=1}^{N} e^{-i\Delta u\cdot\Delta x\cdot(k-1)(n-1)}\phi(u_n)\Delta u\right].$$

By selecting the discretisations such that $\Delta u \cdot \Delta x = 2\pi/N$ (that is we choose $\Delta u = \frac{2\pi}{N}\frac{N-1}{x_U - x_L}$), we see that $f(x_k)$ is now in the desired form for the FFT algorithm.

2.5.4 On the simulation of stable and tempered stable variates

In this section we focus on simulating variates from the stable and tempered stable Lévy distributions in a similar manner as presented in [3]. In the methods discussed, we shall only consider the spectrally positive versions of these distributions however the algorithms can be extended to consider the full bilateral case. For more information on the stable distribution see for e.g [62, 77] and for the tempered stable distribution see for e.g [68, 69, 78].

2.5.5 Sampling from the stable distribution

The distribution of a spectrally positive stable process, denoted as $S(\alpha, a) := \mathcal{L}(L_1^{(s)})$, can be simulated exactly thanks to a result from Chambers et al. in [4] which gives the following representation

$$S(\alpha, a) = (-a\Gamma(-\alpha)\cos(\pi\alpha/2))^{1/\alpha} \frac{\sin(\alpha U + \theta)}{(\cos U \cos\theta)^{1/\alpha}} \left(\frac{\cos((1-\alpha)U - \theta)}{E} \right)^{\frac{1-\alpha}{\alpha}}, \quad (2.13)$$

where $\theta = \arctan(\tan(\pi\alpha/2))$, U is a uniform random variable on $(-\pi/2, \pi/2)$, and E is a standard exponential random variable independent of U. Note that by the scaling property, we can generalise this result to any time using $S(\alpha, a) := \mathcal{L}(t^{-1/\alpha} L_1^{(s)})$ which implies that

$$S(\alpha, ta) = \mathcal{L}\left(t^{-1/\alpha} f_{S(\alpha,a)}(t^{-1/\alpha} x) \right).$$

2.5.6 Acceptance-rejection sampling

Before we discuss sampling from the tempered-stable distribution, a necessary preliminary that will first be discussed is the acceptance-rejection sampling method or simply rejection sampling. For more information on acceptance-rejection sampling; e.g., see [79, 80].

The idea behind acceptance-rejection sampling is that instead of simulating some distribution F with PDF $f(x)$ that might be difficult to sample from, we find an alternate distribution G with density function $g(x)$ that we can effectively sample from and then accepting the sample from G with probability $f(x)/(cg(x))$. Ideally the proposal distribution $g(x)$ is 'close' to $f(x)$ and moreover the ratio $f(x)/g(x)$ is bounded by the constant $c > 0$; $\sup_x \{f(x)/g(x)\} \leq c$. That is we would like c to be as close to 1 as possible. This acceptance-rejection algorithm can be summarised in Algorithm 3 below.

Algorithm 3 Acceptance-rejection sampling

1: Fix $c > 0$
2: **repeat**
3: $\quad Y \sim \mathbf{Law}(G)$
4: $\quad U \sim \text{Uniform}(0, 1)$
5: **until** $U \leq \frac{f(Y)}{cG(Y)}$
6: **return** $X \leftarrow Y$

The unconditional acceptance probability of the above algorithm can be shown to be $p = \mathbb{P}(U \leq \frac{f(Y)}{cg(Y)}) = 1/c$. Hence the further G is from F, the more likely we are to reject Y. We can also show through a series of computations that the conditional distribution of Y given $U \leq \frac{f(Y)}{cg(Y)}$ as follows

$$\begin{aligned}
\mathbb{P}(Y \leq y | U \leq \frac{f(Y)}{cg(Y)}) =& \frac{\mathbb{P}(U \leq \frac{f(Y)}{cg(Y)} | Y \leq y)\mathbb{P}(Y \leq y)}{\mathbb{P}\left(U \leq \frac{f(Y)}{cg(Y)}\right)}, \\
=& \frac{\mathbb{P}(U \leq \frac{f(Y)}{cg(Y)}, Y \leq y)}{G(y)} \cdot \frac{G(y)}{1/c}, \\
=& c \int_{-\infty}^{y} \mathbb{P}(U \leq \frac{f(Y)}{cg(Y)}, w \leq y) g(w) \mathrm{d}w, \\
=& c \int_{-\infty}^{y} \frac{f(w)}{cg(w)} g(w) \mathrm{d}w, \\
=& c \int_{-\infty}^{y} f(w) \mathrm{d}w, \\
=& F(y).
\end{aligned}$$

As an example of the acceptance-rejection sampling method, consider the special case of the beta distribution

$$f(x) = bx^n(1-x)^n, \quad b = \left[\int_0^1 x^n(1-x)^n.\right]^{-1}.$$

We now set $g(x) = 1$ for $x \in \mathrm{U}(0, 1)$. Then $f(x)/g(x) = f(x)$ and $c = \sup_x f(x)$ which is found at $x = 1/2$ and thus $c = b(1/4)^n$. The acceptance-rejection algorithm is thus as follows

1. Generate U and Y independently from the uniform distribution.
2. If $U \leq 4^n(Y(1-Y))^n$, then set $X = Y$; otherwise go back to 1.

2.5.7 Sampling from the tempered stable distribution

The method of sampling from the tempered stable distribution, unlike the stable distribution, varies according to the value of the stability index α. In the case where $\alpha \in (0,1)$, there exists exact acceptance-rejection sampling schemes in the sense that the variate produced by the scheme is equal in law to that of the tempered stable distribution. However in the case where $\alpha \in (1,2)$ there does not exist a practically exact sampling method in the sense that the currently used methods are only exact under an asymptotic limit which would require an infinite computational time to realise.

We start by looking at an exact sampling method for the tempered stable process that is when $\alpha \in (0,1)$. Setting the notation of the tempered stable process as $TS(\alpha, a, \lambda)$ and the tempered stable subordinator as $TS'(\alpha, a, \lambda)$ where the tempered stable subordinator is obtained by adding back the centering term as follows: $TS'(\alpha, a, \lambda) = \mathcal{L}\left(L_1^{ts} + \Gamma(1-\alpha)a\lambda^{\alpha-1}\right)$. Since we can express the density of the tempered stable subordinator in terms of the density of the stable process (e.g., see [81, 82]) as

$$f_{TS'(a,\alpha,\lambda)}(x) = \frac{e^{\lambda x}}{\mathbb{E}\left[e^{-\lambda L_1^{(S)}}\right]} f_{S(a,\alpha)}(x) = e^{-\lambda x - a\Gamma(-\alpha)\lambda^{\alpha}} f_{S(a,\alpha)}(x), \tag{2.14}$$

it follows that the density of the tempered stable process can be expressed in terms of the stable density as

$$\begin{aligned} f_{TS(a,\alpha,\lambda)}(x) =& f_{TS'(a,\alpha,\lambda)}(x - \Gamma(1-\alpha)a\lambda^{\alpha-1}), \\ =& e^{-\lambda x - a(\alpha+1)\Gamma(-\alpha)\lambda^{\alpha}} f_{S(\alpha,a)}(x - \Gamma(1-\alpha)a\lambda^{\alpha-1}). \end{aligned}$$

Since we can sample from the stable distribution thanks to the result in Eq. (2.13), we can then apply the acceptance-rejection algorithm discussed in Section 2.5.6. Setting

$$\begin{aligned} c =& \sup_x \left\{ \frac{f_{TS'(a,\alpha,\lambda)}(z)}{f_{S(a,\alpha)}(x)} \right\}, \\ =& \sup_x \left\{ e^{-\lambda x - \alpha\Gamma(-\alpha)\lambda^{\alpha}} \right\}, \\ =& e^{-\alpha\Gamma(-\alpha)\lambda^{\alpha}}, \end{aligned}$$

since $\lambda > 0$, the acceptance probability $\frac{f(Y)}{cg(Y)}$ in the acceptance-rejection algorithm is then given by

$$\begin{aligned}\frac{f(Y)}{cg(Y)} =& \frac{f_{TS'(a,\alpha,\lambda)}(Y)}{cf_{S(a,\alpha)}(Y)},\\ =& \frac{e^{-\lambda Y - \alpha\Gamma(-\alpha)\lambda^\alpha}}{e^{-\alpha\Gamma(-\alpha)\lambda^\alpha}},\\ =& e^{-\lambda Y},\end{aligned}$$

where Y is a random variate from the stable distribution $S(\alpha, \Delta a)$. By accounting for the difference due the centering term between $TS(\alpha, a, \lambda)$ and $TS'(\alpha, a, \lambda)$, we can simulated variates from the tempered stable distribution as outlined in Algorithm 4.

Algorithm 4 Sample from the law of $TS(\alpha, \Delta a, \lambda)$, $\alpha \in (0, 1)$

1: **repeat**
2: sample $U \sim \text{Uniform}(0, 1)$
3: sample $Y(\Delta) \sim \textbf{Law}(S(\alpha, \Delta a))$
4: **until** $U \leq \exp(-\lambda Y(\Delta))$
5: **return** $Y_1 = Y(\Delta) - \Delta\Gamma(1-\alpha)a\lambda^{\alpha-1}$

As we previously mentioned, Algorithm 4 is exact in the sense that the variates produced by the algorithm, Y_1, are equal in distribution to $TS(\alpha, \Delta a, \lambda)$. However in the case that $\alpha \in (1, 2)$ this is no longer the case as the support is now the real line $\mathbb{R}$, as opposed to $\mathbb{R}_+$, upon which the exponential tempering $e^{-\lambda z}$ explodes at either positive or negative infinity. In the case that $\alpha \in (1, 2)$, the law of $TS(\alpha, \Delta a, \lambda)$ can be approximately simulated through a method similar to Algorithm 4 through the introduction of a truncation of the real line to the domain upon which the exponential tempering $e^{-\lambda z}$ is performed. This truncation is denoted by the parameter c. This method as discussed in [81] is outlined in Algorithm 5.

Algorithm 5 Sample from the law of $TS(\alpha, \Delta a, \lambda)$, $\alpha \in (1, 2)$

1: Fix $c > 0$
2: **repeat**
3: sample $U \sim \text{Uniform}(0, 1)$
4: sample $Y \sim \textbf{Law}(S(\alpha, \Delta a))$
5: **until** $U \leq \exp(-\lambda(Y(\Delta) + c))$
6: **return** $Y_2 = Y - \Delta\Gamma(1-\alpha)a\lambda^{\alpha-1}$

From [81] and [3], the acceptance rate of Algorithm 5 denoted as $p(\Delta, c)$ can be shown to be

$$p(\Delta, c) = \mathbb{E}\left[e^{-\lambda(Y(\Delta))+c}; Y(\Delta) > -c\right] + \mathbb{P}(Y(\Delta) \leq -c).$$

The density of Y_2 can then be given in terms of the acceptance rate $p(\Delta, c)$ through the following expression

$$f_{\mathcal{L}(Y_2(\Delta,c))} = \begin{cases} p(\Delta,c)^{-1} f_{S(\alpha,\Delta a)} & \text{if } x \in (-\infty, -c], \\ p(\Delta,c)^{-1} e^{-\lambda(x+c)} f_{S(\alpha,\Delta a)} & \text{if } x \in (-c, \infty). \end{cases}$$

We note that as $c \to \infty$, the density $f_{\mathcal{L}(Y_2(\Delta,c))}$ converges to that of the tempered stable subordinator density $f_{TS'(\alpha,\Delta a,\lambda)}(z)$

$$\begin{aligned} \lim_{c\to\infty} f_{\mathcal{L}(Y_2(\Delta,c))}(x) &= \lim_{c\to\infty} \frac{e^{-\lambda(x+c)}}{\mathbb{E}\left[e^{-\lambda(Y(\Delta))+c}; Y(\Delta) > -c\right] + \mathbb{P}(Y(\Delta) \leq -c)} f_{S(\alpha,\Delta a)}(x), \\ &= \frac{e^{-\lambda(x)}}{\mathbb{E}\left[e^{-\lambda(Y(\Delta))}\right]} f_{S(\alpha,\Delta a)}(x), \\ &= f_{TS(\alpha,\Delta a,\lambda)}(x), \end{aligned}$$

where in the second line we have used the identity from Eq. (2.14).

Optimal selection of c concerns the maximisation of the acceptance rate $p(\Delta, c)$ and the minimisation of the Kolmogorov-Smirnov distance $D_{KS}(\mathcal{L}(Y_2(\Delta, c)), TS(\alpha, \Delta a, \lambda))$. Since optimal selection can be a difficult task, in practise we select c from the 90% quartile of the tempered stable distribution, that is we choose

$$c = Q_{TS}(0.9, \alpha, \Delta a, \lambda),$$

where Q_{TS} is the quartile function of the tempered stable process $TS(\alpha, \Delta a, \lambda)$. The quartile function can be approximated through the density function obtain from an inversion of the characteristic function as discussed in Section 2.5.3.

We close off this section by noting that for the simulation of the bilateral tempered stable distribution, that is the process $L_t^{(TS)}$ whose characteristic function is given by

$$\mathbb{E}\left[e^{iuL_t^{(TS)}}\right] = \exp\left[t \int_{\mathbb{R}_0} (e^{iux} - 1 - iux)\left(a_+ \frac{e^{-\lambda_+ x}}{x^{\alpha_+ + 1}} \mathbf{1}_{\{x>0\}} + a_- \frac{e^{-\lambda_- |x|}}{x^{\alpha_- + 1}} \mathbf{1}_{\{x<0\}}\right) \mathrm{d}x\right],$$

where $\mathbf{1}$ is the indicator function, Algorithms 4 and 5, in the cases of $\alpha \in (0, 1)$ and $\alpha \in (1, 2)$ respectively, need to be implemented twice, once for the positive component, and the other for the negative component.

2.6 WHMC simulation of Lévy processes

In this section we present the Wiener-Hopf Monte Carlo (WHMC) algorithm for the simulation of Lévy processes originally proposed by Kuznetsov et al. in [5] and recently extended to the simulation of SDEs in [83]. For simplicity, we now describe only the Lévy process case. The WHMC method allows for the simulation of the joint law of the position and running of a Lévy process through the application of a path decomposition in terms of the Wiener-Hopf factors of the process, that is, it can simulate from the law of $(L_{\boldsymbol{g}}, \overline{L}_g)$ for some random time $\boldsymbol{g}$ such that the distribution can be constructed arbitrarily close to t.

The WHMC relies upon the construction of a so-called stochastic grid whose creation can be attributed to Carr in [84]. To create this stochastic grid, we first consider a sequence of i.i.d exponentially distributed random variable with unit mean given by $\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n$ that exists on a common product space with law $\mathbb{Q}$ orthogonal to the space $(\Omega, \mathcal{F}, \mathbb{P})$, that is $\mathbb{Q}$ is orthogonal to the space upon which the Lévy process exists. Due to the law of strong numbers we have the result that for all $t \geq 0$, as $n \uparrow \infty$

$$\mathbf{g} := \mathbf{g}(n, n/t) = \sum_{i=1}^{n} \frac{t}{n} \mathbf{e}_i \to t, \quad \mathbb{Q}\text{-almost surely.} \tag{2.15}$$

Since $\mathbf{g}$ is the sum of exponential random variables, it is itself equal in law to a Gamma random variable with shape n and scale n/t. The set of random points

$$\{0 = \mathbf{g}(0, n/t), \mathbf{g}(1, n/t), \ldots, \mathbf{g}(k, n/t)\},$$

where $\mathbf{g}(i, n/t) < \mathbf{g}(i+1, n/t)$ for $i \in [0, k-1]$ then forms the 'stochastic grid' in time. The distance between mesh points in the temporal grid form a sequence of i.i.d. exponentially distributed random variables. Using this 'stochastic grid' for a sufficiently large n, the law $\mathbb{P}\left(L_t \in \mathrm{d}x, \overline{L}_t \in \mathrm{d}y\right)$ can be approximated by $\mathbb{P} \times \mathbb{Q}\left(L_{\mathbf{g}} \in \mathrm{d}x, \overline{L}_{\mathbf{g}} \in \mathrm{d}y\right)$.

A path decomposition theory of Lévy processes, known as the Wiener-Hopf factorisation, tells us that for any Levy process L and some $q > 0$, we have that the random variables $\overline{L}_{e(q)}$ and $L_{e_q} - \overline{L}_{e(q)}$ are independent and that $L_{e_q} - \overline{L}_{e(q)} \overset{d}{=} \underline{\mathrm{L}}_{e(q)}$. The WHMC method utilities the Wiener-Hopf factorisation upon this 'stochastic grid' by introducing the random variables $V(n, n/t)$ and $J(n, n/t)$ on a new probability space whose law can be expressed in terms of the laws of $\overline{L}_{e(n/t)}$ and $\underline{\mathrm{L}}_{e(n/t)}$. The introduction of these two new random variables gives rise to the approximation that, for any n, it can be shown that (see for e.g [5])

$$\left(L_{\mathbf{g}(n,n/t)}, \overline{L}_{\mathbf{g}(n,n/t)}\right) \overset{d}{=} \left(V(n, n/t), J(n, n/t)\right),$$

where $V(k, n/t)$ and $J(k, n/t)$ are defined iteratively through

$$V(k, n/t) = V(k-1, n/t) + S^{(k)}_{n/t} + I^{(k)}_{n/t}, \tag{2.16}$$

$$J(k, n/t) = \max\left(J(k-1, n/t), V(k-1, n/t) + S^{(k)}_{n/t}\right), \tag{2.17}$$

for $k \geq 1$, $V(0, n/t) = J(0, n/t) = 0$ and $S^{(0)}_{n/t} = I^{(0)}_{n/t} = 0$ where the two i.i.d sequences of random variables $\left\{S^{(j)}_{n/t} : j \geq 1\right\}$ and $\left\{I^{(j)}_{n/t} : j \geq 1\right\}$ are distributed as $\overline{L}_{\frac{t}{n}\mathbf{e}_1}$ and $\underline{\mathrm{L}}_{\frac{t}{n}\mathbf{e}_1}$ respectively. From this we can make the observation that we need only simulate i.i.d copies of the random variables $S^{(1)}_{n/t}$ and $I^{(1)}_{n/t}$ in order to produce a realisation of $(L_{\mathbf{g}}, \overline{L}_{\mathbf{g}})$.

For some integrable function F, standard Monte Carlo techniques then give

$$\mathbb{E}\left[F(L_t, \overline{L}_t)\right] \simeq \frac{1}{M}\sum_{m=1}^{M} F\left(V^{(m)}(n, n/t), J^{(m)}(n, n/t)\right),$$

where $\left(V^{(m)}(n, n/t), J^{(m)}(n, n/t)\right)_{1\leq m\leq M}$ are i.i.d copies of $(V(n, n/t), J(n, n/t))_{1\leq m\leq M}$ and $M > 0$ is a large integer. By the strong law of large numbers, the right-hand side converges almost surely to $\mathbb{E}\big[F(L^{\uparrow}_t, \overline{L}_t)\big]$ as $n, M \to \infty$.

Algorithm 6 WHMC algorithm for sampling from the law of $(L_t, \overline{L}_t)$

Choose $n \geq 1$ so that $\mathbf{g} := \sum_{i=1}^{n} \frac{t}{n}\mathbf{e}_i \approx t$ by SLLN.
$\lambda \leftarrow n/t$
$L(0) \leftarrow 0$, $Z(0) \leftarrow 0$
for $i = 1, \ldots, n$ **do**
    sample $S \sim \mathbf{Law}(\overline{L}_{\mathbf{e}_1/\lambda})$
    sample $I \sim \mathbf{Law}(\underline{L}_{\mathbf{e}_1/\lambda})$
    $L(i) \leftarrow L(i-1) + S + I$
    $Z(i) \leftarrow \max(Z(i-1), L(i))$
return $(L(n), Z(n))$

The astute reader will notice that Algorithm 6 is only interesting if one can sample from the laws of $\overline{L}_t$ and $\underline{\mathrm{L}}_t$ in a computationally tractable manner. Such laws are sometimes achievable through Wiener-Hopf factorisation and there exist several families of processes for which explicit identities of the Wiener-Hopf factors are known (e.g., [5, 70, 71, 85]). We shall focus upon the the β-family of Lévy processes as discussed in Section 2.4.2.2 in the application of the WHMC method. We also note that this algorithm can be further refined through the extension to the multi-level Monte Carlo setting as discussed in Appendix A.4.

2.6.1 Random walk approximation of the WHMC method

In addition to providing a means to sample from the law of $(L_t, \overline{L}_t)$, the WHMC can also produce a reasonable approximation of the random walk of the process L_t. We first note that in the stochastic grid constructed in the WHMC method, in addition to the convergence result in Eq. (2.15) , we have that for any $s \in [0,t]$, by the SLLN

$$\mathbf{g}(k(n), n/t) \to s, \quad \mathbb{Q}\text{-almost surely,}$$

for some sequence $k(n) \in \{0, \ldots, n\}$. Thus we can view the stochastic grid as being dense in the interval $[0,t]$ as $n \to \infty$. and hence we can easily show that constructing the law of $(L_t, \overline{L}_t)$ in the WHMC algorithm produces an approximation of the vector $\big((L_0, \overline{L}_0), (L_{1/n}, \overline{L}_{1/n}), \ldots, (L_t, \overline{L}_t)\big)$.

Lemma 2.2 *[WHMC random walk approximation]*
Let L_t be a Lévy process, $\{\mathbf{g}(k, n/t), k \in [0,n]\}$ be the stochastic grid constructed as in Eq. (2.15) and V and J be sequences defined in Eq. (2.16) and Eq. (2.17) respectively. Then the sequence $\{(V(k,n/t), J(k,n/t))\}_{k\in[0,n]}$ approximates a random walk of $(L_t, \overline{L}_t)$ in the sense that

$$\begin{aligned}&\big((L_{\mathbf{g}(0,n/t)}, \overline{L}_{\mathbf{g}(0,n/t)}), \ldots, (L_{\mathbf{g}(n,n/t)}, \overline{L}_{\mathbf{g}(n,n/t)})\big) \overset{d}{=} \\ &((V(0,n/t), J(0,n/t)), \ldots, (V(n,n/t), J(n,n/t)))\,.\end{aligned}$$

We can see that Lemma 5.3 holds through an induction argument. When $k = 1$, this identity holds due to the Wiener-Hopf factorisation identity $(L_{\mathbf{e}(q)}, \overline{L}_{\mathbf{e}(q)}) \overset{d}{=} (\overline{L}_{\mathbf{e}(q)} + \underline{\mathrm{L}}_{\mathbf{e}(q)}, \overline{L}_{\mathbf{e}(q)})$. For some $k \geq 2$, we have that

$$\begin{aligned}\big(L_{\mathbf{g}(k,n/t)}, \overline{L}_{\mathbf{g}(k,n/t)}\big) &= \left(L_{\mathbf{g}(k-1,n/t)} + L_{\mathbf{e}_{(n/t)}}, \sup_{0\leq u\leq \mathbf{g}(k-1,n/t)} \overline{L}_u \bigvee \sup_{\mathbf{g}(k-1,n/t)\leq u\leq \mathbf{g}(k,n/t)} \overline{L}_u\right), \\ &\overset{d}{=} \left(L_{\mathbf{g}(k-1,n/t)} + L_{\mathbf{e}_{(n/t)}}, \sup_{0\leq u\leq \mathbf{g}(k-1,n/t)} \overline{L}_u \bigvee \big(L_{\mathbf{g}(k-1,n/t)} + \overline{L}_{\mathbf{e}(n/t)}\big)\right), \\ &\overset{d}{=} \Bigg(V(k-1,n/t) + S_k^{(n/t)} + I_k^{(n/t)}, \\ &\qquad J(k-1,n/t) \bigvee (V(k-1,n/t) + S_k^{(n/t)})\Bigg), \\ &= (V(k,n/t), J(k,n/t)),\end{aligned}$$

where we have used the stationary and independent increment property of the process L_t, the identity $L_{\mathbf{e}_{(n/t)}} = S_k^{(n/t)} + I_k^{(n/t)}$ and the induction assumption that

$$\left(L_{\mathbf{g}(k-1,n/t)}, \overline{L}_{\mathbf{g}(k-1,n/t)}\right) \stackrel{d}{=} (V(k-1,n/t), J(k-1,n/t)).$$

2.6.2 WHMC simulation for stable and tempered stable processes

We conclude this section on the simulation of Lévy processes by applying the WHMC algorithm from Section 2.6 in the context of simulating tempered stable process followed by a numerical comparison of the aforementioned algorithm with the acceptance-rejection method from Section 2.5.6 and the characteristic inversion method from Section 2.5.3.

Due to the versatility of the β-family of Lévy processes, we can obtain many of the more widely used Lévy processes as parametric limits of the the β-family. In the following Lemma, we consider how the stable and tempered stable processes can be obtain through the β-family [70].

Lemma 2.3 *[Convergence of a β-family process to a tempered stable process.] Let $\nu_{TS}(x)$ and $\nu_\beta(x)$ be the Lévy measures of a tempered stable process and a β-family process respectively where the measures are defined by*

$$\nu_{TS}(x) = a_+ \frac{e^{-\lambda_+ x}}{x^{\alpha_+ + 1}} \mathbf{1}_{\{x>0\}} + a_- \frac{e^{\lambda_- x}}{|x|^{\alpha_- + 1}} \mathbf{1}_{\{x<0\}},$$
$$\nu_\beta(x) = c_1 \frac{e^{-\alpha_1 \beta_1 x}}{(1-e^{-\beta_1 x})^{\lambda_1}} \mathbf{1}_{\{x>0\}} + c_2 \frac{e^{\alpha_2 \beta_2 x}}{(1-e^{\beta_2 x})^{\lambda_2}} \mathbf{1}_{\{x<0\}},$$

where the β-family has the parametric constraints $\alpha_{1,2} > 0$, $\beta_{1,2} > 0$, $c_{1,2} \geq 0$, $\lambda_{1,2} \in (0,3)$ and the tempered stable process has the parametric constraints $\alpha_{+,-} \in (0,2]\setminus\{1\}$, $\lambda_{+,-} \geq 0$, $a_{+,-} \geq 0$. If in the β-family we set $c_1 = a_+ \beta^{(1+\alpha_+)}$, $c_2 = a_- \beta^{(1+\alpha_-)}$, $\alpha_1 = \lambda_+ \beta^{-1}$, $\alpha_2 = \lambda_- \beta^{-1}$, $\lambda_1 = 1+\alpha_+$, $\lambda_2 = 1+\alpha_-$, $\beta_1 = \beta_2 = \beta$, then in limit as $\beta \downarrow 0$, $\nu_\beta(x)$ converges to $\nu_{TS}(x)$.

First considering the spectrally positive case, under the new parameterisation the measure of the β-family process is given by

$$\nu_\beta(x) = a_+ \frac{\beta^{1+\alpha_+} e^{-\lambda_+ x}}{(1-e^{-\beta x})^{1+\alpha_+}} \mathbf{1}_{\{x>0\}}.$$

Now taking the limit $\beta \to 0$

$$\begin{aligned}
\lim_{\beta \downarrow 0} \nu_\beta(x) =& a_+ e^{-\lambda_+ x} \lim_{\beta \downarrow 0} \left(\frac{\beta}{1 - e^{-\beta x}} \right)^{1+\alpha_+} \mathbf{1}_{\{x>0\}}, \\
=& a_+ \frac{e^{-\lambda_+ x}}{x^{\alpha_+ + 1}} \lim_{\beta^* \uparrow 0} \left(\frac{\beta^*}{1 - e^{\beta^*}} \right)^{1+\alpha_+} \mathbf{1}_{\{x>0\}}, \\
=& a_+ \frac{e^{-\lambda_+ x}}{x^{\alpha_+ + 1}} \mathbf{1}_{\{x>0\}}, \\
=& \nu_{TS}(x),
\end{aligned}$$

where in the second line we have made the transformation $\beta = -\beta^*/x$ and in the third line we have used the limit identity $\lim_{x \to 0} \frac{1-e^x}{x} = 1$. Thus in the spectrally positive case, the measure of the β-family process converges to that of the tempered stable process. A similar result holds for the spectrally negative case.

We can now apply the WHMC algorithm from Section 2.6 to simulate tempered stable processes in addition to the β-family processes. On Figure 2.5 we present a comparison of the distribution of the bilateral variate $L_{0.1}^{TS}$ approximated through the acceptance-rejection method, the characteristic inversion method and the WHMC method. In the simulation of the tempered stable variate we use the parameters $\alpha = 1.5$, $\lambda = 0.5$, $a = 1$, $\Delta = 0.1$ and $\beta = 0.3$ for the parametric limit in the WHMC method. In terms of simulation speed, the acceptance-rejection method was found to be the fastest of the three however the WHMC method has the advantage in that it provides an approximation of the path of the process up until the terminal time. We conclude by noting that in the application of these methods, we are essentially taking an approximation of an approximation thus resulting in the slight divergence between the densities of the three methods. In the WHMC method, numerical precision quickly becomes an issue in the limit as $\beta \to 0$, thus leading to the choice of $\beta = 0.3$ in the direct application of the parametric limit of the β-family. In the acceptance-rejection method, the choice of the truncation parameter c is limited to the 90th quartile of the tempered stable distribution. However in the limits $\beta \to 0$ and $c \to \infty$, these distributions converge.

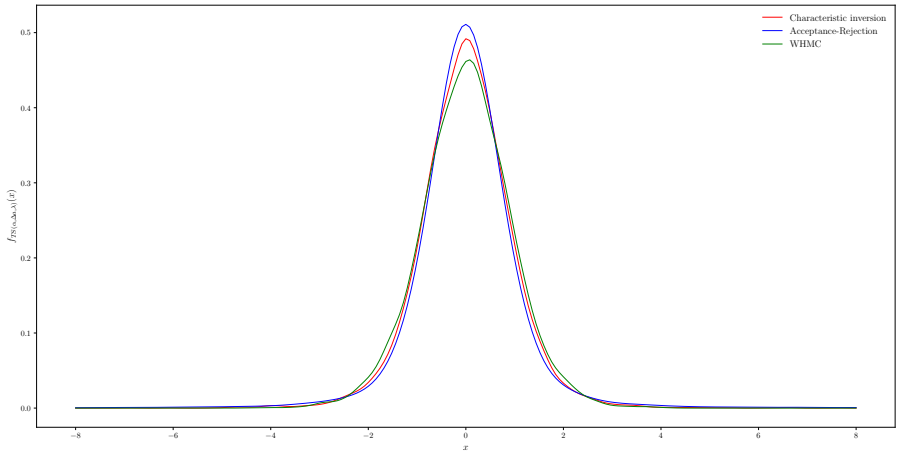


FIGURE 2.5: Distribution of the tempered stable variate $TS(\alpha, a\Delta, \lambda)$ under the acceptance-rejection method, the characteristic inversion method and the WHMC method for $\alpha = 1.5$, $\lambda = 0.5$, $a = 1$, $\Delta = 0.1$ and $\beta = 0.3$ in the WHMC method. The densities were obtained using a kernel density estimator from 10000 simulations.

Chapter 3

Two Network Kuramoto-Sakaguchi Model Under Tempered Stable Lévy Noise

In this chapter we examine a model of two interacting populations of phase oscillators labelled 'Blue' and 'Red'. To this we apply tempered stable Lévy noise, a generalisation of Gaussian noise where the heaviness of the tails parameterised by a power law exponent α can be controlled by a tempering parameter λ. This system models competitive dynamics, where each population seeks both internal phase synchronisation and a phase advantage with respect to the other population, subject to exogenous stochastic shocks. This work was done collaboratively with the DSTG which lead to the paper titled 'Two Network Kuramoto-Sakaguchi Model Under Tempered Stable Lévy Noise' in [22]. In this chapter, we present the paper in full. My personal contribution to this collaborative work is in the development of the numerical methods used to simulate and analyse the heavy-tailed versions of the model, the implementation of Bayesian optimisation and the 0-1 test and the production of plots throughout the paper and in the editing of the manuscript.

We study the system from an analytic and numerical point of view to understand how the phase lag values and the shape of the noise distribution can lead to steady or noisy behaviour. Comparing the analytic and numerical studies shows that the bulk behaviour of the system can be effectively described by dynamics in the presence of tilted ratchet potentials. Generally, changes in α away from the Gaussian noise limit, $1 < \alpha < 2$, disrupts the locking between Blue and Red, while increasing λ acts to

restore it. However we observe that with further decreases of α to small values, $\alpha \ll 1$, with $\lambda \neq 0$, locking between Blue and Red may be restored. This is seen analytically in a restoration of metastability through the ratchet mechanism, and numerically in transitions between periodic and noisy regions in a fitness landscape using a measure of noise. This non-monotonic transition back to an ordered regime is surprising for a linear variation of a parameter such as the power law exponent and provides a novel mechanism for guiding the collective behaviour of such a complex competitive dynamical system.

3.1 Introduction

Many biological, social and economic examples of complex systems contain both cooperative and competitive phenomena in tension that lead to both ordered and disordered behaviours. Various stylised features of these systems have been studied through the Kuramoto model [1] of synchronising oscillators and its many extensions; see the review articles [35–39].

In this chapter, we study synchronisation in what we call the 'Blue vs Red' model [6, 45] under the effect of heavy-tailed stochastic perturbations. This model extends the Kuramoto-Sakaguchi oscillator model (e.g., see [37–39]) to the case of two interacting populations of oscillators attempting to synchronise with each other under various frustrations. More precisely, we consider N blue agents in a undirected network given by adjacency matrix $\mathcal{B}_{ij}$, $(i,j \in \{1,\dots,N\} = \mathcal{B})$, and M red agents in a undirected network given by adjacency matrix $\mathcal{R}_{ij}$, $(i,j \in \{1,\dots,M\} = \mathcal{R})$. Let $\beta_i \in \mathbf{R}$ denote the phase of the blue agent $i \in \mathcal{B}$ and $\rho_i \in \mathbf{R}$ the phase of the red agent $i \in \mathcal{R}$. The nonlinear dynamical system we consider is

$$\begin{aligned} \dot{\beta}_i &= \omega_i - \zeta_B \sum_{j\in\mathcal{B}} \mathcal{B}_{ij} \sin(\beta_i - \beta_j) + f_{BR}(\beta_i, \rho_i) + \xi_i, \quad i \in \mathcal{B} \\ \dot{\rho}_i &= \nu_i - \zeta_R \sum_{j\in\mathcal{R}} \mathcal{R}_{ij} \sin(\rho_i - \rho_j) + f_{RB}(\rho_i, \beta_i) + \tilde{\xi}_i, \quad i \in \mathcal{R}, \end{aligned} \tag{3.1}$$

where $\xi_i, \tilde{\xi}_i$ are time-varying independent heavy-tailed white noises and f_{BR} and f_{RB} are interaction terms between the red and blue networks. These nonlinear interaction terms are given by

$$\begin{aligned} f_{BR}(\beta_i, \rho_i) &= -\zeta_{BR} \sum_{j\in\mathcal{R}} \mathcal{A}^{(BR)}_{ij} \sin(\beta_i - \rho_j - \phi) \\ f_{RB}(\rho_i, \beta_i) &= -\zeta_{RB} \sum_{j\in\mathcal{B}} \mathcal{A}^{(RB)}_{ij} \sin(\rho_i - \beta_j - \psi), \end{aligned}$$

where the matrices $\mathcal{A}^{(BR)}$ and $\mathcal{A}^{(RB)}$ of size $N \times M$ and $M \times N$ respectively are adjacency matrices representing the undirected and unweighted edges from 'Blue to Red' (BR) and 'Red to Blue' (RB). We assume that $\mathcal{A}^{(BR)} = \mathcal{A}^{(RB)^T}$ so that the edges between blue and red populations are symmetric. These interaction terms contain temporal lags ϕ and ψ, often called frustrations, between the populations similar to the Sakaguchi variation [86–89] of the Kuramoto model. This typically drives a tendency to disorder in tension with the ordering tendency of synchronisation.

Extending the Kuramoto model to multi-networks offers a viable approach to modelling competitive dynamics between ecological, social and organisational populations [6, 40–45, 90]; hence our use of the labels 'red' and 'blue'. It has been shown that richer behaviour in the dynamics of such models are observed under various forms of Gaussian white or coloured noise perturbations [36, 46–51]. Recent studies have also shown relationships between the heaviness of the noise distribution's tails and synchronisation behaviour of the Kuramoto model [52, 53]. The case of tempered stable noise, which we consider here, is particularly interesting due to its relation with tempered fractional diffusion [54, 55, 91–94] and are part of broader class of processes called Lévy processes. At the core is so-called stable Lévy noise where the stochastic process arises from a distribution with a heavy tail and scale law parametrised by an exponent α. As α is decreased from $\alpha = 2$ the noise deviates more strongly from Gaussian noise, and the tails become heavier with significant changes to the distribution below $\alpha = 1$. From a modelling perspective, *tempered* stable noise offers a number of useful attributes such as a small number of parameters, finite moments of all order, and a parameter λ that controls the tempering of the distribution's tails. Moreover, our approximations hold in a regime at the interface between order and noisy behaviour, which is often where competing populations seek to operate.

A key question will be how the tempered stable noise changes dynamics observed for the deterministic [45] and Gaussian stochastic [6] versions of the Blue-Red system. As in these works, we explore the dynamics of our model from both an analytic and numerical point of view. Indeed, in [45] we showed that much of the full dynamics of the system could be understood through linearisation around a fixed point describing respectively internally synchronised Blue and Red clusters and examining the system in the vicinity of the regime where Blue and Red lose relative phase locking, and the role of a projection of the system into zero and normal modes of Laplacians of the networks. In the locked configuration, whichever of Blue or Red is ahead in phase by an amount less than π may be deemed to have a competitive advantage over the other. The system is thus poised at this interface between order and noisy behaviour, with a transition through periodic behaviour where Blue and Red remain internally synchronised but drifting in relation to each other. It turns out that this regime corresponds to a particle in a

ratchet potential, which is a superposition of an oscillatory potential on a linear slope. In such systems there is an interplay of the size of the wells and the slope determining if the particle may become trapped or rolls down the slope. In the Blue-Red system this maps to the property that Blue and Red clusters lock, respectively drift, with respect to each other; drift would mean that Blue and Red rotate and thus the overall dynamics is periodic in time. In [6] we went on to show how this structure figured in the stochastic system, such that noise could trigger metastability, namely situations where even though deterministically Blue and Red may be locked with respect to each other, noise would enable drift and thus periodic behaviour. Nevertheless this analysis showed that the deterministic fixed point system leaves its fingerprint on the stochastic dynamics. Nevertheless as the type of noise becomes more exotic the potential for disruption of global synchronisation in the pure Kuramoto model, or locking for multiple populations, increases.

In this chapter we uncover a genuinely novel type of behaviour that is not merely incremental to previous insights, namely transitions between order, periodicity and noisy behaviour within the aforementioned nonlinear mechanism of metastability. Specifically, we discover a novel mechanism for restoring locking between Blue and Red in the presence of noise, or even diminishing the drift between the two when the system is deterministically periodic in behaviour.

The key results of this paper are several: that we may analytically derive transitions to and away from disruption of locking between Blue and Red for particular parameter ranges of the tempered stable Lévy noise in a regime where the deterministic system shows locked behaviour (in Fig.3.3); that in numerical computation for the full nonlinear dynamics both individual paths (Figs.3.4,3.5) of the fitness landscape for frustration parameters ϕ and ψ the imprint of the deterministic dynamics (Fig.3.7) survives for weak noise strength across a range of values for the power and tempering indices of the tempered stable Lévy noise (Fig.3.8); and that this shows a transition from noisy behaviour back to locking (Fig.3.11) in the regime predicted in the analytic approach. What is genuinely surprising is that this restoration of locking or suppression of drift in the presence of noise occurs for $\alpha \ll 1$ where naively the heavy tails of the stable noise case are extreme; in the absence of tempering decreasing α only destroys locking with monotonically increasing severity (Fig.3.10). Both the generalisation of such a two network frustrated Kuramoto model to tempered stable Lévy noise, and the detection of transitions to/from noisy behaviour of a nonlinear model of multiple competing populations of synchronising and frustrated oscillators is, to our understanding, novel.

To this end, this work is organised in the following manner. Section 3.2 provides the

preliminary details of our setup, in particular this part gives details about the networks interactions and how we model the heavy-tailed noise injected into the system. Section 3.3 provides an analytical approximation of the dynamics of the system near certain locked states for the deterministic version of the model and details the choice of networks used later in the numerical analysis. Section 3.4 discusses analytical results for the stochastic model in some special cases. Here we establish the result that the zero modes of the system are described by dynamics in the presence of tilted ratchet potentials manifesting metastability in regimes that would otherwise have been expected to be noisy. Section 3.5 focuses on the full nonlinear dynamics of the system using numerical simulations with aim to identify optimal internal synchronisation and phase lag of the Blue network against the Red network. We achieve these results using Bayesian optimisation to maximise an appropriately chosen objective function. We also study the presence of noisy paths given the choice of parameters in our noise through the implementation of the 0-1 test. Section 3.6 provides a discussion on the extension of our numerical optimisation to a wider parameter set.

3.2 Preliminaries

3.2.1 Network interactions and coupling

The $(N+M)\times(N+M)$ adjacency matrix for the corresponding external Blue-Red connections, labeled $\mathcal{M}$ has the following block off-diagonal form,

$$\mathcal{M} = \begin{pmatrix} 0 & \mathcal{A}^{(BR)} \\ \mathcal{A}^{(RB)} & 0 \end{pmatrix}.$$

The quantities ω_i, ν_i give the natural frequencies of the associated Blue and Red agents respectively; they may be drawn from some probability distribution. Finally, $\zeta_B, \zeta_R, \zeta_{BR}, \zeta_{RB}$ are coupling constants, respectively, for intra-Blue, intra-Red, Blue to Red and Red to Blue interactions. Asymmetry between Blue and Red potentially exists in the coupling constants and frustrations: ζ_{BR} and ζ_{RB}, and ϕ and ψ need not be equal.

3.2.2 Definitions for the linearised Blue-Vs-Red model

In this part we bring together the various elements needed to specify and analyse the model, from the deterministic system, the noise and the specific examples we will use to illustrate behaviours in the model. As discussed more generally in Section 2.1, we define the degree matrix d and the Laplacian matrix L for the blue and red networks in

a similar manner. The degree of Blue agent at node $i \in \mathcal{B}$ is the number of links from node i to other Blue agents,

$$d_i^{(B)} \equiv \sum_{j \in \mathcal{B}} \mathcal{B}_{ij}, \quad i \in \mathcal{B}.$$

The corresponding Blue *degree-matrix* $\mathcal{D}^{(B)}$ is a diagonal matrix with the degrees $d_i^{(B)}$ inhabiting the diagonal entries

$$\mathcal{D}_{ij}^{(B)} \equiv d_i^{(B)} \delta_{ij}, \quad \{i, j\} \in \mathcal{B}.$$

The matrices L in Eq. (3.21) constitute a *graph Laplacian*, where the Laplacian for the Blue population is given by the expression

$$L_{ij}^{(B)} \equiv \mathcal{D}_{ij}^{(B)} - \mathcal{B}_{ij}, \quad \{i, j\} \in \mathcal{B}.$$

Equivalent definitions for $L^{(R)}$ apply to the Red network.

Addressing the corresponding cross-network quantities, we define the degree with which a Blue agent at node i connects to Red agents as,

$$d_i^{(BR)} \equiv \sum_{j \in \mathcal{B} \cup \mathcal{R}} \mathcal{M}_{ij} = \sum_{j \in \mathcal{R}} \mathcal{A}_{ij}^{(BR)}, \quad i \in \mathcal{B}.$$

We note that in this instantiation of the model $\mathcal{A}^{(BR)}$ is the transpose of $\mathcal{A}^{(RB)}$, and vice-versa (it need not be so, however), so there is symmetry between the total number of degrees between the Blue and Red networks:

$$d_T^{(BR)} = d_T^{(RB)} = \sum_{i \in \mathcal{B}} d_i^{(BR)} = \sum_{i \in \mathcal{R}} d_i^{(RB)}.$$

The diagonal degree matrix $\mathcal{D}^{(BR)}$ which encodes all the Blue to Red links is given by,

$$\mathcal{D}_{ij}^{(BR)} \equiv d_i^{(BR)} \delta_{ij}, \quad i \in \mathcal{B}, \quad j \in \mathcal{B} \cup \mathcal{R},$$

where the final M diagonal entries of $\mathcal{D}^{(BR)}$ are zero. This finally leads to the cross-network Laplacian from Blue to Red

$$L_{ij}^{(BR)} \equiv \mathcal{D}_{ij}^{(BR)} - \mathcal{M}_{ij}, \quad i \in \mathcal{B}, \quad j \in \mathcal{B} \cup \mathcal{R}.$$

Similar considerations lead to an equivalent expression for the Red to Blue cross-network Laplacian $L^{(RB)}$.

The Blue and Red network Laplacians obey the following eigenvalue equations,

$$\sum_{j\in\mathcal{B}} L_{ij}^{(B)} e_j^{(B,r)} = \lambda_r^{(B)} e_i^{(B,r)} \quad i \in \mathcal{B}, \quad \sum_{j\in\mathcal{R}} L_{ij}^{(R)} e_j^{(R,r)} = \lambda_r^{(R)} e_i^{(R,r)} \quad i \in \mathcal{R}.$$

The graph eigenvalues are well-studied objects [33]. For instance, the *zeroth*-eigenvalue is always zero valued, and the remaining eigenvalues are all real, positive, semi-definite and can be ordered as follows,

$$0 = \lambda_0^{(B)} \leq \lambda_1^{(B)} \leq \lambda_2^{(B)} \leq \cdots \leq \lambda_N^{(B)}$$

where we have used the Blue network as an example. The normalised zero eigenvectors are

$$e_i^{(B,0)} = \frac{1}{\sqrt{N}} \quad i \in \mathcal{B}, \quad e_i^{(R,0)} = \frac{1}{\sqrt{M}} \quad i \in \mathcal{R}.$$

They provide an alternate expression for the Blue and Red centroids:

$$B = \frac{1}{\sqrt{N}} \sum_{i\in\mathcal{B}} \beta_i e_i^{(B,0)}, \quad P = \frac{1}{\sqrt{M}} \sum_{j\in\mathcal{R}} \rho_j e_j^{(R,0)}.$$

Projections onto the eigenvectors are:

$$\omega^{(s)} = \sum_{i\in\mathcal{B}} \omega_i e_i^{(B,s)}, \quad \bar{\omega} = \frac{1}{N} \sum_{i\in\mathcal{B}} \omega_i, \quad d_s^{(BR)} = \sum_{i\in\mathcal{B}} d_i^{(BR)} e_i^{(B,s)},$$

and equivalent expressions hold for the relevant Red quantities.

Fluctuations b_i and p_j are expanded in the following non-zero *normal-modes*:

$$b_i = \sum_{r\in\mathcal{B}^E/\{0\}} x_r e_i^{(B,r)} \quad i \in \mathcal{B}, \quad p_j = \sum_{r\in\mathcal{R}^E/\{0\}} y_r e_j^{(R,r)} \quad j \in \mathcal{R}.$$

3.2.3 Tempered stable Lévy noise

The stochastic terms in our system can be rewritten by considering a stochastic differential equation formulation of our system by formally multiply both sides of Eq.(3.1) by dt and assume $\xi_i \, \mathrm{dt} = \mathrm{d}L_i(t)$ where $L_i = (L_i(t))_{t\geq 0}$ is a tempered stable Lévy process, for $i = 1, \ldots, N$. We do the same for the $\tilde{\xi}_i$ for $i = 1, \ldots, M$ terms which are simply independent copies. This gives us $N + M$ noise terms that we relabel L_i for $i = 1, \ldots, N+M$. These processes have 3 parameters governing the distribution of their increments: the stability parameter α, the tempering parameter λ, and the asymmetry

parameter θ. From the Lévy-Khintchine formula, the characteristic exponent Λ satisfying $\mathbf{E}[e^{kL_i(t)}] = e^{tk\Lambda(k)}$ of the process is given by $\Lambda(k) := C(k)$ for $0 < \alpha < 1$ and $\Lambda(k) := C(k) - 2ik\alpha\theta\lambda^{\alpha-1}$ for $1 < \alpha < 2$ where

$$C(k) := -\frac{1}{2\cos(\alpha\pi/2)}\left((1+\theta)(\lambda+ik)^\alpha + (1-\theta)(\lambda-ik)^\alpha - 2\lambda^\alpha\right).$$

The probability density $p(x,t)$ of each process L_i for $i = 1,\ldots,N+M$ satisfies a tempered fractional diffusion equation $x = L_i(t)$,

$$\partial_t p(x,t) = c\partial_x^{\alpha,\theta,\lambda} p(x,t),$$

where the constant $c > 0$ for $1 < \alpha < 2$ and $c < 0$ for $0 < \alpha < 1$ and the tempered-fractional-diffusion operator $\partial_x^{\alpha,\theta,\lambda}$ is given explicitly as [93, 94]

$$\partial_x^{\alpha,\theta,\lambda} = \mathcal{D}_x^{\alpha,\theta,\lambda} + v^{\alpha,\theta,\lambda}\frac{\partial}{\partial x} + \nu^{\alpha,\lambda}$$

where $v^{\alpha,\theta,\lambda}$ and $\nu^{\alpha,\lambda}$ are additional drift and source/sink terms given by,

$$v^{\alpha,\theta,\lambda} = \begin{cases} 0, & \alpha \in (0,1) \\ \frac{\alpha\theta\lambda^{\alpha-1}}{|\cos(\pi\alpha/2)|} & \alpha \in (1,2) \end{cases}, \quad \nu^{\alpha,\lambda} = \frac{\lambda^\alpha}{\cos(\pi\alpha/2)}.$$

Here $\alpha \in (0,1)\cup(1,2]$ is the fractional power or stability parameter, governing the heavy-tail law of the distribution generating the noise, $\theta \in [-1,1]$ is an asymmetry generating skew in the distribution, and $\lambda \in (0,\infty)$ is a tempering parameter that exponentially suppresses large jumps in the process η. Importantly, we take the noise as *centred* in the Gaussian limit (zero mean). For stability parameter $\alpha = 2$, the Lévy noise becomes Gaussian and the tempering λ no longer has any effect. Importantly, the value $\alpha = 1$ is excluded because it corresponds to the pathological Cauchy process. The operator $\mathcal{D}_x^{\alpha,\theta,\lambda}$ is called the λ-truncated fractional derivative of order α, given by,

$$\mathcal{D}_x^{\alpha,\theta,\lambda} = l(\theta)e^{-\lambda x}\,{}_{-\infty}D_x^\alpha e^{\lambda x} - r(\theta)e^{\lambda x}\,{}_xD_\infty^\alpha e^{-\lambda x}$$

where the operators ${}_{-\infty}D_x^\alpha$ and ${}_xD_\infty^\alpha$ are the Riemann-Liouville derivatives defined as [94]

$$e^{-\lambda x}\,{}_{-\infty}D_x^\alpha e^{\lambda x} f(x) = \frac{e^{-\lambda x}}{\Gamma(m-\alpha)}\frac{\partial^m}{\partial x^m}\int_{-\infty}^x \frac{d\zeta e^{\lambda\zeta}}{(x-\zeta)^{\alpha+1-m}} f(\zeta)$$
$$e^{\lambda x}\,{}_xD_\infty^\alpha e^{-\lambda x} f(x) = \frac{(-1)^m e^{\lambda x}}{\Gamma(m-\alpha)}\frac{\partial^m}{\partial x^m}\int_x^\infty \frac{d\zeta e^{-\lambda\zeta}}{(\zeta-x)^{\alpha+1-m}} f(\zeta)$$

for $\alpha - 1 < m < \alpha$. These can be equivalently defined in Fourier space [95, 96]

$$\mathcal{F}\left[e^{-\lambda x}\,{}_{-\infty}D_x^{\alpha}e^{\lambda x}f(x)\right] = (\lambda - ik)^{\alpha}\hat{f}(k)$$
$$\mathcal{F}\left[e^{\lambda x}\,{}_{x}D_{\infty}^{\alpha}e^{-\lambda x}f(x)\right] = (\lambda + ik)^{\alpha}\hat{f}(k)$$

where

$$\mathcal{F}\left[f(x)\right] = \int_{-\infty}^{\infty} dx e^{ikx} f(x) = \hat{f}(k)$$
$$\mathcal{F}^{-1}\left[\hat{f}(k)\right] = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{-ikx}\hat{f}(k) = f(x).$$

The weighting factors

$$l(\theta) = \frac{\theta - 1}{2\cos(\pi\alpha/2)}, \quad r(\theta) = \frac{\theta + 1}{2\cos(\pi\alpha/2)}$$

determine the asymmetry imposed on each of the Riemann-Liouville derivatives. It is also useful to take the Fourier transform of the fractional derivative $\partial_x^{\alpha,\theta,\lambda}$ giving

$$\mathcal{F}\left[\partial_x^{\alpha,\theta,\lambda}f(x)\right] = \underbrace{\left\{l(\theta)(\lambda - ik)^{\alpha} - r(\theta)(\lambda + ik)^{\alpha} - ikv^{\alpha,\theta,\lambda} + \nu^{\alpha,\lambda}\right\}}_{=\Lambda(k)/\sigma^2}\widehat{f}(k) \quad (3.16)$$

where $\Lambda(k)$ is the characteristic exponent (given earlier) and and σ^2 becomes the variance of the noise in the Gaussian limit. The centred noise is specified by four parameters α,σ,θ, λ, generalising from the one parameter of mean-zero Gaussian noise, σ, that we obtain in the limit as $\alpha \to 2$.

3.2.4 Measures of synchronisation

To measure the degree of synchronisation within a given population we use local forms of Kuramoto's order parameter [1] given by

$$O_B = \frac{1}{N}\left|\sum_{j\in\mathcal{B}} e^{i\beta_j}\right|, \quad O_R = \frac{1}{M}\left|\sum_{j\in\mathcal{R}} e^{i\rho_j}\right|. \quad (3.17)$$

We emphasise that in many of our examples in the following the total system ('global') order parameter $1/(N+M)|\sum_{i\in\mathcal{B}} e^{i\beta_i} + \sum_{j\in\mathcal{R}} e^{i\rho_j}|$ will be far from the value one.

3.3 Bulk dynamics near locked states

In this section we make various approximations to understand the bulk dynamics analytically when the system is near certain locked states. In Section 3.5 we return to the studying the full system (without approximations), albeit numerically.

There are various ways that the system can be deemed locked. We refer to the phase locking within Blue, that is $\beta_i \approx \beta_j \ \forall \ i,j \in \mathcal{B}$, and within Red $\rho_i \approx \rho_j \ \forall \ i,j \in \mathcal{R}$, as *internal* or *local locking*, and the phase locking of Blue externally with Red, $\beta_i \approx \rho_j \ \forall \ \{i,j\} \in \{\mathcal{B},\mathcal{R}\}$, as *external* or *global phase locking*. If we denote by B and P the centroids of the Blue and Red phases given by

$$B = \frac{1}{N}\sum_{i\in\mathcal{B}} \beta_i, \qquad P = \frac{1}{M}\sum_{i\in\mathcal{R}} \rho_i. \tag{3.18}$$

and consider the difference

$$\Delta(t) \equiv B(t) - P(t), \tag{3.19}$$

then if $\Delta \neq 0$ is time-independent then we may speak of external *frequency locking*.

3.3.1 The two cluster ansatz

In general, the system in Eq. (3.1) can only be solved numerically. To gain analytic insight in a region of greatest relevance, given that Blue and Red may deem internal phase synchronisation ideal, we explore a fixed point given by the ansatz

$$\beta_i(t) = B(t) + b_i(t), \ \ \rho_j = P(t) + p_j(t), \ \ \forall \ \{i,j\} \in \{\mathcal{B},\mathcal{R}\}, \tag{3.20}$$

with b_i, p_j small so that $b_i^2 \approx 0$, $p_j^2 \approx 0$. This leads to $\sum_{i\in\mathcal{B}} b_i = \sum_{i\in\mathcal{R}} p_i = 0$.

We linearise the system using Eqs. (3.20) and (3.19), keeping only terms first order in a Taylor expansion in b_i and p_j giving

$$\begin{aligned}
\dot{b}_i + \dot{B} &\approx \Omega_i - \zeta_B \sum_{j\in\mathcal{B}} L_{ij}^{(B)} b_j - \zeta_{BR}\cos(\Delta-\phi) \sum_{j\in\mathcal{B}\cup\mathcal{R}} L_{ij}^{(BR)} v_j, \\
\dot{p}_i + \dot{P} &\approx \Omega_i - \zeta_R \sum_{j\in\mathcal{R}} L_{ij}^{(R)} p_j - \zeta_{RB}\cos(\Delta+\psi) \sum_{j\in\mathcal{B}\cup\mathcal{R}} L_{ij}^{(RB)} v_j,
\end{aligned} \tag{3.21}$$

where

$$v_i = \begin{cases} b_i & i \in \mathcal{B} \\ p_i & i \in \mathcal{R} \end{cases}, \quad \Omega_i = \begin{cases} \omega_i - \zeta_{BR} d_i^{(BR)} \sin(\Delta-\phi) \\ \nu_i + \zeta_{RB} d_i^{(RB)} \sin(\Delta+\psi). \end{cases}$$

The quantities d and L, with superscripts $B,\ R,\ BR$ and RB, represent respectively the degree and corresponding Laplacian for the Blue, Red, Blue-Red and Red-Blue networks using the matrix $\mathcal{M}$; see Appendix A. Note that the Laplacian is defined as $L := D - A$. See [33, 90, 97] for the significance of the Laplacian spectrum.

3.3.2 Decoupling the dynamics

We now assume that

$$\sum_{j\in\mathcal{B}\cup\mathcal{R}} L_{ij}^{(BR)} v_j \approx \sum_{j\in\mathcal{B}\cup\mathcal{R}} L_{ij}^{(RB)} v_j \approx 0, \tag{3.22}$$

so that the equations for the fluctuations b_i and p_j in Eq. (3.21) may be decoupled through the use of the graph Laplacians. More specifically, the intra-population Laplacians $L^{(B)}$ and $L^{(R)}$ both contain a complete spanning set of orthonormal eigenvectors for the N and M dimensional subspaces. These are labelled as $e^{(B,r)}$ for $r = 0, 1, \ldots, N-1 \in \mathcal{B}_E$ and $e^{(R,r)}$ for $r = 0, 1, \ldots, M-1 \in \mathcal{R}_E$. The associated eigenvalues are denoted by $\lambda^{(B)}$ and $\lambda^{(R)}$. We distinguish between indices in node space $\{\mathcal{B}, \mathcal{R}\}$, and those in eigenmode space $\{\mathcal{B}^E, \mathcal{R}^E\}$ (which has the same dimensionality) by reserving labels $\{i, j\}$ for expressions involving graph nodes, and labels $\{r, s\}$ for expressions involving Laplacian eigenmodes. We assume that Blue and Red networks each consist of one component. Up to normalisation, the corresponding zero eigenvectors $\bar{e}^{(B,0)}$ and $\bar{e}^{(R,0)}$ consists of all unit valued entries [33]. Combined with Eq. (3.18), we see that B and P are the zero-mode projections of the phases β_i and ρ_j. Analogously, we denote by x_r and y_s the projections of b_i and p_i on the Blue and Red non-zero eigenvectors.

Using Eq. (3.22) and the eigenvector projections in Eq. (3.21) combined with the orthonormality of the eigenvectors gives

$$\begin{aligned}
\dot{x}_r &= \omega^{(r)} - \zeta_B \lambda_r^{(B)} x_r - \zeta_{BR} d_r^{(BR)} \sin(\Delta - \phi), \\
\dot{y}_s &= \nu^{(s)} - \zeta_R \lambda_s^{(R)} y_s + \zeta_{RB} d_s^{(RB)} \sin(\Delta + \psi), \\
\dot{B} &= \bar{\omega} - \frac{\zeta_{BR} d_T^{(BR)}}{N} \sin(\Delta - \phi), \\
\dot{P} &= \bar{\nu} + \frac{\zeta_{RB} d_T^{(BR)}}{M} \sin(\Delta + \psi),
\end{aligned} \tag{3.23}$$

for $r \in \mathcal{B}^E/\{0\}$ and $s \in \mathcal{R}^E/\{0\}$. The terms $\omega^{(r)}$ and $d_r^{(BR)}$ are the projections onto the $r-$th eigenvector and $\bar{\omega}$ is the average over $\mathcal{B}$ with $d_s^{(RB)}$ the projections onto the $s-$th eigenvector and $\bar{\nu}$ the average frequency for Red. The difference of $\dot{B}$ and $\dot{P}$ gives then

$$\dot{\Delta} = -V'(\Delta), \tag{3.24}$$

where

$$\begin{aligned} V(\Delta) =& -\mu\Delta - \sqrt{S^2+C^2}\cos\left(\Delta-\varrho\right) \\ \mu =& (\bar{\omega}-\bar{\nu}) \\ \varrho \equiv& \tan^{-1}\left(S/C\right), \\ C \equiv& d_T^{(BR)}\left(\frac{\zeta_{BR}\cos\phi}{N}+\frac{\zeta_{RB}\cos\psi}{M}\right), \\ S \equiv& d_T^{(BR)}\left(\frac{\zeta_{BR}\sin\phi}{N}-\frac{\zeta_{RB}\sin\psi}{M}\right). \end{aligned} \tag{3.25}$$

Thus Δ may be solved first from Eq.(3.24), and then used in the forcing terms of the normal mode equations in Eqs.(3.23) which are otherwise solvable in their own right.

The substitution, $\Delta = \varrho + 2\tan^{-1}\left(\vartheta + \frac{\sqrt{S^2+C^2}}{\mu}\right)$, leads to the solution

$$\Delta(t) = \varrho + 2\tan^{-1}\left\{\sqrt{\frac{S^2+C^2}{\mu^2}} + \sqrt{\frac{\mathcal{K}}{\mu^2}}\tanh\left(\frac{\sqrt{\mathcal{K}}}{2}(\text{const}-t)\right)\right\} \tag{3.26}$$

where $\text{const} = \frac{2}{\sqrt{\mathcal{K}}}\tanh^{-1}\left\{\frac{\mu}{\sqrt{\mathcal{K}}}\vartheta_0\right\}$, and

$$\mathcal{K} = C^2 + S^2 - \mu^2. \tag{3.27}$$

The solutions to the normal modes are given in Appendix A. Ultimately their dynamics depends on the behaviour of $\Delta(t)$ which in turn is governed by the potential $V(\Delta)$ in Eq. (3.24). This is commonly referred to as a tilted periodic [98–100] or tilted Smoluchowski-Feynman ratchet [101] potential. Importantly, we have periodicity, $V'(\Delta) = V'(\Delta+2\pi)$, and the tilt refers to the constant forcing term in $V'(\Delta)$ given by the difference of frequency averages μ. The sign of $\mathcal{K}$ is critical. If $\mathcal{K} > 0$ the solution asymptotes to the value $\varrho + \sin^{-1}\left(\frac{\mu}{\sqrt{S^2+C^2}}\right) \bmod 2\pi$. If $\mathcal{K} < 0$ then Eq. (3.26) gives oscillatory behaviour with period $\frac{2\pi}{\sqrt{|\mathcal{K}|}}$. Alternately, for $\mathcal{K} > 0$, $V(\Delta)$ is a series of local maxima with unstable fixed points at $\Delta = \pi + \varrho - \sin^{-1}\left(\frac{\mu}{\sqrt{S^2+C^2}}\right) + 2\pi n,\ n \in \mathbb{Z}$) and local minima with stable fixed points at $\Delta = \varrho + \sin^{-1}\left(\frac{\mu}{\sqrt{S^2+C^2}}\right) + 2\pi n,\ n \in \mathbb{Z}$ on a landscape which has an overall slope according to the sign of μ. For $\mathcal{K} = 0$ the hills and valleys of the potential become points of inflection, and hence unstable fixed points. For $\mathcal{K} < 0$, the potential loses all of its fixed points, even the unstable ones. We may also speak of *drift* when there is no stable fixed point. This implies that Blue and Red clusters will be in motion with respect to each other.

3.3.3 Measures of synchronisation and the fixed point approximation

Something absent in the works [6, 45] is the centroid representation for order parameters

$$O_B = \frac{1}{N}\sqrt{\sum_{i,j\in\mathcal{B}} \cos^2(\beta_i - \beta_j)} = \frac{1}{N}\sqrt{\sum_{i,j\in\mathcal{B}} \cos^2(b_i - b_j)}, \tag{3.28}$$

$$O_R = \frac{1}{M}\sqrt{\sum_{i,j\in\mathcal{R}} \cos^2(\rho_i - \rho_j)} = \frac{1}{M}\sqrt{\sum_{i,j\in\mathcal{R}} \cos^2(p_i - p_j)}. \tag{3.29}$$

In the fixed point approximation when $\mathcal{K} > 0$ (giving Δ constant), b_i, p_i may be determined from the stationary solutions for x_r, y_r in Eqs. (3.23) which may be substituted into Eqs. (3.28) and (3.29). Alternatively, and reintroducing the L^{BR}, L^{RB} matrices, Eqs. (3.21) in the static limit may also be compactly written in terms of a 'super-Laplacian' matrix $\mathcal{L}$,

$$\vec{\Omega} - \mathcal{L}\vec{v} = 0, \tag{3.30}$$

where the off-diagonal parts of the matrix $\mathcal{L}$ incorporate the terms ignored thus far, Eqs. (3.22).

The super-Laplacian, given explicitly in [45] is not symmetric but has zero column sum and therefore a single zero right eigenvalue. Thus Eq. (3.30) is invertible through the pseudo-inverse of the super-Laplacian (even though it may only be computed numerically, in general) and used in Eqs. (3.28) and (3.29).

Importantly, this brings together information about the network structure through the Laplacian, but also on the nature of the threshold in $\mathcal{K}$ through the implicit dependence on Δ. Thus any deterministic optimisation of O_B (for example) in relation to parameters ϕ and ψ will encode the transition to periodic behaviour at $\mathcal{K} = 0$. We will show a numerical example of this below when we consider the fitness landscape of the deterministic system.

3.3.4 Example networks, frequencies and couplings

To illustrate our otherwise general solutions and compare to numerical simulation we consider an example of Blue agents in a hierarchy and Red agents on a random network. As in [6, 45], for $\mathcal{B}$ we consider a complete 4-ary tree, thus $N = 21$. For $\mathcal{R}$ we use a random Erdös-Rényi network also of $N = 21$ with link of probability 0.4.

Blue-to-Red interactions are arranged such that each leaf-node of Blue ($i = 6, \ldots, 21$) interacts with the correspondingly labelled Red node ($i = 27, \ldots, 42$), shown as open

circles in Fig. 3.1. Thus $d_i^{(BR)} = d_i^{(RB)} = 1$ or 0, for agents engaged, respectively not engaged, with a competitor, but $d_T^{(BR)} = 16$. Given the significance of the graph Laplacians we show their spectra and the frequencies of each agent in Fig. 3.2 where the specific frequencies of each agent are drawn from a uniform distribution $\omega_i, \nu_j \in [0, 1]$. In this instance the average frequencies are $\bar{\omega} = 0.503, \bar{\nu} = 0.551$, giving $\mu = -0.048$. If cross-couplings were set to zero the Red population would lap Blue over time. The sign of μ here implies a negative slope for the tilted ratchet potential $V(\Delta)$. Note here that there are many more low lying eigenvalues for $L^{(B)}$ compared to $L^{(R)}$ - a consequence of the comparatively poor connectivity of the tree graph (left, Fig. 3.2).

We choose couplings [6, 45]

$$\zeta_B = 8, \zeta_R = 0.5, \zeta_{BR} = \zeta_{RB} = 0.4.$$

For these network instances these give high internal phase synchronisation (note that $\zeta_B > \zeta_R$, reflecting the difficulty of hierarchies to synchronise) but allow for some changes in dynamics as the frustrations are varied; specifically we set $\psi = 0$ and manipulate ϕ. Then $\mathcal{K}$ changes sign from positive to negative at $\phi_c = 0.9498\pi$. In the vicinity of this point, in the absence of noise, for $\phi < \phi_c$ Blue locks to a fixed phase ahead of Red, and for $\phi \geq \phi_c$ Blue and Red remain internally phase synchronised but lap each other with a period that decreases as ϕ gets larger. Thus we shall be interested in behaviours slightly below and above this point, for example $\phi = 0.94\pi$ and $\phi = 0.96\pi$.

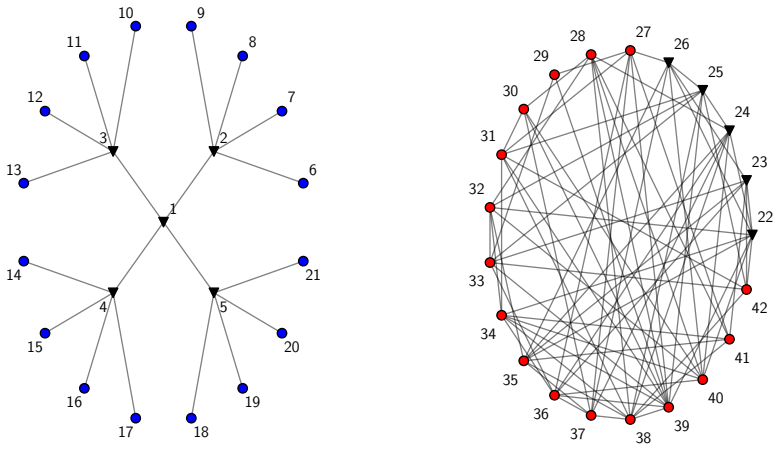


FIGURE 3.1: Left: The Blue network $\mathcal{B}$ created from a complete 4-ary tree with $N = 21$. Right: The Red network $\mathcal{R}$ created according to the Eröds-Rényi model with $N = 21$ and a link probability of 0.4. The node i in the tree graph denoted by a circle interacts with a corresponding number, node i + 21; all other nodes are represented by a filled triangle.

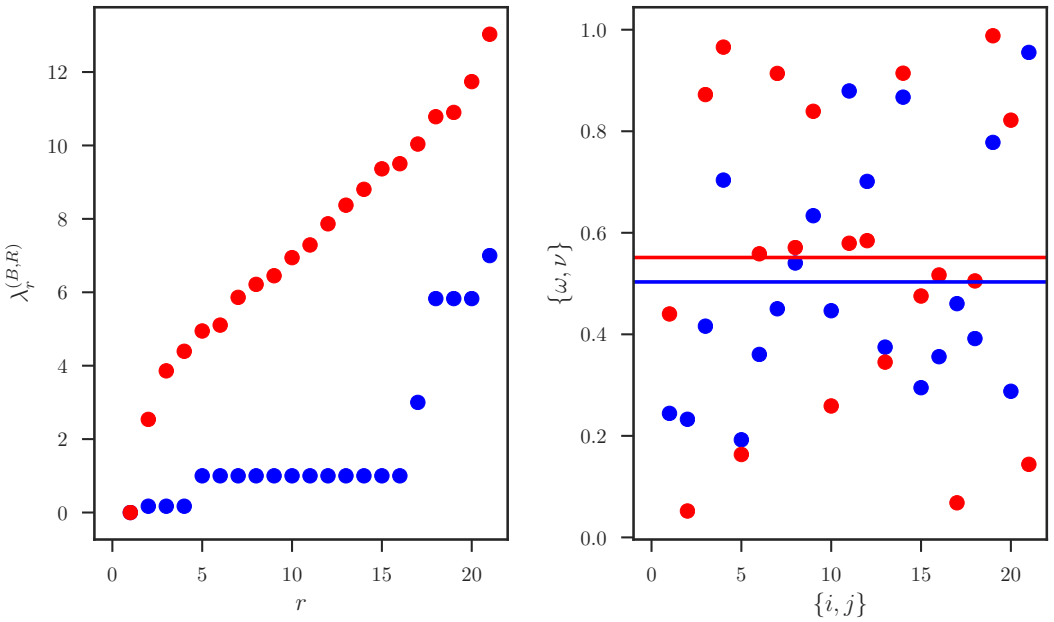


FIGURE 3.2: Left: The spectrum of the graph Laplacian for Blue and Red networks coloured respectively Blue and Red. Right: The frequencies for Blue and Red agents according to the node labelling again coloured blue and red respectively, with solid lines indicating their corresponding means.

3.4 Analytical results

In this section, we derive some analytic results when the system is perturbed by exogenous stochastic shocks. We seek to understand the bulk behaviour through a projection onto an expansion in terms of the Red/Blue Laplacians.

3.4.1 Projecting the noise

For every small increment of time, we apply a small increment of our processes L_i to the Blue and Red network nodes. To aid the notation, we denote by $L_i^{(B)}(t)$ and $L_i^{(R)}(t)$ the processes on the Blue and Red network nodes, respectively. At every time point t, the increments of the processes may be decomposed in terms of eigenvectors of the Blue/Red Laplacians as

$$\begin{aligned} \mathrm{d}L_i^{(B)}(t) &= \sum_{r\in\mathcal{B}_E} \gamma_r^{(B)} e_i^{(B,r)}\, \mathrm{d}\eta_r^{(B)}(t), \quad i\in\mathcal{B}, \\ \mathrm{d}L_i^{(R)}(t) &= \sum_{r\in\mathcal{R}_E} \gamma_r^{(R)} e_i^{(R,r)}\, \mathrm{d}\eta_r^{(R)}(t), \quad i\in\mathcal{R} \end{aligned}$$

where $\eta_r^{(B)}$ and $\eta_r^{(R)}$ are tempered stable Lévy processes with a change of parameters. We stress that this does not mean a decoupling of the noise, but a decomposition of general tempered stable Lévy noise under a complete basis. In particular, we note that the projections onto Laplacian eigenvectors in Section 3.4.1 mean that the parameters $\alpha, \sigma, \theta, \lambda$ of the noise of the original process L_i are not necessarily the same as those of the projected system η.

In [53] the general expression for projection onto an N-dimensional vector u_j of the tempered stable Lévy process η_i was derived as is given by

$$\mathrm{d}L_{\mathrm{proj}} = \sum_{j=1}^{N} u_j\, \mathrm{d}\eta_j.$$

For the noise η this is defined as a function of the parameters $\alpha, \sigma, \theta, \lambda$, as in Eq. (3.16). Then for the projected process L_{proj} the characteristic exponent is

$$\begin{aligned} \Lambda_{\mathrm{proj}}(k) =& \sigma^2 \sum_{j=1}^{N} \big[(1+\theta)(\lambda + iku_j)^\alpha \\ &+(1-\theta)(\lambda - iku_j)^\alpha - iku_j v^{\alpha,\theta,\lambda} + \nu^{\alpha,\lambda}\big]. \end{aligned} \tag{3.32}$$

Essentially, the 'momentum' k is multiplied by the component of the vector u_j.

This gives the Langevin or Ornstein-Uhlenbeck system, written in SDE form as

$$\mathrm{d}x_r(t) = q_r^{(B)}(x_r(t), \Delta)\,\mathrm{d}t + \gamma_r^{(B)}\,\mathrm{d}\eta_r^{(B)}(t), \tag{3.33}$$

$$\mathrm{d}y_s(t) = q_s^{(R)}(y_s(t), \Delta)\,\mathrm{d}t + \gamma_s^{(R)}\,\mathrm{d}\eta_s^{(R)}(t),$$

$$\mathrm{d}\Delta(t) = -V'(\Delta)\,\mathrm{d}t + \gamma_0^{(B)}\,\mathrm{d}\eta_0^{(B)}(t) - \gamma_0^{(R)}\,\mathrm{d}\eta_0^{(R)}(t), \tag{3.34}$$

for $r \in \mathcal{B}^E/\{0\}$, $s \in \mathcal{R}^E/\{0\}$, and where the variables and parameters are as used in Eq. (3.23).

We test the consequences of noise applied to zero or normal modes via the selections:

- $\gamma_0^{(B)} = \gamma_0^{(R)} = 0$, $\gamma_r^{(B)} = \gamma_s^{(R)} = 1$, $\{r, s\} \in \{\mathcal{B}^E/\{0\}, \mathcal{R}^E/\{0\}\}$ has noise applied to normal modes;
- $\gamma_0^{(B)} = \gamma_0^{(R)} = 1$, $\gamma_r^{(B)} = \gamma_s^{(R)} = 0$, $\{r, s\} \in \{\mathcal{B}^E/\{0\}, \mathcal{R}^E/\{0\}\}$ has noise applied to zero modes.

These choices then specify the random variables in the stochastic system. In the full problem of noise applied generally, there are three sets of random variables, two vector and one scalar, corresponding to the deterministic variables $(\vec{x}, \vec{y}, \Delta)$. These correspond to Blue and Red Laplacian projections and the phase difference between centroids.

3.4.2 Noisy normal modes subject to tempered stable noise

We consider first the case $\gamma_0^B = \gamma_0^R = 0$ where Lévy noise is introduced on normal mode fluctuations on the background of the deterministic Δ. The latter is governed by the solution Eq. (3.26) which tends to a fixed value as $t \to \infty$ or is periodic, according to the sign of $\mathcal{K}$ in Eq. (3.27). We show in the following how with tempered Lévy noise this devolves to a system we have solved elsewhere [102]; we will not give the solution here but explain the behaviour it implies and why it is less interesting than the zero mode case.

As Δ is no longer a random variable, in this case, all the normal mode Langevin equations in Eq. (3.33) are independent. We may therefore consider the fractional version of a Langevin equation for the Blue modes x_s - the corresponding equations for y_s are easily translated from Eq. (3.33)

$$\dot{x}_s = q_s^{(B)}(x_s, \Delta) + \dot{L}_s^{\alpha,\theta,\lambda}(t) \tag{3.35}$$

where $L^{\alpha,\theta,\lambda}(t)$ is a tempered stable Lévy process in time described by parameters α, θ, λ. For ease of notation we will first take the noise as if applied within the normal mode

projected system, and translate back to the original basis subsequently. Because, in the approximation that off-diagonal entries of the super-Laplacian are taken to be zero, the u_j in Eq. (3.32) are, for the Blue system, the eigenvector components $e_j^{(B,r)}, j = 1, \ldots, N$.

Thus the corresponding Fokker-Planck equation can be decomposed into a product of $N + M - 2$ densities,

$$\mathcal{P}(\vec{x}, \vec{y}, t) = \prod_{r \in \mathcal{B}^E/\{0\}} \mathcal{P}^{(B,r)}(x_r, t) \prod_{s \in \mathcal{R}^E/\{0\}} \mathcal{P}^{(R,s)}(y_s, t)$$

leading to a decoupling into separate fractional Fokker-Planck equations for each mode.

The probability density $\mathcal{P}(x, t)$ associated with the tempered-stable Lévy process given by Eq. (3.35) satisfies the fractional Fokker-Planck equation [103],

$$\begin{aligned} \frac{\partial}{\partial t}\mathcal{P}^{(B,r)}(x,t) =& \left\{ \sigma^2 \partial_{x_r}^{\alpha,\theta,\lambda} - \frac{\partial}{\partial x_r} q_r^{(B)}(x_r, t) \right\} \mathcal{P}^{(B,r)}(x_r, t), \\ \mathcal{P}^{(B,r)}(x_r, 0) =& \delta(x_r - x_r'), \end{aligned} \tag{3.36}$$

where $\sigma^2 \in (0, \infty)$ is the diffusion constant - which becomes the variance of the process in the Gaussian limit. Unpacking $q_r^{(B)}$ gives the explicit form of the tempered-fractional-Fokker-Planck equation (TFFP) for the normal mode projection of the Blue-vs-Red system

$$\begin{aligned} \frac{\partial}{\partial t}\mathcal{P}^{(B,r)}(x_r, t) = \{ & \sigma^2 \mathcal{D}_{x_r}^{\alpha,\theta,\lambda} + \Big[\zeta_B \lambda_r^{(B)} x_r + \sigma^2 v^{\alpha,\theta,\lambda} \\ & - (\omega^{(r)} - \zeta_{BR} d_r^{(BR)} \sin(\Delta(t) - \phi)) \Big] \frac{\partial}{\partial x_r} \\ & + \left[\zeta_B \lambda_r^{(B)} + \sigma^2 \nu^{\alpha,\lambda} \right] \} \mathcal{P}^{(B,r)}(x_r, t). \end{aligned} \tag{3.37}$$

For the Gaussian limit, i.e. $\partial_x^{\alpha,\theta,\lambda} = \frac{\partial^2}{\partial x^2}$, section 1.8.3.6 of [104] gives a set of nonlinear transformations which result in the Gaussian equivalent of Eq. (3.37) becoming the standard heat equation. Indeed, this technique enabled the analytical investigation of clustering effects in the frustrated Kuramoto model under the influence of Gaussian white noise [6].

As alluded several times, we have solved a version of this problem in [102] where the sign of $\mathcal{K}$ and the shape of the underlying noise distribution interplay in determining the solution. For $\mathcal{K} > 0$, where Δ asymptotes to a constant value, the probability density broadens from a delta function to a tempered stable density of corresponding characteristics given by α, θ, λ. For $\mathcal{K} < 0$, where Δ becomes oscillatory, the Fokker-Planck density evolves from the delta function to a tempered stable density of *fixed*

shape (again, according to α, θ, λ) but whose centre oscillates according to the dynamics of Δ. This is a natural generalisation of the behaviour seen for the Gaussian case in [6].

These behaviours will result in corresponding dynamics for the order parameters $O_B(t)$ and $O_R(t)$ for weak values of σ. For $\mathcal{K} > 0$, these will exhibit noisy behaviour (showing tempered-heavy-tails) around a steady-state value, while for $\mathcal{K} < 0$ there will be similar noisy behaviour around an oscillatory background. Importantly, if noise is applied across all nodes independently, we expect a residual imprint of the sign of $\mathcal{K}$ in the dynamics. In this respect, there is little that is surprising in the normal modes and so we turn to the potentially more interesting case of the zero modes.

3.4.3 Noisy zero modes

Taking $\gamma_i = 0$ and $\gamma_0 = 1$, we consider the tempered fractional Fokker-Plank equation with the tilted ratchet potential from Langevin form Eq. (3.34) and wrap their densities $\mathcal{P}(x,t)$ onto $\mathbb{S}^1$. For more details regarding the analytic results of this Section please refer to [105].

From Eq. (3.36) we get that each population density can be rewritten in terms of the *probability current* $\mathcal{J}(x,t)$ as

$$\begin{aligned} &\frac{\partial}{\partial t}\mathcal{P}(x,t) + \frac{\partial}{\partial x}\mathcal{J}(x,t) = 0 \\ &\Rightarrow \mathcal{J}(x,t) = -\left\{\sigma^2 \partial_x^{\alpha,\theta,\lambda} + \frac{\partial}{\partial x}V'(x)\right\}\mathcal{P}(x,t). \end{aligned} \tag{3.38}$$

We now restrict the support of x to $\mathbb{S}^1$ by constructing the so-called *reduced density* $\mathcal{P}^{(reduc)}(x,t)$ and *reduced probability current* $\mathcal{J}^{(reduc)}(x,t)$ through

$$\begin{aligned} \mathcal{P}^{(reduc)}(x,t) &\equiv \sum_{n=-\infty}^{\infty} \mathcal{P}(x + 2\pi n, t), \\ \mathcal{J}^{(reduc)}(x,t) &\equiv \sum_{n=-\infty}^{\infty} \mathcal{J}(x + 2\pi n, t). \end{aligned}$$

Due to the linearity of the Fokker-Planck equation, the reduced density and reduced probability current also obey Eq. (3.38) but with the new boundary and normalisation conditions

$$\mathcal{P}^{(reduc)}(-\pi, t) = \mathcal{P}^{(reduc)}(\pi, t), \quad \int_{-\pi}^{\pi} \mathcal{P}^{(reduc)}(x,t)\,\mathrm{dx} = 1. \tag{3.39}$$

Just like the Gaussian case [101], we expect that for tempered stable Lévy noise the steady state equivalent of Eq. (3.36) with vanishing boundary conditions at the natural boundaries $x \to \pm\infty$ will be non-normalisable due to metastability. The steady state

density $\mathcal{P}_{st}^{(reduc)}(x)$ on $\mathbb{S}^1$ satisfies

$$\left\{\sigma^2\partial_x^{\alpha,\theta,\lambda} + \frac{\partial}{\partial x}V'(x)\right\}\mathcal{P}_{st}^{(reduc)}(x) = 0 \tag{3.40}$$

with boundary and normalisation conditions given by Eq. (3.39).

The solution is obtained by Fourier transformation of Eq. (3.40) revealing the expression,

$$\begin{aligned} \widehat{Q}(k+1) &= -f_k\widehat{Q}(k) + \widehat{Q}(k-1), \\ \text{where } \widehat{Q}(k) &\equiv e^{-ik\rho}\widehat{\mathcal{P}}_{st}^{(reduc)}(k), \\ \text{and } f_k &\equiv -\frac{2}{k\sqrt{S^2+C^2}}[\Lambda(k) + ik\mu], \end{aligned} \tag{3.41}$$

and we note that because of periodicity, the Fourier variables k are integer-valued. We can then construct the corresponding reduced probability density via the discrete inverse Fourier transform

$$\mathcal{P}_{st}^{(reduc)}(x) = \frac{1}{2\pi}\left\{\widehat{Q}(0) + 2\mathbb{R}\sum_{n=1}^{\infty}e^{-in(x-\rho)}\widehat{Q}(n)\right\}, \tag{3.42}$$

where $\widehat{Q}(0) = 1$ from the normalisation condition in Eq. (3.39). The remaining $\widehat{Q}(n)$ may be solved via the nonlinear transform

$$S_{k+1} = \frac{\widehat{Q}(k+1)}{\widehat{Q}(k)},$$

which transforms the linear three-term recurrence relation given by Eq. (3.41) into the following nonlinear two-term recurrence relation,

$$S_k = \frac{1}{f_k + S_{k+1}}. \tag{3.43}$$

Eq. (3.43) can be solved iteratively with the application of continued fractions,

$$S_k = \cfrac{1}{f_k + \cfrac{1}{f_{k+1}+S_{k+2}}} = \cfrac{1}{f_k + \cfrac{1}{f_{k+1}+\cfrac{1}{f_{k+2}+\cfrac{1}{\ddots}}}}$$

whose values can be used to reconstruct the corresponding $\widehat{Q}(k)$ expressions using,

$$\widehat{Q}(k) = S_kS_{k-1}\ldots S_2S_1\widehat{Q}(0)$$

for insertion into Eq. (3.42).

In this paper we mainly focus on the average velocity $\langle\dot{\Delta}\rangle$, which is defined through the reduced steady state probability current $\mathcal{J}_{st}^{(reduc)}(x)$ with x replaced by Δ as follows

$$\langle\dot{\Delta}\rangle = \frac{d}{dt}\langle\Delta\rangle = \int_{-\pi}^{\pi} \mathcal{J}_{st}^{(reduc)}(\Delta)d\Delta.$$

Non-zero values of $\langle\dot{\Delta}\rangle$ indicate that through a combination of drift and noise there is a rotation of the centroid of Blue with respect to that of Red; positive values mean that Blue advances and negative values mean that Red advances. The average velocity $\langle\dot{\Delta}\rangle$ is solved for giving

$$\langle\dot{\Delta}\rangle = \begin{cases} -\frac{\sigma^2\alpha\theta\lambda^{\alpha-1}}{\cos\left(\frac{\pi\alpha}{2}\right)} + \mu - 2\pi\gamma\mathbb{I}\left\{\widehat{Q}(1)\right\} & 0<\alpha<1 \\ \mu - 2\pi\gamma\mathbb{I}\left\{\widehat{Q}(1)\right\} & 1<\alpha<2. \end{cases} \tag{3.44}$$

We now need to take into account the Laplacian projection of the noise. It is straightforward to see that as the equation for Δ arises from the difference $B-P$, the $N+M$ components of u_j in Eq. (3.32) for this purpose are:

$$u_j = \begin{cases} \frac{1}{\sqrt{N}}, j=1,\ldots,N \\ -\frac{1}{\sqrt{M}}, j=N+1,\ldots,N+M \end{cases}.$$

Eq. (3.32) thus gives

$$\begin{aligned}\Lambda_0(k) =& \sigma^2\left\{(1+\theta)\left[N\left(\lambda+\frac{ik}{\sqrt{N}}\right)^{\alpha} + M\left(\lambda-\frac{ik}{\sqrt{M}}\right)^{\alpha}\right]\right. \\ &+ (1-\theta)\left[N\left(\lambda-\frac{ik}{\sqrt{N}}\right)^{\alpha} + M\left(\lambda+\frac{ik}{\sqrt{M}}\right)^{\alpha}\right] \\ &\left.- ik\left(\sqrt{N}-\sqrt{M}\right)v^{\alpha,\theta,\lambda} + (N+M)\nu^{\alpha,\lambda}\right\}.\end{aligned}$$

Noting that in the explicit example we consider, $N=M$, this leads to

$$\Lambda_0(k) = 2\sigma^2 N\left[\left(\lambda+\frac{ik}{\sqrt{N}}\right)^{\alpha} + \left(\lambda-\frac{ik}{\sqrt{N}}\right)^{\alpha} + \nu^{\alpha,\lambda}\right].$$

This can be further simplified to

$$\Lambda_0(k) = 2\sigma^2 N^{1-\alpha/2}\left[\left(\sqrt{N}\lambda+ik\right)^{\alpha} + \left(\sqrt{N}\lambda-ik\right)^{\alpha} + \nu^{\alpha,\sqrt{N}\lambda}\right].$$

We see that for the mode Δ, the parameters of the noise H for the zero mode projected system are a simple rescaling of those of the noise η on the original system:

$$\begin{aligned}\sigma^2 &\to \sigma^2 N^{1-\alpha/2}\\ \lambda &\to \sqrt{N}\lambda\\ \theta &\to 0.\end{aligned}$$

Significantly, both the tempering and noise strength are *enhanced*, and the skew is *eliminated*. The interplay of these for various α and λ are not easily read off from Eq. (3.44). In Figure 3.3 we show the λ-dependence of $\langle\dot{\Delta}\rangle$ for various α from Eq. (3.44) for the couplings in Eq. (3.31) and with $\phi = 0.96\pi, \psi = 0$ (top, for which $\mathcal{K} < 0$) and $\phi = 0.94\pi, \psi = 0$ (bottom, for which $\mathcal{K} > 0$).

In both cases, we see that values of the average velocity are generally small and negative (indicating that Red will advance ahead of Blue), and the near Gaussian case (red curve) is correctly λ independent. The curves on the top, for $\phi = 0.96\pi$, represent generally larger absolute values of $\langle\dot{\Delta}\rangle$, consistent with $\mathcal{K} < 0$ so that deterministic dynamics already imply $\dot{\Delta} \neq 0$. For small tempering λ the average velocity is larger in absolute magnitude for smaller α (though difficult to discern on this scale), but there is a critical value of λ at which this behaviour reverses, with small α values seeing more suppression. However the effect eventually flattens out. Thus, even when deterministically there is drift between clusters, tempered stable noise with small α and large tempering λ can cancel some, but not all, of this drift. This effect is more pronounced for $\phi = 0.94\pi$ (bottom), where $\mathcal{K} > 0$ and thus deterministically $\dot{\Delta} = 0$. Thus the effect is naturally more exaggerated since the underlying deterministic dynamics places the system in steady-state. The black curve here for $\alpha = 0.3$ genuinely approaches zero. Thus, for small α and sufficient tempering λ, the noise induced drift between Blue and Red may be cancelled entirely.

To summarise the key insight from this section: for small α in the presence of tempering, the noise may drive a tendency for the Blue and Red clusters to drift less even when deterministically they might otherwise be in relative motion, while for deterministically locked scenarios noise may be cancelled entirely. Though we have not yet drawn attention to it, this behaviour is independent of the graph structure of Blue and Red since it arises purely from the structure of the fixed point as represented in the potential $V(\Delta)$ in Eq. (3.25).

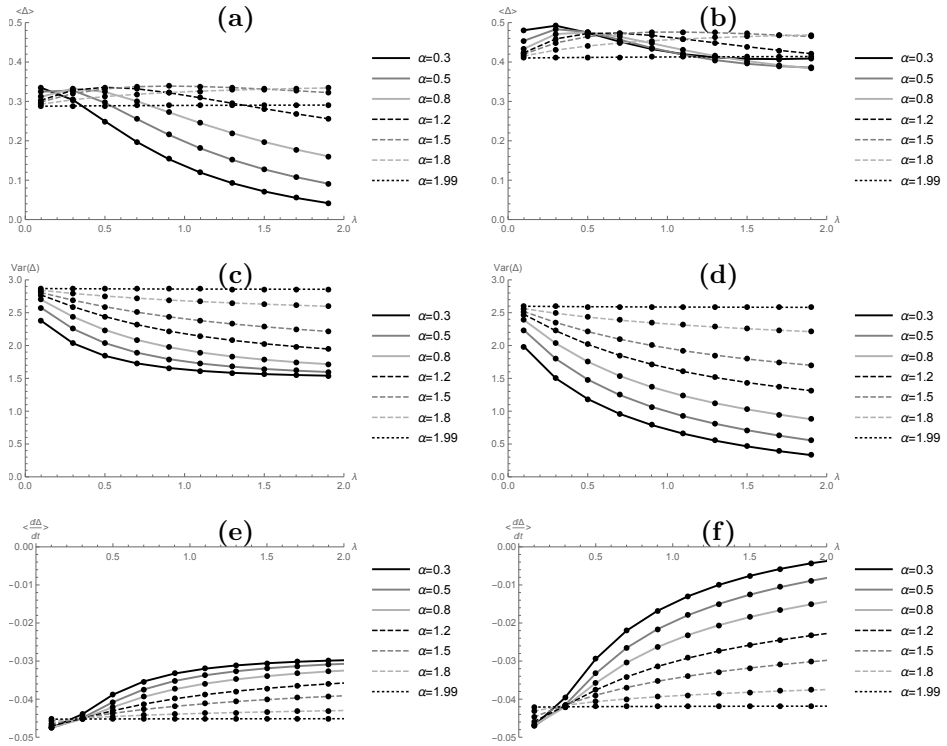


FIGURE 3.3: Plots of the average ((a) and (b)) and variance ((c) and (d)) of Δ and the velocity $\langle\dot{\Delta}\rangle$ ((e) and (f)) given by Eq. (3.44) as a function of the tempering parameter λ for different values of the stability parameter α. The couplings are as in Eqs. (3.31) and the noise strength $\sigma = 0.1$. The left hand plots show when the frustration parameters are $\phi = 0.96\pi, \psi = 0$ (when Blue and Red deterministically drift) and the right hand the case for $\phi = 0.94\pi, \psi = 0$ (where Blue and Red deterministically lock).

3.5 Noisy behaviour in the full dynamics

We now focus on investigating the full nonlinear dynamics of the stochastic Blue vs Red model using numerical simulations. The aim of this numerical investigation is to identify the optimal internal synchronisation and phase lag of the Blue network against the Red network and to understand how the noise parameters α, σ, and λ lead to either steady, periodic, or noisy pathwise behaviour.

3.5.1 Implementation

To simulate the full tempered fractional stochastic Blue-v-Red model we implemented code in C++ for maximum performance and uses the Messaging Passing Interface (MPI) to parallelise the simulation across an arbitrary number of CPU cores. The simulation is performed using an Euler method to simulate the dynamics at the N vertices. This algorithm was chosen as it allows the code to easily switch between the cases of

(1) no noise, (2) Gaussian noise, (3) stable noise, or (4) tempered stable noise without changing the underlying algorithm. The code takes the following parameters: the noise distribution parameters; the number of paths to simulate; the number of time steps in each path; and the filename containing the graph $\mathcal{G}$ stored in graph6 format [106]. The code also includes a number of hand-implemented random sampling algorithms for stable and tempered stable distributions. The evaluation of the expected values of the order parameter is approximated through a simple Monte Carlo procedure, that is we approximate the expected value of the order parameter upon the discretisations $\mathbb{E}[(O_B)_i] = \mathbb{E}[(O_B(t_i)], t_i \in [0,T], i = 1,\ldots,N$ through the mean $(\bar{O}_B)_i = \frac{1}{M}\sum_{m=1}^{M}(O_B)_i^{(m)}$ where $(O_B)_i^{(m)}$ is the i'th discretised value of the m'th simulated path of the order parameter O_B. From the CLT we have that

$$(\bar{O}_B)_i \sim \mathcal{N}\left(\mathbb{E}[(O_B)_i], \frac{s}{\sqrt{M}}\right), \tag{3.45}$$

where $\mathcal{N}$ is the normal distribution and the standard deviation of the order parameter estimate is given as $s = \frac{1}{M-1}\sum_{m=1}^{M}((O_B)_i^{(m)} - (\bar{O}_B)_i)^2$. This is similarly done for $(O_B)_i$ and Δ_i.

3.5.2 Pathwise dynamics

In the following we show the pathwise dynamics, namely individual but characteristic instances, for tempered stable Lévy noise choices for cases where we have previously examined the deterministic [45] and Gaussian noise [6] systems. We first focus on the dynamics for specific values of σ, α, λ while always using the one-sided case $\theta = 1$. We generate time-series of the order parameters O_B, O_R and Δ. We choose values where both deterministic and stochastic properties are evident, in other words where chaos is not dominant. In most cases, this limits us to choices $\alpha > 1$ so we use $\alpha = 1.5$ as a value sufficiently far from Gaussian. Our choice of points in the (ϕ, ψ) landscape will be both below and above the threshold $\mathcal{K} = 0$. These should be compared to corresponding plots for the Gaussian case in [6]. We note that whereas in that work we plotted box-whisker charts for the data at each time-step, here we simply plot the mean and the standard deviation on a dual axis. Thus the fluctuations in O_B and O_R may appear to exceed unity - this is an artefact of how we represent the deviations.

To begin with we consider stable noise for $\sigma = 0.05$ and $\alpha = 1.5$ in Figure 3.4.

We show examples for four points in the (ϕ, ψ) plane, two below, one at the threshold of $\mathcal{K} = 0$ and the other above. We observe, firstly, that levels of synchronisation in O_B and O_R are high in all cases, however the nature of the fluctuations can be quite different. While $\mathcal{K} > 0$, as for the top two cases, the fluctuations are quite random about

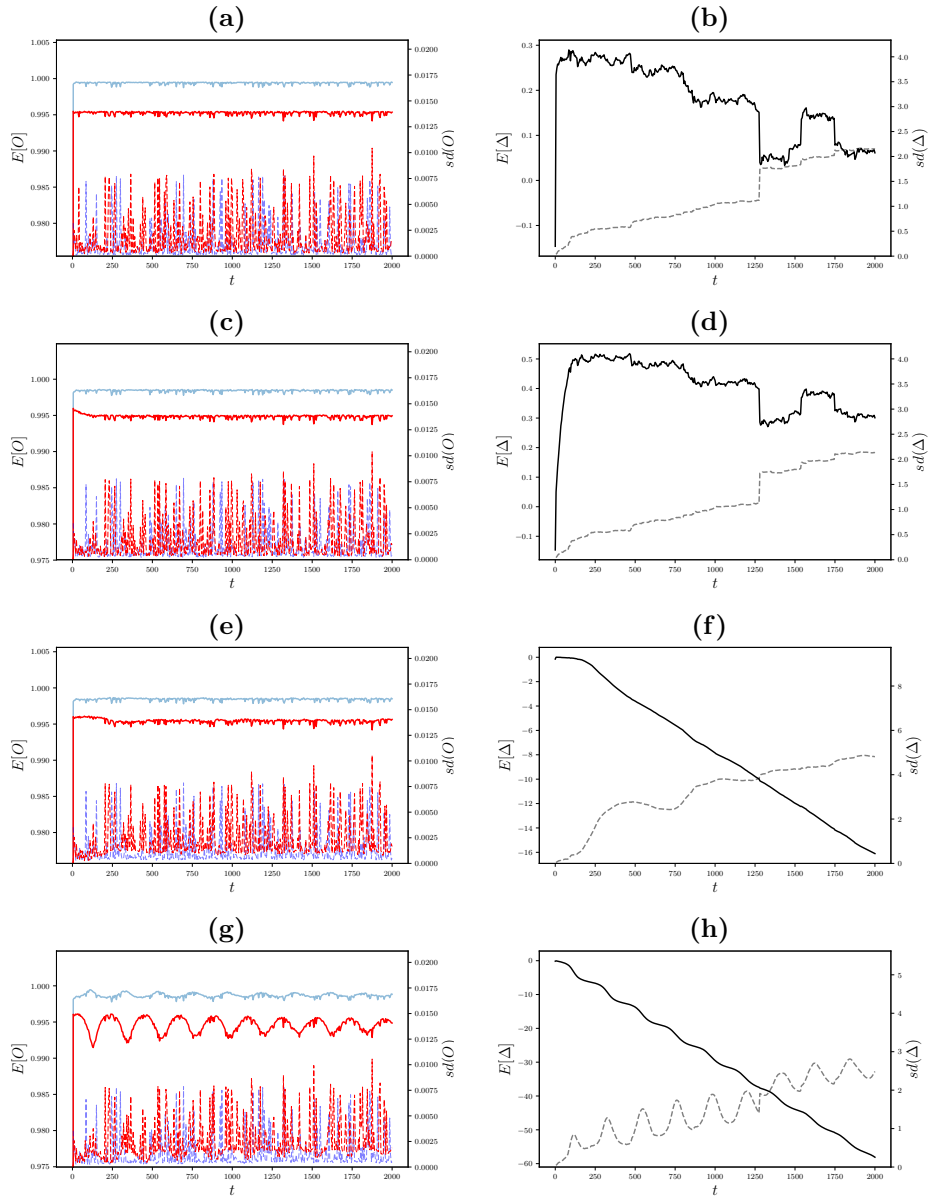


FIGURE 3.4: Time series plots in the stable Lévy noise case with $\sigma = 0.05$ and $\alpha = 1.5$. Left axis: Time series plots for the average order parameters and Δ. Right axis: Time series plots (dashed lines) for the standard deviation of the Monte Carlo estimates of the order parameters and Δ. Left column: order parameters for the blue and red networks as given by the blue (light grey) and red (grey) lines respectively. Right column: Δ as denoted by the black lines. (a) and (b): $(\phi, \psi) = (0.2\pi, 0)$, (c) and (d): $(\phi, \psi) = (0.94\pi, 0)$,(e) and (f): $(\phi, \psi) = (0.95\pi, 0)$,(g) and (h): $(\phi, \psi) = (0.96\pi, 0)$

a constant mean. The heavy tail jumps are more evident in Δ in contrast to O_B, O_R. Below $\mathcal{K} = 0$, as for $\phi = 0.96\pi, \psi = 0$ a clear periodicity is evident in all measures, however, for O_B and O_R the superimposed jumps are also marked. These should be contrasted with a far more incremental stochasticity in the Gaussian case shown in [6]. The behaviour of Δ in the fourth instance is quite telling: the periodicity is evident, however after each cycle there is a distortion in the period that incrementally increases.

The important point to emphasise is that we are applying noise here to all nodes so that we are stimulating in essence both the zero and normal modes according to the Laplacian. Nevertheless, we can distinguish the impact on the two classes of mode quite cleanly. Firstly, the dominance of the periodicity when $\mathcal{K} < 0$ with noise superimposed is a consequence of the noise on *normal modes* but the underlying deterministic behaviour of the zero mode, namely $\Delta(t)$, leaving its mark. Secondly, the stochasticity in the periodicity itself of Δ when $\mathcal{K} < 0$, is a consequence of the noise on the zero mode itself. Specifically, the Lévy ratchet potential manifests itself in the probabilistic nature of how long the system stays in the metastable well when subject to noise. We recall from [6] that when Gaussian noise is only applied to normal modes the behaviour of $\Delta(t)$ in the $\mathcal{K} < 0$ regime is strictly periodic.

In light of this discussion, understanding the tempered case is straightforward. Shown in Figure 3.5, we observe essentially the same characteristics but with moderation of the stochasticity. Tempering here is at $\lambda = 0.75$ so that the jumps in O_B and O_R are significantly smaller and sparser, both for constant and periodic behaviours. In the behaviour of Δ the broadening for $\mathcal{K} > 0$ is smoother. For the periodic regime, $\mathcal{K} < 0$, the period is far more well-defined as a consequence of the stabilisation of the system in the ratchet potential.

We have avoided thus far to show results for $0.5 < \alpha < 1$ as there is little structure distinguishable due to the heaviness of the tails. However, in the presence of tempering with $\alpha = 0.5$ recognisable structures emerge as shown in Figure 3.6. We see here the behaviour predicted analytically, that with tempering and small values of α a smoother dynamics is restored.

3.5.3 Objective function for deterministic system

We now focus on the fitness landscape, where fitness refers to optimal *internal synchronisation and phase lag* for one population with respect to the other; we choose Blue for

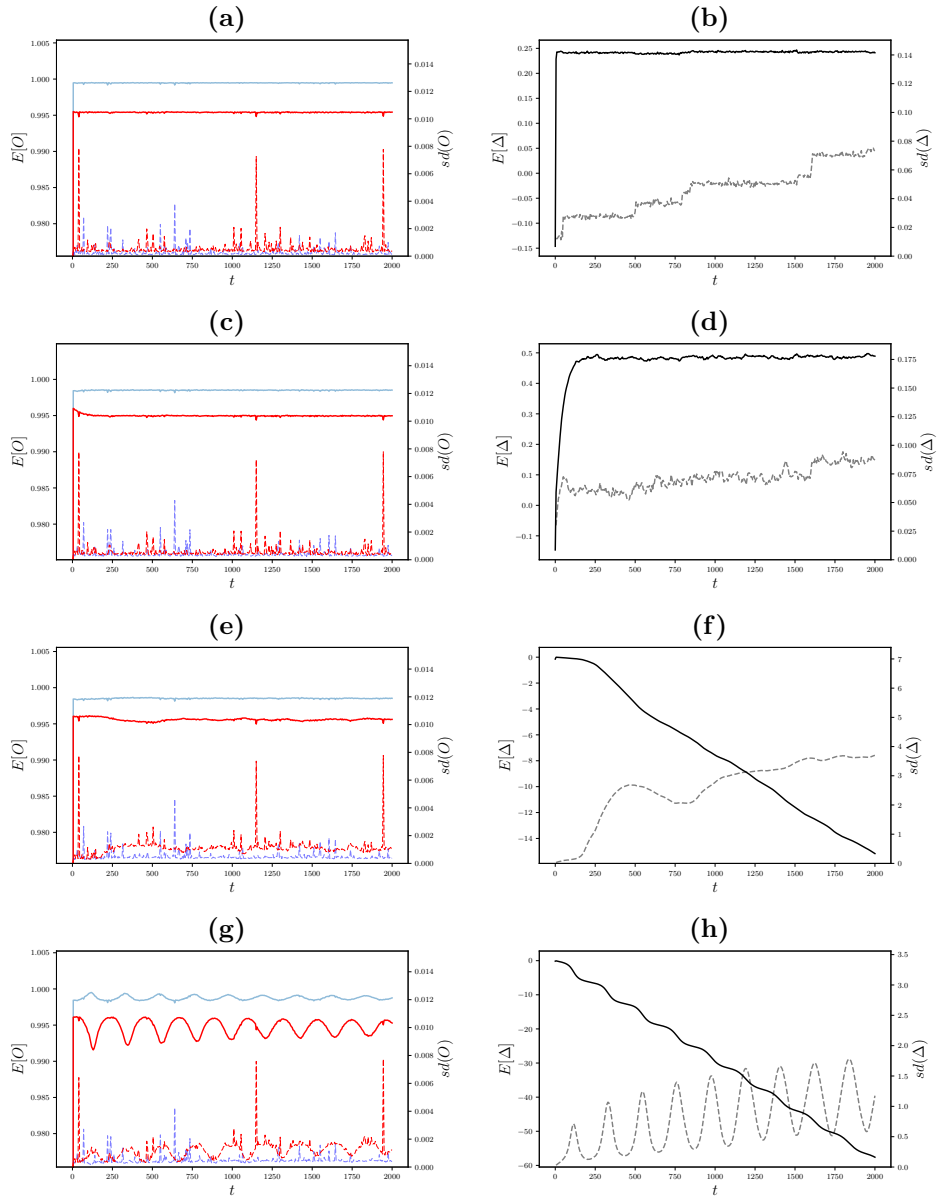


FIGURE 3.5: Time series plots in the stable Lévy noise case with $\sigma = 0.05$ and $\alpha = 1.5$ and $\lambda = 1$. Left axis: Time series plots for the average order parameters and Δ. Right axis: Time series plots (dashed lines) for the standard deviation of the Monte Carlo estimates of the order parameters and Δ. Left column: order parameters for the blue and red networks as given by the blue (light grey) and red (grey) lines respectively. Right column: Δ as denoted by the black lines. (a) and (b): $(\phi, \psi) = (0.2\pi, 0)$, (c) and (d): $(\phi, \psi) = (0.94\pi, 0)$,(e) and (f): $(\phi, \psi) = (0.95\pi, 0)$,(g) and (h): $(\phi, \psi) = (0.96\pi, 0)$

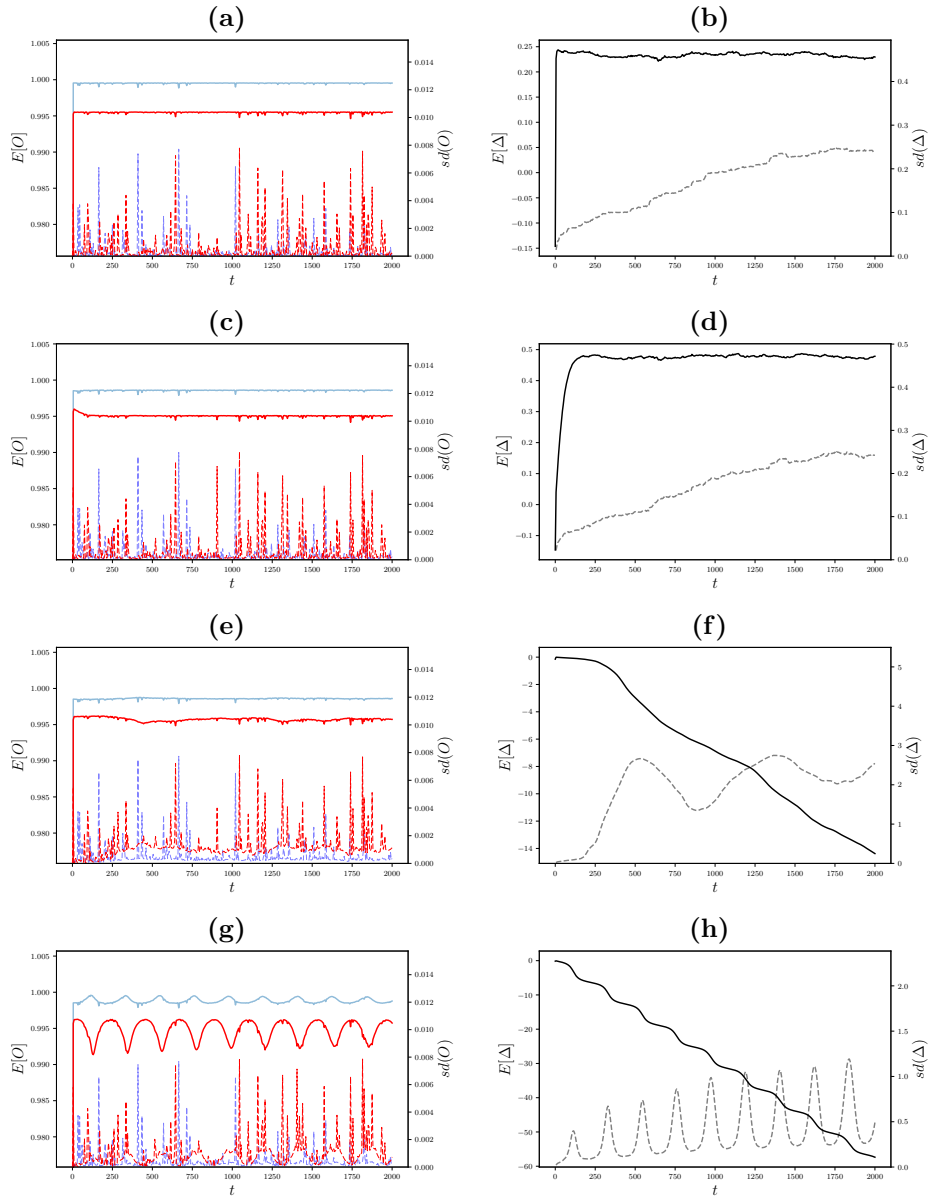


FIGURE 3.6: Time series plots in the stable Lévy noise case with $\sigma = 0.05$ and $\alpha = 0.5$ and $\lambda = 0.75$. Left axis: Time series plots for the average order parameters and Δ. Right axis: Time series plots (dashed lines) for the standard deviation of the Monte Carlo estimates of the order parameters and Δ. Left column: order parameters for the blue and red networks as given by the blue (light grey) and red (grey) lines respectively. Right column: Δ as denoted by the black lines. (a) and (b): $(\phi, \psi) = (0.2\pi, 0)$, (c) and (d): $(\phi, \psi) = (0.94\pi, 0)$,(e) and (f): $(\phi, \psi) = (0.95\pi, 0)$,(g) and (h): $(\phi, \psi) = (0.96\pi, 0)$

this purpose. To this end we choose an objective function given by

$$U(\phi,\psi) = \left(\sum_{i=1}^{3} w_i U_i\right)^{-\frac{1}{4}} \tag{3.46}$$

where w_i are the weights of sub-objective functions U_i. These sub-objective functions, upon the discretisations $(O_B)_i = O_B(t_i), t_i \in [0,T], i = 1,\ldots,N$, and similarly for $(O_R)_i$ and α_i, are given by

$$\begin{aligned} U_1 =& \sqrt{\sum_{i=1}^{N}((O_B)_i - 1)^2} + \sqrt{\sum_{i=1}^{N}(\frac{d}{dt}(O_B)_i - 0)^2} \\ U_2 =& \sqrt{\sum_{i=1}^{N}(\Delta_i - \phi)^2} \\ U_3 =& \sqrt{\sum_{i=1}^{N}(\frac{d}{dt}(O_R)_i - 0)^2} \end{aligned} \tag{3.47}$$

where U_1,U_2 and U_3 corresponds to O_B being constant and close to 1, Δ being close to ϕ and O_R being constant respectively. We indicate the value of 0 in Eq. (3.47) to emphasise that objective function seeks constant O_B and O_R. This choice of combination of U_1, U_2 and U_3 captures the trade-offs required by the Blue agents to maintain both stable internal synchronisation $O_B = 1, \dot{O}_B = 0$, and maintain a lead in the cycle ϕ over the competitor, Red, namely $\Delta = \phi$. We also include a requirement that Red have static value of the order parameter O_R, without seeking it to be close to 1. In other words, Blue prefers to be internally well-synchronised, and Red at best frequency synchronised but to the extent that Blue agents interacting with Red may achieve a phase advantage of ϕ. The choice of the power $-\frac{1}{4}$ in Eq. (3.46) is for visualisation purposes to enable better resolution of peaks and troughs in the landscape, as will be seen shortly.

3.5.4 Bayesian optimisation

Instead of performing an exhaustive grid search across the parameters ϕ and ψ over the region $[0,2\pi]\times[0,2\pi]$ we approach the problem of finding a global optimum of the various objective functions using the theory of constrained global optimisation built upon the ideas of Bayesian inference and Gaussian processes [7, 107]. In other words, we seek

$$\max_{\phi,\phi} \{U(\phi,\psi) : \phi,\psi \in [0,2\pi]\},$$

using the minimal number of probes of $U(\phi, \psi)$. One of the consequences of this approach is that we may obtain a view of the fitness landscape by interpolating over the intermediate values obtained by the algorithm as it searches for the global maximum.

As discussed in Appendix A.1, Bayesian optimisation is different to other kinds of optimisation in that it constructs a probabilistic model for seeking the maximum of a function $f(x)$ on some bounded set $\mathcal{X} \subset \mathbb{R}^d$ and then exploits this model to make decisions about where in $\mathcal{X}$ to evaluate next. It uses all of the information available from previous evaluations of $f(x)$ and not simply local gradient and Hessian approximations. This gives an algorithm that finds the maximum of non-convex functions with relatively few evaluations at the cost of performing more computation to determine the next point to try. This approach is particularly well suited when the evaluation of $f(x)$ is expensive to perform (as in our case). When performing this approach in practice there are two major implementation choices that need to be made. The first is a model for the prior over functions that express assumptions about the function being optimised. We assume here a Gaussian process prior for the influence of ϕ and ψ on $\mathcal{K}$. Second, the implementation requires a choice of an 'acquisition function' used to construct a utility function from the model posterior and allows the algorithm to determine the next point to evaluate.

To improve the performance of the Bayesian optimisation algorithm and, in the case of the stochastic system, to account for the standard error of our estimate in Eq. (3.45), we perform data smoothing on a coarser version of the original data set through the application of a symmetric moving average filtering algorithm. The symmetric moving average filter that we used is of window length $2n + 1$ and is given by

$$(\hat{O}_B)_i = \sum_{j=-n}^{n} \varpi_j \mathbb{E}\left[(O_B)_i\right], \quad n < i < N - n \tag{3.48}$$

where the weights are chosen such that $\sum_j \varpi_j = 1$ and the expectations are approximated through Eq. (3.45). In our application of this algorithm we make the simple selection of $\varpi_j = \frac{1}{2n+1}$. For the n observations at the beginning and end of the time series, we use an asymmetric moving average filter to preserve all observations. Unless otherwise stated, we use $n = 10$ in our window length.

3.5.5 Fitness landscape for deterministic system

In the top panel of Fig. 3.7 we show the objective function for the deterministic system in the ϕ, ψ plane. We compare it below with a plot of the objective function but with only the term corresponding to w_1 included, where we evaluate $O_B(\phi, \psi)$ based on the

truncation to lowest order in the fixed point approximation, Eq. (3.28), as discussed earlier. In addition, we overlay on this latter plot the lines (in black) where $\mathcal{K} = 0$.

In the top panel of Fig. 3.7 we observe an overall 'ridge' of optimality in the bright strip on the diagonal, and, with same alignment, dark narrow troughs far from optimality. This ridge however sits at the top of a similarly aligned region that drops off orthogonally to the bright ridge down to the darkly coloured troughs. Correspondingly, the overall symmetry of the objective function on the top is displayed in the truncated form of O_B below. We may infer that the negatively sloped diagonal of maxima and minima in the top panel are a consequence completely of the behaviour of $\mathcal{K}$ shown in the lower plot. Because of the specific small choice of μ in our example, the region where $\mathcal{K} < 0$ (where the angle between Blue and Red centroids, Δ, becomes time-dependent) is very narrow - in the narrow gap between the black lines in the bottom plot. This is, therefore, responsible for the narrow troughs in the objective function on the top. Similarly, where the objective function takes its maximum coincides at points with where $\mathcal{K}$ achieves its maximum positive value, as well as O_B itself as seen periodically along the central diagonal of the lower plot.

However, something difficult to detect in a single quantity such as used in O_B on the bottom but evident in the full objective function on the top is that it is impossible for Blue to achieve the *combined objective* of optimal internal synchronisation *and* large frustration in relation to Red, ϕ, for a fixed value of Red's frustration in relation to Blue. This is visible in the top plot as follows. On the one hand, following the bright ridge of the plot, positive values of ϕ require negative values of ψ to maintain high optimality. This ridge corresponds to Red being 'compliant' with Blue's intent: large positive phase ϕ for Blue is consistent with large negative phase, $-\phi$, for Red. However these bands do not extend indefinitely along the diagonals of the (ϕ, ψ) plane but are truncated. This is a consequence of the presence of U_2 in the objective function Eq. (3.47) where optimality requires a sign of Δ: Blue seeks to be a certain phase *ahead* of Red. This truncation of the diagonal band is only seen partially in the bottom plot in that the central cell is a square rather than a rectangle in the upper plot.

Once values of (ϕ, ψ) off the diagonal are chosen, the degree of optimality drops off into rectangular bands that align with the diagonal ridge. This drop off is a consequence of the combination of internal synchronisation *and* achievement by Blue of maintaining close to the phase difference ϕ ahead of Red, but where Red chooses 'non-compliant' values ψ. The larger is $\psi > 0$, the smaller the range of $\phi > 0$ available before instability occurs, namely a dark ridge in the plot is encountered. Thus, this landscape representation summarises the trade-offs in internal synchronisation and ability to synchronise in a stable manner ahead of the competitor available to one side or the other given the

choices of the other. The question we address next is how much of this structure is maintained in the presence of stable and tempered stable noise.

3.5.6 Fitness landscape for stochastic system

To study the fitness landscape in the presence of noise, we use the objective function of Eq. (3.46) but with sub-objective functions using the discretisations $(\hat{O}_B)_i = \hat{O}_B(t_i), t_i \in [0,T], i = 1,\ldots,N$, and similarly for $(\hat{O}_R)_i$ and $(\hat{\Delta})_i$, are given in terms of the smoothed expected values as defined in Eq. (3.48). The sub-objective functions then are as follows

$$
\begin{aligned}
U_1 =& \sqrt{\sum_{i=1}^{N}((\hat{O}_B)_i - 1)^2} + \sqrt{\sum_{i=1}^{N}(\frac{d}{dt}(\hat{O}_B)_i - 0)^2}, \\
U_2 =& \sqrt{\sum_{i=1}^{N}((\hat{\Delta})_i - \phi)^2}, \\
U_3 =& \sqrt{\sum_{i=1}^{N}(\frac{d}{dt}(\hat{O}_R)_i - 0)^2},
\end{aligned}
$$

where U_1,U_2 and U_3 corresponds to $\hat{O}_B$ being constant and close to 1, $\hat{\Delta}$ being close to ϕ and $\hat{O}_R$ being constant respectively.

In Fig. 3.8 we show four cases of the objective function, comparing the deterministic, Gaussian (top left and right) and, importantly, the stable case for $\alpha = 1.5$ (bottom left) and tempered stable case with $\alpha = 1.5, \lambda = 0.5$. In particular, for the two non-Gaussian cases we have used the Bayesian optimisation approach.

We observe firstly that the Gaussian case (top right) maintains to a high degree the integrity of the landscape from the deterministic result. However, stable noise (bottom left), as a consequence of the heavy tails, significantly smears the landscape - we note to this end the change in the colour gradations for this case showing that overall values of the objective function are suppressed in comparison to deterministic and Gaussian cases. Also, the accumulation of points searched by the Bayesian optimisation underscores the relative sub-optimality of (ϕ, ψ) values away from the origin. Nevertheless the structure of the landscape is maintained. With suppression of the heavy tails by tempering (bottom right), the landscape is closer to the Gaussian case. In Fig. 3.9 we plot numerical evidence that the optimal fitness landscape results picks out optimal values by plotting the time series of the various order parameters at a point within the optimal band.

In conclusion for this stage, the overall imprint of the deterministic dynamics is retained in non-Gaussian noise cases. In particular, the partition of the (ϕ, ψ) plane according to the sign of $\mathcal{K}$ remains key for determining regions of stochastic optimality.

3.5.7 Determining the presence of regular and noisy path behaviour

In order to analyse for which values of $(\alpha, \sigma, \lambda)$ the property of periodicity holds we introduce a new objective function

$$U^{(01)}(\alpha, \sigma) = 1 - K \tag{3.49}$$

where K is the result of the application of the '0-1 test' upon the $\hat{O}_R$ time series as defined in Eq. (3.48). The 0-1 test is the means for determining the presence of chaos in time series data that is a computationally efficient alternative to computing a Lyapunov exponent [9, 10]. In this work, we shall utilise the 0-1 test to differentiate between regular and noisy path behaviour of $\hat{O}_R$. To apply the 0-1 test upon the time series $\hat{O}_R$, we select an observable from the solution $\psi(\hat{O}_R)$ and some parameter $c > 0$. The method used in this test is independent of the form of the function ψ. We now define

$$\theta(t) = ct + \int_0^t \psi(\hat{O}_R) ds,$$
$$p(t) = \int_0^t \psi(\hat{O}_R) \cos(\theta(s)) ds.$$

If the underlying dynamics are regular then $p(t)$ is bounded otherwise $p(t)$ behaves asymptotically like Brownian motion. To identify the growth of $p(t)$, the mean-square displacement is used:

$$M(t) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \left(p(t+\tau) - p(\tau)\right)^2 d\tau.$$

If the behaviour $p(t)$ is Brownian, then $M(t)$ will grow linearly in time; otherwise if $p(t)$ is bounded, then $M(t)$ will be bounded. To distinguish between these two properties, the asymptotic growth rate of the mean-square displacement is calculated:

$$K = \lim_{t} \frac{\log(M(t))}{\log t}.$$

This can be calculated through a linear regression of $\log(M(t))$ and $\log t$ and will be near 1 if the underlying dynamics are chaotic or noisy and 0 otherwise. Note that the choice of $1 - K$ in Eq. (3.49) is the reverse of the typical test, used in this way here for consistency with the landscape plots wherein large values (periodic behaviour in the 0-1 test and desired dynamics in the previous objective functions) are 'good', and low

scoring values (noisy behaviour in the 0-1 test and undesired dynamics in the previous objective functions) are 'bad'. More details of the 0-1 test are provided in Appendix A.2.

We now conduct straightforward grid search through the range of parameters in presenting landscape plots. We show two examples of computation of $U^{(01)}$ for stable ($\lambda = 0$) and tempered stable (varying across λ), in Figs. 3.10 and 3.11, where yellow indicates ordered regions and blue noisy regions. We choose a case of $\phi = 0.96\pi, \psi = 0$ where $\mathcal{K} > 0$ and therefore we know that the deterministic dynamics should be constant in quantities such as O_B, O_R and Δ.

For the stable case, we show both one-sided (top panel) and two-sided (bottom panel) cases in Fig. 3.10. We observe that across all values of α, for $\sigma > 0.05$ there is an onset of noise with the strongest departures to noise at low α consistent with the property that tails in the noise distribution become very heavy. Evidently the role of skew in the noise plays a marginal role.

Contrastingly, we now choose $\sigma = 0.05$ and explore the tempered stable case in Fig. 3.11 for $\phi = 0.96\pi$ for both the one-sided ($\theta = 1$) (top) and symmetric ($\theta = 0$) (bottom) cases. We see that close to the Gaussian regime $\alpha \approx 2$, there is an absence of noisy behaviour across all values of λ, as to be expected. We also see that for $\lambda = 0$, going vertically down the left hand edge of the plot, there is a transition from steady, to periodic behaviour, to noisy behaviour and then long transients, reflecting the structure of the landscape for the stable case in Fig. 3.10. However, once $\lambda \neq 0$ this monotonic trend from steady state to noisy behaviour changes. For any fixed $0 < \lambda < 0.3$ as α decreases there is a transition to noisy behaviour and then back to periodic behaviour (yellow goes to blue which goes back to yellow). Thus there is a *restoration* of ordered behaviour for all λ for *small* values of α, typically $\alpha \leq 0.25$.

In other words, Blue and Red will exhibit smoother periodic behaviour in relation to each other. This is broadly consistent with the pattern observed analytically for the dynamics of Δ using the ratchet potential in the zero mode projection of the Blue-vs-Red system. Recall three significant effects were seen to occur in that approach: the skew drops out, there is a greater suppression of variance in Δ for small α for non-zero λ, and there is a reduction of the average drift between Blue and Red clusters. That the skew does not play a role can be seen in the result for the landscape of for $U^{(01)}$ for zero skew (namely symmetric two-sided noise) in the bottom panel, with $\phi = 0.96\pi$ again. Once more we observe the suppression of noisy behaviour at small α. Though the onset of noisy behaviour is more general in the full system, applying between and within Blue and Red. However, the analytical result for the variance of Δ and the average value of $\dot{\Delta}$ shows that the generation, or suppression, of variability in the zero mode of the system is *consistent* with this behaviour. Of course there is need for caution, as the

Noise	U	Parameters			
		ψ	ϕ	ζ_{BR}	ζ_{RB}
None	57.8598	1.675π	$0.1944\ \pi$	0.8997	0.1628
Gaussian	98.6311	1.3747π	0.0033π	0.8170	0.0668
Stable	10.6854	$0.2007\ \pi$	$0.0083\ \pi$	0.8289	0.7124
T. Stable	42.0838	1.2693π	0.0143π	0.9932	0.1017

TABLE 3.1: Bayesian optimisation results for $U(\psi, \phi, \xi_{BR}, \xi_{RB})$ where $\psi, \phi \in [0, 2\pi]$ and $\xi_{BR}, \xi_{RB} \in (0, 1]$ for 500 iterations and T = 2000. $\sigma = 0.05$, $\alpha = 1.5$ and $\lambda = 0.75$.

analytic approach is not valid in the deep noisy regime; the *onset* of drift and variation from the noise is a precursor to the stronger regime of noisy behaviour. In this sense, we may say there is qualitative agreement between numerical and analytical results on the impact of small α in suppressing the tendency to noisy behaviour of the nearly locked Blue-Red system in the presence of tempering; increasing the exponent of the tail of the distribution does not disrupt Blue and Red locking with monotonically increasing strength. The observation of this, and the broad consistency between analytical and numerical approaches are the main results of this work.

3.5.8 Optimisation across multiple parameters

We have seen throughout the work that the non-linear nature of the interactions in the Blue-v-Red model mean that, even deterministically, it is not generally possible for one population, say Blue, to achieve the desired phase ahead of the other, for example ϕ. We now use the Bayesian approach to search for optima *for Blue* across the four cross-population parameters $\phi, \psi, \zeta_{BR}, \zeta_{RB}$ given the Blue and Red networks and internal couplings ζ_B, ζ_R as previously given. We conduct 500 iterations running the simulation out to $T = 2000$ and consider four cases: deterministic, Gaussian (as used in [6]), stable ($\alpha = 1.5$) and tempered stable ($\alpha = 1.5, \lambda = 0.75$). The results are shown in Table 3.1.

We note that the Gaussian case obtained a higher scoring optimal value than the deterministic case as a result of either a relativity small number of optimisation iterations for a 4-dimensional parameter space and the algorithm favouring a local maximum. The optimisation algorithm was unable to find a decently scoring set of parameters in the stable Lévy noise case. We observe nevertheless a broad consistency between the three other cases in that $\psi > \phi$ and $\zeta_{BR} > \zeta_{RB}$. In the interests of space we forego here showing plots of the corresponding time-series for $O_{B,R}$ and $\Delta(t)$ but simply comment that O_B shows generally higher values than O_R, and $\Delta(t)$ is a positive constant for deterministic, ($\Delta \approx 0.6$), and Gaussian ($\Delta \approx 0.01$) cases. For the tempered stable case, the expected value of Δ is approximately 0.06, however the extreme cases are time-varying both positively and negatively, not untypical for $1 < \alpha < 2$. The requirement

that Δ is close to ϕ is part of the objective function U_2 ensures that Blue comes as close as possible to its objective by not seeking to be as ambitious in its phase ahead of Red as Red seeks in its relation to Blue. The actual values of Δ match ϕ for the deterministic and Gaussian cases. On the other hand, to achieve this Blue agents must seek to couple to Red agents more strongly than Red to Blue; the nature of the internal networks intuitively plays an important role in this, just as Blue's hierarchy requires stronger internal coupling to achieve its level of internal synchronisation. Stable noise is generally destructive of such optima outside the narrow range where one population complies with the intent of the other (for example, $\phi = -\psi$), as previous results have emphasised. These results are simply a taste of what is possible with such a model. Future work may consider optimisation across more parameters and ultimately across different networks.

3.6 Conclusions

We have explored the two networked populations 'Blue-vs-Red' frustrated Kuramoto model subject to tempered stable noise. Using approximations in the region where the two populations are internally synchronised but seek to lock with respect to each other, we have solved the stochastic dynamics using the tempered fractional Fokker-Planck approach. The zero-mode projection under the graph Laplacian of this reveals that, whereas for $1 < \alpha < 2$, noise generally causes an otherwise deterministically locked system to degrade leading to drift between Blue and Red clusters, for small $\alpha < 1$ improved locking between Blue and Red takes place. In the situation where deterministically the system shows locking, the noise may be such that stochastically induced drift between Blue and Red is entirely cancelled. When the deterministic system exhibits drift between Blue and Red clusters, such tempered stable noise can be used to moderate the stochastic drift.

Through our simulations we see a similar pattern: noise of higher strength or stronger tails can disrupt the locking between Blue and Red in cases where deterministically the system may be achieving high levels of global synchronisation - with both sets of agents well synchronised internally and having a fixed frustration in relation to each other. Tempering, naturally, has the effect of restoring locking by moderating the heavy tails. The surprising result we find is that with tempering in the regime where $\alpha < 1$ the system may be become less noisy, and thus showing smoother periodic, and approaching steady-state, behaviour. Again, this is consistent with Blue and Red either smoothly drifting in relation to each other or even locking. There is only the caveat that the analytical result does not apply in the noisy regime, but provides an indicator or precursor to it.

Thus both analytically and numerically we obtain a surprising result, that the boundaries of regions of different dynamics (steady-state, periodic, noisy behaviour) of the stochastic system are not linear in variation of the parameter α governing the tail of the noise once tempering is introduced. In contrast, for purely stable noise reducing α only destroys locking between Blue and Red with monotonically increasing severity.

Intuitively, we propose that the underlying reason for this behaviour lies in an interplay of the ratchet mechanism with tempering and the fundamental change in the stable noise probability density for $\alpha < 1$. As illustrated in [53], for $\alpha < 1$ the density is no longer centred around zero but with a heavy tail that is increasingly suppressed as λ increases. This centering becomes the dominant effect with tempering, and seemingly acts to counter the deterministic drift in the ratchet potential.

Importantly, much of the structure of the landscape of the stochastic system, as determined through simulation, can be predicted through the analytically accessible thresholds for instability and periodic behaviour from the deterministic system at least with weak noise strength. Thus analytic understanding of the landscape in the absence of noise remains an important guide to the overall structure of the landscape of dynamics in the regime of weak noise, close to the boundary with noisy behaviour. These insights are then complemented by examining the Laplacian zero mode projection of the linearised system, where the properties of ratchet potentials under tempered stable noise plays an important, and analytically tractable, role in predicting the weak noise strength dynamics. This nonlinear mechanism, with tilted ratchet potentials, has not been articulated before in the context of synchronisation on multi-networks.

Interpreting these results in the context of the competitive system of Blue and Red agents is now interesting. Intuitively it is clear that if two populations both seek to maintain a certain phase advantage over the other it is not possible for both to simultaneously achieve their aims. For different network structures and couplings one side or the other may be able to achieve close to their result within a certain bound. Indeed, our results show that in such interactions 'less is more', namely that one side seeking less ambitious phase advantage in relation to the other may be able to achieve more optimal results. The landscape method shown here enables determination of these boundaries before instability is triggered, and how close to optimality may be achieved. The added dimension of noise, in general, makes this optimal behaviour even more difficult. However, for certain forms of non-Gaussian noise with tempering λ for large jumps and far-from-Gaussian power law index α better approach to optimal behaviour may be achieved. Indeed, such behaviour becomes attractive as a mechanism for guiding the behaviour of otherwise complex competitive stochastic systems or informing an understanding of the capacity for resilience or robustness of the system. As alluded already, future work may

broaden the scope of the optimisation to cover the networks themselves. For example, for a given topology, or statistical parametrisation thereof, of Red what is the optimal network within and across for Blue? Returning to our overall aim to apply such models to competitive distributed decision making or social processes, investigations like this with the two-network frustrated Kuramoto model open the door for testing many rather qualitative hypotheses in the organisational, social and management science literature.

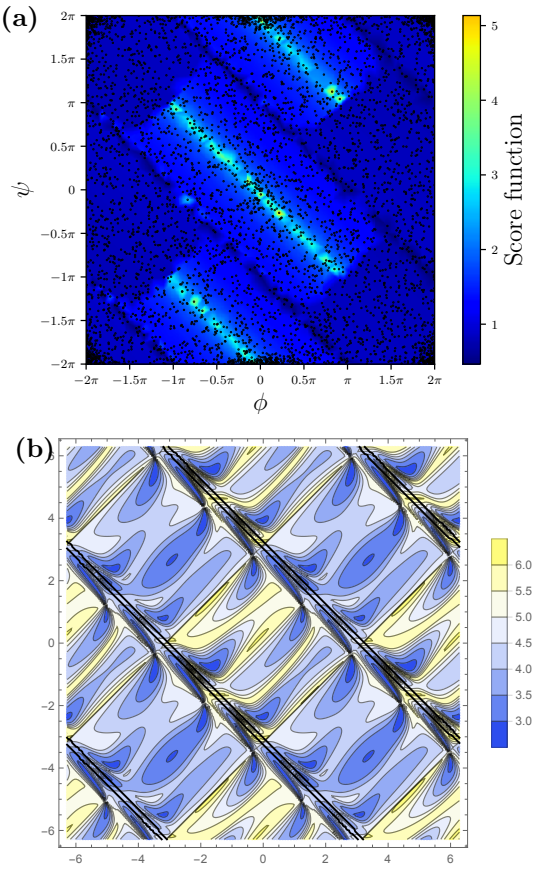


FIGURE 3.7: (a): Objective score function for $U(\phi, \psi)$ in the deterministic case with 500 iterations and T = 1000. The objective weights are given by $w_1 = 1$, $w_2 = 0.1$ and $w_3 = 1$. High scoring (red) values correspond to O_B constant and close to 1 and O_R constant whilst low scoring values (blue) correspond to either periodicity or noisy behaviour in the times series or O_B far away from 1. Note that the optimal value near $(1.1\pi, 1.85\pi)$ in this fitness landscape corresponds to a periodic time series with a very long periodicity, beyond that of T=1000. The subsequent plot repeats the construction of the fitness landscape for T=2000. (b): the objective function including only the analytical evaluation of O_B truncating to leading order in the fixed point. This is overlaid in black with the lines where $\mathcal{K} = 0$.

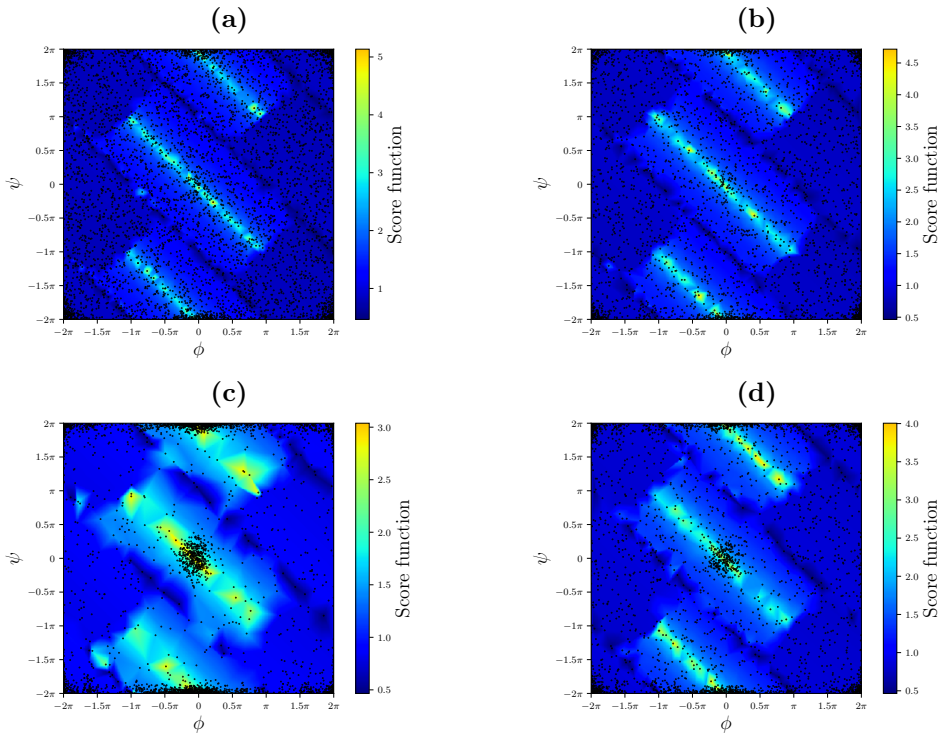


FIGURE 3.8: Objective score function for $U(\phi, \psi)$ for different noise cases: (A) deterministic , (b) Gaussian (top right), (c) stable with $\sigma = 0.05, \alpha = 1.5$, and (d) tempered stable with $\sigma = 0.05, \alpha = 1.5, \lambda = 0.5$, with 500 iterations and $T = 1000$. The objective weights are given by $w_1 = 1$, $w_2 = 0.1$ and $w_3 = 1$. High scoring (red) values correspond to O_B constant and close to 1 and O_R constant whilst low scoring values (blue) correspond to either periodicity or noisy behaviour in the times series or O_B far away from 1. Note that the optimal value near $(1.1\pi, 1.85\pi)$ in this fitness landscape corresponds to a periodic time series with a very long periodicity, beyond that of $T = 1000$. The subsequent plot repeats the construction of the fitness landscape for $T = 2000$.

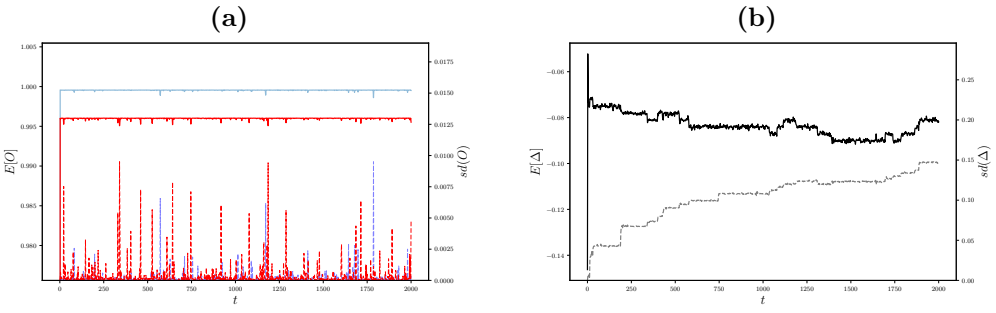


FIGURE 3.9: Time series plots in the optimal regime of $U(\phi, \psi)$ in the tempered stable Lévy noise case with $\sigma = 0.05$, $\alpha = 1.5$, $\lambda = 0.5$ and $(\phi, \psi) = (0, 0)$. Left axis (solid lines): Time series plots for the average order parameters and Δ. Right axis (dashed lines): Time series plots for the standard deviation of the Monte Carlo estimates of the order parameters and Δ. (a) order parameters for the blue and red networks as given by the blue (light grey) and red (grey) lines respectively. (b) Δ as denoted by the black lines.

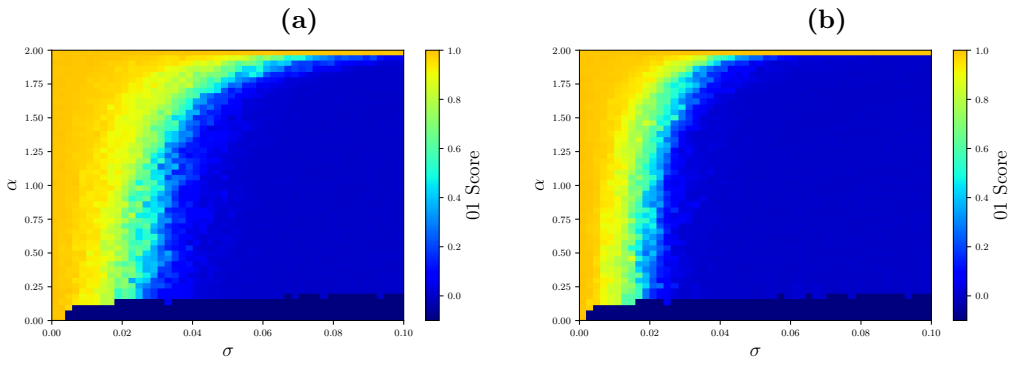


FIGURE 3.10: Heat maps of the objective score function for $U^{(01)}(\alpha, \sigma)$ in the stable Lévy noise case with $\phi = 0.96\pi$, $\psi = 0.$ and 100 iterations for T = 6000; (a) is the one-sided and (b) is the two-sided case. High scoring (dark yellow) values correspond to O_R steady or periodic behaviour, low scoring non-zero values (light blue) correspond to noise, and negative values (dark blue) correspond to long transients in the times series.

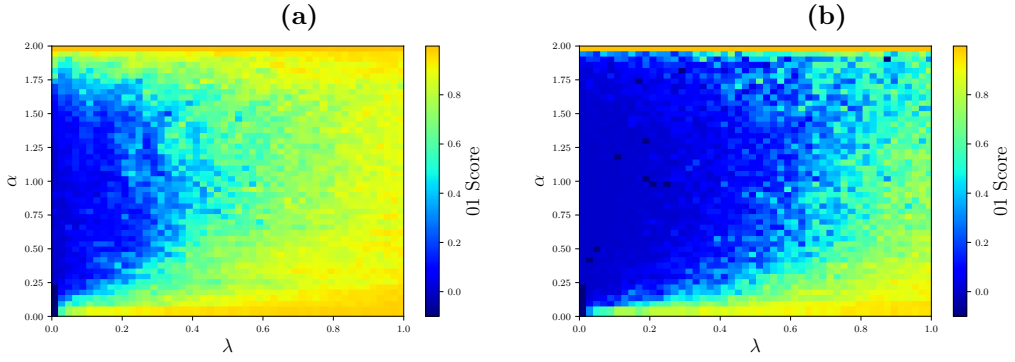


FIGURE 3.11: Heat maps of the objective score function for $U^{(01)}(\alpha, \lambda)$ in the tempered stable Lévy noise case, with $\psi = 0.$, $\sigma = 0.05$ and 100 iterations for T = 6000; top for the one-sided $\phi = 0.96\pi$ and bottom for the two-sided case with $\phi = 0.96\pi$. High scoring (dark yellow) values correspond to O_R steady or periodic behaviour, low scoring non-zero values (light blue) correspond to noise, and negative values (dark blue) correspond to long transients in the times series.

Chapter 4

Swarming Adversarial Models under Tempered Stable Noise

In this chapter we seek to extend the model introduced by Eq. (3.1) in Chapter 3 to include spatial dynamics in a similar manner as described in the swarming framework as was recently introduced by O'Keeffe et al. in [12]. The aim is to take this so-called Blue vs Red swarming model and apply it in the context of attrition modelling which seeks to model direct combat between a blue and a red network. Historically, attrition models have a long standing basis in the so-called Lanchester equations [13] derived from Lanchester's laws which analyses the relative strengths of opposing military forces. These attrition equations have been recently investigated in the context of the Blue vs Red Kuramoto model by Kalloniatis in [11]. The first attrition model that we propose is a model in which the strength of each node is described by a process whose drift is negatively affected by the variance of the phases of the nodes within the respective networks and driven by a spectrally negative Lévy noise. The second of the two models seeks to incorporate the Lanchester equations into the attrition model in a similar manner to [11].

4.1 Attrition models in the Lanchester setting

An early example of an attrition model based upon the Lanchester equations was examined in [13] and further in [108, 109]. The idea behind these models is to describe the time dependent strengths of the two armies, the blue force B and the red force R,

through the following system of equations

$$\begin{aligned}\dot{B}(t) =&\alpha_S B(t) - \alpha_{DA} R(t),\\ \dot{R}(t) =&\beta_S R(t) - \beta_{DA} B(t),\end{aligned} \tag{4.1}$$

where α_S and β_S are the rate of resupply for the blue and red forces respectively, α_{DA} and β_{DA} are the rate of attrition for the blue and red forces.

Using the Lanchester equations in Eq. (4.1) as a basis, an example of an attrition model in the Blue vs Red setting was recently proposed in [11] to model direct combat between the two opposing forces. This Blue vs Red Lanchester type attrition model is given by the following system of equations

$$\begin{aligned}\dot{\beta}_i =&\omega_i - \sigma_B \frac{B(t)}{B(0)} \sum_{j=1}^{N} \mathcal{B}_{ij} \sin(\beta_i - \beta_j),\\ \dot{\rho}_i =&\nu_i - \sigma_R \frac{R(t)}{R(0)} \sum_{j=1}^{M} \mathcal{R}_{ij} \sin(\rho_i - \rho_j),\end{aligned} \tag{4.2}$$

where the variables $B(t), R(t)$ are governed by the Lanchester equations as follows

$$\begin{aligned}\dot{B}(t) =&\alpha_S O_B(t) B(t) - \alpha_{DA} O_R(t) R(t),\\ \dot{R}(t) =&\beta_S O_R(t) R(t) - \beta_{DA} O_B(t) B(t).\end{aligned}$$

where O_B and O_R are the order parameters for the blue and red networks respectively and the parameters $\alpha_S, \beta_S, \alpha_{DA}$ and β_{DA} are similarly defined as in the Lanchester equations in Eq. (4.1). The idea behind this model is that good sychronisation in the blue network positively affects its ability to resupply itself whilst enhancing its ability to attrit the red network. Conversely, poor sychronisation adversely affects the networks ability to resupply itself and weakens its ability to attrit the red network. The variables $B(t)$ and $R(t)$ in the system Eq. (4.2) directly affect the ability of the respective networks to couple within itself.

4.2 Swarming models

One of the aims of this chapter, is to extend our Blue vs Red model to include spatial dynamics. The incorporation of spatial dynamics, specifically in the context of combat modelling, has an origin in [110] which proposes an extension of the classical modelling of combat through the Lanchester equations as in Eq. (4.1) to include one-dimensional spatial effects. This models extends the Lanchester equations, which in general form,

are given by

$$\begin{aligned} B(t) =& f_1(t, \lambda; B, R), \\ R(t) =& f_2(t, \lambda; B, R), \end{aligned}$$

to the more general PDE given by the following equations

$$\begin{aligned} B(t) =& F_1(t, x, \lambda; B, R, B_x, R_x, B_{xx}, R_{xx}, \dots), \\ R(t) =& F_2(t, x, \lambda; B, R, B_x, R_x, B_{xx}, R_{xx}, \dots), \end{aligned}$$

where f_1, f_2 (and F_1, F_2, respectively) represent the nonlinear interaction laws and λ is the set of all parameters specific to the model dynamics.

Since our goal is to further generalise the spatial dynamics of a combat model, specifically a combat model driven by the Kuramoto equation, we require some framework in which such a generalisation is possible. Our proposed model is based upon the generalised swarming model introduced in [12]. The so-called swarming model is named for the property that the oscillators have phase and spatial dynamics which are coupled. The generalised dynamics of the system of swarmalators is given by the following equations

$$\begin{aligned} \dot{\boldsymbol{x}}_i &= \boldsymbol{v}_i + \frac{1}{N}\sum_{j=1}^{N}\left[\boldsymbol{I}_{\text{att}}(\boldsymbol{x}_j - \boldsymbol{x}_i)F(\theta_j - \theta_i) - \boldsymbol{I}_{\text{rep}}(\boldsymbol{x}_j - \boldsymbol{x}_i)\right], \\ \dot{\theta}_i &= \omega_I + \frac{K}{N}\sum_{j=1}^{N} H_{\text{att}}(\theta_j - \theta_I)G(\boldsymbol{x}_j - \boldsymbol{x}_i), \end{aligned} \tag{4.3}$$

where $\boldsymbol{x}_i$ is the position of the i'th swarmalator, θ_i, ω_i and $\boldsymbol{v}$ are its phase, natural frequency and self-propulsion velocity. $\boldsymbol{I}_{\text{att}}$ and $\boldsymbol{I}_{\text{rep}}$ are the spatial attractions and repulsions respectively, F measures the influence of phase similarity on spatial attraction, H_{att} is the phase interaction and G is the influence of spatial proximity on the phase attraction.

Based upon the swarming framework in Eq. (4.3), the Blue vs Red model previously considered in Section 3 is extended upon to incorporate spatial dynamics by setting the phase interaction H_{att} to the sine function, G is set to the weighted adjacency matrix in which the weighting is described by the Euclidean distance between the nodes, the influence of phase similarity F is described through the cosine function and the spatial attractions and repulsions, $\boldsymbol{I}_{\text{att}}$ and $\boldsymbol{I}_{\text{rep}}$, are described through a inverse power laws and inverse square power laws respectively. These quantities are similarly defined for the inter-networks interactions between the Blue-Red and Red-Blue networks. Using these definitions, the oscillator dynamics in the blue network are governed by the following

system

$$\begin{aligned}\dot{\boldsymbol{x}}_i^{(B)} =& \boldsymbol{v}_i + \frac{1}{N}\sum_{j\in B}\mathcal{B}_{ij}\left[\frac{\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(B)}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(B)}|}(A_B + J_B\cos(\beta_i-\beta_j)) - B_B\frac{\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(B)}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(B)}|^2}\right], \\ &+ \frac{1}{M}\sum_{j\in R}\mathcal{A}_{ij}^{(BR)}\left[\frac{\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(B)}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(B)}|}(A_{BR} + J_{BR}\cos(\beta_i-\rho_j-\phi)) - B_{BR}\frac{\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(B)}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(B)}|^2}\right], \\ \dot{\beta}_i =& \omega_i - \sigma_B\sum_{j\in\mathcal{B}}\frac{\mathcal{B}_{ij}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(B)}|}\sin(\beta_i-\beta_j) - \xi_{BR}\sum_{j\in R}\frac{\mathcal{A}_{ij}^{(BR)}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(B)}|}\sin(\beta_i-\rho_j-\phi),\end{aligned} \tag{4.4}$$

where $i \in \mathcal{B}$ and the following system for the red network

$$\begin{aligned}\dot{\boldsymbol{x}}_i^{(R)} =& \boldsymbol{v}_i + \frac{1}{M}\sum_{j\in R}\mathcal{R}_{ij}\left[\frac{\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(R)}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(R)}|}(A_R + J_R\cos(\rho_i-\rho_j)) - B_R\frac{\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(R)}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(R)}|^2}\right], \\ &+ \frac{1}{N}\sum_{j\in B}\mathcal{A}_{ij}^{(RB)}\left[\frac{\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(R)}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(R)}|}(A_{RB} + J_{RB}\cos(\rho_i-\beta_j-\psi)) - B_{RB}\frac{\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(R)}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(R)}|^2}\right], \\ \dot{\rho}_i =& \omega_i - \sigma_R\sum_{j\in\mathcal{R}}\frac{\mathcal{R}_{ij}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(R)}|}\sin(\rho_i-\rho_j) - \xi_{RB}\sum_{j\in B}\frac{\mathcal{A}_{ij}^{(RB)}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(R)}|}\sin(\rho_i-\beta_j-\psi),\end{aligned} \tag{4.5}$$

where $i \in \mathcal{R}$. The parameters J_B, J_R, J_{BR} and J_{RB} measure the extent to which phase similarity supports spatial attraction where J_B and J_R affect the internal blue and red network attractions respectively and J_{BR} and J_{RB} affect the attraction between the blue to red networks and the red to blue networks respectively. In this system, the dynamics of the phase of the oscillators are similar to that of the Blue vs Red model as previously discussed in Chapter 3 however the adjacency matrix is now weighted by the distance between the two oscillators. The spatial dynamics of the system are governed in three parts namely the self-propulsion velocity, the intra-network (i.e Blue-Blue) attractions and repulsion and the inter-network (i.e Blue-Red) attractions and repulsions respectively. This form of the model is derived in a similar fashion as proposed in [12] with the addition of our inter-network terms and adjacency matrices to account for the fact that our network is not one-to-one.

A convenient assumption in the swarming model is to consider the case of constant

magnitude and direction of the swarmalator propulsion velocity, however the inclusion of alignment into the model can be obtained by setting $\boldsymbol{v}_i = v_0^{(i)}\hat{\boldsymbol{n}}_i$ where $\hat{\boldsymbol{n}}_i = (\cos\beta_i, \sin\beta_i)$. The addition of alignment in the model introduces additional dynamics through the addition of the angular variables $\boldsymbol{\beta}$. One possible description for the dynamics of $\boldsymbol{\beta}$ is given by the Vicsek type interactions from [111]. These interactions describe a relationship between $\boldsymbol{\beta}$ and $\boldsymbol{x}$ and are governed by the equations

$$\dot{\beta}_i = -\beta_i + \frac{1}{|\Lambda_i|}\sum_{j\in\Lambda_i}\beta_j, \tag{4.6}$$

where Λ_i is defined as the set of swarmalators within a distance δ of the i'th swaramalator and $|\Lambda_i|$ is the count of all these neighbours.

The extension of the Blue vs Red swarmalator model in Eq. (4.4) and Eq. (4.5) to the stochastic version can be easily implemented through the addition of additive noise to each of the equations for $\dot{\boldsymbol{x}}_i^{(B)}$, $\dot{\boldsymbol{x}}_i^{(R)}$, $\dot{\beta}_i$, $\dot{\rho}_i$ and similarly for the self-propulsion $\dot{\beta}_i$ as in Eq. (4.6). In our initial analysis of the swarming model, we shall consider the deterministic case and only consider the extension to a stochastic system when we consider our attrition models.

To measure synchronisation in the swarmalator model, we can use a similar set of order parameters as defined by Eq. (3.17) in Section 3 however due to the addition of spatial dynamics, we can extend these order parameters to take into account the positioning of the oscillators. The new order parameters for the swarmalator model extends the usual Kuramoto order parameter to include spatial angle of the oscillators ϕ_i where $\phi_i = \tan^{-1}(y_i/x_i)$ is the spatial angle of the i-th oscillator. These order parameters are given by the following equations

$$O_B^{(\pm)} = S_B^{(\pm)} e^{i\Psi_B^{(\pm)}} = \frac{1}{N}\sum_{j\in\mathcal{B}} e^{i(\phi_j \pm \beta_j)}, \tag{4.7}$$

$$O_R^{(\pm)} = S_R^{(\pm)} e^{i\Psi_R^{(\pm)}} = \frac{1}{M}\sum_{j\in\mathcal{R}} e^{i(\phi_j \pm \rho_j)}, \tag{4.8}$$

where we then take $S_B = \max(S_B^{(+)}, S_B^{(-)})$ as our order parameter for the blue network (and similarly for the red network). The need for the inclusion of the spatial angle of the oscillators in the order parameters arises to account for the potential correlations between ϕ and either β or ρ in the various clustering scenarios that may arise (see [12] for more details.)

One of the issues with the implementation of the spatial angle that is not present in the oscillator phase, is that the spatial angle of the oscillators ϕ_i is dependent upon the choice of reference frame. In its introduction in [12], the reference frame was the origin

(0,0) and this posed no issue as the expected centroid was the origin. This might no longer be the case in our scenario. To fix this, we propose a shift of the reference frame to the centroids of the blue and red networks respectively, thus the corrected spatial angle is given by $\phi_i(x_c, y_c) = \tan^{-1}((y_i - y_c)/(x_i - x_c))$ where (x_c, y_c) is the centroid of the blue or red networks.

4.2.1 Numerical investigation of the Blue vs Red swarming model

In this initial investigation of the swarming model, we consider a simplified version of the model through the following parameter choices. We consider the case of identical oscillators with $\omega_i = \omega$ and $\boldsymbol{\nu}_i = \boldsymbol{\nu}$ where the self propulsion velocity $\boldsymbol{v}$ is assumed to have constant magnitude and direction. Through a choice of reference frame, we can thus set $\omega = 0$ and $\boldsymbol{\nu} = \mathbf{0}$. By rescaling time and space we can set $A_B = A_B = B_B = B_R = 1$. Fixing the coupling strengths $\sigma_B = 8K/(N+M)$, $\sigma_R = 0.5K/(N+M)$ and $\xi_{BR} = \xi_{RB} = 0.4K/(N+M)$, where $K \in \mathbb{R}$ is some scaling parameter, the frustration parameters $\phi = \psi = 0$ and setting $J_B = J_R = J$ and $J_{BR} = J_{RB} = \{0.1J, J\}$ for some $J \in \mathbb{R}$, the two free parameters of the model are then (J, K). The choice of $J_{BR} = J_{RB} = J$ corresponds to to equal inter-network and intra-network spatial attraction based upon phase similarity. The alternative paramerterisation of $J_{BR} = J_{RB} = 0.1J$ was selected based upon the assumption that agent movements is primarily governed upon the phases within their own networks and only a vast difference in phase between the competing networks will resulting in the inter-network term exceeding the magnitude of the intra-network term. We shall initially investigate both parameterisations, however in the later numerical analysis of the model, we shall solely use the later. The first part of our numerical analysis of the Blue vs Red swarming model is given on Figures 4.1 and 4.2. Here we plot some of the states of the system for $(J, K) \in \{(1, -42), (1, -1), (1, 1), (1, 42)\}$ for both $J_{BR} = J_{RB} = \{0.1J, J\}$ after the system is allowed to run until a time $T = 1000$. In these plots, we look to observe the variety of spatial configurations that this swarming model can exhibit. In the case of $J_{BR} = J_{RB} = J$ for $(J, K) \in \{(1, -1), (1, 1)\}$, we obtain similar spatial configurations as observed in [12], however when we increase the coupling constant we note that for large K no stable state is reached. Although no stable state is reached, the spatial configuration present are nonetheless interesting, for example in the plot with $(J, K) = (1, 42)$ we obverse that the blue network attempts to surround the red network whilst the red network attempts to cluster. In the case where we now set $J_{BR} = J_{RB} = \{0.1J, J\}$, additional spatial configuration emerge. When $(J, K) = (1, -42)$ there is no steady state however and in this scenario the blue network attempts to cluster whilst the red networks attempts to surround it. Additionally there is asynchronisation amongst the phases of the red network. In the case of $(J, K) = (1, -1)$,

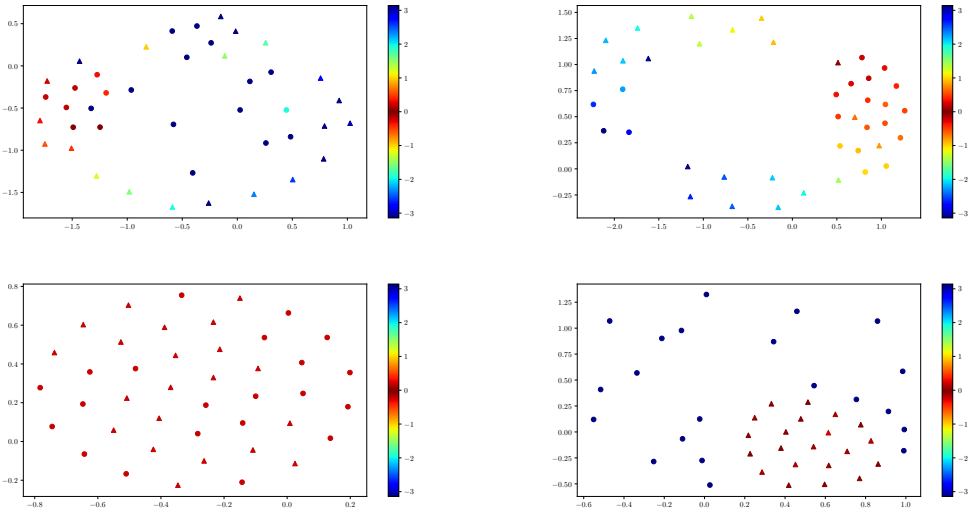


FIGURE 4.1: Scatter plots of the final network states at $T = 1000$ in the adversarial swarming model. Circles correspond to the blue network, triangles correspond to the red network and the colour of the nodes correspond to the phase of the respective node $\mod 2\pi$. We consider the cases with $J_{BR} = J_{RB} = J$ with $J = 1$ and various values of K. The plots are ordered as follows: Top left $(J, K) = (1, -42)$, top right $(J, K) = (1, -1)$, bottom left $(J, K) = (1, 1)$ and bottom right $(J, K) = (1, 42)$.

a steady state is reached where the blue network clusters whilst the red network attempts to surround it however unlike the previous cases the blue network separates into two separate clusters with the red network focusing on surrounding the larger of the two clusters. When $(J, K) = (1, 1)$, the networks separate into two distinct opposing clusters and in the final plots with $(J, K) = (1, 42)$, the blue network attempts to surround the red network whilst the red network attempts to cluster.

We attempt to summarise network cohesion as a function of the coupling constant K on Figure 4.3 where we plot the order parameters for the blue and red networks as given by Eq. (4.7) and Eq. (4.8) respectively and the order parameter for the full system against the scaling parameter K. On this plot we look at the case of $J_{BR} = J_{RB} = 0.1J$ and will continue with this parameterisation from here onward. From these plots we observe that for positive values of K, the blue and red networks have similar order parameters however at a threshold at roughly about $K = 12$, the blue network order parameter is much smaller than the of the red network, corresponding the situation where at this threshold the blue network attempts to surround the red network whilst the red network remains clustered. A similar scenario occurs for negative K with the roles reversed, however in this case the order parameter for the blue network is always greater than that of the red network. Additionally for negative K the magnitude of the order parameter is weaker than that of the respective positive K, whoever this is expected as a negative K promotes phase repulsion between oscillators of similar phase.

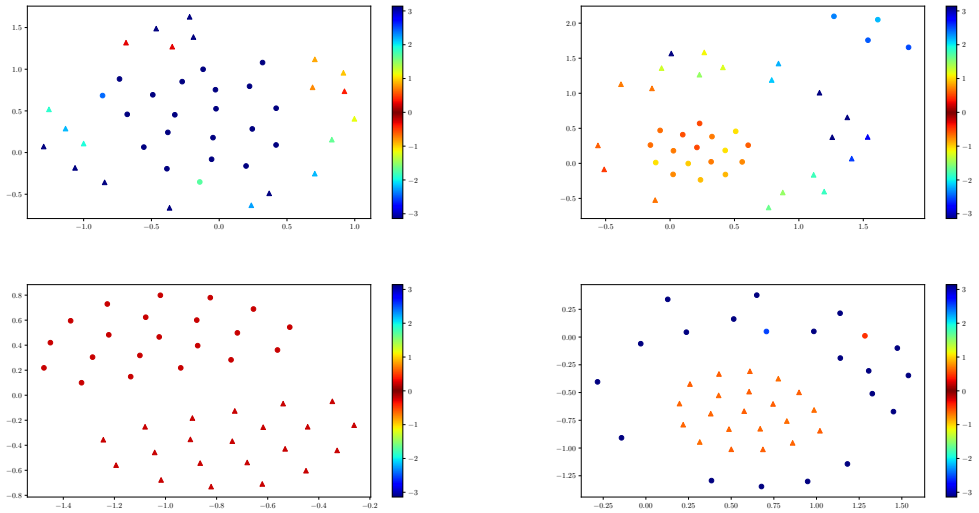


FIGURE 4.2: Scatter plots of the final network states at $T = 1000$ in the adversarial swarming model. Circles correspond to the blue network, triangles correspond to the red network and the colour of the nodes correspond to the phase of the respective node mod 2π. We consider the cases with $J_{BR} = J_{RB} = 0.1J$ with $J = 1$ and various values of K. The plots are ordered as follows: Top left $(J, K) = (1, -42)$, top right $(J, K) = (1, -1)$, bottom left $(J, K) = (1, 1)$ and bottom right $(J, K) = (1, 42)$.

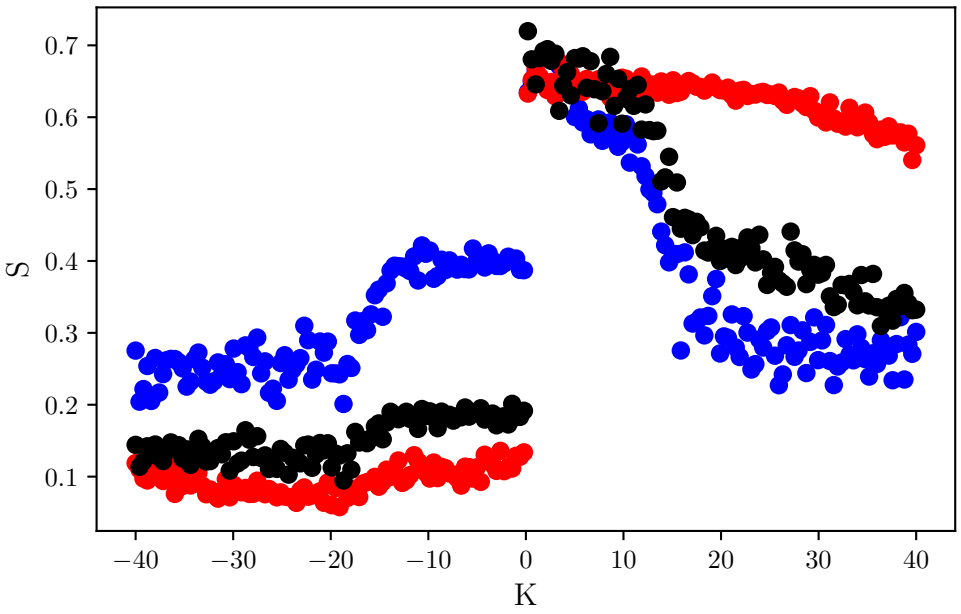


FIGURE 4.3: Order parameters for the blue and red networks as given in Eq. (4.7) and Eq. (4.8) respectively and the full system (black) for $J_{BR} = J_{RB} = 0.1J$ with $J = 1$.

4.3 Swarming adversarial models with attrition

4.3.1 The initial attrition model

The first of the two attrition models that we consider in this section constructs a process which governs the dynamics of the strength of each of the nodes in the Blue vs Red swarming model as defined in 4.4 and 4.5. We define the strength of each node of the blue and red networks as $\mu^{(B)}$ and $\mu^{(R)}$ respectively which all of the strengths are bounded above by $\mu_{\max}$ which we define as the maximum life of each node. When $\mu^{(B)} \leq 0$ or $\mu^{(R)} \leq 0$, that is the strength of a node is 0 or less, the individual node is removed from the network. The dynamics of the node strength are then governed by three terms, a resupply terms which we treat as a constant factor for each node, an attrition term whose strength growths when the relevant network is in a state of asynchronisation which is modelled through the variance of the phases and in addition to the variance, the magnitude of the attrition rate is affected by the relative difference in numbers between the two groups. The final term in each process is an additive, spectrally negative tempered stable Lévy noise which models the constant fire that each node undergoes in a combat scenario. The dynamics of this version of the Blue vs Red swarming model are thus governed by the following equations

$$\begin{aligned}\dot{\mu}_i^{(B)} &= \delta_i^{(B)} - \left(\frac{M}{N}\right)^{\nu}\left[\alpha \cdot \operatorname{var}\left(\tilde{\boldsymbol{\beta}} \bmod 2\pi\right) + \eta_i^{(B)}\right] + \dot{L}_i, \quad i \in \mathcal{B}, \\ \dot{\mu}_j^{(R)} &= \delta_j^{(R)} - \left(\frac{N}{M}\right)^{\nu}\left[\alpha \cdot \operatorname{var}\left(\tilde{\boldsymbol{\rho}} \bmod 2\pi\right) + \eta_j^{(R)}\right] + \dot{L}_j, \quad j \in \mathcal{R},\end{aligned} \tag{4.9}$$

where $\tilde{\boldsymbol{\beta}} \subseteq \boldsymbol{\beta}$,$\tilde{\boldsymbol{\rho}} \subseteq \boldsymbol{\rho}$ are the subsets of the blue and red networks respectively, α is a scaling factor for the impact of the variability on the drift, N and M are the current number of alive nodes in $\mathcal{B}$ and $\mathcal{R}$ respectively, ν is the magnitude of the numbers advantage effect, $\delta^{(B)}, \delta^{(R)}$ are the positive drift components corresponding to the resupply rates of the respective nodes, $\eta^{(B)}, \eta^{(R)}$ are the attrition rates and L is a spectrally negative tempered stable Lévy process. We note that we need to include mod 2π in the variance operator as the phases of the nodes are equivalent mod 2π.

4.3.2 Numerical investigation of the initial attrition model

With the attrition model established in Eq. (4.9), we now ask the question: "What is the optimal strategy for the blue network across the (ϕ, ψ) parameter space?". To define what we mean by an optimal strategy for the blue network across the pair of frustration parameters, we say that the blue network has an optimal strategy if it maximises its expected win rate over a set simulated of war games whilst minimising its expected

losses and, in the event that the blue network losses, it maximises the expected losses of the red network. To mathematically define this optimal strategy, in a similar manner as our work in Section 3.5, we define the objective function:

$$\begin{aligned} U_{(\text{Attr})} = \mathbb{E}(\text{blue win rate}) * 100 + \mathbb{E}(\#\text{ of surviving blue nodes}) \\ - \mathbb{E}(\#\text{ of surviving red nodes}). \end{aligned} \tag{4.10}$$

With the objective function defined as in Eq. (4.10) and the fact that in our example we set $N = M = 21$, it follows that the range is given by $U_{(\text{Attr})} \in [-21, 121]$, thus the best possible strategy in our case has a score of 121. More generally, this range is given by $U_{(\text{Attr})} \in [-M, 100 + N]$. To optimise this objective over the frustration parameters, we employ a Bayesian optimisation algorithm as discussed in Appendix A.1 and used in the context of the Blue vs Red model from Section 3.5. Using this Bayesian optimisation algorithm across the $(\phi, \psi) \in [-2\pi, 2\pi]$ parameter space with the objective function $U_{(\text{Attr})}$ from Eq. (4.10), we recreate the fitness landscape for both the cases of $(J, K) = (-1, 1)$ and $(J, K) = (1, 1)$ with $J_{BR} = J_{RB} = 0.1J$. We set $\mu_{max} = 100$, $\alpha = 1$, $\nu = 1$, $\eta_{i,j} = 1$ for all i, j and sample the resupply rates for both networks from $\delta \sim \text{U}(0.1, 0.2)$. The spectrally negative Lévy noise is taken from the tempered stable distribution with $\alpha = 1.5$, $\lambda = 0.5$ and $\sigma = 0.1$ with the usual parameterisation used throughout Chapter 3. For the initial conditions of the two networks, we set $(x, y)^{(B)} \sim \text{U}(1, 2)$ and $(x, y)^{(R)} \sim \text{U}(-1, -2)$. We then allow the system to run up till a time $T = 200$ with the attrition terms fixed, that is $\dot{\mu}^{(B,R)} = 0$ for $t \in [0, T]$, so that the system is first allowed to configure into its steady state formulation as presented in the plots on Figure 4.3 in the case of $(\phi, \psi) = (0, 0)$.

The results of our Bayesian optimisation routine is provided on Figure 4.4. We begin out analysis of these plots by first noting the score ranges across the fitness parameter landscape for the two given cases. In the case where $K = 1$, the range of the objective function is approximately given by $U_{(\text{Attr})}^{(K=1)} \in (4, 86)$ with an approximate mean of 59 across all values whilst in the case of $K = -1$, the range is given by $U_{(\text{Attr})}^{(K=1)} \in (61, 97)$ with an approximate mean of 83. Thus the blue network generally performs better in the case of $K = -1$. This is not a surprising result as since the variance drives the attrition of the model and given that we established the result from Figure 4.3 in the case of $(\phi, \psi) = (0, 0)$ we have that $U_{(\text{Attr})}^{(K=-1)} \approx 71$ and this follows from the result that the magnitude of the order parameter is greater in the blue network which directly governs the variance of the network. In the case of $K = 1$, at the point $(\phi, \psi) = (0, 0)$, we have that $U_{(\text{Attr})}^{(K=1)} \approx 45$ which agrees with the magnitudes of the order parameters of the blue and red networks being similar in Figure 4.3. Similarly to the fitness landscape constructed in the model from Section 3, there exists periodic bands of optimal strategies

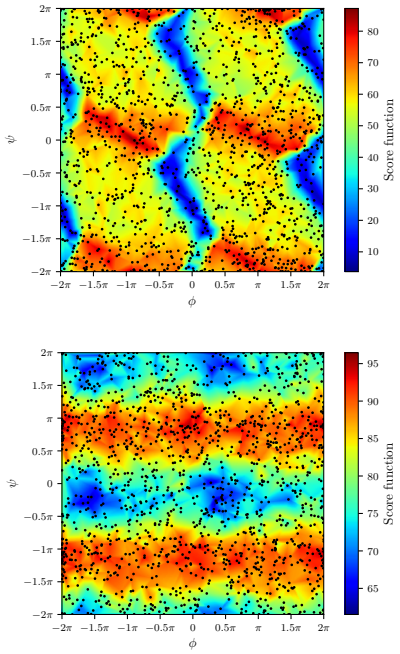


FIGURE 4.4: Bayesian optimisation of the objective/scoring function $U_{(\text{Attr})}$ over the across the $(\phi, \psi) \in [-2\pi, 2\pi]$. Top: $K = 1$. Bottom : K = -1.

for the blue network in the (ϕ, ψ) parameters space as opposed to a singular point. In the case of $K = 1$, there two separate periodic, with periodicity 2π, and diagonal bands of optimality between the ranges of $\phi \in [-2\pi, 0]$ and $\phi \in [0, 2\pi]$ whereas in the case of $K = -1$, there is a single but wider and non-diagonal periodic band of optimality with a periodicity of π. We conclude by noting that the relevant magnitude of the order parameters in the swarming model from Eq. (4.4) and Eq. (4.5) has the potential to be a good predictor for success in the wargame governed by the attrition model in Eq. (4.9).

4.3.3 A swarming model with Lanchester-type attrition

We conclude our discussion on adversarial models with a proposed extension of the Blue vs Red Lanchester attrition model from Eq. (4.2) to include spatial dynamics as described in the swarming model from Eq. (4.4) and incorporate the strength dynamics from the Lanchester equations individually for each node as opposed to the whole network in a similar manner as done in our initial model in Eq. (4.9).

In this proposed model, we implement the idea from Blue vs Red Lanchester attrition model from Eq. (4.2) which models the coupling constant as its own dynamical process however in our implementation, the dynamics of the coupling constant is unique to each oscillator and will be governed by a process inspired by the Lanchester equations. The implication of formulating the coupling dynamics in such a manner is that the strength

of the node directly impacts the ability of a node to sychronisation its phase within it's network and stay ahead of the opposing network in respect to its respective frustration parameter. Additionally to the phase, the strength of the node will also directly impact the ability of the node to move in either the direction of nodes with similar phases in its own network or to move to take advantage of the appropriate difference in phases in itself and opposing nodes.

We thus model the dynamics of the spatial movements and phases of the blue network through the following system of equations

$$\begin{aligned}
\dot{\boldsymbol{x}}_i^{(B)} =& \boldsymbol{v}_i + \frac{B_i(t)}{B_i(0)}\left(\frac{1}{N}\sum_{j\in B}\left[\frac{\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(B)}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(B)}|}(A_B + J_B\cos(\beta_i-\beta_j)) - B_B\frac{\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(B)}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(B)}|^2}\right],\right.\\
&\left.+\frac{1}{M}\sum_{j\in R}\left[\frac{\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(B)}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(B)}|}(A_{BR} + J_{BR}\cos(\beta_i-\rho_j-\phi)) - B_{BR}\frac{\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(B)}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(B)}|^2}\right]\right),\\
\dot{\beta}_i =& \omega_i - \frac{B_i(t)}{B_i(0)}\left(\sigma_B\sum_{j\in\mathcal{B}}\frac{\mathcal{B}_{ij}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(B)}|}\sin(\beta_i-\beta_j)\right.\\
&\left.+\ \xi_{BR}\sum_{j\in R}\frac{\mathcal{A}_{ij}^{(BR)}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(B)}|}\sin(\beta_i-\rho_j-\phi)\right),
\end{aligned} \tag{4.11}$$

where $i\in\mathcal{B}$ and the dynamics of the red network are governed by the system of equations

$$\begin{aligned}
\dot{\boldsymbol{x}}_i^{(R)} =& \boldsymbol{v}_i + \frac{R_i(t)}{R_i(0)}\left(\frac{1}{M}\sum_{j\in R}\left[\frac{\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(R)}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(R)}|}(A_R + J_R\cos(\rho_i-\rho_j)) - B_R\frac{\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(R)}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(R)}|^2}\right],\right.\\
&\left.+\frac{1}{N}\sum_{j\in B}\left[\frac{\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(R)}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(R)}|}(A_{RB} + J_{RB}\cos(\rho_i-\beta_j-\psi)) - B_{RB}\frac{\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(R)}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(R)}|^2}\right]\right),\\
\dot{\rho}_i =& \omega_i - \frac{R_i(t)}{R_i(0)}\left(\sigma_R\sum_{j\in\mathcal{R}}\frac{\mathcal{R}_{ij}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(R)}|}\sin(\rho_i-\rho_j)\right.\\
&\left.+\xi_{RB}\sum_{j\in B}\frac{\mathcal{A}_{ij}^{(RB)}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(R)}|}\sin(\rho_i-\beta_j-\psi)\right),
\end{aligned} \tag{4.12}$$

where $i\in\mathcal{R}$. The spatial and phase dynamics in this model are thus similar to that of the swarming model as presented in Eq. (4.4) and Eq. (4.5) but with the addition of the strength of the individual nodes at a given time t, denoted as $B_i(t)$ and $R_i(t)$ for the blue and red networks respectively where $B_i(0)$ and $R_i(0)$ are the initial strengths of the respective networks. The dynamics of B and R are modelled as processes inspired

by the Lanchester equations and are given by the following system of equations

$$\begin{aligned}\dot{B}_i(t) =&\alpha_S O_{B_i}(t)\sum_{j\in\mathcal{B}}\frac{\mathcal{B}_{ij}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(B)}|}B_j(t)\\ &-\alpha_{DA}O_{R_i}(t)\sum_{j\in\mathcal{R}}\frac{\mathcal{A}_{ij}^{(BR)}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(B)}|}R_j(t)+\sum_{j\in\mathcal{R}}\frac{\dot{L}_i^{(j)}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(B)}|},\quad i\in\mathcal{B},\end{aligned}\tag{4.13}$$

$$\begin{aligned}\dot{R}_i(t) =&\beta_S O_{R_i}(t)\sum_{j\in\mathcal{R}}\frac{\mathcal{R}_{ij}}{|\boldsymbol{x}_j^{(R)}-\boldsymbol{x}_i^{(R)}|}R_j(t)\\ &-\beta_{DA}O_{B_i}(t)\sum_{j\in\mathcal{B}}\frac{\mathcal{A}_{ij}^{(RB)}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(R)}|}B_j(t)+\sum_{j\in\mathcal{B}}\frac{\dot{L}_i^{(j)}}{|\boldsymbol{x}_j^{(B)}-\boldsymbol{x}_i^{(R)}|},\quad j\in\mathcal{R}.\end{aligned}\tag{4.14}$$

where $\dot{L}_i^{(j)}$ are versions of the (white-noise) Lévy process $\dot{L}_i$. Similarly to our model in Eq. (4.9), the strengths are bounded above by $B_i^{(\max)}$ and $R_i^{(\max)}$ which we define as the maximum lives of the respective network node and when $B_i \leq 0$ or $R_i \leq 0$, that is the strength of a node is 0 or less, the individual node is removed from the network. We note that in this model, there doesn't exist supply conservation however this can be included through modification. In this system, the parameters α_S and β_S are the rate of resupply for the blue and red networks respectively, α_{DA} and β_{DA} are the rate of attrition for the blue and networks. The order parameters $O_{B_i}(t)$ and $O_{R_i}(t)$ are the standard Kuramoto order parameters for the phase and can be either local or global order parameters in the sense that

$$O_{B_i}(t)=\frac{1}{|B_i'|}\left|\sum_{j\in B_i'}e^{i\beta_j(t)}\right|,\quad O_{R_i}(t)=\frac{1}{|R_i'|}\left|\sum_{j\in R_i'}e^{i\rho_j(t)}\right|.$$

When we consider the case of the global order parameters we have that $B_i'=B$, $R_i'=R$, $|B_i'|=N$ and $|R_i'|=M$. In the case of a local order parameter we set $B_i'\subset B$, $R_i'\subset R$ which are the sets of nodes within some distance δ of the i'th node and $|B_i'|$ and $|R_i'|$ are the sizes of the respective local networks.

Similarly to the Lanchester equations, the first two terms in these processes correspond to the resupply and attrition rates of the system however our models seeks to incorporate spatial dynamics into these processes in the sense that the proximity of notes directly impacts their ability to resupply and attrit one another. That is, if a node is close to another node within its own network, it is able to effectively resupply the friendly node and similarly if it is close to nodes in the opposing network it is effectively able to attrit that rival node. The implementation of the spectrally negative tempered stable is done in a manner to model the phenomenon that each node is under constant fire by every other node in the opposing network and that the magnitude of this fire is inversely proportional to the distance between the competing nodes.

With the model laid out by the system of equations in Eq. (4.11), Eq. (4.12), Eq. (4.13) and Eq. (4.14), we have laid the basis for future work in developing a model that effectively captures a wargame scenario that captures decision making based upon the OODA loop in which agents are free to move about the battlefield and directly eliminate opposing forces from the battlefield through a Lévy driven process.

Chapter 5

Simulation of Conditioned Stochastic Processes

In this chapter we extend our discussion on the simulation of stochastic processes to the case of conditioned processes. We begin in Section 5.1 with exploring the simulation of bridge diffusion processes through an application of the Doob h-transform which results in an SDE that specifies the dynamics of the conditioned process, a simulation method that makes use of the time-reversibility of a diffusion process and a naive acceptance-rejection scheme. We then extend our work to the simulation of conditioned Lévy processes beginning in Section 5.2 where we present a new simulation technique based upon the WHMC algorithm for the simulation of Lévy processes conditioned to to stay positive. In Section 5.3 we explore the simulation of Lévy bridges. Our work on Lévy bridges explores a new technique based upon the WHMC algorithm and the time reversed diffusion bridge. We also derive the double saddle-point approximation of Lévy bridges for the β-family. Finally in Section 5.4 we investigate the simulation of Lévy processes conditioned on their extrema through a novel application of the WHMC algorithm.

5.1 Conditioned diffusion processes

5.1.1 The Doob h-transform and conditioned diffusion processes

We begin our investigation on the simulation of conditioned processes by looking at applying the Doob h-transform to diffusion processes. To start, we consider X_t to be a diffusion process given by the SDE in Eq. (2.5) which exists in the state space S and is

conditioned on the event $X_T \in A$ for some $A \subset S$.

The transition density of the unconditioned process X_t is given by $p(T, y|t, x)$ and defined in Eq. (2.8). We now define the h function to be the probability that the process is in the conditioned state at time T, that is

$$h(t, x) := \mathbb{P}(X_T \in A | X_t = x).$$

The conditioned transition density $\hat{p}(t + s, y|t, x)$ for some $s > 0$ can then be expressed in terms of the unconditioned transition density p and the h-function as

$$\begin{aligned}
\hat{p}(t + s, y|t, x) =& \mathbb{P}(X_{t+s} = y | X_t = x, X_T \in A), \\
=& \frac{\mathbb{P}(X_{t+s} = y, X_T \in A | X_t = x)}{\mathbb{P}(X_T \in A | X_t = x)}, \\
=& \mathbb{P}(X_{t+s} = y | X_t = x) \frac{\mathbb{P}(X_T \in A | X_{t+s} = y)}{\mathbb{P}(X_T \in A | X_t = x)}, \\
=& p(t + s, y|t, x) \frac{h(t + s, y)}{h(t, x)}.
\end{aligned}$$

In general the h function is never known, however there exist a small number of cases in which this can be useful. The following theorem allows for an approximation of the h-function to construct bridge diffusions.

Theorem 5.1 *[Doob's h-transform for diffusion processes] Let X_t be a 1-dimensional diffusion given by the SDE*

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t$$

on $[0, T]$ and conditioned on the event $X_T \in A$. Then the conditioned diffusion Z_t satisfies the SDE

$$\mathrm{d}Z_t = \left(\mu(Z_T) + \sigma(Z_t)^2 \frac{\partial_x h(t, Z_t)}{h(t, Z_t)} \right) \mathrm{d}t + \sigma(Z_t)\mathrm{d}W_t.$$

where $h(t, x) = \mathbb{P}(X_T \in A | X_t = x)$. Furthermore, if we take $A = \{y\}$ for some $y \in S$, then $h(t, x) \approx p(T, y|t, x)$.

Recall the definition of the generator from Eq. (2.6). For the conditioned diffusion at some time t, the generator is given by

$$\begin{aligned}\mathcal{G}^{(t)}f(x) &= \lim_{s\downarrow 0}\frac{\mathbb{E}\left[f(X_{t+s})\,|X_t = x, X_T \in A\right] - f(x)}{s},\\ &= \lim_{s\downarrow 0}\frac{\mathbb{E}\left[f(X_{t+s})\mathbf{1}_{\{X_T\in A|X_t=x\}}\,|X_t = x\right] - f(x)h(t,x)}{s\,h(t,x)},\\ &= \lim_{s\downarrow 0}\frac{\mathbb{E}\left[f(X_{t+s})h(t+s,X_{t+s})\,|X_t = x\right] - f(x)h(t,x)}{s\,h(t,x)}.\end{aligned}$$

We can then express the generator $\mathcal{G}^{(t)}$ in terms of the generator of the original diffusion $\mathcal{L} = \mu\partial_x + \frac{1}{2}\sigma^2\partial^2_{xx}$ as

$$\mathcal{G}^{(t)}f = \frac{(\partial_t + \mathcal{L})(hf)}{h}.$$

Since $(\partial_t + \mathcal{L})h = 0$, we can write

$$\begin{aligned}\mathcal{G}^{(t)}f &= \frac{1}{h}\left[f\partial_t h + h\partial_t f + h\mu\partial_x f + f\mu\partial_x h + \frac{1}{2}\sigma^2\left(f\partial_{xx}h + 2\partial_x f\partial_x h + h\partial_{xx}f\right)\right],\\ &= \left(\partial_t f + \mu\partial_x f + \frac{1}{2}\sigma^2\partial_{xx}f\right) + \frac{f}{h}\left(\partial_t h + \mathcal{L}h\right) + \sigma^2\frac{\partial_x h}{h}\partial_x f,\\ &= \mathcal{L}^{(t)}f + \sigma^2\frac{\partial_x h}{h}\partial_x f.\end{aligned}$$

Thus the dynamics of the conditioned diffusion are equivalent to that of the unconditioned diffusion with a modified version of the drift. This means that the conditioned diffusion, denoted as Z_t, follows the SDE

$$\mathrm{d}Z_t = \left(\mu(Z_T) + \sigma(Z_t)^2\frac{\partial_x h(t,Z_t)}{h(t,Z_t)}\right)\mathrm{d}t + \sigma(Z_t)\mathrm{d}W_t.$$

Note that if we take $A = \{y\}$, the probability $P(X_T = y\,|X_t = x)$ is identically equal to 0. However by considering the approximation

$$\mathbb{P}(X_T \in (y, y+\mathrm{d}y)\,|X_t = x) \approx p(T,y|t,x)\,\mathrm{d}y + O(\mathrm{d}y),$$

we can then condition on the event $X_T = y$ by taking $h(t,x)$ to be the transition density $p(T,y|t,x)$.

We now consider three applications of Theorem 5.1 in deriving the bridge dynamics for three different diffusion processes conditioned on $X_T = y$. Specifically we consider a Brownian motion with drift, an Ornstein-Uhlenbeck process, and a Cox-Ingersoll-Ross process. In each case we give the SDE for the unconditioned process denoted by X_t, the transition density $p(T,y|t,x)$ and the SDE for the bridge process Z_t derived from the application of $p(T,y|t,x)$ in Theorem 5.1.

- Brownian bridge with drift:

$$\begin{aligned} \mathrm{d}X_t =& \mu \mathrm{d}t + \sigma \mathrm{d}W_t, \\ p(T, y|t, x) \propto& \exp\left[-\frac{(x - y - \mu(T-t))^2}{2(T-t)\sigma^2}\right], \\ \mathrm{d}Z_t =& (\mu + \frac{y - X_t + \mu(T-t)}{T-t})\mathrm{d}t + \sigma \mathrm{d}W_t. \end{aligned}$$

- Ornstein-Uhlenbeck bridge:

$$\begin{aligned} \mathrm{d}X_t &= -\beta(X_t - \alpha)\mathrm{d}t + \sigma \mathrm{d}W_t, \\ p(T, y|t, x) &\propto \exp\left[-\frac{(y - \alpha - (x-\alpha)e^{-\beta(T-t)})^2}{\sigma^2\left(1 - e^{-\beta t}\right)/\beta}\right], \\ \mathrm{d}Z_t &= \Big(-\beta(X_t - \alpha) + 2\frac{[y - \alpha - (x-\alpha)\exp\left(-\beta\left(T-t\right)\right)]}{-\exp\left(-2\beta\left(T-t\right)\right)/\beta} \\ &\quad \times \exp\left(-\beta\left(T-t\right)\right)\Big)\mathrm{d}t + \sigma \mathrm{d}W_t. \end{aligned} \tag{5.2}$$

- Cox-Ingersoll-Ross bridge:

$$\begin{aligned} \mathrm{d}X_t &= -\theta(X_t - \alpha)\mathrm{d}t + \sigma\sqrt{X_t}\mathrm{d}W_t \quad \theta > 0, \alpha > 0, \sigma > 0, \\ p(T, y|t, x) &\propto \left(\frac{y}{x}\right)^{\frac{1}{2}\nu} \exp\left[\frac{\beta(x+y)}{\exp(\theta(T-t)) - 1}\right] I_\nu\left(\frac{\beta\sqrt{xy}}{\sinh\left(\frac{1}{2}\theta(T-t)\right)}\right), \\ \mathrm{d}Z_t &= \Bigg(-\theta(X_t - \alpha) - \frac{\sigma^2 y\nu}{X_t}\left(\frac{y}{X_t}\right)^{(\nu-1)} - \frac{\sigma^2 X_t \beta}{\exp(\theta(T-t) - 1)} \\ &\quad + \frac{\sigma^2 X_t}{4}\frac{\beta}{\sinh(\frac{1}{2}\theta t)}(X_t y)^{-\frac{1}{2}}\frac{I_{\nu-1}(\eta) + I_{\nu+1}(\eta)}{I_\nu(\eta)}\Bigg)\mathrm{d}t + \sigma \mathrm{d}W_t, \end{aligned}$$

  where $\beta = 2\theta\sigma^{-2}, \nu = \beta\alpha - 1$, I_ν is a modified Bessel function of the first kind with index ν and

$$\eta = \frac{\beta\sqrt{X_t y}}{\sinh(\frac{1}{2}\theta(T-t))}.$$

Since the SDEs for the bridge processes are effectively modifications of the unconditioned diffusion processes through an additional term in the drift, if we are interested in simulating $\mathbb{E}(f(Z_t))$ for some function $f(\cdot)$, we are able to take Algorithm 1 with the modified drift and simulate approximations of $\mathbb{E}(f(Z_t))$ through this MC scheme. On Figure 5.1 we look at applying the Euler-Maruyama method from Section 2.3.1 to simulated 5 paths of an Ornstein-Uhlenbeck bridge.

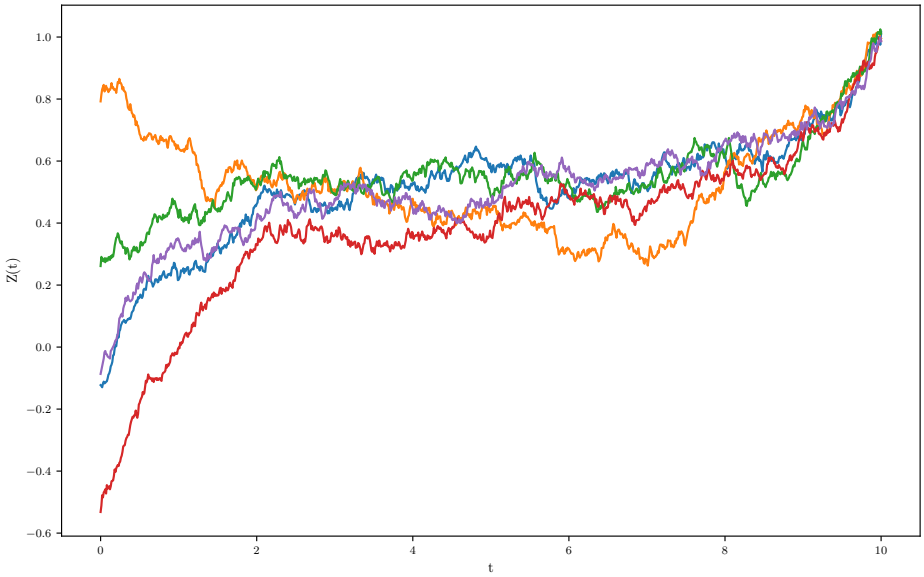


FIGURE 5.1: 5 sample paths of the Ornstein-Uhlenbeck bridge simulated according to the SDE Eq. (5.2) derived from the Doob h-transform with $\beta = 1$, $\alpha = 0.5$, $\sigma = 0.1$, $T = 10$ and $y = 1$.

5.1.2 Simulating diffusion bridges

We finish this section by extending our investigation on the simulation of diffusion bridges to two different methods, namely a time reversal bridge method and a naive acceptance-rejection sampling method. The first method makes use of the time reversibility property of one-dimensional diffusions with finite speed measure that was originally proposed by Bladt and Sørensen in [14]. This method of simulating bridges is simple to implement and whose construction plays a role in the proposed algorithms for simulating conditioned Lévy processes in a forthcoming section. We start by considering two diffusions given by

$$\begin{aligned} \mathrm{dX}_t^1 &= \mu(t, X_t^1)\mathrm{d}t + \sigma(t, X_t^1)\mathrm{d}W_t^1, \\ \mathrm{dX}_t^1 &= \mu(t, X_t^2)\mathrm{d}t + \sigma(t, X_t^2)\mathrm{d}W_t^2, \end{aligned}$$

where W_t^1 and W_t^2 are independent Wiener processes and the initial conditions are given by $X_0^1 = x$ and $X_0^2 = y$. The idea behind this method is then to simulate the process X^1 forward in time and X^2 backward in time and if the paths intersect, then they can be combined into a realisation of the bridge process. That is, given the simulated paths $\hat{X}_{n\Delta t}^{(1)}$ and $\hat{X}_{n\Delta t}^{(2)}$ with $n = 0, 1, \ldots N$ and $\Delta t = T/N$ such that the paths cross, the

process which approximates the diffusion bridge is defined by

$$Z_{n\Delta t} = \begin{cases} \hat{X}^1_{n\Delta t} & \text{for } n = 0, 1, \ldots, \nu - 1, \\ \hat{X}^2_{n\Delta t} & \text{for } n = \nu, \nu + 1, \ldots, N, \end{cases}$$

where

$$\nu = \begin{cases} \min\left\{n \in \{1, \ldots, N\} \,|\hat{X}^1_{n\Delta t} \leq \hat{X}^2_{(N-n)\Delta t}\right\} & \text{if } \hat{X}^1_0 \geq \hat{X}^2_{\Delta t}, \\ \min\left\{n \in \{1, \ldots, N\} \,|\hat{X}^1_{n\Delta t} \geq \hat{X}^2_{(N-n)\Delta t}\right\} & \text{if } \hat{X}^1_0 \leq \hat{X}^2_{\Delta t}. \end{cases}$$

The procedure for simulating diffusion bridges according to the time reversal method is given below in Algorithm 7

Algorithm 7 Time reversal simulation of diffusion bridges

Choose T, x, y, N
$\Delta t \leftarrow T/N$
repeat
    $X^1(0) \leftarrow x, X^2(0) \leftarrow y$
    for $j = 0, \ldots, N-1$ **do**
        sample $W^1, W^2 \sim \mathcal{N}(0, \Delta t)$
        $X^1(j+1) \leftarrow X^1(j) + \mu(\Delta t \cdot j, X(j)) \cdot \Delta t + \sigma(\Delta t \cdot j, X(j)) \cdot W^1$
        $X^2(j+1) \leftarrow X^2(j) + \mu(\Delta t \cdot j, X(j)) \cdot \Delta t + \sigma(\Delta t \cdot j, X(j)) \cdot W^2$
until $\nu \neq \emptyset$
$X^2 \leftarrow X^2[:: -1]$
$Z \leftarrow [X^1[0 : \nu - 1], X^2[\nu : N]]$
return Z

The application of this algorithm is only valid when the speed measure of the diffusion is finite, that is given the density of the speed measure defined by

$$m(x) = \frac{1}{\sigma^2(x)} \exp\left(2 \int_z^x \frac{\alpha(y)}{\sigma^2(y)} \mathrm{d}y\right),$$

for some arbitrary point z and x in the state space S, the following finiteness condition holds

$$M = \int_{x \in S} m(x) \mathrm{d}x < \infty.$$

For more information about the speed measure of a diffusion see for e.g., [112]. The following theorem from Bladt and Sørensen [14] then proves that this method approximates the bridge diffusion and is summarised below in Theorem 5.2.

Theorem 5.2 *[Time-reversible approximation for diffusion processes.]*
Let $\tau = \inf\{0 \leq t \leq T | X_t^1 = X_{T-t}^2\}$ *with* $\inf \emptyset = +\infty$ *and define*

$$Z_t = \begin{cases} X_t^1 & \text{if } 0 \leq t \leq \tau, \\ X_{T-t}^2 & \text{if } \tau \leq t \leq T. \end{cases}$$

Now assume that the density of the speed measure is finite, $M < \infty$*. Then the distribution of* $\{Z_t\}_{0 \leq t \leq T}$ *conditioned on* $\{\tau \leq T\}$ *is equal to that of a diffusion bridge.*

We now present a naive acceptance-rejection sampling method for the bridge process $\mathbb{P}(X_T \in y | X_t = x) = 0$. The practical implementation of this proposed algorithm requires that the bridge be conditioned on an interval as opposed to a point, that is we simulated the diffusion X_t conditioned on the the event $X_T \in [y - \epsilon, y + \epsilon]$ for some $\epsilon > 0$. The idea behind the algorithm is to simulate paths of the unconditioned process and reject all paths that fail to meet the conditioning criteria. Though this procedure can be generalised to any type of conditioning, we present the following version for the case of a diffusion bridge below:

Algorithm 8 Naive acceptance rejection sampling of diffusion bridges.

1: Choose $\epsilon > 0, T, x, y, N$
2: $\Delta t \leftarrow T/N$
3: **repeat**
4: $\quad X(0) \leftarrow x$
5: $\quad$ **for** $j = 0, \ldots, N-1$ **do**
6: $\quad\quad$ sample $W \sim \mathcal{N}(0, \Delta t)$
7: $\quad\quad X(j+1) \leftarrow X(j) + \mu(\Delta t \cdot j, X(j)) \cdot \Delta t + \sigma(\Delta t \cdot j, X(j)) \cdot W$
8: **until** $|X(N) - y| < \epsilon$
9: **return** X

The smaller the value of ϵ, the more accurate the algorithm becomes in approximating the diffusion bridge in the sense the Kolmogorov-Smirnov distance, D_{KS}, is minimised that is $\lim_{\epsilon, \Delta T \to 0} D_{KS}(\mathcal{F}(\hat{Z}_t(\epsilon, \Delta T)), \mathcal{F}(Z_t)) = 0$ where $\mathcal{F}(\hat{Z}_t(\epsilon, \Delta T))$ and $\mathcal{F}(Z_t)$ are the laws of the approximation and actual bridge process. However, a small ϵ adversely effects the acceptance rate of the algorithm, $R(\cdot)$, since $R(\epsilon) = 1/\mathbb{P}(X_T \in [y - \epsilon, y + \epsilon] | X_t = x)$. Thus practical simulation of Algorithm 8 requires a trade off between accuracy and computational time. This trade off between the acceptance rate R and the Kolmogorov-Smirnov distance D_{KS} will be a commonly recurring theme in the sections to come. In general, the run time of this algorithm will be poor however due to its simple 'brute-force' approach to simulating a conditioned process for any type of conditioning or process, it is a useful tool to analyse the plausibility of any newly proposed simulation technique.

We close off this section with a numerical comparison between the naive acceptance-rejection sampling in Algorithm 8, the Doob h-transform driven bridge from Theorem 5.1 and the time reversal bridge from Algorithm 7. Fixing $t = 0$, $T = 10$, $x = 0$, $y = 1$, we compare the distributions of the various bridges near the time of the terminal value and the expected value of the process across the time interval. That is, for each of the three methods, we present a comparison of the distributions of Z_9 and the expected value $E[Z_t]$ obtained from a MC simulation with 10000 replications. The total time taken for each of these methods is also record with the time-reversal method found to be the fastest of the three. Results are shown on Figure 5.2.

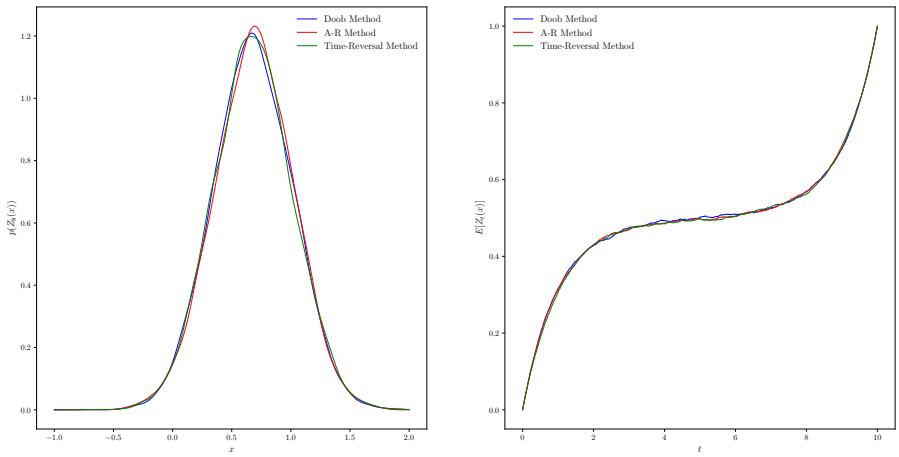


FIGURE 5.2: Numerical comparison of the density of $Z_9(x)$ (Left) and the $\mathbb{E}[Z_t(x)]$ (Right) using the Doob h-transform method, the time-reversal method and the acceptance-rejection method for the simulation of an Ornstein-Uhlenbeck bridge with $\beta = 1$, $\alpha = 0.5$, $\sigma = 0.5$, $T = 10$ and $y = 1$. The distribution of $Z_5(x)$ for each case is plotted using a standard kernel density estimation procedure. The total times for each of the methods in order of speed are: Time-reversal method (41 seconds), Doob h-transform method (73 seconds), acceptance-rejection method (2391 seconds).

5.2 Lévy processes conditioned to stay positive

We begin our discussion on the simulation of conditioned Lévy processes by considering a Lévy processed conditioned to stay positive. We denote the measure of the conditioned process starting at 0 by $\mathbb{P}^{\uparrow}$, where the transition densities of the conditioned processes can be define through the Doob h-transform as discussed in Section 5.1.1. We introduce the following notation to describe the conditioned process in terms of the unconditioned process L_t:

$$L_t^{\uparrow} := \left\{ L_t \middle| L_t \geq 0, t \geq 0 \right\}.$$

In this section we shall present an acceptance-rejection type simulation scheme based upon the WHMC approach proposed in [5] that was presented in Section 2.6. For more information the general theory on Lévy processes conditioned to stay positive see for example, [16, 17, 113].

5.2.1 Simulating processes conditioned to stay positive

Our algorithm for simulating Lévy process conditioned to stay positive is based upon an adaptation of the WHMC method from Algorithm 6. Similarly to the WHMC case, this method is based upon the construction of a 'stochastic grid' upon which the joint law of $(L_t^{\uparrow}, \overline{L}_t^{\uparrow})$ is approximated. That is we suppose that $\mathbf{e}_1, \mathbf{e}_2, \ldots$ is an i.i.d sequence of exponentially distributed random variables on a common product space with law $\mathbb{Q}$ orthogonal to the space $(\Omega, \mathcal{F}, \mathbb{P}^{\uparrow})$ then $\mathbf{g}(n,\lambda) = \sum_{i=1}^{n} \mathbf{e}_i/\lambda$ is equal in law to a Gamma random variable. By the strong law of large numbers, as $n \to \infty$, $\mathbf{g}(n, n/t) \to t$, $\mathbb{Q}$-almost surely. Using this stochastic grid for a sufficiently large n, the law $\mathbb{P}^{\uparrow}\left(L_t^{\uparrow} \in \mathrm{d}x, \overline{L}_t^{\uparrow} \in \mathrm{d}y\right)$ can be approximated by $\mathbb{P}^{\uparrow} \times \mathbb{Q}\left(L_{\mathbf{g}}^{\uparrow} \in \mathrm{d}x, \overline{L}_{\mathbf{g}}^{\uparrow} \in \mathrm{d}y\right)$.

Given the Wiener-Hopf factors of the process L_t, the WHMC method produces the approximation $\left(L_{\mathbf{g}(n,n/t)}, \overline{L}_{\mathbf{g}(n,n/t)}\right) \stackrel{d}{=} (V(n,n/t), J(n,n/t))$. From this approximation, we now present an acceptance-rejection adaption of the WHMC simulation technique for Lévy processes conditioned to stay positive given in Theorem 5.3 below.

Theorem 5.3 *[WHMC for Lévy processes conditioned to stay positive.] For all $n \in \mathbb{N}$, define $\boldsymbol{g}(n,\lambda) := \sum_{i=1}^{n} \boldsymbol{e}_i/\lambda$ and $\lambda = n/t$. Then*

$$\left(L^{\uparrow}_{\boldsymbol{g}(n,\lambda)}, \overline{L}^{\uparrow}_{\boldsymbol{g}(n,\lambda)}\right) \stackrel{d}{=} \left(V^{\uparrow}(n,\lambda), J^{\uparrow}(n,\lambda)\right),$$

where $V^{\uparrow}(n,\lambda)$ and $J^{\uparrow}(n,\lambda)$ are defined iteratively for $k \in \mathbb{N}$ as

$$\begin{aligned} V^{\uparrow}(k,\lambda) =& V^{\uparrow}(k-1,\lambda) + S_{\lambda}^{(k)\uparrow} + I_{\lambda}^{(k)\uparrow}, \\ J^{\uparrow}(k,\lambda) =& \bigvee_{j=1}^{k} \left(V^{\uparrow}(j-1,\lambda) + S_{\lambda}^{(j)\uparrow}\right), \end{aligned}$$

and $V^{\uparrow}(0,\lambda) = J^{\uparrow}(0,\lambda) = 0$. The i.i.d sequences of random variables $\left\{S_{\lambda}^{(j)\uparrow} : j \geq 1\right\}$ and $\left\{I_{\lambda}^{(j)\uparrow} : j \geq 1\right\}$ are drawn from the distributions of $\overline{L}_{\boldsymbol{e}_1/\lambda}$ and $\underline{L}_{\boldsymbol{e}_1/\lambda}$ respectively under the condition that $S_{\lambda}^{(j)\uparrow} + I_{\lambda}^{(j)\uparrow} > -V^{\uparrow}(j-1,\lambda)$ for all $j \in \mathbb{N}$.

To see that this holds in the conditioned version of the WHMC, we follow a similar line of reasoning as provided in [5]. Fixing $n \geq 1$ and letting $\mathcal{F}^{L}_{\mathbf{g}(n-1,\lambda)}$ be the filtration

generated by the process L_t up until the time $\mathbf{g}(n-1,\lambda)$, we define $\overline{L}_{s,t} = \sup_{s\le u\le t} L_u$ and then note that

$$\left(L_{\mathbf{g}(n,\lambda)}, \overline{L}_{\mathbf{g}(n,\lambda)}\right) = \left(L_{\mathbf{g}(n-1,\lambda)} + \left(L_{\mathbf{g}(n,\lambda)} - L_{\mathbf{g}(n-1,\lambda)}\right), \overline{L}_{\mathbf{g}(n-1,\lambda)} \vee \overline{L}_{\mathbf{g}(n-1,\lambda),\mathbf{g}(n,\lambda)}\right),$$

where $\mathbf{g}(0,\lambda) = 0$. By defining $L_t^{(n)} = L_{\mathbf{g}(n-1,\lambda)+t} - L_{\mathbf{g}(n-1,\lambda)}$ and $\overline{L}^{(n)}_{\mathbf{e}_n/\lambda} = \sup_{s\le \mathbf{e}_n/\lambda} L_s^{(n)}$, it then follows that

$$\left(L_{\mathbf{g}(n,\lambda)}, \overline{L}_{\mathbf{g}(n,\lambda)}\right) = \left(L_{\mathbf{g}(n-1,\lambda)} + L^{(n)}_{\mathbf{e}_n/\lambda}, \overline{L}_{\mathbf{g}(n-1,\lambda)} \vee \overline{L}_{\mathbf{g}(n-1,\lambda)} + \overline{L}^{(n)}_{\mathbf{e}_n/\lambda}\right).$$

Now consider conditioning the process to remain positive at each n, then from 5.2.1 at each iteration we have that

$$\begin{aligned}\left(L^{\uparrow}_{\mathbf{g}(n,\lambda)}, \overline{L}^{\uparrow}_{\mathbf{g}(n,\lambda)}\right) &= \left\{\left(L_{\mathbf{g}(n,\lambda)}, \overline{L}_{\mathbf{g}(n,\lambda)}\right) \middle| L_{\mathbf{g}(n,\lambda)} \ge 0, \mathcal{F}^L_{g(n-1,\lambda)}\right\} \\ &= \left\{\left(L_{\mathbf{g}(n-1,\lambda)} + L^{(n)}_{\mathbf{e}_n/\lambda}, \overline{L}_{\mathbf{g}(n-1,\lambda)} \vee \left(\overline{L}_{\mathbf{g}(n-1,\lambda)} + \overline{L}^{(n)}_{\mathbf{e}_n/\lambda}\right)\right)\right. \\ &\qquad \left. \middle| L_{\mathbf{g}(n,\lambda)} \ge 0, \mathcal{F}^L_{g(n-1,\lambda)}\right\}.\end{aligned}$$

Since the process $L^{(n)}$ is independent of $\{L_s : s \le \mathbf{g}(n-1,\lambda)\}$, has law $\mathbb{P}$, and noting that the Wiener-Hopf factorisation tells us that $\overline{L}_{\mathbf{e}_1/\lambda}$ and $L_{\mathbf{e}_n/\lambda} - \overline{L}_{\mathbf{e}_1/\lambda}$ are independent and the latter is equal in distribution to $\underline{\mathrm{L}}_{\mathbf{e}_1/\lambda}$, it follows that

$$\begin{aligned}\left(L^{\uparrow}_{\mathbf{g}(n,\lambda)}, \overline{L}^{\uparrow}_{\mathbf{g}(n,\lambda)}\right) = \Big\{&\Big(L_{\mathbf{g}(n-1,\lambda)} + S^{(n)}_{\lambda} + I^{(n)}_{\lambda}\ , \\ &\overline{L}_{\mathbf{g}(n-1,\lambda)} \vee \left(\overline{L}_{\mathbf{g}(n-1,\lambda)} + S^{(n)}_{\lambda}\right)\Big) \Big| L_{\mathbf{g}(n,\lambda)} \ge 0, \mathcal{F}^L_{g(n-1,\lambda)}\Big\},\end{aligned}$$

where $S^{(n)}_{\lambda}$ and $I^{(n)}_{\lambda}$ are random variables with distributions equal to that of $\overline{L}_{\mathbf{e}_1/\lambda}$ and $\underline{\mathrm{L}}_{\mathbf{e}_1/\lambda}$ for all n respectively.
Now noting that $L_{\mathbf{g}(n,\lambda)} \overset{d}{=} L_{\mathbf{e}_1/\lambda} + L_{g(n-1,\lambda)}$, we have that

$$\begin{aligned}\left(L^{\uparrow}_{\mathbf{g}(n,\lambda)}, \overline{L}^{\uparrow}_{\mathbf{g}(n,\lambda)}\right) = \Big(&L_{\mathbf{g}(n-1,\lambda)} + \left\{\left(S^{(n)}_{\lambda} + I^{(n)}_{\lambda}\right) \middle| L_{\mathbf{e}_1/\lambda} \ge -L_{g(n-1,\lambda)}\right\}, \\ &\overline{L}_{\mathbf{g}(n-1,\lambda)} \vee \left(\overline{L}_{\mathbf{g}(n-1,\lambda)} + \left\{S^{(n)}_{\lambda} \middle| L_{\mathbf{e}_1/\lambda} \ge -L_{g(n-1,\lambda)}\right\}\right)\Big).\end{aligned}$$

Since $L_{g(n-1,\lambda)} \overset{d}{=} V(n-1,\lambda)$ and $L_{\mathbf{e}_1/\lambda} \overset{d}{=} S^{(n)}_{\lambda} + I^{(n)}_{\lambda}$, if at each iteration we choose increments $S^{(n)}_{\lambda} + I^{(n)}_{\lambda}$ from $\overline{L}_{\mathbf{e}_1/\lambda} + \underline{\mathrm{L}}_{\mathbf{e}_1/\lambda}$ under the condition that $S^{(n)}_{\lambda} + I^{(n)}_{\lambda} \ge -V(n-1)$, then we have that

$$\left(L^{\uparrow}_{\mathbf{g}(n,\lambda)}, \overline{L}^{\uparrow}_{\mathbf{g}(n,\lambda)}\right) \overset{d}{=} \left(V^{\uparrow}(n,\lambda), J^{\uparrow}(n,\lambda)\right),$$

where $V^{\uparrow}(n,\lambda)$ and $J^{\uparrow}(n,\lambda)$ are defined as in the theorem above.

Our proposed acceptance-rejection algorithm to sample from the law of $(L_t^{\uparrow}, \overline{L}_t^{\uparrow})$ is given in Algorithm 9.

Algorithm 9 Sample from the law of $(L_t^{\uparrow}, \overline{L}_t^{\uparrow})$

Choose $n \geq 1$ so that $\mathbf{g} := \sum_{i=1}^{n} \frac{t}{n}\mathbf{e}_i \approx t$ by SLLN.
$\lambda \leftarrow n/t$
$L(0) \leftarrow 0$, $Z(0) \leftarrow 0$
for $i = 1, \ldots, n$ **do**
    repeat
        sample $S \sim \mathbf{Law}(\overline{L}_{\mathbf{e}_1/\lambda})$
        sample $I \sim \mathbf{Law}(\underline{L}_{\mathbf{e}_1/\lambda})$
        $L(i) \leftarrow L(i-1) + S + I$
    until $L(i) \geq 0$.
    $Z(i) \leftarrow \max(Z(i-1), L(i))$
return $(L(n), Z(n))$

For an integrable function F, standard Monte Carlo techniques then gives

$$\mathbb{E}\left[F(L_t^{\uparrow}, \overline{L}_t^{\uparrow})\right] \simeq \frac{1}{M}\sum_{m=1}^{M} F\big(L_{(m)}^{\uparrow}, \overline{L}_{(m)}^{\uparrow}\big),$$

where $\big(L_{(m)}^{\uparrow}, \overline{L}_{(m)}^{\uparrow}\big)_{1\leq m\leq M}$ are sampled using Algorithm 9 and $M > 0$ is a large integer. By the strong law of large numbers, the right-hand side converges almost surely to $\mathbf{E}\big[F(L_t^{\uparrow}, \overline{L}_t^{\uparrow})\big]$ as $n, M \to \infty$.

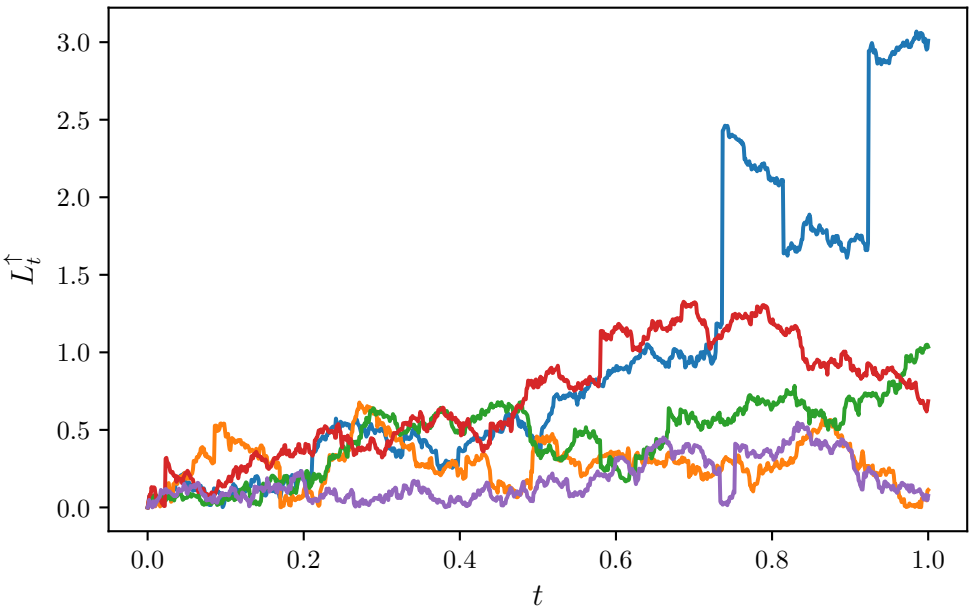


FIGURE 5.3: 5 sample paths of a β-family Lévy process conditioned to stay positive.

5.2.2 Path functionals of processes conditioned to stay positive

In this section we turn our attention to an extension of the acceptance-rejection WHMC method to path functionals of Lévy processes conditioned to stay positive. The results in this section are based upon the work done by Ferreiro-Castallia and van Schaik in [114], where we now consider the results in the context of a Lévy processes conditioned to stay positive. The path functionals that we are interested in are the first passage time of the process above a certain level $u \geq 0$ and the overshoot, undershoot and last maximum before this first passage time where the first passage time of the process $L_t^{\uparrow}$ over the level $u \geq 0$ is defined by

$$\tau_u := \inf\left\{t > 0 \middle| L_t^{\uparrow} > u\right\}.$$

and $L_{\tau_u}^{\uparrow} - u$, $u - L_{\tau_u^-}^{\uparrow}$ and $u - \overline{L}_{\tau_u^-}^{\uparrow}$ are the overshoot, undershoot and last maximum before first passage time respectively. Applications of these particular quantities are particularly relevant in the context of ruin process in insurance modelling, for further information, see for example [65, 115, 116].

We begin our discussion of path functionals by first noting we can construct a random walk approximation of the process $L_t^{\uparrow}$ through the WHMC method in the manner as discussed in Section 2.6.1. That is the WHMC algorithm produces an approximation of the vector $\left((L_0^{\uparrow}, \overline{L}_0^{\uparrow}), (L_{1/n}^{\uparrow}, \overline{L}_{1/n}^{\uparrow}), \ldots, (L_t^{\uparrow}, \overline{L}_t^{\uparrow})\right)$. Due to the sequence $\left\{L_{\mathbf{g}(k,n/t)}^{\uparrow}\right\}_{k\geq 0}$ being a suitable random walk approximation of the Lévy process $L_t^{\uparrow}$, we can evaluate the path functionals of the process according to the following theorem.

Theorem 5.4 *[Path functionals for Lévy processes conditioned to stay positive.]*
Let $L^{\uparrow}$ be any Lévy process conditioned to stay positive. Fix some $t > 0$, $u > 0$ and define $V^{\uparrow}$ and $J^{\uparrow}$ as in Theorem 5.3. Set for all $n \in \mathbb{N}$

$$\kappa_u^{(n)} := \inf\left\{k \in \{0, \ldots, n\} \,|J^{\uparrow}(k, \lambda) > u\right\}$$

Then we have as $n \to \infty$

$$\left(\frac{t}{n}(\kappa_u^{(n)} \wedge n), V^{\uparrow}(\kappa_u^{(n)} \wedge, \lambda) - u, u - V^{\uparrow}((\kappa_u^{(n)} - 1) \wedge n, \lambda), u - J^{\uparrow}((\kappa_u^{(n)} - 1) \wedge n, \lambda)\right)$$
$$\xrightarrow{d} \left(\tau_u \wedge t, L_{\tau_u \wedge t}^{\uparrow} - u, u - L_{(\tau_u \wedge t)^-}^{\uparrow}, u - \overline{L}_{(\tau_u \wedge t)^-}^{\uparrow}\right).$$

The proof of Theorem is provided in Appendix 5.4. As in Theorem 5.3, the application of Theorem 5.4 only requires the simulation of i.i.d copies of the distributions of $S_\lambda = S_\lambda^{(1)}$ and $I_\lambda = I_\lambda^{(1)}$ from which a functional transformation then gives about the required

realisation of the random variables $(\tau_u \wedge t, L^{\uparrow}_{\tau_u \wedge t} - u, u - L^{\uparrow}_{(\tau_u \wedge t)^-}, u - \overline{L}^{\uparrow}_{(\tau_u \wedge t)^-})$. The implementation of Theorem 5.4 is given below in Algorithm 10.

Algorithm 10 Sample from the law of $(\tau_u \wedge t, L^{\uparrow}_{\tau_u \wedge t} - u, u - L^{\uparrow}_{(\tau_u \wedge t)^-}, u - \overline{L}^{\uparrow}_{(\tau_u \wedge t)^-})$

```
Choose n ≥ 1 so that g := Σ_{i=1}^{n} (t/n) e_i ≈ t by SLLN.
λ ← n/t
X(0) ← 0, Z(0) ← 0
for i := 1, ..., n do
    repeat
        sample S ~ Law(L̄_{e_1/λ})
        sample I ~ Law(L_{e_1/λ})
        L(i) ← L(i − 1) + S + I
    until L(i) ≥ 0.
    Z(i) ← max(Z(i − 1), L(i))
    if Z(i) > u then
        tau ← t/n · i
        overshoot ← L(i) − u
        undershoot ← u − L(i)
        LM ← u − Z(i − 1)
        break
    if i = n then
        tau ← t
        overshoot ← L(n) − u
        undershoot ← u − L(n)
        LM ← u − Z(n − 1)
return (tau, overshoot, undershoot, LM)
```

5.2.3 Numerical Results

In this section we present a numerical analysis of the WHMC for Lévy processes conditioned to positive and path functionals of these processes. Similarly to the work done in [5], our computations are performed with the following parameter sets

$$(a, \sigma, \alpha_1, \beta_1, \lambda_1, c_1, \alpha_2, \beta_2, \lambda_2, c_2) = (a, 0.4, 1, 1.5, 1.5, 1, 1, 1.5, 1.5, 1).$$

where set 1 has a driftless process $a = 0$ and set 2 has a linear drift where a is chosen such that $\Psi(-i) = -r$ with $r = 0.05$ which amounts to the process $\{\exp(L_t - rt) : t \geq 0\}$ being a martingale. In the cases of both parameter sets, the resulting processes are of unbounded variation as $\sigma \neq 0$ holds and have infinite activity in the jump as the conditions $\lambda_1, \lambda_2 \in (1, 3)$ hold.

The simulations of the acceptance-rejection WHMC method are verified against that of a naive increment-by-increment Euler simulation. As discussed in Section 2.5, our Euler simulation is achieved through a Fast Fourier Transform of the characteristic function of

the β-family process whose characteristic exponent was given in Eq. (2.11). These naive simulations are run with a step size of $h = 0.01$ and we run a total number of $M = 1000$ simulations in this classic Monte Carlo approach. The results of these simulations are presented on Figures 5.4-5.6. We note that in all figures present, the resulting densities produced by naive increment-by-increment simulations and the WHMC simulation are generally in agreement with one-another. In the case of Figure 5.5, the presence of the bimodal curve is due to the event that $t > \tau_u$ occurring for some paths in these simulations.

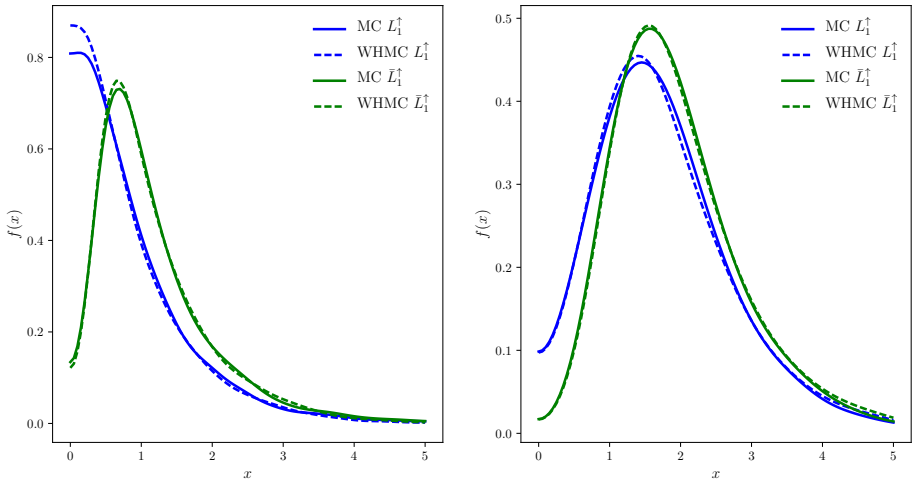


FIGURE 5.4: Comparisons of Algorithm 9 (WHMC) and a naive increment-by-increment simulation (MC) for the distributions of $L_1^{\uparrow}$ and $\bar{L}_1^{\uparrow}$ for parameter set 1 (left) and parameter set 2 (right).

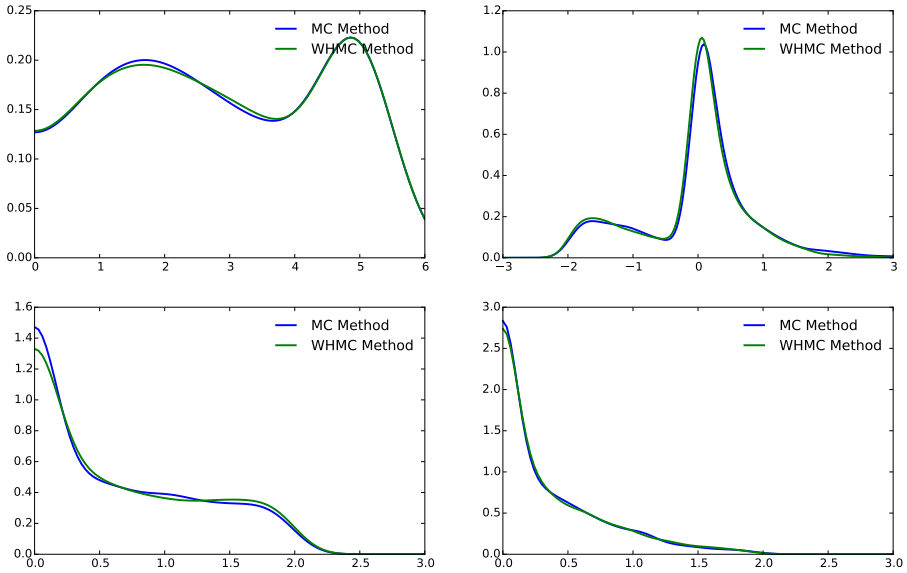


FIGURE 5.5: Comparisons of Algorithm 10 (WHMC) and a naive increment-by-increment simulation (MC) for the distributions of $\tau_u \wedge t$ (top left), $L^{\uparrow}_{\tau_u \wedge t} - u$ (top right), $u - L^{\uparrow}_{(\tau_u \wedge t)^-}$ (bottom left) and $u - \overline{L}^{\uparrow}_{(\tau_u \wedge t)^-}$ (bottom right) for parameter set 1 with $t = 5$ and $u = 2$.

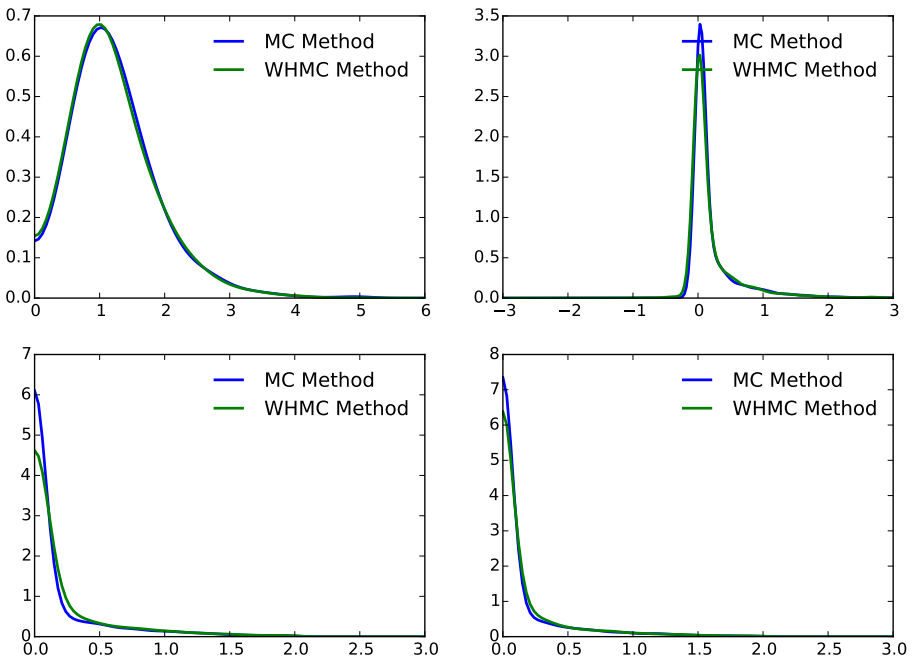


FIGURE 5.6: Comparisons of Algorithm 10 (WHMC) and a naive increment-by-increment simulation (MC) for the distributions of $\tau_u \wedge t$ (top left), $L^{\uparrow}_{\tau_u \wedge t} - u$ (top right), $u - L^{\uparrow}_{(\tau_u \wedge t)^-}$ (bottom left) and $u - \overline{L}^{\uparrow}_{(\tau_u \wedge t)^-}$ (bottom right) for parameter set 2 with $t = 5$ and $u = 2$.

5.3 Lévy bridges

In this section we investigate a variety of methods for the simulation of Lévy bridges, that is we are interested in simulating the Lévy process L_t starting from 0 and terminating at some value $b \in \mathbb{R}$. In this section, we shall describe the Lévy bridge in terms of the unconditioned process as

$$L_t^{Br_b} := \left\{ L_t \middle| L_T = b \right\}.$$

Similarly to 5.2, we shall first consider an acceptance-rejection type simulation scheme based upon the WHMC method that was presented in Section 2.6. For more information the general theory on Lévy bridges, see for example [117–119].

5.3.1 Conditioning on the terminal value

In this section we propose a variation of the WHMC method to simulate a Lévy bridge. This method is based upon the scheme from Bladt and Sørensen [14] as previously discussed in Section 5.1.2 for the case of diffusion bridges. The starting point of this method is to consider the two independent Lévy processes $L_t^{(1)}$ and $L_t^{(2)}$ with the initial conditions $L_0^{(1)} = 0$ and $L_0^{(2)} = b$ for some $b > 0$. We then simulate a random walk approximation of the process $L^{(1)}$ forward in time and $L^{(2)}$ backward in time through the WHMC algorithm and if the paths intersect, then they can be combined into a realisation of the bridge process $L_t^{Br_b}$. From this procedure, we have the following theorem for the simulation of Lévy bridges under the WHMC setting.

Theorem 5.5 *[WHMC for Lévy bridges.]*
Let L_t be a time reversible Lévy process. For all $n \in \mathbb{N}$, define $\boldsymbol{g}(n,\lambda) := \sum_{i=1}^{n} \boldsymbol{e}_i/\lambda$ and $\lambda = n/t$ and suppose that $\left\{(V^{(i)}(k,\lambda), J^{(i)}(k,\lambda))\right\}_{k\in[0,n]}$ approximates the random walks of $(L_t^{(i)}, \bar{L}_t^{(i)})$ for $i \in \{1,2\}$ where $L_0^{(1)} = 0$, $L_0^{(2)} = b$ for some $b > 0$. Assume that $k^b < \infty$ where

$$k^b := \inf\left\{ k \in \{0,\dots,n\} \,|L^{(1)}_{\mathbf{g}(k,\lambda)} \geq L^{(2)}_{\mathbf{g}(n-k,\lambda)} \right\}$$

Then

$$\left(L^{Br_b}_{\boldsymbol{g}(n,\lambda)}, \bar{L}^{Br_b}_{\boldsymbol{g}(n,\lambda)}\right) \stackrel{d}{=} \left(V^B(n,\lambda), J^B(n,\lambda)\right),$$

where for $k \in \mathbb{N}$,

$$\left(V^b(k,\lambda), J^b(k,\lambda)\right) = \begin{cases} \left(V^{(1)}(k,\lambda), J^{(1)}(k,\lambda)\right) & \text{if } 0 \leq k \leq k^b, \\ \left(V^{(2)}(n-k,\lambda), J^{(1)}(k^b,\lambda) \vee J^{(2)}_{k^b}(n-k,\lambda)\right) & \text{if } k^b < k \leq n. \end{cases}$$

The quantities $(V^{(1)}(k,\lambda), J^{(1)}(k,\lambda))$ *are constructed through Eq.* (2.16) *and Eq.* (2.17) *and for the reversed process we first backwards construct* $V^{(2)}(k,\lambda)$ *and then work forwards from the increments of* $\left\{V^{(2)}(k,\lambda)\right\}$ *to obtain* $J^{(2)}_{k^b}(k,\lambda)$*)*

$$V^{(2)}(n-(k+1),\lambda) = V^{(2)}(n-k,\lambda) - S^{(n-k)}_{\lambda} - I^{(n-k)}_{\lambda},$$

$$J^{(2)}_{k^b}(k,\lambda) = \begin{cases} 0 & \text{if } k \leq k^b, \\ \max\left(J^{(2)}_{k^b}(k-1,\lambda), V^{(2)}(k-1,\lambda) + S^{(k-1)}_{\lambda}\right) & \text{if } k > k^b. \end{cases}$$

for $k \geq 1$, $V^{(2)}(n,\lambda) = b$ *and the two i.i.d sequences of random variables* $\left\{S^{(j)}_{\lambda} : j \geq 1\right\}$ *and* $\left\{I^{(j)}_{\lambda} : j \geq 1\right\}$ *are distributed as* $\overline{L}_{\frac{t}{n}\mathbf{e}_1}$ *and* $\underline{L}_{\frac{t}{n}\mathbf{e}_1}$ *respectively.*

Let $L^{(3)}_t$ be a Lévy process whose triple is identical to $L^{(1)}_t$ and the distribution of $L^{(3)}_0$ has the density ν. Let ρ be the first time that the process $L^{(3)}_t$ hits the sample path of $L^{(1)}_t$ and define the process

$$L^{(4)}_t = \begin{cases} \begin{cases} L^{(1)}_t & \text{if } 0 \leq t \leq \rho, \\ L^{(3)}_t & \text{if } \rho < t \leq T \end{cases} & \text{if } \rho \leq T, \\ L^{(1)}_t & \text{otherwise.} \end{cases}$$

By the strong Markov property for Lévy processes, $L^{(4)}$ has the same distribution as $L^{(1)}$. Thus under the conditioning $L^{(3)} = b$ we have that

$$\mathbb{P}(L^{(4)}_t \in \cdot | L^{(3)}_T = b, \rho \leq T) = \mathbb{P}(L^{(4)}_t \in \cdot | L^{(4)}_T = b, \rho \leq T).$$

From the time reversibility of the process L_t, we have that $\left\{L^2_{T-t}\right\}_{0\leq t\leq T}$ is approximately equal in distribution to that of $\left\{L^{(3)}_t | L^{(3)}_T = b\right\}_{0\leq t\leq T}$, thus

$$\mathbb{P}(L^{(4)}_t \in \cdot | L^{(3)}_T = b, \rho \leq T) \approx \mathbb{P}(\hat{L}^{Br_b} = b \in \cdot | \rho \leq T).$$

where the bridge approximation $\hat{L}^{Br_b}_t$ is given by

$$\hat{L}^{Br_b}_t = \begin{cases} L^{(1)}_t & \text{if } 0 \leq t \leq \tau \\ L^{(2)}_{T-t} & \text{if } \tau < t \leq T \end{cases}$$

for some $\tau = \inf\left\{0 \leq t \leq T | L^{(1)}_t = L^{(2)}_{T-t}\right\}$.

Defining k^b as in the theorem above and assume that $\left\{k^b < \infty\right\}$ holds. Then in the limit as $n \to \infty$ by the SLLN and using the fact that the process L does not jump at

the fixed times $\mathbf{g}(k,\lambda)$ we have that

$$\begin{aligned} k^b &= \inf\left\{k \in \{0,\ldots,n\} \,|L^{(1)}_{\mathbf{g}(k,\lambda)} \geq L^{(2)}_{\mathbf{g}(n-k,\lambda)}\right\} \\ &\to \inf\left\{0 \leq t \leq T | L^{(1)}_t = L^{(2)}_{T-t}\right\} \quad \mathbb{Q}\text{-almost surely} \\ &= \tau. \end{aligned}$$

Then through the WHMC construction, we have that

$$(L^{(1)}_{\mathbf{g}(k,\lambda)}, \overline{L}^{(1)}_{\mathbf{g}(k,\lambda)}) \stackrel{d}{=} (V^{(1)}(k,\lambda), J^{(1)}(k,\lambda)) \text{ for } 0 \leq k \leq k^b \text{ and } 0 \leq t \leq \tau.$$

For $k > k^b$, by the definition of k^b we have that

$$\begin{aligned} (L^{(2)}_{\mathbf{g}(n-k^b,\lambda)}, \overline{L}^{(2)}_{\mathbf{g}(n-k^b,\lambda)}) &\stackrel{d}{=} (V^{(1)}(k^b,\lambda), J^{(1)}(k^b,\lambda)) \\ &= (V^{(2)}(n-k^b,\lambda), J^{(1)}(k^b,\lambda) \vee J^{(2)}_{k^b}(n-k^b,\lambda)), \end{aligned}$$

where the second equality follows by construction of $J^{(2)}_{k^b}$. Since V^2 is time reversed WHMC algorithm starting from b, we have that

$$L^2_{\mathbf{g}(n-k,\lambda)} \stackrel{d}{=} V^{(2)}(n-k) \text{ for } k^b \leq k \leq n.$$

Thus as $n \to \infty$, $L^b_t \to L^b_{\mathbf{g}(k,\lambda)} \stackrel{d}{=} V^b(k,\lambda)$ $\mathbb{Q}$-almost surely. For the supremum, we first note that we can write the approximation $\hat{\overline{L}}^b$ in terms of $L^{(1)}_t$ and $L^{(2)}_{T-t}$ as

$$\hat{\overline{L}}^{Br_b}_t = \begin{cases} \overline{L}^{(1)}_t & \text{if } 0 \leq t \leq \tau, \\ \overline{L}^{(1)}_\tau \vee \max_{\tau' \in (\tau,t]} L^{(2)}_{T-\tau'} & \text{if } \tau < t \leq T. \end{cases}$$

We have already shown that $\hat{\overline{L}}^b_t \to J^b(k,\lambda)$ for $0 \leq t \leq \tau$. For $t > \tau$, by the SLLN as $n \to \infty$

$$\begin{aligned} \overline{L}^{(1)}_t \vee \max_{\tau' \in (\tau,t]} L^{(2)}_{T-\tau'} &\to \overline{L}^{(1)}_{\mathbf{g}(k^b,\lambda)} \vee \max_{k' \in (k^b,k]} L^{(2)}_{\mathbf{g}(n-k^B,\lambda)} \quad \mathbb{Q}\text{-almost surely} \\ &\stackrel{d}{=} J^{(1)}(k^b,\lambda) \vee J^{(2)}_{k^b}(n-k,\lambda)), \end{aligned}$$

where the second equality follows by construction of $J^{(2)}_{k^b}$. Thus as $n \to \infty$, $\overline{L}^{Br_b}_t \to \overline{L}^{Br_b}_{\mathbf{g}(k,\lambda)} \stackrel{d}{=} J^b(k,\lambda)$ $\mathbb{Q}$-almost surely which completes the proof.

The procedure for simulating diffusion bridges according to the time reversal method is given in Algorithm 7.

Algorithm 11 Sample from the law of $(L_t^{Br_b}, \bar{L}_t^{Br_b})$

Choose $n \geq 1$ so that $\mathbf{g} := \sum_{i=1}^{n} \frac{t}{n}\mathbf{e}_i \approx t$ by SLLN.
$\lambda \leftarrow n/t$
repeat
    $L^1(0) \leftarrow 0$, $Z^1(0) \leftarrow 0$
    $L^2(n) \leftarrow B$, $Z^2(0) \leftarrow 0$
    for $i = 1, \ldots, n$ **do**
        sample $S^1, S^2(n-(i-1)) \sim \mathbf{Law}(\overline{L}_{\mathbf{e}_1/\lambda})$
        sample $I^1, I^2(n-(i-1)) \sim \mathbf{Law}(\underline{L}_{\mathbf{e}_1/\lambda})$
        $L^1(i) \leftarrow L^1(i-1) + S^1 + I^1$
        $Z^1(i) \leftarrow \max(Z^1(i-1), L^1(i) + S^1)$
        $L^2(n-i) \leftarrow L^2(n-(i-1)) - S^2(n-(i-1)) - I^2(n-(i-1))$
    $k^b \leftarrow \inf\left\{k | L^1(k) \geq L^2(k)\right\}$
until $k^b < \infty$.
$Z^2(k^b) \leftarrow Z^1(k^b)$
for $i = k^b + 1, \ldots, n$ **do**
    $Z^2(i) \leftarrow \max(Z^2(i-1), L^2(i-1) + S^2(i))$
return $([L^1(1:k^b), L^2(k^b+1:n)], [Z^1(1:k^b), Z^2(k^b+1:n)])$

We present a numerical simulation of the acceptance-rejection WHMC algorithm for Lévy bridges on Figure 5.7 below. Similarly as in the case for Lévy processes conditioned to stay positive in section 5.2.3, we use the β-family of Lévy processes with the parameterisation

$$(a, \sigma, \alpha_1, \beta_1, \lambda_1, c_1, \alpha_2, \beta_2, \lambda_2, c_2) = (0, 0.1, 1, 1.5, 1.5, 1, 1, 1.5, 1.5, 1).$$

For numerical verification we include the densities of the bridge process under the acceptance-rejection WHMC algorithm and a naive MC algorithm at the time $T = 5$. In the case of the MC algorithm, we are simulating the approximation to conditioned process $\{L_t | L_T \in [b-\epsilon, b+\epsilon]\}$ with $\epsilon = 0.1$. In this case, we note that the densities obtained from the MC and WHMC algorithms roughly align with one another.

5.3.2 Path functionals of Lévy bridges

Similar to the case of Lévy processes conditioned to stay positive, we can adapt the path functional results from [114] to the case of Lévy bridges. In this scenario, the first passage time of the process $L_t^{Br_b}$ over the level $u \geq 0$ is defined by

$$\tau_u := \inf\left\{t > 0 \middle| L_t^{Br_b} > u\right\}.$$

We then have the following theorem for path functionals of Lévy bridges.

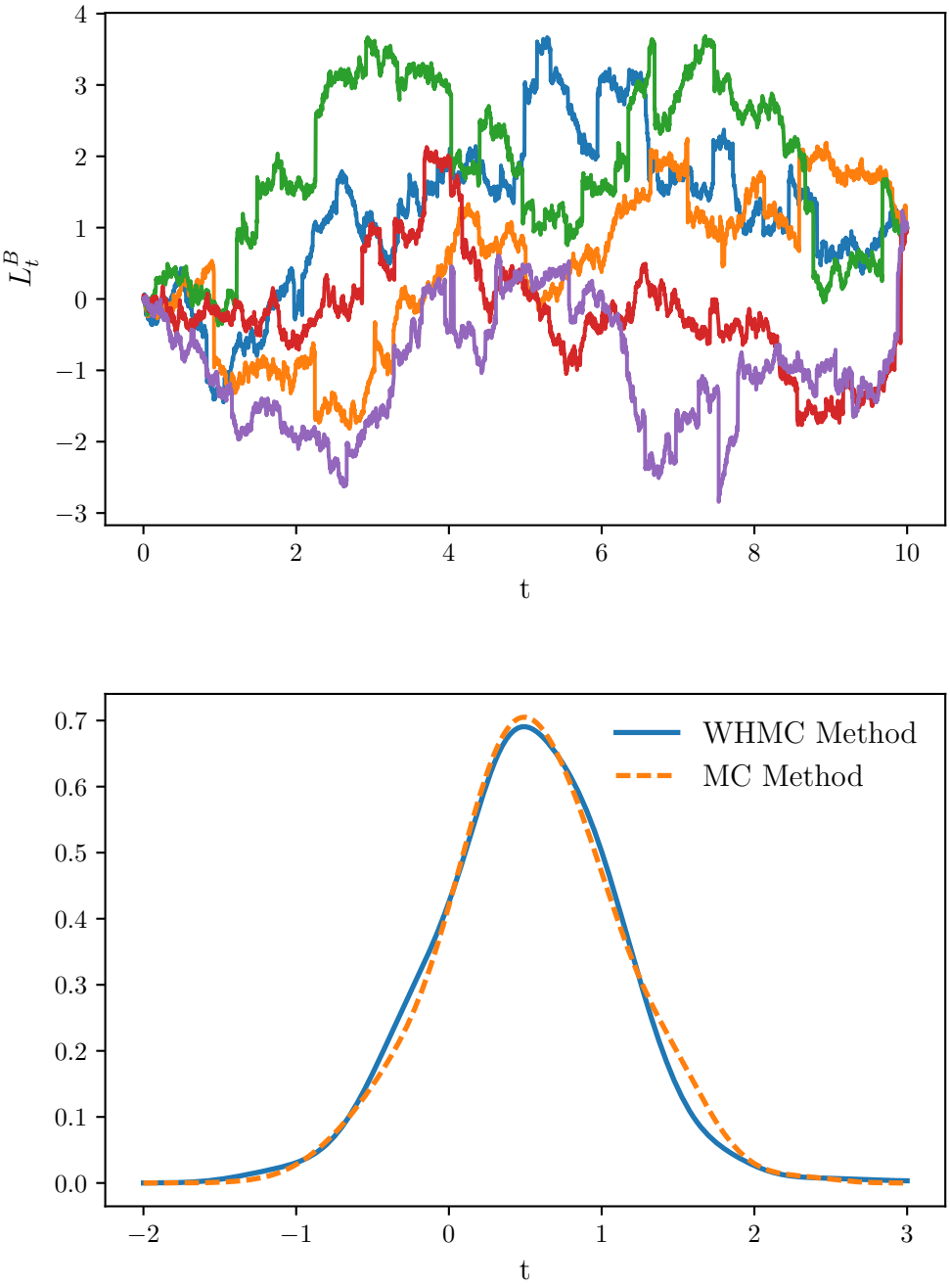


FIGURE 5.7: Top: 5 sample paths of the acceptance-rejection WHMC Lévy bridge algorithm. Bottom: Comparisons of the acceptance-rejection WHMC Lévy bridge algorithm and a naive approximate simulation (MC) for the distribution of L_5^b. Simulations are run with $b = 1$ and $T = 10$.

Theorem 5.6 *[Path functionals for Lévy bridges.]*
Let $L_t^{Br_b}$ be a Lévy process conditioned on its terminal value. Fix some $t > 0$, $u > 0$ and define V^b and J^b as in Theorem 5.5. Set for all $n \in \mathbb{N}$

$$\kappa_u^{(n)} := \inf\left\{k \in \{0, \dots, n\} \,|\, J^B(k, \lambda) > u\right\}$$

Then we have as $n \to \infty$

$$\left(\frac{t}{n}(\kappa_u^{(n)} \wedge n), V^b(\kappa_u^{(n)} \wedge n, \lambda) - u, u - V^b((\kappa_u^{(n)} - 1) \wedge n, \lambda), u - J^b((\kappa_u^{(n)} - 1) \wedge n, \lambda)\right) \xrightarrow{d} \left(\tau_u \wedge t, L^b_{\tau_u \wedge t} - u, u - L^b_{(\tau_u \wedge t)^-}, u - \overline{L}^b_{(\tau_u \wedge t)^-}\right).$$

The proof of Theorem 5.6 follows simply through a repeated application of the proof from Theorem 5.4 in with the definitions of J^b and V^b from Theorem 5.5.

5.3.3 Double saddle-point approximation of Lévy bridges

In this section we present a method for sampling from a Lévy bridges based upon the double saddlepoint approximation. In this section we present the result by Kim and Kim in [120] who used the double point saddle approximation to derive the approximated conditional PDF for a tempered stable Lévy process with stability index less than one and then we will derive the approximate conditional PDF for the spectrally one-sided β-process.

The saddlepoint approximation [121, 122] is a method in statistics to approximate the conditional PDF of the mean of n_x i.i.d random variables $\mathbf{X} = (X_1, \dots X_{n_x})$ where each component X_i has the density $f(\cdot)$, moment generating function $M(\mathbf{u}) = \mathbb{E}\left[\mathbf{u}^T\mathbf{X}\right]$ for $\mathbf{u} = (u_!, \dots, u_{n_x})$ and cumulant generating function $K(\mathbf{u}) = \log M(\mathbf{u})$. The saddlepoint approximation to the density of $f(\mathbf{X})$ is then given as

$$\hat{f}(\mathbf{x}) = \frac{1}{(2\pi)^{n_x/2}}\left(\frac{1}{|K''(\hat{\mathbf{u}})|}\right)^{1/2} \exp\left(K(\hat{\mathbf{u}}) - \hat{\mathbf{u}}^T\mathbf{x}\right),$$

where $\hat{\mathbf{u}}$ is the unique solution to the n-dimensional saddlepoint equation $K'(\mathbf{u}) = \mathbf{x}$.

The saddlepoint approximation can be used to obtain an approximation of a conditional PDF through the use of two saddlepoint approximations, one for the joint PDF and one for the marginal PDF [122, 123]. Introducing the random vector of n_y i.i.d random variables $\mathbf{Y} = (Y_1, \dots, Y_{n_y})$ and the $n_x + n_y = n$ dimensional vector $(\mathbf{X}, \mathbf{Y})$ with joint PDF $f(\mathbf{x}, \mathbf{y})$, the double saddle point approximation for the conditional PDF $f(\mathbf{x}|\mathbf{y})$ is

given by

$$\hat{f}(\mathbf{x}|\mathbf{y}) = \frac{1}{(2\pi)^{n/2}} \left(\frac{\left|K''_{\mathbf{vv}}(0,\hat{\mathbf{v}}_0)\right|}{|K''(\hat{\mathbf{u}},\hat{\mathbf{v}})|} \right)^{1/2} \exp\left(K(\hat{\mathbf{u}},\hat{\mathbf{v}}) - \hat{\mathbf{u}}^T\mathbf{x} - \hat{\mathbf{v}}^T\mathbf{y} - (K(0,\hat{\mathbf{v}}_0) - \hat{\mathbf{v}}_0^T\mathbf{y})\right), \tag{5.3}$$

where $\hat{\mathbf{v}}_0$ is the unique solution to the n-dimensional saddlepoint equation $K'(0,\hat{\mathbf{v}}_0) = \mathbf{y}$ and $(\hat{\mathbf{u}},\hat{\mathbf{v}})$ is the solution to $K'(\hat{\mathbf{u}},\hat{\mathbf{v}}) = (\mathbf{x},\mathbf{y})$. We denote the gradient as K' and Hessian by K'' where $K''_{\mathbf{vv}}$ is the Hessian with respect to the components $\mathbf{v}$.

5.3.4 Double saddle-point approximation for the one-sided tempered stable Lévy bridge

In this section we present the result from Kim and Kim [120] for the double saddlepoint approximation for the one-sided tempered stable Lévy process. This approximation is given in the next theorem.

Theorem 5.7 *[Double saddlepoint approximation for a tempered stable Lévy bridge.] Let $L^{(TS)}_t$ be a tempered stable subordinator with $\alpha \in [0,1)$ and parameterisation $\lambda = (1-\alpha)/\kappa$ and $a = \lambda^{1-\alpha}/\Gamma(1-\alpha)$ for some $\kappa > 0$. Then the double saddlepoint approximation to the conditional PDF $f_{L^{(TS)}_t|L^{(TS)}_T}(x|y)$ for $\alpha \in (0,1)$ and $0 < t < T$ is given by*

$$\begin{aligned}\hat{f}_{L^{(TS)}_t|L^{(TS)}_T}(x|y) =& \frac{1}{N}\frac{1}{\sqrt{2\pi\kappa}} \left(\frac{t(T-t)}{T}\right)^{\frac{1}{2(1-\alpha)}} \left(\frac{y}{(y-x)x}\right)^{\frac{2-\alpha}{2(1-\alpha)}} \\ &\cdot \exp\left[-\frac{(1-\alpha)^2}{\alpha\kappa}\left(\frac{t^{1/(1-\alpha)}}{x^{\alpha/(1-\alpha)}} + \frac{(T-t)^{1/(1-\alpha)}}{(y-x)^{\alpha/(1-\alpha)}} - \frac{T^{1/(1-\alpha)}}{y^{\alpha/(1-\alpha)}}\right)\right]\end{aligned}$$

where N is some normalisation constant. For $\alpha = 0$ and some normalising constant N', the double saddlepoint approximation is given by

$$\hat{f}_{L^{(TS)}_t|L^{(TS)}_T}(x|y) = \frac{1}{N'} \cdot \frac{T^{\frac{T}{\kappa}+\frac{1}{2}}}{(T-t)^{\frac{T-t}{\kappa}+\frac{1}{2}}} \cdot t^{\frac{t}{\kappa}+\frac{1}{2}} \cdot \frac{1}{\sqrt{2\pi\kappa}} (y-x)^{\frac{T-t}{\kappa}-1} \cdot x^{\frac{t}{\kappa}-1} \cdot y^{1-\frac{T}{\kappa}}.$$

Under the parameterisations $\alpha = 0$ and $\alpha = 0.5$, we obtain the special cases of the so called gamma and inverse Gaussian bridges respectively. In these cases the exact conditional PDF of these bridges are known from [124, 125] and are given as

$$f^G_{L^{(TS)}_t|L^{(TS)}_T}(x|y) = \frac{1}{B(\frac{t}{\kappa}, \frac{T-t}{\kappa})} \cdot (y-x)^{\frac{T-t}{\kappa}-1} \cdot y^{1-\frac{T}{\kappa}},$$

and

$$f^{IG}_{L_t^{(TS)}|L_T^{(TS)}}(x|y) = \frac{1}{\sqrt{2\pi\kappa}}\frac{t(T-t)}{T}\left(\frac{y}{(y-x)x}\right)^{\frac{3}{2}} \cdot \exp\left[-\frac{1}{2\kappa}\left(\frac{t^2}{x}+\frac{(T-t)^2}{y-x}-\frac{T^2}{y}\right)\right],$$

respectively.

The known densities of the gamma and inverse Gaussian bridges allow us to formulate an acceptance-rejection algorithm for the sampling of tempered stable Lévy bridges. In the case of $\alpha \in (0, 0.5)$, we have that the ratio of the double saddlepoint approximation to the gamma bridge which is given by

$$\frac{\hat{f}_{L_t^{(TS)}|L_T^{(TS)}}(x|y)}{f^{G}_{L_t^{(TS)}|L_T^{(TS)}}(x|y)} \propto g_1(x), \tag{5.4}$$

where

$$g_1(x) = x^{-\frac{t}{\kappa}-\frac{p}{2}}(y-x)^{-\frac{T-t}{\kappa}-\frac{p}{2}}\exp\left[-\frac{1}{p(1+p)\kappa}\left(\frac{t^{1+p}}{x^p}+\frac{(T-t)^{1+p}}{(y-x)^p}\right)\right],$$

with $p = \alpha/(1-\alpha)$ is uniformly bounded on $(0, y)$. Similarly in the case where $\alpha \in (0.5, 1)$, we have that the ratio of the double saddlepoint approximation to the inverse Gaussian bridge which is given by

$$\frac{\hat{f}_{L_t^{(TS)}|L_T^{(TS)}}(x|y)}{f^{IG}_{L_t^{(TS)}|L_T^{(TS)}}(x|y)} \propto g_2(x), \tag{5.5}$$

where

$$g_2(x) = x^{-\frac{p-1}{2}}(y-x)^{-\frac{p-1}{2}}\exp\left[-\frac{1}{p(1+p)\kappa}\left(\frac{t^{1+p}}{x^p}+\frac{(T-t)^{1+p}}{(y-x)^p}\right)+\frac{1}{2\kappa}\left(\frac{t^2}{x}+\frac{(T-t)^2}{y-x}\right)\right],$$

with $p = \alpha/(1-\alpha)$ is uniformly bounded on $(0, y)$. Letting C_{g_1} and C_{g_2} be the upper bounds on the ratio from Eq. (5.4) and Eq. (5.5) respectively, the acceptance-rejection sampling scheme for tempered stable Lévy bridges is then given in Algorithm 12.

In the case $0.5 < \alpha < 1$ we are required to sample from the inverse Gaussian bridge however a sampling algorithm can be obtained from the work by Ribeiro and Weber in [124]. Algorithm 12 additionally requires the computation of the relevant upper bounds of C_{g_1} or C_{g_2} which can be obtained through simple numerical approximation.

Algorithm 12 Acceptance-rejection sampling for tempered stable Lévy bridges for $\alpha \in (0,1)$

if $0 < \alpha < 0.5$ **then**
  Choose y, t, T
  Compute C_{g_1}
  repeat
    Sample $B \sim B(\frac{t}{\kappa}, \frac{T-t}{\kappa})$
    Sample $U \sim \text{Uniform}(0,1)$
    $Y \leftarrow zB$
  until $U \leq g_1(Y)/C_{g_1}$
  return $L \leftarrow Y$
else if $0.5 < \alpha < 1$ **then**
  Choose y, t, T
  Compute C_{g_2}
  repeat
    Sample $Y \sim f^{IG}_{L_t|L_T}(x|y)$
    Sample $U \sim \text{Uniform}(0,1)$
  until $U \leq g_2(Y)/C_{g_2}$
  return $L \leftarrow Y$

5.3.5 Double saddle-point approximation for the one-sided β-family Lévy bridge

In this part we propose a double saddle-point approximation for the spectrally positive β-family process $L_t^{(\beta)}$. We begin by noting that the MGF of $L_t^{(\beta)}$ is given by $M(u) = \mathbb{E}[e^{uL_t^{(\beta)}}] = e^{tl(u)}$ where $l(u)$ is Laplace exponent of which is given by

$$l(u) = -au - \frac{1}{2}\sigma^2 u^2 + \frac{c}{\beta}\left[B(\alpha, 1-\lambda) - B(\alpha + \frac{u}{\beta}, 1-\lambda)\right]$$

where $c, \beta, \alpha, \lambda > 0$ and $a, \sigma \in \mathbb{R}$. For times $0 < t < T$, the two-dimensional vector $(L_T^{(\beta)}, L_t^{(\beta)})$, where the cumulant generating function of each component is given by $T \log l(u)$ and $T \log l(u)$ respectively, has the joint cumulant generating function $K(u,v)$

$$\begin{aligned}
K(u,v) =& \log(\mathbb{E}\left[uL_t^{(\beta)} + vL_T^{(\beta)}\right]),\\
=& \log(\mathbb{E}\left[(u+v)L_t^{(\beta)} + vL_{T-t}^{(\beta)}\right]),\\
=& tl(u+v) + (T-t)l(v),\\
=& t\left(-a(u+v) - \frac{1}{2}\sigma^2(u+v)^2 + \frac{c}{\beta}\left[B(\alpha,1-\lambda) - B(\alpha + \frac{(u+v)}{\beta}, 1-\lambda)\right]\right)\\
&+(T-t)\left(-av - \frac{1}{2}\sigma^2 v^2 + \frac{c}{\beta}\left[B(\alpha,1-\lambda) - B(\alpha + \frac{v}{\beta}, 1-\lambda)\right]\right)
\end{aligned}$$

where we have used the stationary increment property of $L_t^{(\beta)}$ in the second equality and the independent increment property in the third equality.

We now are required to solve the equations $K'(\hat{u},\hat{v}) = (x,y)$ and $K'(0,\hat{v}_0) = y$. The components of the gradient are given by

$$\begin{aligned} K'_u(u,v) =& t\left[-a-\sigma^2(u+v)+\xi(u+v)\right], \\ K'_v(u,v) =& t\left[-a-\sigma^2(u+v)+\xi(u+v)\right]+(T-t)\left[-a-\sigma^2 v+\xi(v)\right], \end{aligned}$$

where

$$\xi(x) := -\frac{c}{\beta^2}B(\alpha+\frac{x}{\beta},1-\lambda)\left[\psi(\alpha+\frac{x}{\beta})-\psi(\alpha+\frac{x}{\beta}+1-\lambda)\right],$$

and $\psi(x)$ is the digamma function. Setting $F(x) = -a-\sigma^2 x+\xi(x)$, the equations for $\hat{u}$ and $\hat{v}$ are then given by the following

$$\begin{aligned} tF(\hat{u}+\hat{v})+(T-t)F(\hat{v}) =& y, \\ tF(\hat{u}+\hat{v}) =& x. \end{aligned}$$

We can thus solve $F(\hat{v}) = (y-x)/(T-t)$ numerically for $\hat{v}$ through a standard root finding procedure and then solve $F(\hat{u}+\hat{v}) = x/t$ for $\hat{v}$. For the $\hat{v}_0$ quantity, the problem reduces to numerically solving the equation $F(\hat{v}_0) = y/T$. The components of the Hessian $K''(u,v)$ are given by

$$\begin{aligned} K_{uv} &= t\left[-\sigma^2+\xi^{(1)}(u+v)\right], \\ K_{vv} &= t\left[-\sigma^2+\xi^{(1)}(u+v)\right]+(T-t)\left[-\sigma^2+\xi^{(1)}(v)\right], \\ K_{uu} &= t\left[-\sigma^2+\xi^{(1)}(u+v)\right], \end{aligned}$$

where $\xi^{(1)}(x)$ is defined by

$$\begin{aligned} \xi^{(1)}(x) := -\frac{c}{\beta^3}B(\alpha+\frac{x}{\beta},1-\lambda)\Bigg[&\left(\psi(\alpha+\frac{x}{\beta})-\psi(\alpha+\frac{x}{\beta}+1-\lambda)\right)^2 \\ &+\psi^{(1)}(\alpha+\frac{x}{\beta})-\psi^{(1)}(\alpha+\frac{x}{\beta}+1-\lambda)\Bigg], \end{aligned}$$

and $\psi^{(1)}(x)$ is the first order polygamma function.

We can now calculate the determinant in terms of the quantities $\hat{u},\hat{v}$ and $\hat{u}_0$ as

$$\begin{aligned} |K''(\hat{u},\hat{v})| &= t(T-t)\left[-\sigma^2+\xi^{(1)}(\hat{u}+\hat{v})\right]\left[-\sigma^2+\xi^{(1)}(\hat{v})\right], \\ |K''_{vv}(-,\hat{v}_0)| &= t\left|-\sigma^2+\xi^{(1)}(\hat{v}_0)\right|. \end{aligned}$$

From Eq. (5.3), the double saddlepoint approximation for the spectrally positive β-family process is then given by

$$\hat{f}_{L_t^{(\beta)}|L_T^{(\beta)}}(x|y) = \frac{1}{\sqrt{2\pi}} \left(\frac{\left|-\sigma^2 + \xi^{(1)}(\hat{v}_0)\right|}{(T-t)\left[-\sigma^2 + \xi^{(1)}(\hat{u}+\hat{v})\right]\left[-\sigma^2 + \xi^{(1)}(\hat{v})\right]} \right)^{1/2} \cdot \exp\left[tl(\hat{u}+\hat{v}) + (T-t)l(\hat{v}) - \hat{u}x - \hat{v}y - (Tl(\hat{v}_0) - \hat{v}_0 y)\right].$$

5.4 Lévy processes conditioned on their extrema

To conclude our discussion on the simulation of Lévy processes, we shall investigate an extension of the WHMC algorithm to the problem of conditioning a Lévy process on its extrema.

5.4.1 Conditioning on the extrema

The focus of this section is the proposition of an algorithm for the simulation of Lévy processes conditioned on their extrema. That is, we are interested in simulation of the Lévy process of the form

$$L_t^{\sup_f, \inf_g} := \left\{ L_t \middle| \overline{L}_T \overset{d}{=} f(\cdot), \underline{L}_T \overset{d}{=} g(\cdot) \right\}.$$

for some PDFs $f(x)$ and $g(x)$, which corresponding to conditioning the process L_t such that the supremum is distributed as $f(\cdot)$ and the infimum is distributed as $g(\cdot)$.

In this section, will shall observe the simple case of conditioning the supremum of the process to hit a point value as given by the following process

$$L_t^{\sup_u} := \left\{ L_t \middle| \overline{L}_T = u \right\},$$

for some $u \geq 0$. Introducing the hitting time

$$\tau_u = \inf\left\{ t \in T | L_t^{\sup_u} = u \right\},$$

we then observe for some new process $\tilde{L}_t$ whose Lévy triple is identical to that of L_t, the conditioned process $\left\{ \tilde{L}_t | A, 0 \leq t \leq T \right\}$ where the set A is defined as

$$A = \left\{ \mathbf{1}_{0 \leq t < \tau_u} \left\{ \tilde{L}_t < \tilde{L}_{\tau_u} = u \right\}, \mathbf{1}_{\tau_u \leq t \leq T} \left\{ \tilde{L}_t \leq u \right\} \right\}, \tag{5.6}$$

i.e. the process on $[0, \tau_u)$ is conditioned to have a first hitting time of τ_u and on $[\tau_u, T]$ the process is conditioned to stay on or below the level u, is equal in distribution to

$L_t^{\sup_u}$ by construction. From the new formulation of our conditioning in Eq. (5.6) we can define the following process in terms of the path behaviour before and after the hitting time τ_u

$$\tilde{L}_t^{\sup_u} := \begin{cases} L_t^{(1)} := \left\{ \tilde{L}_t | \inf \left\{ t \in T | \tilde{L}_t = u \right\} = \tau_u \right\} & 0 \leq t < \tau_u, \\ L_t^{(2)} := \left\{ \tilde{L}_t | \tilde{L}_{\tau_u} = u, \tilde{L}_t \leq u \right\} & \tau_u \leq t \leq T. \end{cases}$$

where by the strong Markov property, $\tilde{L}_t^{\sup_u} \stackrel{d}{=} L_t^{\sup_u}$. We now note that we can express the process $L_t^{(1)}$ on $t \in (0, \tau_u)$ as

$$\begin{aligned} L_t^{(1)} :=& \left\{ \tilde{L}_t | \inf \left\{ t \in T | \tilde{L}_t = u \right\} = \tau_u \right\}, \\ =& \left\{ \tilde{L}_t | \tilde{L}_t < \tilde{L}_{\tau_u} = u \right\}, \\ =& \left\{ \tilde{L}_t^{Br_u} | \tilde{L}_t^{Br_u} < u \right\}, \end{aligned}$$

that is we can view L_t^1 as a Lévy bridge conditioned to hit u at the time τ_u and remain below the level u on $t \in [0, \tau)$. In the case of $L_t^{(2)}$, we can view this process as being conditioned to remain on or below the level u. If L_t is a symmetric Lévy process whose triplet is identical to $L_t^{(2)}$, we can then express $L_t^{(2)}$ as

$$\begin{aligned} L_t^{(2)} :=& \left\{ \tilde{L}_t | \tilde{L}_{\tau_u} = u, \tilde{L}_t \leq u \right\}, \\ =& \left\{ u - L_t | L_t \geq 0 \right\}, \\ =& u - L_t^{\uparrow}, \end{aligned}$$

where $L_t^{\uparrow}$ is a Lévy process conditioned to stay positive as discussed in Section 5.2. The idea behind the simulation scheme is thus as follows. On the interval $[0, \tilde{T}]$ for some $\tilde{T} \in [0, T]$ we simulate a Lévy bridge as in the WHMC acceptance-rejection method in Section 5.3 under the constraint that the bridge remains below the level u. The bridge approximation is then given by

$$\hat{L}_t^{(1)} = \begin{cases} \left\{ L_t^{(1,1)} | L_t^{(1,1)} < u \right\} & \text{if } 0 \leq t \leq \tau, \\ \left\{ L_{\tilde{T}-t}^{(1,2)} | L_t^{(1,2)} < u \right\} & \text{if } \tau < t < \tilde{T}, \end{cases}$$

for some $\tau = \inf_{t^* \in [0,T]} \inf_{t \in [0,T]} \left\{ L_{T-t}^{(1,2)} = L_{t^*}^{(1,1)} \right\}$ where $\tilde{T}$ is only fixed once we have found τ under the condition that $\tilde{T} \leq T$. We can view τ as the first time that the reversed process $L_{\tilde{T}-t}^{(1,2)}$ hits $L_t^{(1,1)}$ anywhere on the intervals of $[T, 0]$ and $[0, T]$. Once this time τ has been found, we set $\tilde{T} = \tau + (T - t) \leq T$. By construction, $\hat{L}_t^{(1)}$ is an approximation of the process $L_t^{(1)}$, which is in turn the process $\tilde{L}_t^{sup}$ killed at the hitting time τ_u, thus we take $\tilde{T}$ as an approximation to the hitting time τ_u.

In the approximation of the process $L_t^{(2)}$, if L_t is a symmetric Lévy process then we can directly apply Algorithm 9 to simulate $u - L_t^{\uparrow}$ on $(\tilde{T}, T]$, otherwise a simple modification of Algorithm 9 to simulate $\left\{\tilde{L}_t | \tilde{L}_{\tau_u} = u, \tilde{L}_t \leq u\right\}$, that is of a Lévy process conditioned to remain on or below the level u.

In summary, the algorithm for simulating the path of the process $L_t^{sup_u}$ is given below

Algorithm 13 WHMC path simulation of $L_t^{sup_u}$

Choose $n \geq 1$ so that $\mathbf{g} := \sum_{i=1}^{n} \frac{t}{n}\mathbf{e}_t \approx t$ by SLLN.
$\lambda \leftarrow n/t$
$L^{(1)}(0) \leftarrow 0, L^{(2)}(0) \leftarrow u, L^{(3)}(0) \leftarrow u$
$Z^{(1)} \leftarrow 0$, $Z^{(2)} \leftarrow u$
$i^{(1)} \leftarrow 0$, $i^{(2)} \leftarrow 0$
repeat
    for $i = 1, \ldots, n$ **do**
        sample $S^{(1)} \sim \mathbf{Law}(\overline{L}_{e_1/\lambda})$
        sample $I^{(1)} \sim \mathbf{Law}(\underline{\mathrm{L}}_{e_1/\lambda})$
        $L^{(1)}(i) \leftarrow L^{(1)}(i-1) + S^{(1)} + I^{(1)}$
        if $Z^{(1)} > L^{(1)}(i)$ **then**
            $i^{(1)} \leftarrow i$
        $Z^{(1)} \leftarrow \max(Z^{(1)}, L^{(1)}(i) + S^{(1)})$
        repeat
            sample $S^{(2)} \sim \mathbf{Law}(\overline{L}_{e_1/\lambda})$
            sample $I^{(2)} \sim \mathbf{Law}(\underline{\mathrm{L}}_{e_1/\lambda})$
            $L^{(2)}(i) \leftarrow L^{(2)}(i-1) - S^{(2)} - I^{(2)}$
        until $L^{(2)} \leq u$.
        $Z^{(2)} \leftarrow \min(Z^{(2)}, L^{(2)}(i) - S^{(2)})$
        if $Z^{(2)} < L^{(2)}(i)$ **then**
            $i^{(2)} \leftarrow i$
        if $Z^2 \geq Z^1$ **or** $i^{(1)} + i^{(2)} > n$ **then**
            Break
until $Z^2 \geq Z^1$ **and** $i^{(1)} + i^{(2)} \leq n$
$Y^{(1)} \leftarrow L^{(1)}(j)$, $j = 0, 1, \ldots, i^{(1)} - 1$
$Y^{(2)} \leftarrow L^{(2)}(j)$, $j = i^{(2)}, i^{(2)} - 1, \ldots, 0$
$N \leftarrow n - i^{(1)} - i^{(2)}$
for $j = 1, \ldots, N$ **do**
    repeat
        sample $S^{(3)} \sim \mathbf{Law}(\overline{L}_{e_1/\lambda})$
        sample $I^{(3)} \sim \mathbf{Law}(\underline{\mathrm{L}}_{e_1/\lambda})$
        $L^{(3)}(j) \leftarrow L^{(3)}(j-1) + S^{(3)} + I^{(3)}$
    until $L^{(3)} \leq u$.
$Y^{(3)} \leftarrow L^{(3)}(j)$, $j = 1, 2, \ldots, N$
return $[Y^{(1)}, Y^{(2)}, Y^{(3)}]$

We note that Algorithm 13 can easily be extended to the conditioning problem $L_t^{\sup_f} := \left\{L_t \middle| \overline{L}_T \overset{d}{=} f(\cdot)\right\}$ for some PDF $f(x)$ by first sampling u from the distribution $f(x)$ each time we simulate a path of the process.

A numerical simulation of Algorithm 13 is provided on Figure 5.8 below. For numerical verification we include the densities of the conditioned process under the acceptance-rejection WHMC algorithm and a naive MC algorithm at the time $T = 0.5$. In the case of the MC algorithm, we are simulating the approximation to conditioned process $\left\{L_t | \overline{L}_T \in [u - \epsilon, u + \epsilon]\right\}$ with $\epsilon = 0.1$.

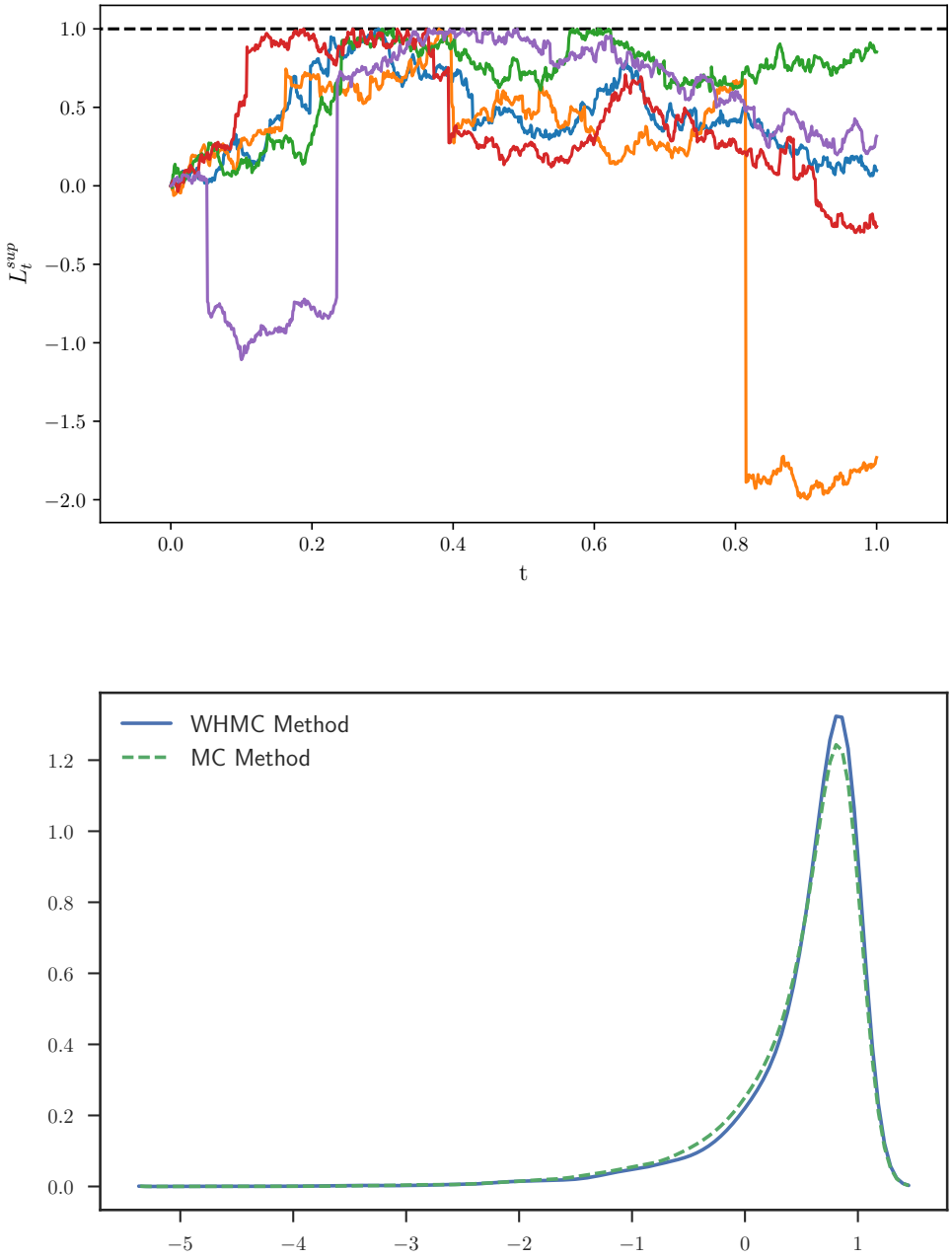


FIGURE 5.8: Top: 5 sample paths of the acceptance-rejection WHMC algorithm for Lévy processes conditioned on their supremum. Bottom: Comparisons of the acceptance-rejection WHMC algorithm for Lévy processes conditioned on their supremum and a naive approximate simulation (MC) for the distribution of $L_{0.5}^{sup}$. Simulations are run with $u = 1$ and $T = 1$.

Chapter 6

The Conditioned Kuramoto Model

In this chapter we consider a standard synchronised Kuramoto system upon an arbitrary network in which a node is chosen to escape from the group synchronised state at a particular time. This model is investigated for both the Gaussian and Lévy noise cases. The major focus of our analysis of this system will be on how the underlying graph structures affects the dynamics of the system before and after the point of conditioning. Specifically, we are interested in investigating if there are potentially any graph properties that can be used as predictors for the dynamics of the conditioned system.

6.1 Problem background

The role that network structure plays in the dynamics of a given model is a complicated question. An example of this problem that has been explored in the literature can be seen in the field of epidemic processes where examples in epidemiological modelling and networked structures are discussed in the context of extinction thresholds (e.g. [126, 127]) and some analytical results are obtained in [128]. More general discussions on network structure in sychronisation models with particular focus upon the degree distribution of the network can be found in [129].

The aim of this chapter is to investigate the dependency of the dynamics of a Kuramoto model upon the graph structure at the times before, on and after the point of conditioning. Our goal is to identify what the relevant topological features of the network are in governing the system dynamics at these times and ideally to use these topological features as predictors for future system states.

6.2 Model overview

The oscillator dynamics in this model are given by the Kuramoto model where we condition one of the oscillators $\theta_n(t)$ for some $n \in \{1, \ldots, N\}$ to hit a specified value θ_c at some time T_c. The dynamics of the conditioned Kuramoto model driven by stochastic noise are given by the following system of equations

$$\begin{aligned} \mathrm{d}\theta_i(t) &= \left(\omega_i - \sigma_G \sum_j^N A_{ij} \sin(\theta_i - \theta_j) \right) \mathrm{d}t + \mathrm{d}{L_t}^{(i)}, \quad i = \{1, \ldots, N\}, \\ \theta_n(T_c) &= \theta_c \text{ for some } |\theta_c| \geq \pi/2, \quad \theta_1(0) = \theta_2(0) = \cdots = \theta_N(0) = 0. \end{aligned} \tag{6.1}$$

where $\theta_i(t)$ is the phase angle for node i at time t, ω_i is the intrinsic frequency of the given node, σ_G is the global coupling constant, $L_t^{(i)}$ is an independent Lévy process at node i, and n is the conditioned node from the set of nodes $n \in \{1, \ldots, N\}$ where n is chosen randomly with a uniform probability from this set. The underlying network, described by the adjacency matrix $\boldsymbol{A}$, is created by sampling from the set of connected graphs for two different cases. The first case that we shall consider is the simple case of 1-connected graphs with $|E| = N$ and in the second case we consider more general graphs that are n-connected, i.e. where the number of edges is $|E| > N$. We note that graphs in the 1-connected case are generally simpler and have Fielder values that can be directly compared between graphs in this set whereas this is not the case for the more general set of n-connected graphs.

6.3 Bridge approximations of the conditioned process

In the numerical simulation of the conditioned Kuramoto model, the fact that the conditioning is now applied upon a multidimensional process means that we will have to resort to some approximation of the conditioned oscillator before any of the previously discussed techniques in Chapter 5 can be applied.

As a first step, we linearise the sine function around 0 and thus can show that the dynamics of $\boldsymbol{\theta}(t)$ satisfy that of an Ornstein-Uhlenbeck process

$$\begin{aligned} \mathrm{d}\theta_i(t) &= \left(\omega_i + \sigma_G \sum_{j=1}^n A_{ij} \sin\left(\theta_i - \theta_j\right) \right) \mathrm{d}t + \mathrm{d}L_t^{(i)}, \\ &\approx \left(\omega_i - \sigma_G \sum_{j=1}^n \mathcal{L}_{ij} \theta_j \right) \mathrm{d}t + \mathrm{d}L_t^{(i)}, \end{aligned}$$

where $\mathcal{L}_{ij} = \begin{cases} A_{ij}, & i \neq j \\ -d_i, & i = j \end{cases}$ and $d_i = \sum_j A_{ij}$.

We thus arrive at the Ornstein-Uhlenbeck process which approximates the unconditioned dynamics of the system of oscillators as

$$\mathrm{d}\boldsymbol{\theta}(t) = (\boldsymbol{\omega} - \boldsymbol{B}\boldsymbol{\theta}(t))\mathrm{d}t + \mathrm{d}\boldsymbol{L}_t,$$

where $\boldsymbol{\omega} = (\omega^1, \ldots, \omega^n)$ and $\boldsymbol{B} = \mathrm{diag}\,(\sigma_G, \ldots, \sigma_G)\,\mathcal{L}$.

Focusing on the Gaussian case, one possibility in further approximating the system in Eq 6.3 is to project the dynamics onto a one-dimension diffusion process using the idea of Markovian projection from Theorem B.1 in Appendix B.3. First, we shall parameterise the model such that $\mathbb{E}[\boldsymbol{\theta}(t)] = \mathbf{0}$ which can be obtained through the selection $\boldsymbol{\omega} = 0$. With this parameterisation we then apply the results of Markovian projection from Theorem B.1 which allows us to write the dynamics of the unconditioned oscillator n as the one-dimensional diffusion process

$$\mathrm{d}\tilde{\theta}_n(t) = -\left(\mathbb{E}\left[\boldsymbol{B}\boldsymbol{\theta}(t) | \boldsymbol{\theta}(t) = \boldsymbol{\theta}\right]\right)_n \mathrm{d}t + \mathrm{d}\boldsymbol{W}_t^{(n)}.$$

Utilising this formulation requires the calculation of the conditional expectation in the drift term which can be difficult to obtain. Thus we shall make a further approximation to the expectation by conditioning on the event $\boldsymbol{\theta}(t) = [0, 0, \ldots, \theta_n(t), \ldots, 0]$. We can think of this approximation as allowing the oscillator n to move freely whilst fixing the other nodes at 0. In this case, the unconditioned dynamics of the oscillator n will be given by the equation whose dynamics explicitly depend on the degree d_n of the node n:

$$\mathrm{d}\tilde{\theta}_n(t) \approx -d_n\sigma_G\theta_n(t) + \mathrm{d}W_t^{(n)}. \tag{6.2}$$

Since Eq. (6.2) is a one-dimensional Ornstein-Uhlenbeck process, we can apply the Doob h-transform from Section 5.1.2 to obtain an SDE for the conditioning $\theta_n(T_c) = \theta_c$.

In the second case where the stochastic noise is a Lévy process, we require an additional approximation in order to apply our techniques for the simulation of conditioned processes. Given the Ornstein-Uhlenbeck approximation to the unconditioned process driven by Lévy noise

$$\mathrm{d}\theta_n(t) \approx (\boldsymbol{B}\boldsymbol{\theta}(t)\mathrm{d}t)_n + \mathrm{d}\boldsymbol{L}_t^{(n)},$$

and the parameterisation of the model such that $\mathbb{E}[\boldsymbol{\theta}(t)] = \mathbf{0}$, by setting $\boldsymbol{\theta}$ to its expected value we have the approximation where the dynamics of the unconditioned oscillator n

are simply given by the Lévy noise

$$\mathrm{d}\theta_n(t) \approx (\boldsymbol{B}\mathbb{E}[\boldsymbol{\theta}(t)]\mathrm{d}t)_n + \mathrm{d}\boldsymbol{L}_t^{(n)} = \mathrm{d}\boldsymbol{L}_t^{(n)}.$$

The conditioned dynamics of the oscillator n are thus simulated through our WHMC algorithm for Lévy bridges in Section 5.3.

In summary, the dynamics of the conditioned oscillator n on the interval $t \in [0, T_c]$ are governed by the following SDEs for the Gaussian and Lévy noise cases:

- Gaussian case: An Ornstein-Ulhenbeck approximation is taken and the bridge dynamics are given by:

$$\begin{aligned}\mathrm{d}\theta_n(t) \approx \Bigg(&- \sigma_G d_n \theta_n(t) + 2\frac{[\theta_c - \theta_n(t)\exp\left(-\sigma_G d_n\left(T_c - t\right)\right)]}{-\exp\left(-2\sigma_G d_n\left(T_c - t\right)\right)/(\sigma_G d_n)} \\ &\times \exp\left(-\sigma_G d_n\left(T_c - t\right)\right)\Bigg)\mathrm{d}t + \sigma \mathrm{d}W_t, \end{aligned} \tag{6.3}$$

- Lévy case: The path is approximated by the WHMC algorithm for Lévy bridges.

$$\mathrm{d}\theta_n(t) \approx \mathrm{d}L_t^{Br_{\theta_c}}.$$

Based upon the approximation of the dynamics of the conditioned process in the Gaussian case in Eq. (6.3), there is explicit dependence on the degree of the conditioned node d_n. With this dependency in mind, we shall now proceed to investigate the dependency of the system dynamics on the network topology numerically. In particular, we shall be interested in investigating what graphs properties aside from the degree of the conditioned node are useful predictors for the dynamics of the conditioned system.

6.4 Numerical investigation of the conditioned model

In the numerical implementation of the system in Eq. (6.1), we simulate the dynamics over the time frame $t \in [0, T]$ where $T = 3$. For the conditioning, we set $\theta_n(T_c) = \frac{\pi}{2}$, where $T_c = 2$. Each node is chosen to have zero natural frequency $\boldsymbol{w} = \boldsymbol{0}$ so that $\mathbb{E}[\boldsymbol{\theta}(t)] = \boldsymbol{0}$ and the coupling constant is chosen to be $\sigma_G = 1$. On the interval $t \in [0, T_c]$, we simulate the conditioned node n according to Eq. (6.3) and on the time $t \in (T_c, T]$, the dynamics are driven by the unconditioned system. That is, in the Gaussian case,

the dynamics of the conditioned node are simulated as

$$\mathrm{d}\theta_n(t) = \begin{cases} \left(-\sigma_G d_n \theta_n(t) + 2\frac{[\theta_c - \theta_n(t)\exp(-\sigma_G d_n(T_c-t))]}{-\exp(-2\sigma_G d_n(T_c-t))/(\sigma_G d_n)} \right. & \text{if } t \in [0, T_c], \\ \left. \quad \times \exp\left(-\sigma_G d_n \left(T_c - t\right)\right)\right)\mathrm{d}t + \sigma \mathrm{d}W_t, & \\ \left(-\sigma_G \sum_j^N A_{ij} \sin(\theta_n(t) - \theta_j(t))\right)\mathrm{d}t + \mathrm{d}W_t & \text{if } t \in (T_c, T], \end{cases}$$

and in the Lévy case the dynamics of the conditioned node are given by

$$\mathrm{d}\theta_n(t) = \begin{cases} \mathrm{d}L_t^{Br_{\theta_c}} & \text{if } t \in [0, T_c], \\ \left(-\sigma_G \sum_j^N A_{ij} \sin(\theta_n(t) - \theta_j(t))\right)\mathrm{d}t + \mathrm{d}L_t & \text{if } t \in (T_c, T]. \end{cases}$$

In the Gaussian case, we choose the volatility $\sigma = 0.1$. In the Lévy case we use the β-family of Lévy processes as discussed in Section 2.4.2.2 where we set

$$(a, \sigma, \alpha_1, \beta_1, \lambda_1, c_1, \alpha_2, \beta_2, \lambda_2, c_2) = (0, 0.1, 1, 1.5, 1.5, 1, 1, 1.5, 1.5, 1).$$

In this numerical investigation, we shall record two different statistics to observe the impact of the graph topology on the system dynamics. Our choice of these two observables is based on our desire to measure the behaviour of the system at a local and global level where we define local as the neighbourhood of the conditioned oscillator n and global as the entire network. We will specifically be interested in observables that measure the overall phase of the system and the synchronisation of the system at the appropriate global and local levels.

The first statistic that we shall observe is the average phase across all oscillators

$$\bar{\theta}(t) = \frac{1}{N}\sum_{i=1} \theta_i(t).$$

This shall serve as a 'global' statistic in our analysis. Since in the unconditioned case we have that $\mathbb{E}(\bar{\theta}_n(t)) = 0$ as a result of our parameterisation, the impact of conditioning $\theta_n(T_c) = \theta_c = \frac{\pi}{2}$ results in the shift $0 < \mathbb{E}[\bar{\theta}(T_c)] < \frac{\pi}{2}$ at time T_c. The magnitude of the shift in $\bar{\theta}$ from 0 shall be the focal point of this statistic.

The second statistic that we shall use is the local order parameter of the conditioned node and this shall serve as our 'local' statistic as it is specific to the dynamics at and near the conditioned oscillator. The local order parameter is given by

$$O^{\theta_n}(t) = \frac{1}{d_n + 1}\left|\sum_{j \in G_n} e^{i\theta_j(t)}\right|,$$

where $G_j \subseteq G$ is a subset of nodes from the full graph G consisting of the conditioned oscillator at node n and all neighbouring nodes of n (for a total of $d_n + 1$ nodes). This will serve as a measure of local sychronisation at the conditioned node, the expectation is that there will be a loss of sychronisation as the conditioned node escapes from the group state at T_c, that is $\mathbb{E}[O^{\theta_n}(t)] > \mathbb{E}[O^{\theta_n}(T_c)]$. The magnitude of the shift in the local order parameter from its synchronised state shall be the focal point of this statistic.

The local and global statistics, $\bar{\theta}(t)$ and $O^{\theta_n}(t)$ respectively, are to be observed at a time prior to conditioning $t = 1$, at the time of conditioning $T_c = 2$ and after conditioning at the terminal time $T = 3$.

In the approximation of the dynamics of the conditioned oscillator, we determined that the degree of the node d_n plays an important role in the behaviour of the oscillator as it escapes and returns to the system. One of the aims of this investigation is to find what other features of the graph can be used as predictors in determining the behaviour of the conditioned node and the overall system in the context of our local and global statistics $\bar{\theta}(t)$ and $O^{\theta_n}(t)$ respectively. To this end, we shall investigate a variety of different graph statistics that are relevant to the entire network and the particular node n on which the conditioning is applied. On Table 6.1 we summarise the graph statistics from Section 2.1.1 that will then be used to extend our numerical analysis to a wider range of graph feature characterisations.

Name	Symbol	Definition	Node/Network Statistic
Degree	d_i	Number of edges connect to node i.	Node
Average Degree	$\langle d \rangle$	Average degree across all nodes	Network
Fiedler value	λ_2	Second smallest eigenvalue of L.	Network
Diameter	D	Maximum shortest path length between any two nodes.	Network
Average shortest path length	$\langle l \rangle$	Average shortest path length over all possible pairs of vertices.	Network
Closeness centrality	g_i	Average distance at node i to all other nodes.	Node
Clustering coefficient	C_i	Local clustering at node i.	Node
Average clustering coefficient	$\langle C \rangle$	Average all of clustering coefficients	Network

TABLE 6.1: Summary of graph statistics used in the numerically analysis of the effect of graph structure on the dynamics of the conditioned Kuramoto model. We list the name, symbolic notation, definition and wherever the statistic is specific to a node or the entire network.

6.4.1 Simulations with 1-connected graphs

As a starting point in our numerical investigation, we shall restrict ourselves to the case of 1-connected graphs, that is the set of graphs where the removal of a single edges results in a disconnected graph hence the number of edges is equal to $N-1$ where we choose $N=20$ for all of our graphs. We start this investigation by observing the degree of node d_n and the Fiedler value λ_2 which describes the overall connectivity of the graph. Thus our initial approach in this investigation is to determine how the overall structure of the network, represented by the Fielder value, impacts the dynamics of the conditioned oscillator. Since we are interested in investigating the impact of the Fiedler value on the system dynamics, we desire that our set of graphs is selected such that we sample uniformly across the allowable range of Fiedler values for each graph. The algorithm in Algorithm 14 is used as a starting point to generate 1-connected graphs.

Algorithm 14 Sampling from the set of 1-connected graphs

1: $N \leftarrow 20$
2: $A \leftarrow \text{zeros}(N, N)$
3: $A[0,1] = A[1,0] \leftarrow 1$
4: **for** $i = 0, \ldots, N-3$ **do**
5: sample $j \in [0, 1, \ldots, i+1]$
6: $A[i+2, j] = A[j, i+2] \leftarrow 1$
7: **return** A

To see that this algorithm samples from the entire set of 1-connected graphs, we introduce the following proposition.

Proposition 6.1 *[The set of 1-connected graphs produced by Algorithm 14.] Let $\mathcal{G}$ be the set of 1-connected graphs produced by Algorithm 14. Then $\mathcal{G}$ is equivalent to the set of all 1-connected graphs.*

Let $G(N, E)$ be an arbitrary graph with $|E| = N-1$ and pick two arbitrary nodes from G indexed i and j. Since G is 1-connected, the shortest path length $l_{ij} \leq N-1$ is the only path between the two nodes. Let the nodes along the path l_{ij} be numbered $[i, n_1, n_2, \ldots, n_{l_{ij}-2}, j]$. It is then sufficient to show that, with non-zero probability, Algorithm 14 can produce a path of length $l_{ij} \leq N$ between nodes i and j. Let $G^{(i)}$ be the graph produced by Algorithm 14 at which node i is initially created, then with probability $p > \frac{1}{i+1}$, an edge will be created between the nodes i and n_1 at some iteration of the algorithm. A lower bound on the probability of Algorithm 14 creating a path of length $l_{ij} \leq N$ between nodes i and j can be obtained by considering the case where the path $[i, n_1, n_2, \ldots, n_{l_{ij}-2}, j]$ is created at subsequent iterations of the algorithm.

The probability of this path occurring after $|l_{ij}|$ iterations is given by $P^{\text{subsequent}}(l_{ij}) = \sum_{n=1}^{|l_{ij}|} \frac{1}{i+n}$. Thus the probability of Algorithm 14 producing a path of length l_{ij} is bounded below by $P(l_{ij}) \geq \sum_{n=1}^{|l_{ij}|} \frac{1}{i+n} > 0$.

One of the issues with Algorithm 14 is that it is biased towards producing graphs with relatively low Fielder values and as such, highly connected graphs with large Fielder values can been seen as rare events of this algorithm. To ensure that we sample graphs evenly across the possible range of Fielder values, we first create the empirical distribution of the Fiedler values from the graphs produced by Algorithm 14. This empirical distribution will then be used to produce an acceptance probability for each graph such that the resulting distribution of Fielder values in the acceptance-rejection scheme is uniformly distributed across the allowable range. The probability of accepting a graph with a Fielder value of λ is calculated by weighting the probability of being within a range of some Δx of the given Fielder value of λ by the inverse ratio of the the density function λ_2 and the value of the desired uniform distribution. This weighting is calculated by

$$l_{\lambda,\Delta x} := l(\lambda - \Delta x \leq x \leq \lambda + \Delta x) = \int_{\lambda-\Delta x}^{\lambda+\Delta x} \left(\frac{f_{\lambda_2}(x)}{\sum_{U_i,L_i \in \mathcal{I}} \frac{1}{U_i - L_i}} \right)^{-1} \mathrm{d}x, \tag{6.4}$$

where $\mathcal{I} = \bigcup_i [L_i, U_i] = \text{supp}(f_{\lambda_2}(x))$ and supp is the support of $f_{\lambda_2}(x)$. The probability of acceptance is then calculated by adjusting the odds of the original probability by the weighting in Eq. (6.4) as follows:

$$P(\lambda - \Delta x \leq x \leq \lambda + \Delta x) = \frac{q_{\lambda,\Delta x}}{1 + q_{\lambda,\Delta x}}, \tag{6.5}$$

where the new odds $q_{\lambda,\Delta x}$ is given by

$$q_{\lambda,\Delta x} = l_{\lambda,\Delta x} \frac{2\Delta x \sum_{U_i,L_i \in \mathcal{I}} \frac{1}{U_i - L_i}}{1 - 2\Delta x \sum_{U_i,L_i \in \mathcal{I}} \frac{1}{U_i - L_i}}.$$

This procedure is summarised below in Algorithm 15 where we use a histogram to approximate the procedure in Eq. (6.4)-Eq. (6.5). A comparison of the probability densities of the Fielder values between the graphs sampled from Algorithm 14 and Algorithm 15 is given in Figure 6.1.

Algorithm 15 Sampling from the set of 1-connected graphs uniformly across Fiedler values

for $i = 0, \ldots$ **do**
    sample G as in Algorithm 14
    $\lambda_2[i] \leftarrow \text{Fiedler}(G)$
$H \leftarrow \text{hist}(\lambda_2)$
$numbins \leftarrow$ number of nonzero bins (H)
$num \leftarrow$ total count (H)
repeat
    sample G as in Algorithm 14
    $\lambda \leftarrow \text{Fiedler}(G)$
    $count \leftarrow H(\lambda)$
    $U \sim \text{Uniform}(0, 1)$
    $l \leftarrow \left(\frac{count}{num/numbins}\right)^{-1}$
    $q \leftarrow l \cdot \frac{1/numbins}{1-1/numbins}$
    $p \leftarrow \frac{q}{1+q}$
until $U \leq p$
return G

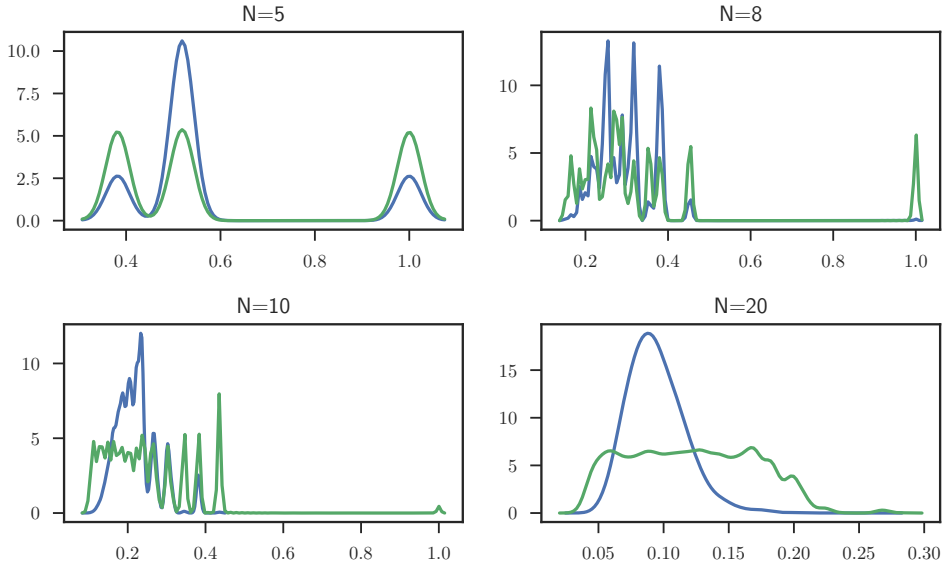


FIGURE 6.1: The probability density of Fiedler values from set of 1-connected graphs sampled from Algorithm 14 (blue) and Algorithm 15 (green) for $N \in \{5, 8, 10, 20\}$.

Using Algorithm 15 we sample from the set of 1-connected graphs G. For each graph, the conditioned oscillator n is chosen randomly from the set of nodes $\{1, 2, \ldots, N\}$ in each graph G. For each graph we record the degree of the conditioned node n and the

Fielder value λ_2. We then record the median values of the local and global statistics, $\bar{\theta}(t)$ and $O^{\theta_n}(t)$ at the times $T = 1, 2, 3$ where the median value is taken over 100 simulations for each graph.

As a baseline scenario, we consider the unconditioned system where none of the oscillators are Figure 6.2. We then consider the Gaussian noise case in Figure 6.3 and the Lévy noise case in Figure 6.4. In each figure we plot the median of the average frequency and local order parameter against the degree of the node d_i and the Fiedler value λ_2.

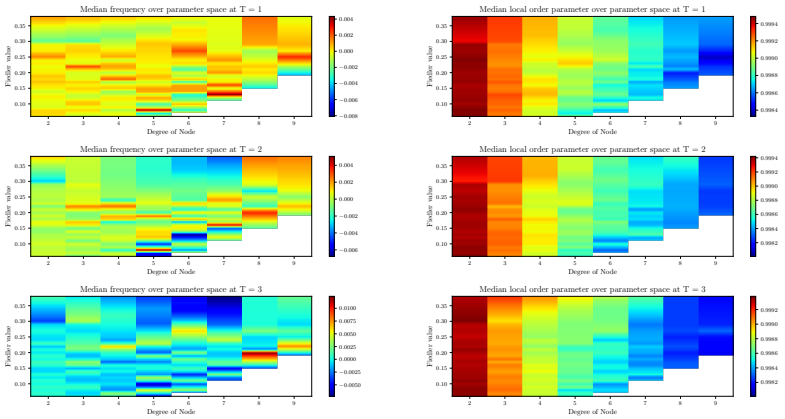


FIGURE 6.2: The baseline scenario (no conditioning). Left column: The expected median frequency. Right: The expected local order parameter at the conditioned node.

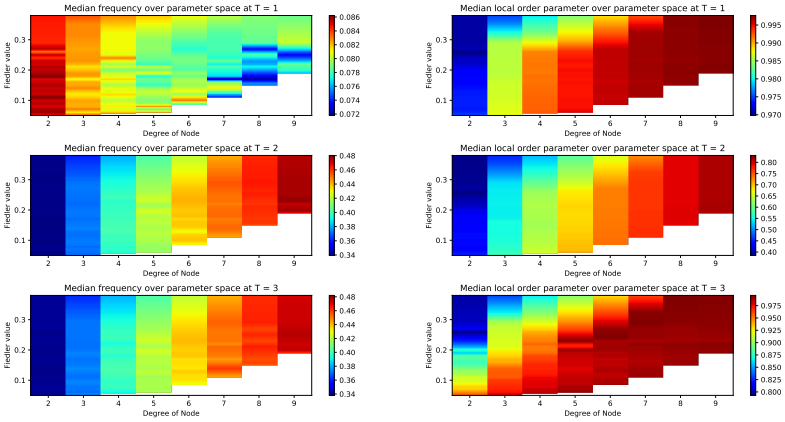


FIGURE 6.3: The conditioned scenario with Gaussian noise. Left column: The expected median frequency. Right: The expected local order parameter at the conditioned node.

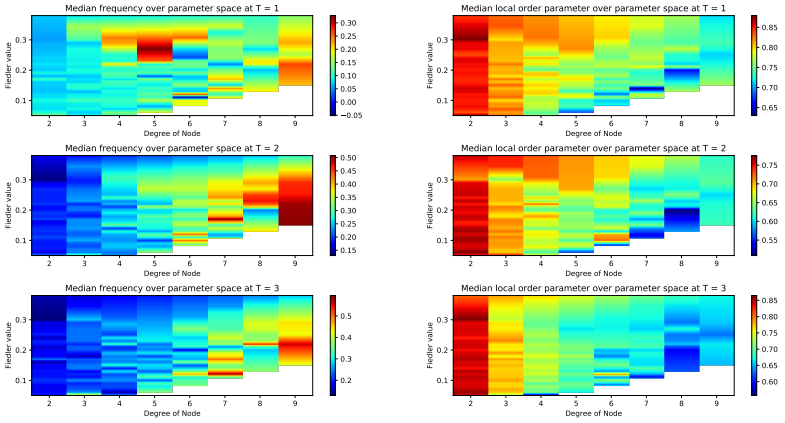


FIGURE 6.4: The conditioned scenario with Lévy noise scenario. Left column: The expected median frequency. Right: The expected local order parameter at the conditioned node.

It is clear from the plots in Figures 6.2-6.4 that the degree of the conditioned node directly impacts the dynamics of the system however the impact of the overall connectivity is less certain. In the unconditioned case on Figure 6.2, graph structure has no effect on the median frequency. The local order parameter shows stronger synchronisation when the degree of the node is smaller however the overall connectivity of the graph does not matter. In the Gaussian case on Figure 6.3, the median frequency is dependent upon the degree of the conditioned node and not the graph structure with a high node degree shifting the median frequency more than the low node degree. This is similarly observed in the local order parameter however, when $T = 3$, a dependence upon the Fielder value emerges. In the Lévy case 6.4, the structure is more blurred. Here the behaviour of the median frequency is similar to the Gaussian case however the behaviour of the local order parameter is similar to that of the unconditioned case.

We now look to extend this analysis to include the various graph statistics from Table 6.1. From this table we shall observe the values of the degree d_i, the Fiedler value λ_2, the diameter D, the average shortest path length $\langle l \rangle$ and the closeness centrality g_i for each graph G as generated through the application of Algorithm 15. In this analysis, we shall observe the differences in the global and local statistics $\bar{\theta}(t)$ and $O^{\theta_n}(t)$ for time $t \in [1, 2)$ and for time $t \in [2, 3)$ which represent the dynamics of the system leading up until the time of conditioning and the dynamics after the time of conditioning respectively. We sample from $N = 1000$ graphs and for each graph we obtain a Monte-Carlo estimate of the expectations through 50 simulations. We then use multiple linear regression to fit a model with the differences of each statistic $\bar{\theta}(\Delta T)$ and $O^{\theta_n}(\Delta T)$ as our response variables and the relevant graph statistics from Table 6.1 as our predictors. We also include the interaction terms between the predictors in the model fit. The Bayesian

information criterion (BIC) is then used in model selection to determine which graphs statistics are potentially useful predictors of system dynamics leading up until the time of conditioning and afterwards. We should however note that many of these graph statistics are dependent upon one another and hence the purpose of this regression should best be interpreted as looking for a best singular statistic to predict dynamics and not a linear combination of statistics. The results of the BIC model selection in the case of 1-connected graphs are given on Figure 6.5 for both the Gaussian and Lévy noise cases.

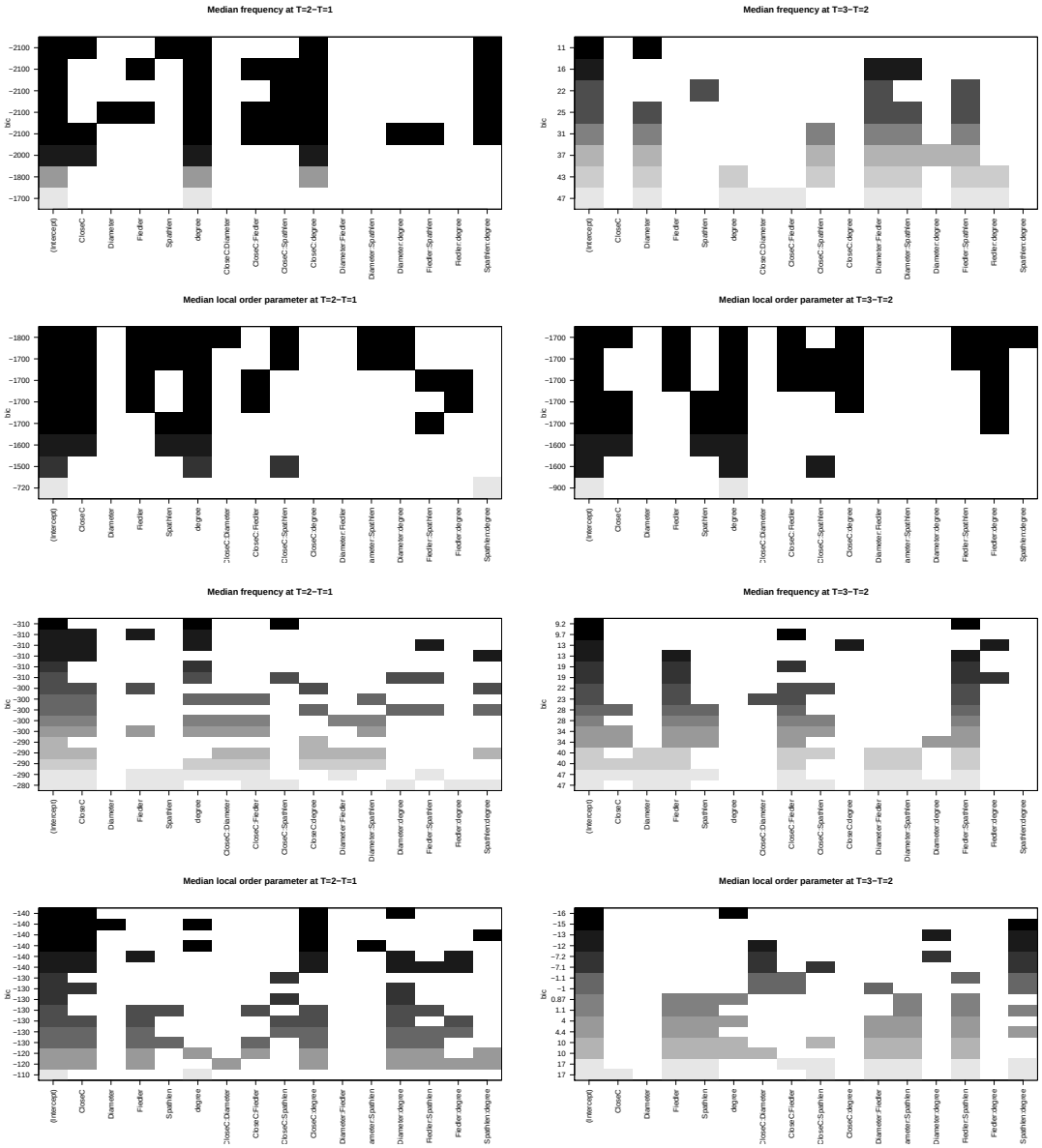


FIGURE 6.5: Model selection for 1-connected graphs under Gaussian noise (top) and Lévy noise (bottom). Here CloseC denotes the closeness centrality, Diameter denotes the Diameter, Fiedler denotes the Fiedler value, Spathlen denotes the shortest path length and degree denotes the degree. The symbol ':' denotes the interaction between the two relevant predictors.

6.4.2 Simulations with n-connected graphs

In the previous part, we only considered 1-connected graphs with N nodes. We now look to extend our numerical investigation to the case of n-connected graphs. The procedure behind our graph generation in this section is based upon Algorithm 15 to ensure that our sampled graph remains connected. Using the 1-connected graph generated from Algorithm 15 as a starting point, we then use the Erdös-Rényi $G(N,p)$ model to generate the additional edges with a link probability of $p = 0.1$. The procedure for generating n-connected graphs is given below in Algorithm 16.

Algorithm 16 Sampling from the set of n-connected graphs

1: $N \leftarrow 20, p \leftarrow 0.1$
2: sample G as in Algorithm 15
3: $H \leftarrow$ Erdös-Renyi(N,p)
4: $G \leftarrow G + H$
5: **return** $G \leftarrow \text{int}(G\,! = 0)$

To see that this algorithm samples from the entire set of n-connected graphs, we introduce the following proposition.

Proposition 6.2 *[The set of n-connected graphs produced by Algorithm 16.] Let $\mathcal{G}$ be the set of n-connected graphs produced by Algorithm 16. Then $\mathcal{G}$ is equivalent to the set of all n-connected graphs.*

Set $n \in [1,\dots,N]$. Let G^1 and G^2 be the graphs produced in steps 2 and 5 of Algorithm 15 respectively. Since G^1 is 1-connected, it follows by step 4 that G^2 is at least 1-connected. Let G be at least an n-connected graph, then it is known, from the property of a Harary graph [130], that G has at least $\lceil nN/2 \rceil$ edges. It is also known that if G is a complete graph, it has $\frac{N(N-1)}{2}$ edges. Thus, if G is at least n-connected, G has $|E|$ number of edges where $\lceil nN/2 \rceil \leq |E| \leq \frac{N(N-1)}{2}$. The probability that the G^2 has $|E|$ edges is thus bounded by

$$\begin{aligned}\binom{N-1}{\lceil nN/2 \rceil} p^{\lceil nN/2 \rceil}(1-p)^{N-1-\lceil nN/2 \rceil} \geq & P(|E|) \\ \geq & \binom{N-1}{\frac{N(N-1)}{2}} p^{\frac{N(N-1)}{2}}(1-p)^{N-1-\frac{N(N-1)}{2}} \\ > & 0.\end{aligned}$$

Since G^2 is at least 1-connected and can have all possible edges up until a complete graph with non-zero probability, it follows that Algorithm 16 can sample from the full set of n-connected graphs.

In our numerical analysis, we shall use all graph statistics present in Table 6.1 however we do note that the Fiedler value is no longer an effective comparative statistic across graphs with a variable number of edges. As before, we shall observe the differences in the global and local statistics $\bar{\theta}(t)$ and $O^{\theta_n}(t)$ between time $t \in [1, 2)$ and time $t \in [2, 3)$. We sample from $N = 1000$ graphs and for each graph we obtain a Monte-Carlo estimate of the expectations through 50 simulations. The results of the BIC model selection in the case of n-connected graphs is given on Figure 6.6 for both the Gaussian and Lévy noise cases.

6.4.3 Observations

A summary of the best fitting models for both the 1-connected graphs in Fig 6.5 and the n-connected graphs in Fig 6.6 is presented on Table 6.2 and further summarised on Table 6.3 where we count the frequency of the explanatory statistics amongst the best fitting models.

We first note that there are poor BIC values for all models for the median of the average frequency for $t \in [2, 3)$. This result is not surprising as all oscillators are set to have zero natural frequency thus we expect that $\mathbb{E}[\bar{\theta}(3)] - \mathbb{E}[\bar{\theta}(2)] = 0$, hence no statistics can be significant predictors for this response variable.

For 1-connected graphs, we observe generally good BIC scores (low) in both the Gaussian and Lévy noise cases. In this case, Gaussian noise models contain generally more explanatory statistics and lower BIC scores than the Lévy noise models. However, when we move to the case of n-connected graphs, the Gaussian models are seen to score similarly to that of the Lévy models. With the additional graph statistics available in the case of n-connected graphs, and the overall increase in connectivity of the graphs when we move from the 1-connected to n-connected case, the Lévy models are found to be more descriptive in the n-connected case.

In terms of the explanatory statistics in the best fitting models, the first observation that can be made is that the degree is relevant in all models with explanatory power and this holds for both 1 and n-connected graphs. This observation confirms the original approximation that was established earlier on in this chapter where we determined that the degree of the graph is a significant predictor of graph dynamics. The closeness centrality of the conditioned node g_i was found to be present in the majority of the

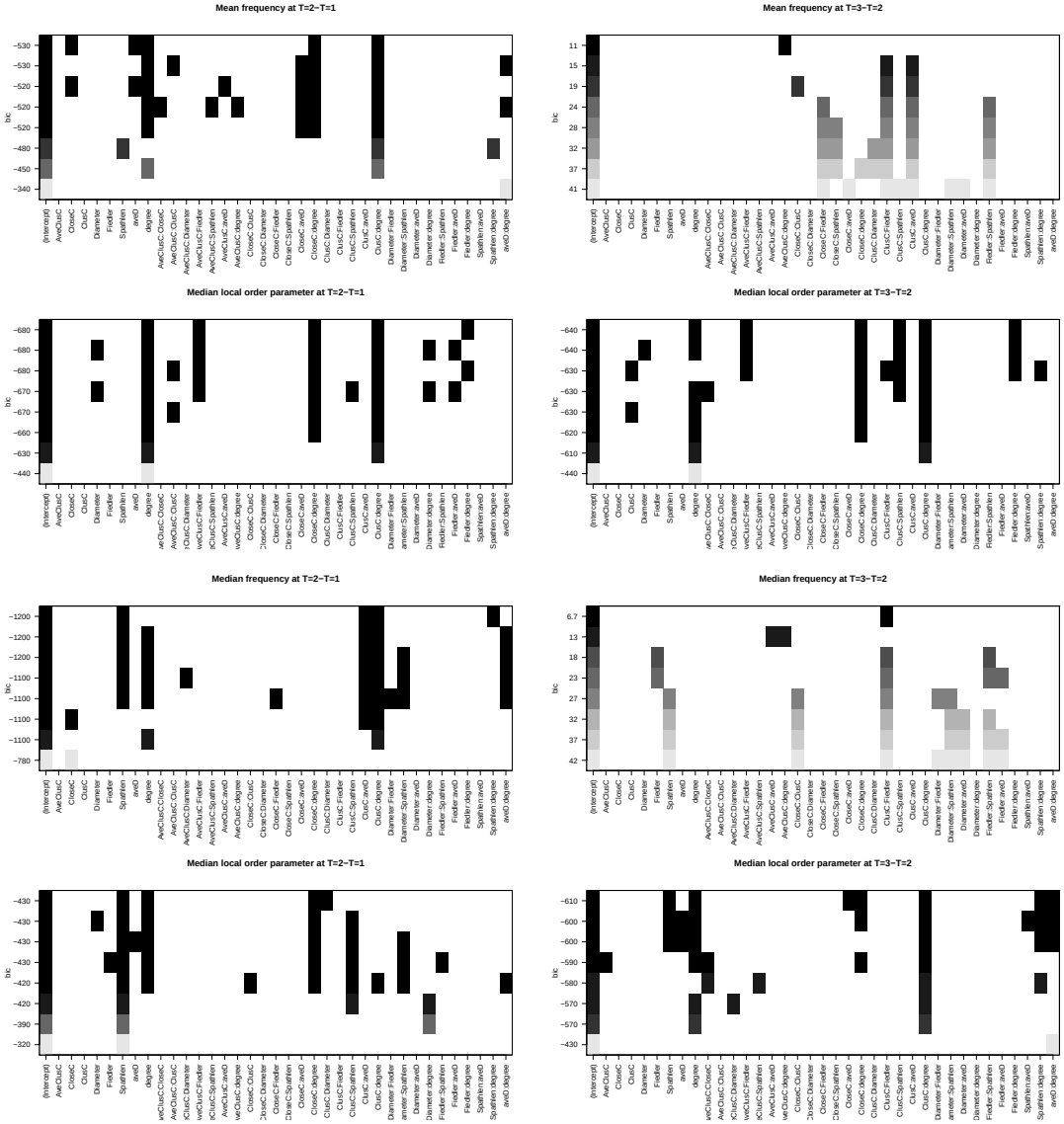


FIGURE 6.6: Model selection for n-connected graphs under Gaussian noise (top) and Lévy noise (bottom). Here AveClusC denotes the average clustering coefficient, CloseC denotes the closeness centrality, ClusC denotes the clustering coefficient, Diameter denotes the Diameter, Fiedler denotes the Fiedler value, Spathlen denotes the shortest path length, aveD denotes the average degree and degree denotes the degree. The symbol ':' denotes the interaction between the two relevant predictors

best fitting models in both the 1 and n-connected cases. In the case of the n-connected graphs, the clustering coefficient of the conditioned node C_i was present in the majority of the best fitting models. It is less clear as to what other statistics are useful as predictors for graph behaviour. The average shortest path length also appeared to be a useful predictor in both cases however this was not as prevalent as the previously mentioned statistics. The remaining statistics, though occasionally present in some of the best fitting models, cannot be confidently declared as useful predictors in an overall sense.

In conclusion, we have shown that the degree of the conditioned node was shown to be a significant predictor in both the approximating SDE of the conditioned processes and the numerical models obtained through our BIC selection procedure. In addition to the degree of the conditioned node, our analysis of the various graph statistics suggest that the closeness centrality and clustering coefficient could be said to be potentially useful predictors.

1-connected graphs			
Response	Time	Noise	Model
median($\bar{\theta}(t)$)	$t \in [1,2)$	Gaussian	$\beta_I + \beta_{g_i} g_i + \beta_{\langle l \rangle} \langle l \rangle + \beta_{d_i} d_i + \beta_{d_i,g_i} d_i \cdot g_i + \beta_{d_i,\langle l \rangle} d_i \cdot \langle l \rangle$
		Lévy	$\beta_I + \beta_{d_i} d_i + \beta_{g_i,\langle l \rangle} g_i \cdot \langle l \rangle$
	$t \in [2,3)$	Gaussian	$\beta_I + \beta_D D$
		Lévy	$\beta_I + \beta_{\langle l \rangle,\lambda_2} \langle l \rangle \cdot \lambda_2$
median($O^{\theta_n}(t)$)	$t \in [1,2)$	Gaussian	$\beta_I + \beta_{g_i} g_i + \beta_{\lambda_2} \lambda_2 + \beta_{\langle l \rangle} \langle l \rangle + \beta_{d_i} d_i + \beta_{g_i,D} g_i \cdot D$ $+\beta_{g_i,\langle l \rangle} g_i \cdot \langle l \rangle + \beta_{D,\langle l \rangle} D \cdot \langle l \rangle + \beta_{D,d_i} D \cdot d_i$
		Lévy	$\beta_I + \beta_{g_i} g_i + \beta_{g_i,d_i} g_i \cdot d_i + \beta_{d_i,D} d_i \cdot D$
	$t \in [2,3)$	Gaussian	$\beta_I + \beta_{g_i} g_i + \beta_{\lambda_2} \lambda_2 + \beta_{d_i} d_i + \beta_{g_i,\lambda_2} g_i \cdot \lambda_2 + \beta_{g_i,d_i} g_i \cdot d_i$ $+\beta_{\lambda_2,g_i} \lambda_2 \cdot g_i + \beta_{\lambda_2,d_i} \lambda_2 \cdot d_i + \beta_{\langle l \rangle,d_i} \langle l \rangle \cdot d_i$
		Lévy	$\beta_I + \beta_{d_i} d_i$
n-connected graphs			
Statistic	Time	Noise	Model
median($\bar{\theta}(t)$)	$t \in [1,2)$	Gaussian	$\beta_I + \beta_{g_i} g_i + \beta_{\langle d \rangle} \langle d \rangle + \beta_{d_i} d_i + \beta_{g_i,d_i} g_i \cdot d_i + \beta_{C_i,d_i} C_i \cdot d_i$
		Lévy	$\beta_I + \beta_{\langle l \rangle} \langle l \rangle + \beta_{C_i,\langle d \rangle} C_i \cdot \langle d \rangle + \beta_{C_i,d_i} C_i \cdot d_i + \beta_{\langle l \rangle,d_i} \langle l \rangle \cdot d_i$
	$t \in [2,3)$	Gaussian	$\beta_I + \beta_{\langle C \rangle,d_i} \langle C \rangle \cdot d_i$
		Lévy	$\beta_I + \beta_{C_i,\lambda_2} C_i \cdot \lambda_2$
median($O^{\theta_n}(t)$)	$t \in [1,2)$	Gaussian	$\beta_I + \beta_{d_i} d_i + \beta_{\langle C \rangle,\lambda_2} \langle C \rangle \cdot \lambda_2$
		Lévy	$\beta_I + \beta_{\langle l \rangle} \langle l \rangle + \beta_{d_i} d_i + \beta_{g_i,d_i} g_i \cdot d_i + \beta_{C_i,D} C_i \cdot D$
	$t \in [2,3)$	Gaussian	$\beta_I + \beta_{d_i} d_i + \beta_{\langle C \rangle,\lambda_2} \langle C \rangle \cdot \lambda_2 + \beta_{g_i,d_i} g_i \cdot d_i$ $+\beta_{C_i\langle l \rangle} C_i \cdot \langle l \rangle + \beta_{C_i,d_i} C_i \cdot d_i + \beta_{\lambda_2,d_i} \lambda_2 \cdot d_i$
		Lévy	$\beta_I + \beta_{\langle l \rangle} \langle l \rangle + \beta_{d_i} d_i + \beta_{d_i} d_i + \beta_{g_i,d_i} g_i \cdot d_i$ $+\beta_{g_i,\langle d \rangle} g_i \cdot \langle d \rangle + \beta_{C_i,d_i} C_i \cdot d_i + \beta_{\langle l \rangle,d_i} \langle l \rangle \cdot d_i + \beta_{\langle d \rangle,d_i} \langle d \rangle \cdot d_i$

TABLE 6.2: Summary of best fitting models for the 1-connected graphs from Fig 6.5 and the n-connected graphs from Fig 6.6. Here we use the following notation: intercept I, degree d_i, average degree $\langle d \rangle$, Fiedler value λ_2, diameter D, average shortest path length $\langle l \rangle$, closeness centrality g_i, clustering coefficient C_i and the average clustering coefficient $\langle C \rangle$.

	Graph Statistic frequency				
Statistic	1-connected Gaussian	n-connected Gaussian	1-connected Lévy	n-connected Lévy	Prevalence over all models
d_i	3	3	3	3	100 %
$\langle d \rangle$	-	1	-	2	50 %
λ_2	2	2	0	0	33 %
D	1	0	1	1	25 %
$\langle l \rangle$	3	1	1	3	66 %
g_i	3	2	2	3	83 %
C_i	-	2	-	3	83 %
$\langle C \rangle$	-	2	-	0	33 %

TABLE 6.3: Graph statistic frequency amongst the best fitting models excluding the median($\bar{\theta}_t$) response at $T = 3 - T = 2$

Chapter 7

Modelling Military Watch Staff Using Stochastic Network Synchronisation

In this chapter we advance an approach to mathematically represent networked human decision makers by adapting the famous Kuramoto model of networked phase oscillators and applying data from a recent study of military headquarters staff. This work was done collaboratively with the DSTG which lead to the paper titled 'Modelling Military Watch Staff Using Stochastic Network Synchronisation' in [23]. In this chapter, we present the paper in full. My personal contribution to this collaborative work is in the development of the numerical methods used to simulate and analyse the stochastic versions of the model, the production of plots throughout the paper and in the editing of the manuscript.

The model that we consider in this work represents a continuous Perception-Action cycle loop of individual agents through a phase at each network node; the network represents the formal and informal communications of human agents and information artefacts they use; a coupling represents the strength of relationship between any two agents; native frequencies represent the individual decision speeds of agents when left to themselves; and stochasticity represent further randomness in the intrinsic temporal behaviour of agents. In particular, heavy tailed noise captures that agents may randomly 'jump' ahead in their decision state in time-stressed conditions where intuition becomes significant. Drawing upon data from a case study where headquarters' networks were mapped for routine business and crisis scenarios, we tune the model using the routine scenario data to match the observation that operations were well synchronised under

such conditions. We then test the impact of the crisis scenario using the corresponding network and stochastic process in the model. We observe a high probability that synchronisation may be lost as a consequence of the high information demand of some agents in the crisis scenario. By comparing with qualitative theories of organisational design, this provides face-validation of the model. We examine the causes of this loss of synchronisation and compare to qualitative approaches in the literature to validate the model. We discuss the model's utility for organisational redesign and technology insertion.

7.1 Introduction

In much of the literature on collective decision making in organisations there is agreement that the structure of interactions between members plays a significant role in the efficiency and effectiveness of the decision-making; by efficiency we refer to the use of resources (here the members), and by effectiveness the outcomes of the decision making. However, real human organisations reflect much more than a structure, but also the heterogeneity of members, such as individual capacities to make constituent decisions, and that these properties may also be dependent on the context of decisions at any one time. In this paper we bring together a number of these elements into a mathematical formulation as coupled differential equations and show how such a model can be used to infer the contribution of these elements to the efficiency of the collective decision making of an organisation. We focus on efficiency here, rather than effectiveness, (because the model is not yet embedded in a representation of the context such that the outcomes can be examined). The model represents networked decision makers as interacting agents engaged in stochastic individual decision cycles. The efficiency of the decision making is then analysed by the degree of synchronisation between the individual decision cycles. Thus the key elements of this mathematical model will be the network and the stochastic dynamical equations. The utility of such a model is in enabling exploration of options for organisational structure in light of the heterogeneity of members of a particular organisation. The context of this work is decision making in military headquarters, a concern in what is called 'Command and Control' (C2), generally the means by which authority is delegated through an organisation to accomplish its missions. We draw upon data collected in a study of C2 within a subgroup in an Australian Defence Force (ADF) headquarters that works somewhere within the spectrum between strategic (based in the capital city) and tactical (resident in individual task groups or elements in a particular geographic area). The actual study from which data was collected, reported in [18], was not concerned specifically with synchronisation of decision making, but it nevertheless afforded an opportunity to collect network data of human-to-human

and human-to-information artefact interactions. This particular staff group supports the twenty-four hour commander's decision making cycle, that is informed by a formal set of briefs at the start of the day. A key group within this are staff that constitute a twenty-four/seven 'watch' who interact with each other, with a range of systems or information artefacts, and other staff who work more regular business hours. There is nothing in this that is necessarily unique to the military. At this level of fidelity many civilian settings, such as emergency management, air-traffic and energy sector control may be modelled in this way. Formal and informal organisational structures may be encoded in the mathematical object of a network, human and technology agents who have varying decision making characteristics may be included, and there are different circumstances, from routine business to specific crisis response, thay may trigger activity. Thus, the utility of our modelling approach transcends military operations. All of these contexts already see a diversity of approaches to improve efficiency or effectiveness. Most common is the Business Process Modelling (BPM) approach[131] where graphically represented workflows are encoded into formal or executable models. Agent Based Modelling (ABM) is also prevalent, where individual organisation member properties and local interactions are represented; the organisational behaviour is then an emergent product of bottom-up interactions. Finally, Social Network Analysis (SNA) is also common, where both formal and information interactions are encoded. For a review of these types of approaches, see [132]. Quite uncommon in this context are stochastic differential equation based models. We present such a model in this paper. Though we argue that the diversity of approaches is absolutely necessary for any validation [133, 134], the modelling approach in this paper has some of the advantages of ABM and SNA over BPM without some of the disadvantages. Bottom-up modelling reflects the emergence of how business is dynamically performed in complex environments. The ABM approach, while extremely flexible, often does not allow for insight; more stylised and abstract mathematical modelling approaches often allow approximations in interesting regimes that are amenable to closed form solution, yielding explicit expressions for the relationship between factors. SNA traditionally fails to capture the dynamical process that occurs on the network. Thus, our approach here, while abstracting significantly, offers a parsimonius but insightful approach to organisational dynamical behaviours. In this paper we seek to validate the model and develop insights into the collective behaviours of the organisation as the style of individual decision making is varied. The data collected in the study [18] mapped out a headquarters' network of interactions for actual routine business and under a hypothetical crisis scenario. In addition to different networks, the decision making style will be different for the two contexts according to cognition models to be discussed below. Thus, to the degree that the stochastic process we employ captures the tendencies within different decision making styles, we

may build such styles into the mathematical model. Specifically, this stochastic process enables representation of a mode of decision making under uncertainty, whereby experienced decision makers will use intuition to jump ahead of an otherwise sequential process. We use the routine business data to tune the model and then run the model for the crisis scenario in order to perform validation. However, because the crisis scenario was not run as an exercise or experiment (this was an operational headquarters) we have no performance data with which to compare the model. Therefore we must use insights from qualitative research as to what the expected outcome may have been. In particular, the phenomenon of uncertainty-reduction behaviour shows that decision makers tend to seek more sources of information in a crisis; a by-product of this is they may expose themselves to information deluge and decision paralysis, on the one side, and/or fail to carry their less connected team members, on the other. We show how the model behaviours in the crisis may be interpreted as reflecting such behaviour, and to this degree we are able to validate the model. From this point we explore how varying the parameters that encode the degree of jumping in decision making further impacts on behaviours. This chapter is organised as follows. First we explain the model in more detail, including its mathematical structure. We explain the data and conduct the model tuning using data around the routine conduct of business. We then conduct the validation by applying data around a crisis scenario, and comparing the behaviours of the model against known properties of organisational behaviours in situations of threat or uncertainty. We then illustrate how the model provides deeper insight into the causes of these behaviours. We conclude with future directions of work.

7.2 The model ingredients

7.2.1 The mathematical model

The key element in our approach is what is known as the stochastic Kuramoto model [46] defined as the equation for the time-derivative $\dot{\beta}_i = \mathrm{d}\beta_i/dt$ of a set of real time-dependent variables $\beta_i = \beta_i(t)$ at time t:

$$\dot{\beta}_i(t) = \omega_i - \sigma_G \sum_{j=1}^{N} A_{ij} \sin(\beta_i(t) - \beta_j(t)) + \gamma \dot{L}_i(t), \qquad i = \{1, \ldots, N\}\,. \tag{7.1}$$

With $\gamma = 0$, Eq. (7.1) is a network generalisation of the classical Kuramoto model [1]. So, A_{ij} is a matrix of ones and zeroes representing the connection, or lack thereof, of nodes labelled by the index i, on a network of size N. Generally undirected networks, where A_{ij} is symmetric, are used; here we apply directed graphs, for which the model

remains consistent but with slightly different properties. The ω_i are constants, often drawn statistically, representing native or intrinsic frequencies or speeds of oscillation. The interaction function as the sine of the phase differences is what makes the β_i true phases: though the β_i can take their values on the real number line, in interactions only the value modulo 2π is relevant. So, the interaction models local synchronisation between connected agents but also allows for anti-synchronisation when agents are an interval π apart. The quantity σ_G is a coupling strength governing the intensity of interaction of connected phases.

Many reviews of the Kuramoto model exist, see [35–39]. But the key property that it exhibits is that at low σ_G phases behave chaotically, at high σ_G they collectively synchronise, and there is some intermediate value at which this behaviour changes over. For undirected networks the collective synchronisation is to the mean of the frequency ensemble, $\theta_i \approx \bar{\omega}t$ for all i. For directed graphs this result receives some corrections from individual degrees of the nodes. Many studies focus on the role of various heterogeneous random network structures in impeding or enhancing synchronisation at lower coupling strengths. Here, structures include the Barabasi-Alberts 'scale free' network [135] studied for the Kuramoto model in [136–139], and the small world network [140] typical of social structures [141].

For $\gamma = 1$, there is the additional term L_i in Eq. (7.1) which is 'noise', basically a function of time whose values at different time-instances satisfy certain probabilistic characteristics. Gaussian-like idealisations of this stochasticity - where the increments of L_i are drawn from a Gaussian distribution and are uncorrelated in time - have been considered [36, 46–51]. However, for reasons to become clear below, heavy-tail distributions [54, 55, 142] are also of value particularly for decision-making under uncertainty, with a first exploration of this for the Kuramoto model in [52, 143].

7.2.2 Mapping the mathematical model to decision making

How does such an equation as Eq. (7.1) model organisational decision making? As originally proposed in [144], we take the phase β_i to represent the point of an individual in a continuous decision, or Perception-Action [145], cycle; the native frequency represents the frequency of that cycle when a decision maker is left to themselves, A_{ij} represents the combined formal-informal-informational network, where we include both human (which have $\omega_i \neq 0$) and information artefacts (which have $\omega_i = 0$, in the absence of autonomy or'self-updating'); σ_G represents the intensity of the relationship between agents; the sine interaction represents the requirement that connected decision makers will adjust their decision state according to that of their connected partner; and noise represents the

property that further heterogeneity between decision makers may also be random and time-dependent. Let us expand a little more on some of these features. While seemingly simplistic, many descriptive cognitive models of individual decision making are fundamentally cyclic in nature and indeed more simplistic in turning a continuous process (as would be reflected in a phase variable) into a repetitive sequence of discrete states; indeed this has been argued to be problematic when seeking to consider hybrid systems of humans and autonomous technological agents [146]. The Neisser model we have cited already [145], but there is also the Endsley model of situation awareness [147] and various Levels or Discretised systems for considering human-autonomy hybrid systems [148]; finally, in the tactical military context there is the Observe-Orient-Decide-Act (OODA) loop model of John Boyd [149, 150], which has become adopted for strategic and broader civilian decision making contexts [151, 152]. The notion of coupling has also deep roots in the organisational/sensemaking literature. Weick [153] in the seventies accumulated some fifteen concepts related to "loose coupling", they may be essentially aggregated to two: low connectivity, and low intensity along connectivity. Many approaches have only taken the first notion of coupling forward, as reflected in network theory, to conduct measurement and experiments in organised decision making settings [154]; a small subset have pursued the importance of intensity of interactions or degree of slackness on a network [155]. In this respect, the Kuramoto model elegantly provides a framework for both notions of coupling.

Stochastic diffusion (Gaussian) processes have already been used to model elementary human decision making [156]. The choice of heavy-tail noise is intended in our work to build in to the mathematical representation models of decision making in complex and uncertain environments, namely so-called Wicked Problems [157] and Recognition Primed, or Naturalistic Decision Making [158, 159]. These respectively hold that in the face of complexity the decision process is (among other things) 'jumpy', and that under uncertainty, experienced decision makers do not incrementally advance through a logical process but observe patterns in the environment that prime their intuition allowing them to leap ahead in the decision process. Heavy tail skewed noise, a variation of fractional non-Gaussian diffusion [54, 55, 142] builds in precisely such forward leaning jumps. Indeed, we employ a form of this noise known as the tempered stable Lévy process [91–94] described by parameters $\sigma, \alpha, \theta, \lambda$ that control the spread (in the Gaussian limit, the variance), scaling, skew and large size tempering of such jumps. This class of processes is attractive from a modelling point of view, compared to the stable process, due to the fact that the process has moments of all order and thus can be more easily fitted to experimental data [160] when available. Thus the use of the tempered stable process has appeared in areas ranging from finance[161], to insurance [162, 163], to biology [164], acoustics [165], and physics [94].

When $\alpha = 2$ the noise is purely Gaussian with variance σ. For $\alpha < 2, \lambda = 0$, so-called stable Lévy noise, the jumps follow a power law tail with jumps of size δ occuring according to a distribution $\delta^{-\alpha}$. An initial study for this type of noise applied to the data from [18] was undertaken in [166]. The parameter λ allows the very large jumps then to be tempered exponentially with distribution $e^{-\delta\lambda}\delta^{-\alpha}$. Intuitively speaking, experienced rational decision makers will jump ahead in a single decision cycle but not over multiple cycles. Thus $\alpha < 2$ and $\lambda > 0$ allow the strength of the jumps to be moderated to potentially reflect realistic intuitive decision making.

7.2.3 Measures of synchronisation

To measure the degree of synchronisation within a given population we use local forms of Kuramoto's global order parameter [1]. The global form is computed as

$$r_G = \frac{1}{N}\left|\sum_{j\in\mathcal{G}} e^{i\beta_j}\right|, \tag{7.2}$$

where N is the total number of nodes on the network $\mathcal{G}$. This measures the amount of synchronisation of the entire network in that when contributing phases are all approximately equal $r \approx 1$. However, the corresponding quantity for sub-networks $\mathcal{G}_s$ of size N_s may also be used to measure local regions of synchronisation:

$$r_{G_s} = \frac{1}{N}\left|\sum_{j\in\mathcal{G}_f} e^{i\beta_j}\right|. \tag{7.3}$$

We specify below the particular sub-groups we examine.

7.2.4 Stable and tempered stable noise parameterisation

In this work we utilise both a totally positively skewed stable process and totally positively skewed tempered stable process. The simulation of the variates of these processes are previously discussed in Section 2.5.4 where the characteristic functions of the total positively skewed stable and tempered stable processes are given in Eq. (2.9) and Eq. (2.10) respectively. In the simulation of these processes, we shall take the parameterisation

$$a := -\frac{1}{2}\sigma^2 \frac{1}{\Gamma(-\alpha)\cos(\pi\frac{\alpha}{2})}$$

so that $\Psi_t(z) \to \exp(-\frac{1}{2}\sigma^2 z^2)$ as $\alpha \to 2$ to recover the Gaussian distribution in the limit.

In Figure 7.1 we plot the probability densities $f(x)$ for two values of α overlaying the cases of different values of λ. We observe in the lower plot the characteristic property that for $\alpha > 1$ the density shows the one-sided heavy tail in the positive direction (seen by zooming in, as in the insets) but that the curve is smooth around the mode at $x = 0$, with near Gaussian tails in the negative direction. Contrastingly, for $\alpha < 1$ a cusp develops at the mode at $x = 0$ with the one-sided heavy tail in the positive direction. Thus for $\alpha > 1$ the mean of the density vanishes, but for $\alpha < 1$ it is non-zero, reflecting the shift in the tempered stable density. In either case, as tempering λ increases the tail is suppressed with the density nevertheless retaining its distinctive shape.

7.3 The context of model and data

7.3.1 On the availability of data for the model

For many of the aspects we have described here there is no presently available quantitative data. While techniques for determining networks are long-standing and decision speed for specific tasks may be measured, many of the other aspects lack quantitative measurement. In particular, as mentioned, though a quantitative model for diffusive decision making exists this is only for elementary binary decisions [156], no fitting of the Recognition Primed model [158] to the parameters of a mathematical stochastic process exists. One motivation for the present work is to encourage Human Factors research to collect such data.

7.3.2 Staff structures and information artefacts

As alluded already, we apply the mathematical formulation to staff and systems in a military headquarters at an echelon within the tactical to strategic span. Organisationally the staff are members of two separate branches labelled using the Common Joint Staff System (CJSS):

- The J2 is concerned with Intelligence, namely the analysis (not collection) of information on potential threats, adversaries and the environment of military operations (including local populations); and

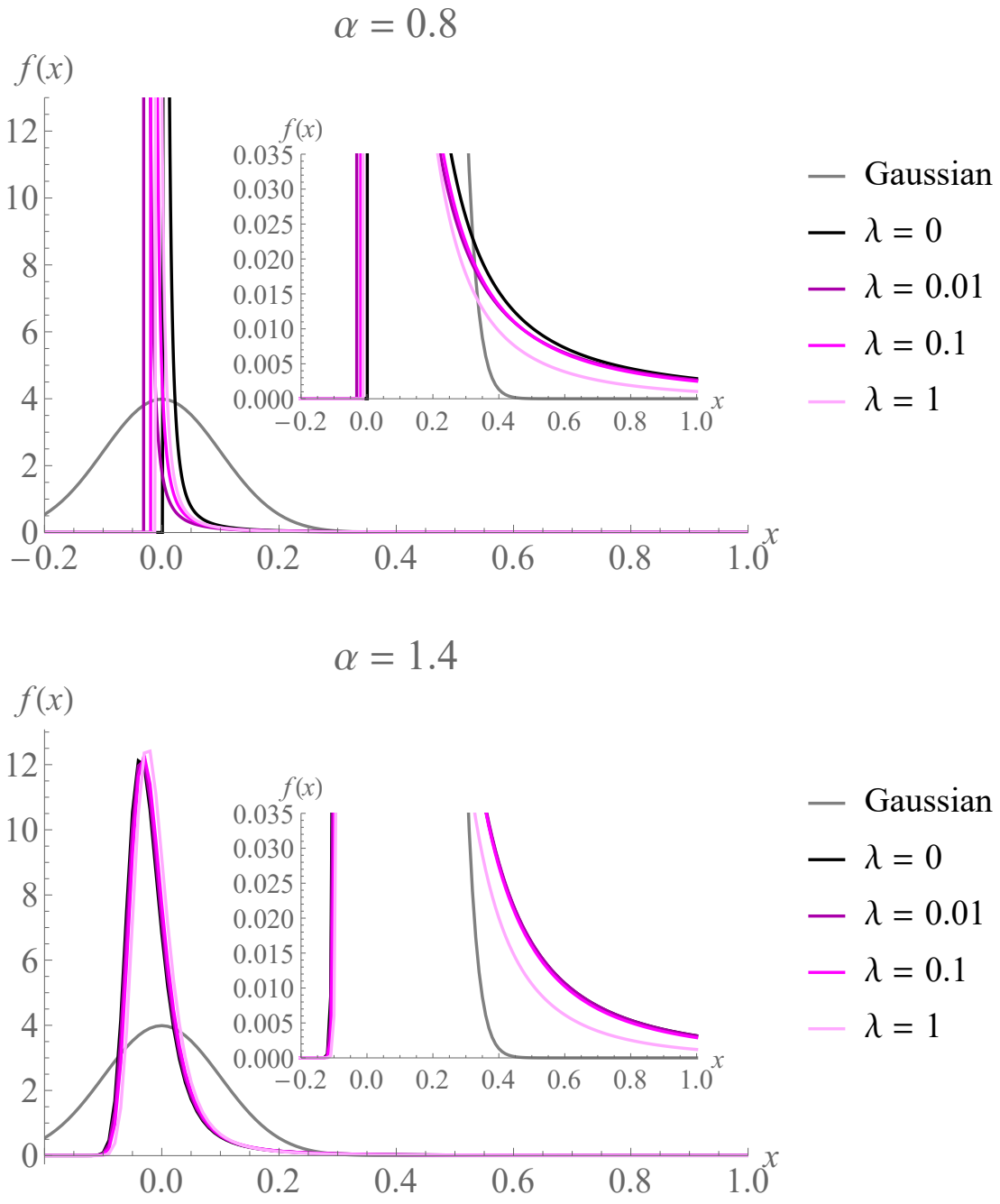


FIGURE 7.1: Plots of the density $f(x)$ for centred noise, top for $\alpha = 0.8$ and bottom for $\alpha = 1.4$ for varying λ and the Gaussian case for $\sigma = 0.1$; in the inset we zoom in close to the horizontal axis to separate the different tail behaviours.

- The J3 is concerned with current operations, including the status of own forces, events in the immediate past, and plans on controlling movements of task units for the immediate future (approximately the next 24-48 hours).

We forego the labels for other branches in this system [167]. Much of what these staff prepare is presented to the Commander and other branch heads at an early morning brief; the shorter term decisions flowing from this brief are then implemented by the J3 staff, or questions requiring further analysis are pursued by the J2 staff. The staff we represent in the model are focused on the twenty-four hour Commander's 'Battle Rhythm' [168]. This means that many of them work within a shift roster with twelve hour shifts while others work roughly business hours conducting deeper analysis but responsive to the daily cycle. Either way, *collective synchronisation to the twenty-four hour cycle is a requirement for efficient operation of this staff.*

In addition to human staff there are a multitude of information systems and artefacts that are fundamental to the work: either extracting information from a system that may be a visual display on a computer monitor, or a text document or a set of briefing slides, or populating such a system with information, preparing or amending said documents and briefing slides. In this context, the systems and artefacts are human-driven: they do not automatically update based on changing sensor data. They must be up-to-date to correctly inform the commander's briefing cycle. They therefore play a role in supporting or undermining the collective synchronisation of the socio-technical system that is the headquarters. In Tables 7.1 and 7.2 we provide a detailed list and description of the staff and systems.

As reported in [18], an instrument was developed to determine the 'sources and sinks' of information of this staff for their work. There we distinguished information pull, when staff draw from another human or from an information system, and information push when they initiate information to another human or populate information into an artefact or system. Regardless of the initiation of the information request, a direction of flow is present in both cases. Here we combine unweighted pull and push networks from [18] by computing a directed adjacency matrix

$$A_{ij} = \lceil A_{ij}^{(\text{pull})} \rceil$$

where $\lceil \cdot \rceil$ represents the ceiling function, and the index order represents flow of information from i to j. For the routine and crisis scenarios we obtained the directed graphs shown in Figure 7.2. Of course, a pure human-to-human network may be derived from this through the technique of multiplying the associated adjacency matrices with their transpose and projecting onto only human agents; this captures human-to-human via

Staff	Description
OCmnd	Commander or branch heads
OJ3W1	Senior staff in J3 Watch
OJ3W2	Junior Officers/NCOs in J3 Watch
OJ3WS	Support Officers in J3 Watch
OJ3S	Support Officers in J3
OJ2W1	Senior Officers in J2 Watch
OJ2W2	Junior Officers/NCOs in J2 Watch
OJ2A	Analysts in J2
OJ2S	Support Officers/NCOs in J2
OStrat	Higher Headquarters staff
OExt	Overseas Organisations
OInt	Other Intelligence Organisations
OJTF	Lower Headquarters
OSS1	Single Service (Army/Navy/Air Force)
OWOG	Non-Defence Government Staff

TABLE 7.1: Staff related to the headquarters; 'support' refers to ancillary functions such as logistics, environmental, personnel management, health and computer systems.

Systems	Description
Pemail	Formal email with information
Pweb	Intranet documents
Pbrief	Fused J3-J2 brief
PCOP	Display of Common Operating Picture
PJ3Brief	Routine briefs from J3 staff
PJ3Ord	Orders issued by J3 Watch
PRps	Formal Reports by military staff
PJ2WBrief	Routine briefs of J2 Watch
Padhoc	Ad hoc document or slides on current events
PJ2AnBrief	Analysis brief by J2
PThrAss	Threat Assessments
POpen	Open source internet information
PEnv	Reports on physical environmental conditions
PExt	Reports from external organisations
PInt	Reports from other Intelligence organisations
PJTF	Brief or Signal from Lower Headquarters
Pdef	Signals from single service units
PWOG	Reports from other government departments

TABLE 7.2: Information Systems/Artefacts related to the headquarters.

the intermediation of an information artefact [169]. However, to the degree that the technology behind office documentation - before the advent of Smartphones and Auto-Correct systems - is not self-updating but entirely human driven, the time-cost of human interaction with such information artefacts is an important factor in the efficiency of organisational work. *We therefore impose vanishing native frequencies on information artefacts.* Their incorporation into the mathematical model however means they will be driven in their decision state through the interaction (namely, information push) of human agents.

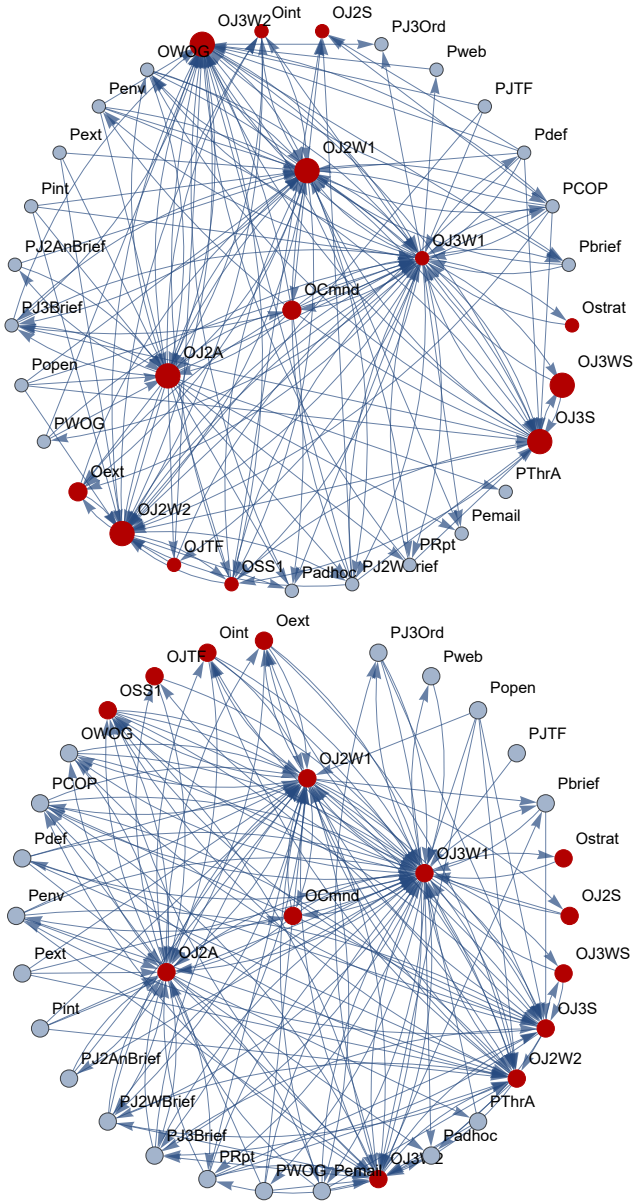


FIGURE 7.2: Networks of human (red nodes, labelled 'O') and information artefacts (grey nodes, labelled 'P') with top the routine scenario, and bottom the crisis scenarios.

7.3.3 Other numerical data

As mentioned, the study in [18] was not specifically focused on synchronisation and therefore data on decision making speeds was not collected. However, given the commonality of the twenty-four hour cycle across all staff in the networks, we may draw ω_i for human agents from uniform distributions and average results over many instances. To that end we normalise to twenty-four hours and draw the specific frequencies of agents in the following way:

- watch staff from a uniform distribution of width $(-0.1, 0.1)$ around $\omega = 1$;
- day staff from a uniform distribution of width $(-0.1, 0.1)$ around $\omega = 0.5$.

This leaves an overall scale that will be fixed by tuning the coupling, as described in the following section.

7.4 Simulation study

7.4.1 Implementation

To simulate the stochastic Kuramoto model we implemented code in C++ for maximum performance and uses the Messaging Passing Interface (MPI) to parallelise the simulation across an arbitrary number of CPU cores. The simulation is performed using an Euler method to simulate the dynamics at the N vertices. This algorithm was chosen as it allows the code to easily switch between the cases of (1) no noise, (2) Gaussian noise, (3) Stable noise, or (4) Tempered Stable noise without changing the underlying algorithm. The code runs on the command line and takes the following parameters: the noise distribution parameters; the number of paths to simulate; the number of time steps in each path; and the filename containing the graph $\mathcal{G}$ stored in graph6 format [106]. In addition, multiple graphs $\mathcal{G}$ can be given so that simulations results can be averaged over ensembles. The code also includes a number of hand-implemented random sampling algorithms for Stable and Tempered Stable distributions. We examine the order parameters for the following groups over multiple paths for different instances of frequencies and noise: the entire networks of Figure 7.2, the J3 and J2 staff separately, including their watch members, and then the watch itself consisting of J3 and J2 watch-keepers. Specifically, we compute expected values, variance and extreme ranges - the difference between maximum and minimum extreme results over time (after discarding an initial transient), after running up to $t = 500$, and over 1000 paths. In the following we use the

routine scenario to tune the model and then explore how these settings impact on the crisis scenario. The principle behind this is as follows. Routine conditions for operations in the headquarters means that staff can adjust the intensity of their interactions over time to achieve acceptable levels of synchronisation within an efficient application of effort; coupling is, after all, an effortful activity [153]. When unanticipated crises strike the initial value of coupling effort will be that achieved during routine conditions; staff will only increase effort in coupling if they detect a loss of synchronisation as they adjust processes to deal with the crisis. It is this potential loss of synchronisation that we use the model to predict.

7.4.2 Routine Scenario: tuning the coupling

We first examine the order parameters Eq. (7.2) and Eq. (7.3) as functions of time. We initially fix the noise parameter $\sigma = 0.1$; we later test for sensitivity to this choice. In Figure 7.3 we show the local order parameter Eq. (7.3) for different staff groupings - the entire staff, the J3 and J2 staff (including their watch sections), and the watch as a whole including both J3 and J2 members - as functions of time for coupling strength $\sigma_G = 0.2$. The solid line in the plots represents the expected value of r over 1000 paths; the orange line, with axis on the right hand side, shows the standard deviation, and the grey region indicates the extreme values. We see that at such a low value of coupling, though the whole organisation and the watch members show reasonably high values of r close to $0.8 - 0.9$, the individual branch groups show considerably more variance. By visual inspection the J3 staff show slightly less synchronisation than the J2 staff. However, comparing these with the relatively high synchronisation of the Watch, we may deduce that the relative lack of synchronisation in J3 and J2 is from the presence of both watch and day staff elements. Increasing the coupling to $\sigma_G = 0.4$ gives a dramatically changed picture in Figure 7.4 Expected values are now consistently in excess of $r = 0.9$. We have zoomed in specifically to better visualise the fluctuations, which are tiny; even the extreme values never drop below $r = 0.85$, indicating a strong consistency in the synchronisation level over time. Thus coupling $\sigma_G = 0.4$ is sufficient to bring all parts of the C2 system to synchronisation. Further increases in σ_G only further increase the average value of r to the value one and shrink the fluctuations.

7.4.3 Sensitivity analysis for tuned coupling value.

We next show an entire landscape of values to test for the sensitivity against the choice of the value of σ by plotting the expected value of $1 - r$ for the various staff groups for various (σ, σ_G) as a surface plot or heat map. This is shown in Figure 7.5. Note that

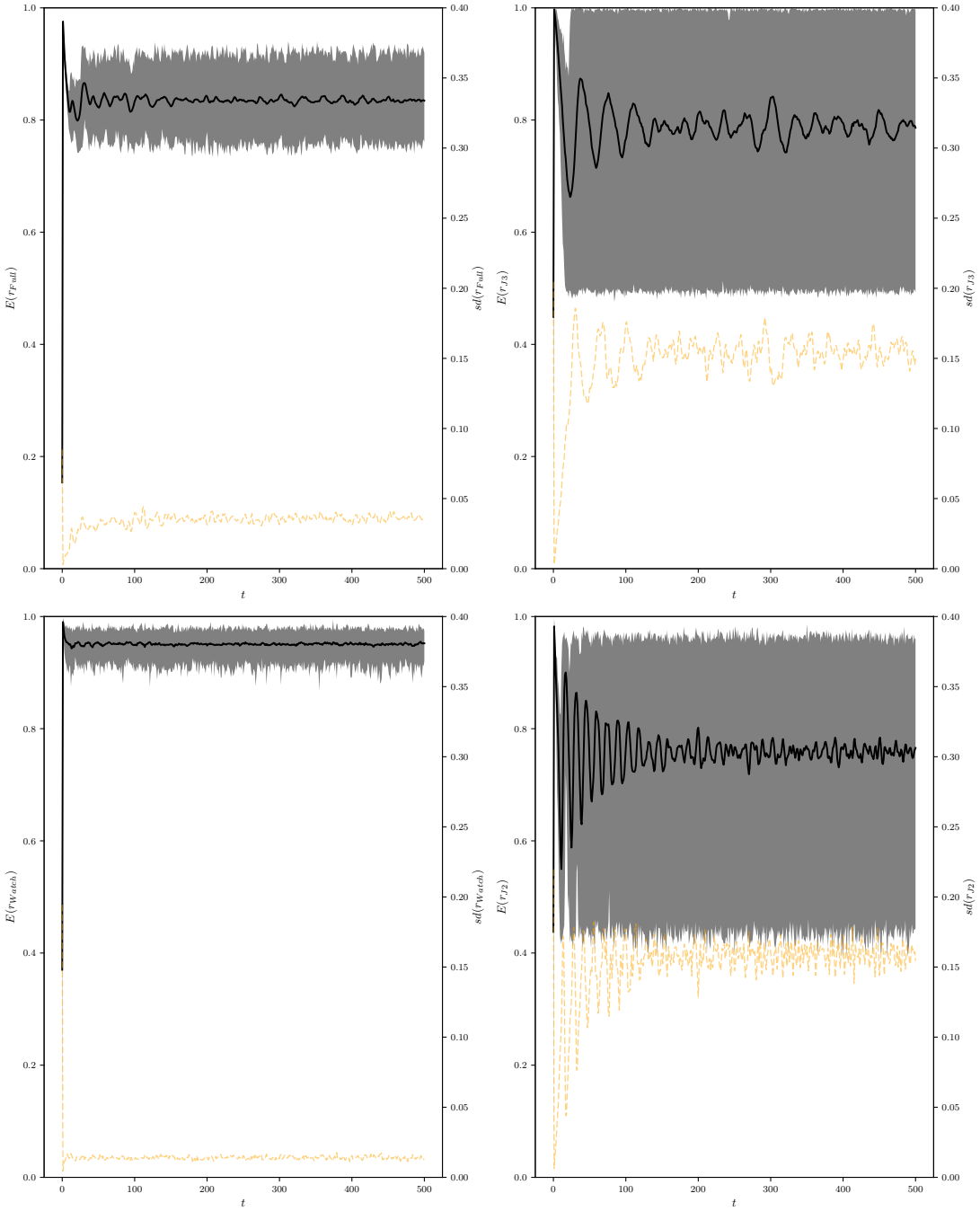


FIGURE 7.3: Order parameter values for the various groups (Full/J3/J2/Watch) in the routine scenario using Gaussian noise for $\sigma = 0.1$ and $\sigma_G = 0.2$. On the left axis we plot the expected value of the order parameter (solid black line) and the extreme values of the order parameter across all paths (grey area). On the right axis we plot the standard deviation of the order parameter across all paths (dashed orange line)

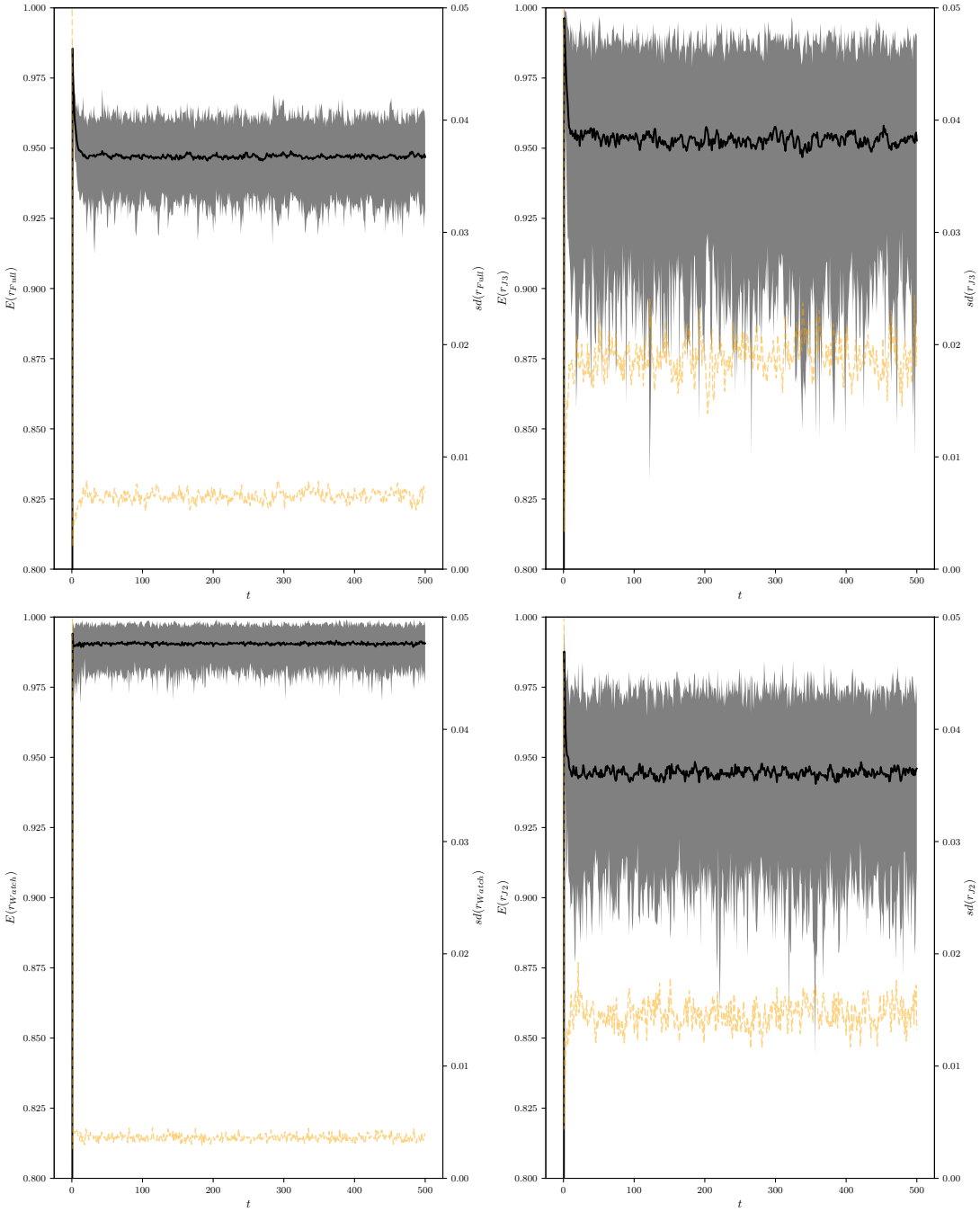


FIGURE 7.4: Order parameter values for the various groups (Full/J3/J2/Watch) in the routine scenario using Gaussian noise for $\sigma = 0.1$ and $\sigma_G = 0.4$. On the left axis we plot the expected value of the order parameter (solid black line) and the extreme values of the order parameter across all paths (grey area). On the right axis we plot the standard deviation of the order parameter across all paths (dashed orange line)

each panel in figures such as this is colour calibrated differently to bring out key features in the separate plots; those comparing colours across panels should be done cautiously. We observe that for $\sigma_G \approx 0.3$ we cross a consistently dark blue horizontal line in the plane for each panel, consistent with $1-r < 0.1$ there words $r > 0.9$. In particular, there is marginal variation in the σ direction. Thus, at least for Gaussian noise, the choice for tuning the coupling σ_G such that synchronisation is achieved in the various sub-groups of the C2 system is indeed insensitive to the variance in the Gaussian noise. So much is based only on the expected value of the order parameters. Normally, the variance over the sample of paths would give a sense of the scale of fluctuations. However, in view of the properties of the Lévy noise to be considered soon, we will instead use the range of the paths, namely the difference between maximum and minimum extreme values. This is shown for the Gaussian noise case in Figure 7.6. For $\sigma_G = 0.4$ we observe that there is somewhat greater sensitivity to the value of σ, with the range increasing with σ. In particular, values of $\sigma_G \geq 0.2$ give ranges in the order parameter greater than 0.4. In particular, for $\sigma_G = 0.4$ as σ varies in the interval $0.1 \geq \sigma \geq 0.5$ the range in r can vary between 0.2 up to 0.5, a considerable variation. At $\sigma_G = 0.6$ this sensitivity is considerably diminished. Thus tuning of the coupling to $\sigma_G = 0.6$ gives both strong synchronisation and demonstrates insensitivity to statistical extreme values under Gaussian noise. To conclude, we tune the model to the routine network data with Gaussian noise with coupling $\sigma_G = 0.6$ and Gaussian noise variance $\sigma = 0.1$ such that the organisation collectively and in its individual groups is synchronised.

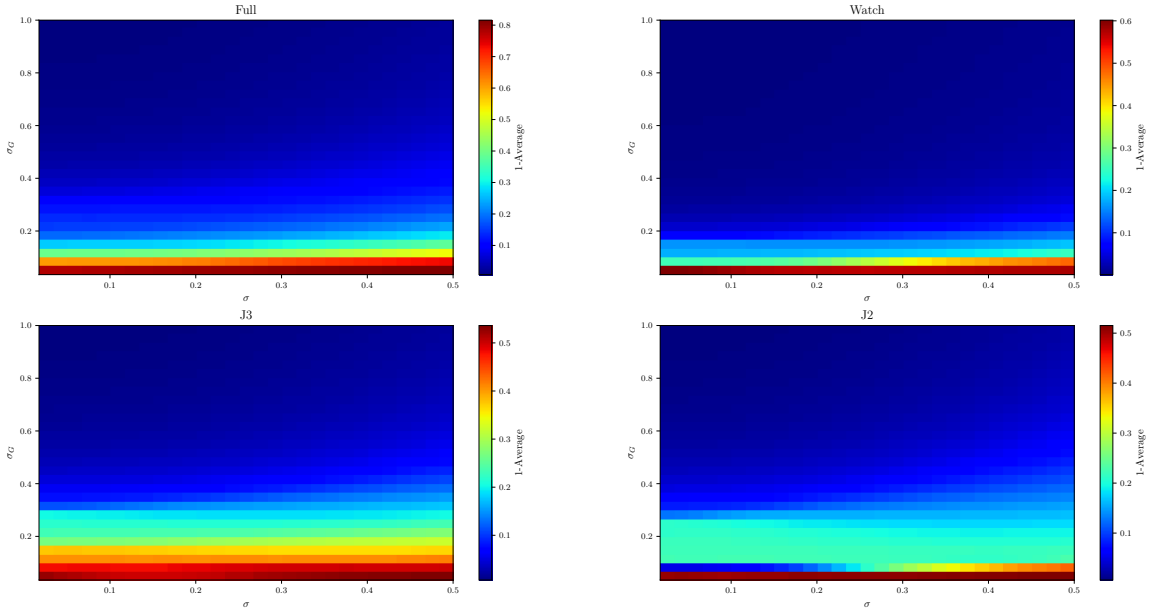


FIGURE 7.5: Variation of the expected value of the average order parameter across the (σ, σ_G) parameter landscape for the Baseline scenario. The simulation was run with 1000 paths up until T=500.

7.4.4 Crisis scenario - Gaussian noise

We now replace the network for routine business with the crisis scenario, namely swap the network in the top panel of Figure 7.2 for that in the bottom panel. Our intention is to also apply tempered stable noise to reflect that in the crisis there is greater uncertainty and therefore individual decision making styles will be more reflective of Recognition Priming [158]. However we first persist with the Gaussian noise for comparison purposes, with the landscape for the expected value of the order parameters as seen in Figure 7.7, where again $1-r$ is plotted. We observe that order parameters values are typically low for low coupling σ_G, particularly for the J3 staff where deviation from perfect synchronisation is strongest. Nonetheless at the tuned value of $\sigma_G = 0.4$, the expected values remain approximately $r = 0.9$. This already suggests that the properties of the network for the crisis scenario are such that synchronisation is less attainable, even with Gaussian noise.

7.4.5 Crisis scenario - tempered stable noise

We now reach the critical test, applying tempered stable noise to the crisis scenario network structure and using the tuned value of coupling $\sigma_G = 0.6$ Here we have fixed σ to its value assumed in the Gaussian case, and set the skewness to its maximum value $\theta = +1$ indicating decision makers jumping forward through the decision cycle.

We must choose a value of α judiciously, though later we will explore the trade-offs between α and λ. To make this choice we use the stable and tempered stable distribution properties discussed in Section 2.4 and plot in Figure 7.8 the probability of a jump x greater than π in one time-step, based on integrating the densities $f(x)$, as a function of the index of the power law, α, and for different values of λ. For the stable case, $\lambda = 0$, we see that the probability increases monotonically with decreasing α, a direct consequence of the changes in the shape of the density reflected in Figure 7.1 We observe that for $\alpha \approx 1.4$ the probability of such jumps are less than one percent. Within the interpretation of the stochastic Kuramoto model as a model for decision makers jumping ahead in their decision cycle we may interpret jumps of more than half a decision cycle as unreasonable. Thus, for the present purpose we take $\alpha = 1.4$. We see with the additional curves in Figure 7.8 that there is a measure of insensitivity of this choice of α to λ for $\lambda < 1$. At $\lambda = 1$ the suppression of the tail is so strong that the probability of large jump is not visible on the scale of the figure. We further probe this insensitivity to λ in the dynamics now by plotting the order parameters for the crisis scenario with purely stable Lévy noise in Figure 7.9. We now see that, though average values and variance remain reasonable, extreme values are considerably larger, reaching down to values of $r \approx 0.5$.

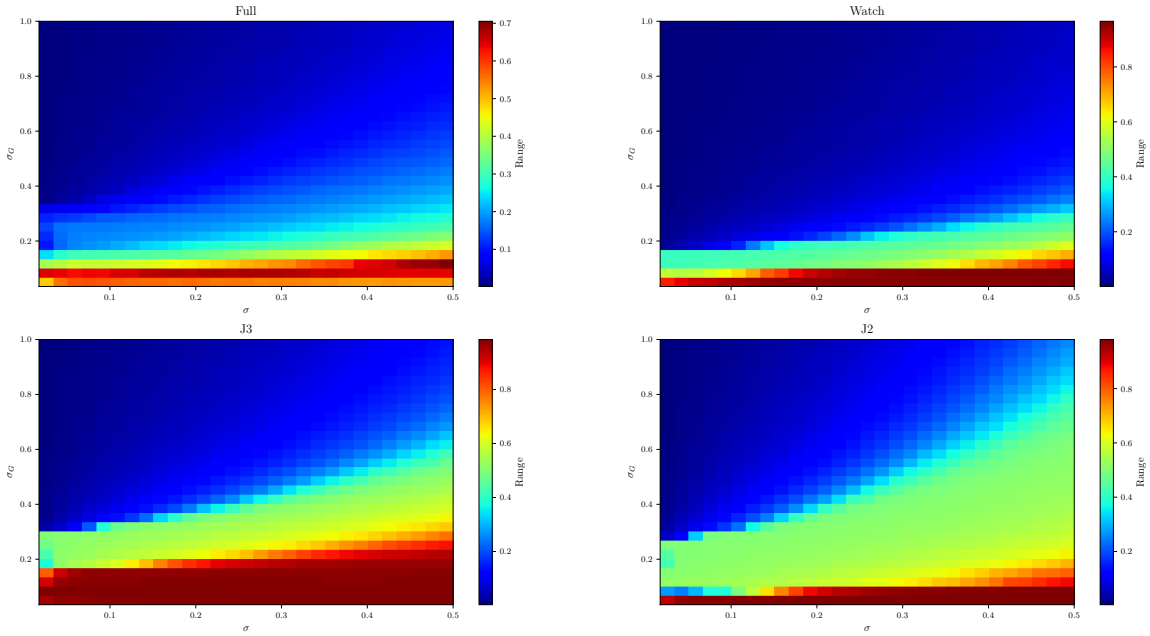


FIGURE 7.6: Variation of the expected value of the range of the order parameter across the (σ, σ_G) parameter landscape for the Baseline scenario. The simulation was run with 1000 paths up until T=500.

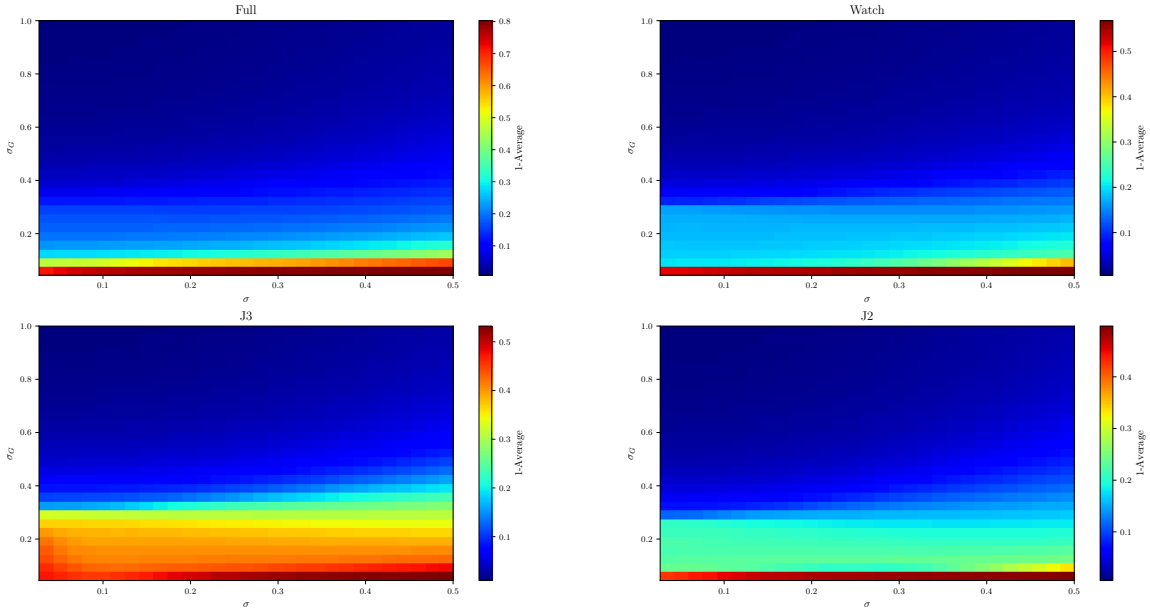


FIGURE 7.7: Average order parameter values for the various groups (Full/J3/J2/Watch) in the crisis scenario using Gaussian noise. The values $1 - r$ are used for the colour scheme.

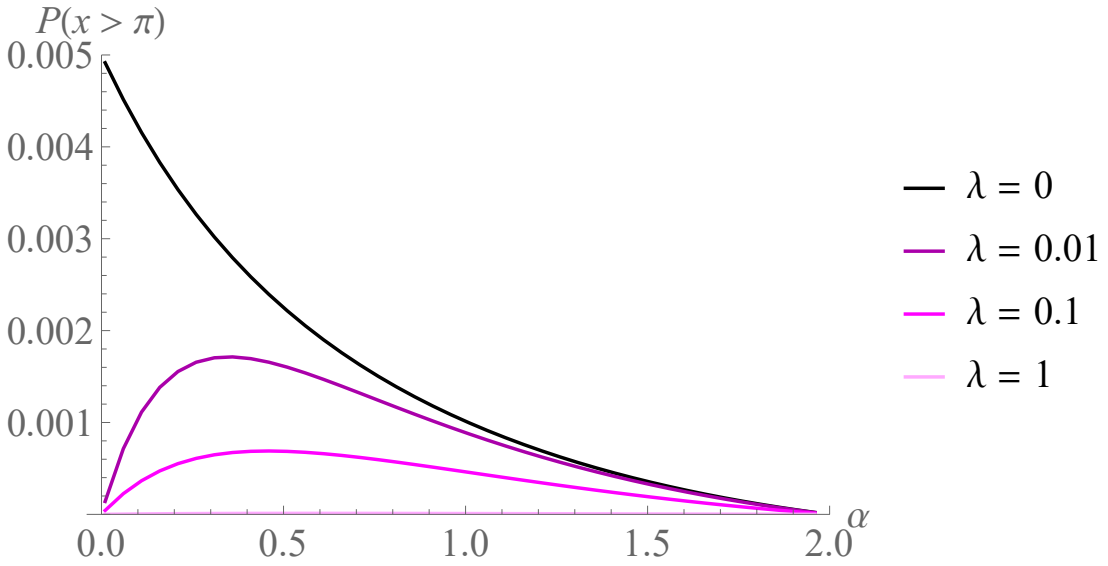


FIGURE 7.8: The probability of a jump of step x greater than π in a single time-step for different tempering.

Thus synchronisation is far from guaranteed from instance to instance. Tempering can moderate the extreme behaviours. For example, with tempered stable noise now for $\lambda = 1$ as shown in Figure 7.10. Here now extreme cases are rarer. Nevertheless, the conclusion stands that given the network structure for the crisis scenario, particularly with individual decision making characterised by skewed heavy- tailed stochastic jumps synchronisation as a robust behaviour is lost at coupling values adopted from routine business.

7.4.6 Sensitivity analysis of Crisis Scenario.

We can further explore the sensitivity of this conclusion around variations of the tempered stable noise parameters α and λ. The result is shown in Figure 7.11. We plot the landscape of $1 - r$ for the different staff groupings for different values of the remaining noise parameters, α and λ. While these show very small deviations from synchronisation for the expected value, a landscape of the range gives a sense of the extreme cases, seen in Figure 7.12. Here the general pattern is that for α far from extreme values 2 and 0 more tempering λ is required for the range to diminish. This is less the case for the entire staff, but for all of the watch, J3 and J2 stuff the range of values can be as large as 0.3 across α,λ values. In other words, order parameter values may be as low as $r = 0.6$, significantly far from synchronisation. We also observe the trade-off between α and λ: as α decreases from 2, higher values of λ are required to temper the extreme deviations from synchronisation. Evidently, for sufficiently large $\lambda > 1$ synchronisation will be recovered, but this is at ranges where the tempered stable noise has been suppressed

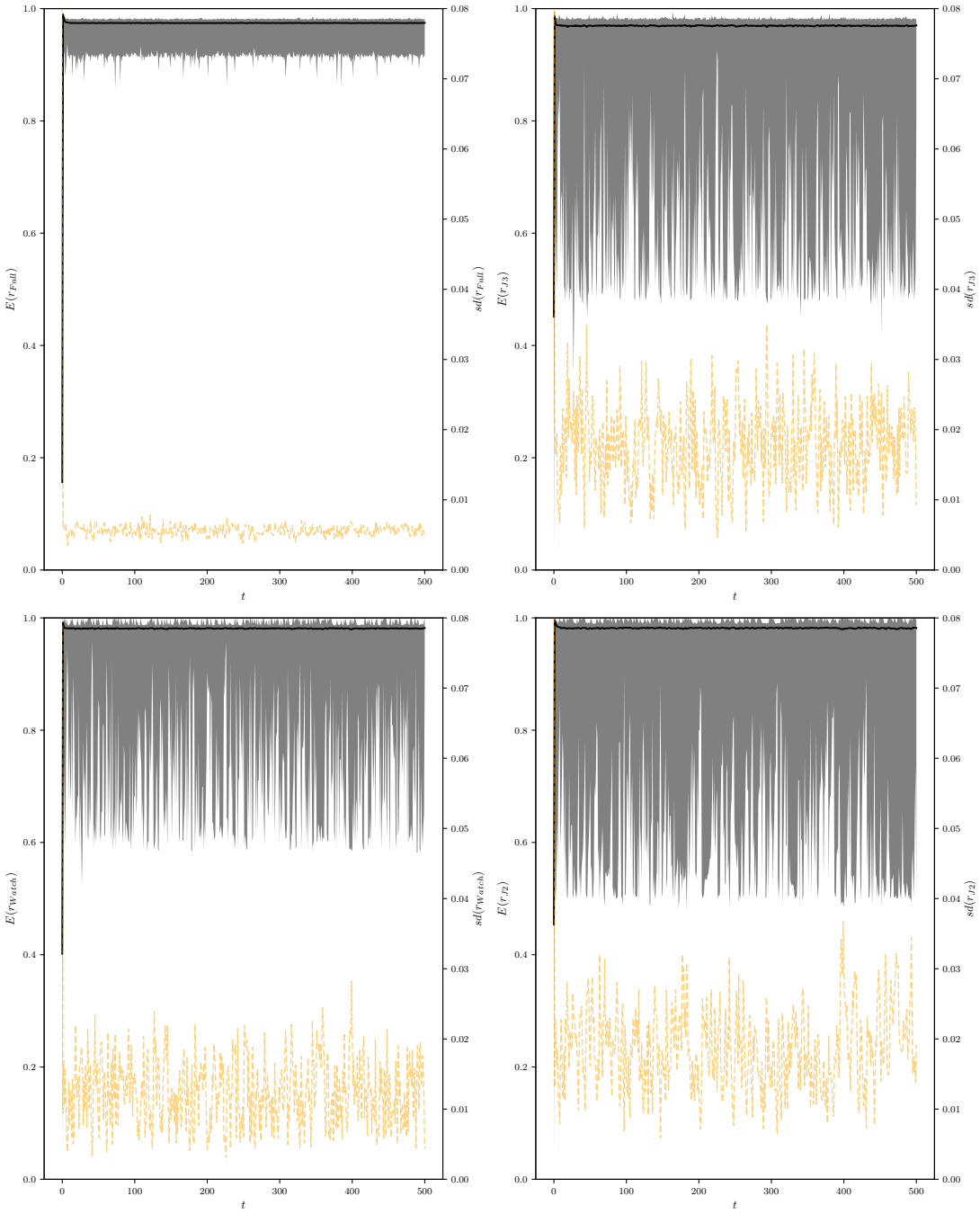


FIGURE 7.9: Order parameters as functions of time for stable noise with $\alpha = 1.4$ and $\sigma_G = 0.6$. See Figure 7.3 for colour scheme.

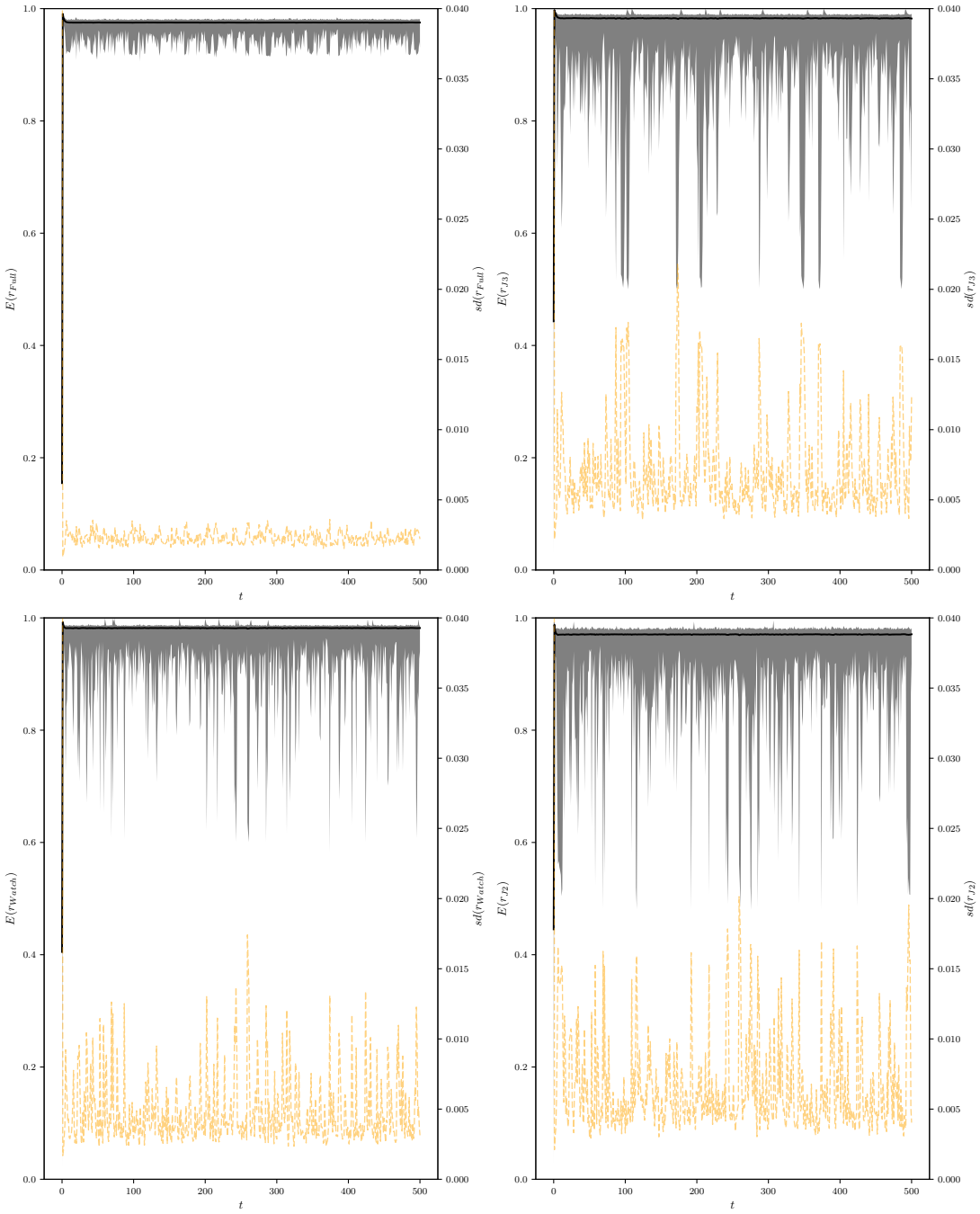


FIGURE 7.10: Order parameters as functions of time for tempered stable noise with $\alpha = 1.4$; $\lambda = 1$ and $\sigma_G = 0.6$. See Figure 7.3 for colour scheme.

to levels close to Gaussian. In conclusion, the combination of altered network structure in the crisis scenario and individual processes for non-Gaussian decision making results in a loss of robust synchronisation for this headquarters staff, as modelled here through synchronising oscillators. This stands as a prediction of the model.

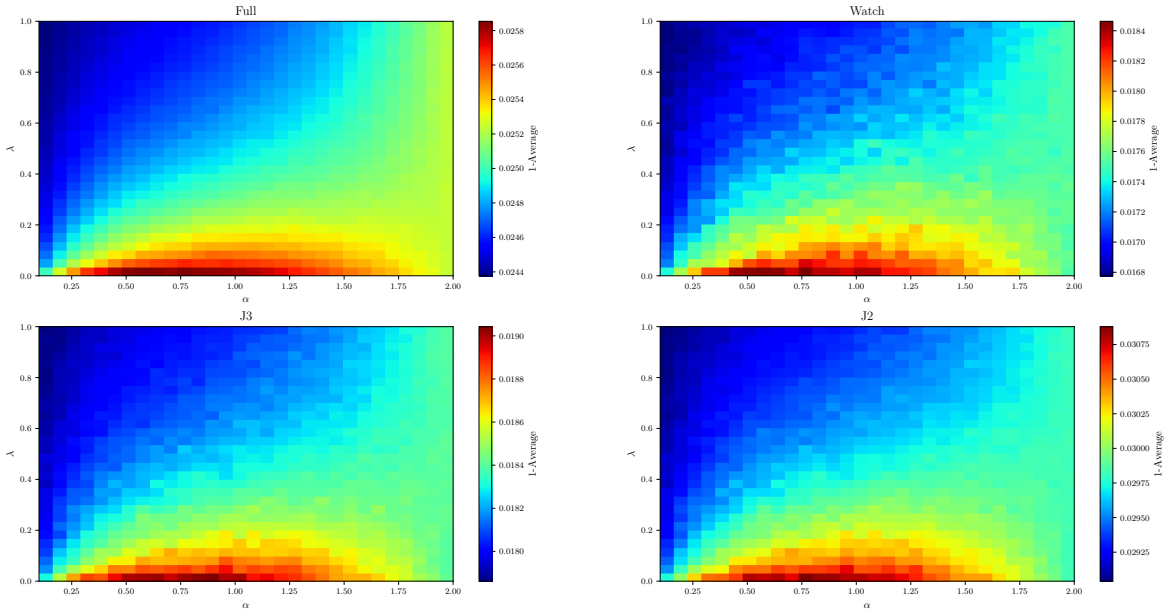


FIGURE 7.11: Average order parameter values for the various groups (Full/J3/J2/Watch) in the crisis scenario using tempered stable noise with $\sigma_G = 0.6$. Colours are determined from the value of $1 - r$.

7.5 Discussion: information, uncertainty reduction and contingency theory

As has been repeated several times, we have no data to empirically confirm the predicted loss of synchronised decision making in this sample of headquarters staff. The means for validation by quantitative historical data or predictive means [170, 171] is not possible. However, qualitative theories of organisational behaviour do permit an initial cross-validation. In a previous work [18] we have already highlighted the extreme centralisation of both the routine and crisis scenario networks by comparing their degree distributions compared to some typical theoretical networks. In Figure 7.13 we repeat this plot, showing degree distributions for two examples of the Barabasi-Alberts (BA) scale free networks, a Small World (SW) and the uniform random Erdös-Renyi (ER) networks of same size. The BA cases are noteworthy because they are characterised by strong hubs and weakly connected periphery, something also seen in hierarchies. We see here that both the headquarter networks show an even more strongly connected

hub, with some nodes having degree $d_i > 25$, far more concentrated than for the BA cases. At the other end, the headquarter networks show a large bulk of nodes with very low degree, spread across a larger range of values than for the BA cases. What distinguishes the crisis scenario network is that it is more centralised than the routine scenario network. As noted in [18], some senior participants in the original study sought to access 140 links to raw data from people and sources for their decisions, despite the role of their staff to do this. This is consistent with what is known as 'Uncertainty Reduction Theory' [172, 173] where decision makers in threatening or uncertain situations seek more information. However, is such an increase in centralisation sound from the perspective of organisational theory? According to a semi-quantitative approach to organisational theory known as Contingency Theory [174], the answer is: no. Without reviewing this significant body of literature, the import is that hierarchical structures are fit-for-purpose for stable, routine contexts but not for unstable, complex, opaque contexts. This is where organic, at or peer-to-peer structures are more appropriate, both because of saving time in waiting for information to move up a chain-of-command and decisions to percolate down, and because information is shared peer-to-peer so that team members may adjust their tasks self-synchronously. A two-dimensional form of this theory goes back to the classical work of Burns and Stalker [175]. Though qualitative, this theory would predict then that there is a probability - non-quantifiable in this approach - that this headquarters arrangement for the crisis would experience incoherent decision making. A model such as we have proposed here can quantify this probability of failure and identify the breakage points. For example, examining individual instances in the case of the stable noise distribution we give a selection of the phase profiles $\beta_i(t)$ for key individuals in the headquarters in Figure 7.14. We observe that consistently, the node to lose synchronisation with the cluster is the lowest degree node of $d_i = 2$, the role J3WS - a set of officers in the J3 part of the watch who play a critical role in providing critical support information (logistics, environmental, health, and computer systems support). That the lowest degree nodes fail to synchronise is the reverse side of the high centralisation of the senior officers in that they fail to carry staff outside the perceived core team. Thus there is an associated risk that critical enabling information, such as availability of equipment, people, medical support, communications networks, or weather changes, may missed by senior staff as they develop responses to the crisis. By enumerating the number of fail-cases in a large sample, the probability of failure, and even the type of failure, may be computed through this model.

7.6 Conclusions

We have proposed and analysed a model for headquarters staff using a stochastic dynamical systems representation of synchronising oscillators. By building in modern statistical properties such as heavy-tail noise, and utilising features of network theory we have shown how empirical data from a case study of networked decision making in a military headquarters may be used to tune the model and gain predictions of organisational behaviour under novel situations. Specifically, we were able to use a skewed form of heavy-tail noise to reflect that individual decision makers would be jumping ahead in their decision processes to anticipate emerging situations in a crisis scenario. In particular, having tuned the model for its behaviour under routine business, we predicted that a crisis would trigger a probability of failure of coherent decision making. This is not because any individual decision makers represented here are disadvantaged by their intuitive leaping forward but that the imbalances in connectivity combined with relative weakness of coupling means that coherent *collective* decision making breaks down.

Due to a lack of data on actual synchronisation of decision making in a crisis scenario in the case study we could not quantitatively validate the model. Nevertheless using qualitative theories of organisational behaviour, we were able to confirm that the prediction is sound - that excessively centralised decision making in crisis situations will lead to instances of failed coherent decision making insofar as some organisational members will be 'left behind', with an associated risk that crucial information may be missing as leaders progress in their decisions. We were thereby able to cross-validate the model and identify the failure nodes and the operational risks that this implies. The lesson out of this is not necessarily that coupling should be stronger; micromanagement by leaders of subordinates is not generally the answer, whose consequences such as degradation of initiative are not represented in this model. Rather, better sharing of information loads in the network via more uniform degree distributions of networks is appropriate when crises emerge and intuitive judgement becomes critical.

The model evidently now allows for empirical studies. Firstly, the use of the tempered stable Lévy process to build the stochastic dynamical model means that data can be better incorporated in the model, permitting quantitative validation. This provides an impetus for quantitative data collection on distributed human decision making. This in turn creates scope for experimentation on intervention strategies to mitigate the risks in the crisis that we have identified. For example, it is straighforward to undertake variations in the network structure such increasing links between subordinates while moderating the degree of centralisation of the senior staff. Alternately, we may consider training scenarios to tighten, for example, the frequency distributions of the various groups, or moving staff between watch and regular business hour positions. Finally,

we may explore the impact of autonomy in such a system, for example considering scenarios where individual information artefacts are replaced or aggregated into 'apps' enabled to self-update based on automatically feed dynamical data. This was initially studied for pure stable Lévy noise in [166]. A typical concern with the proliferation of autonomy in socio-technical systems is the mis-match in decision speeds between humans and technology. A model of this type enables exploration of the performance of the headquarters system as a coherent whole given such a hybrid nature.

As we have argued, a model such as this is applicable in a wide variety of settings outside of the military, from emergency response, to energy and traffic control organisations. With this variety of applications, and focusing, as it does, on speeds of individual decision making and organisational coherence of such processes, we trust this spurs human factors researchers to refine methods for data collection to further validate and operationalise such a model.

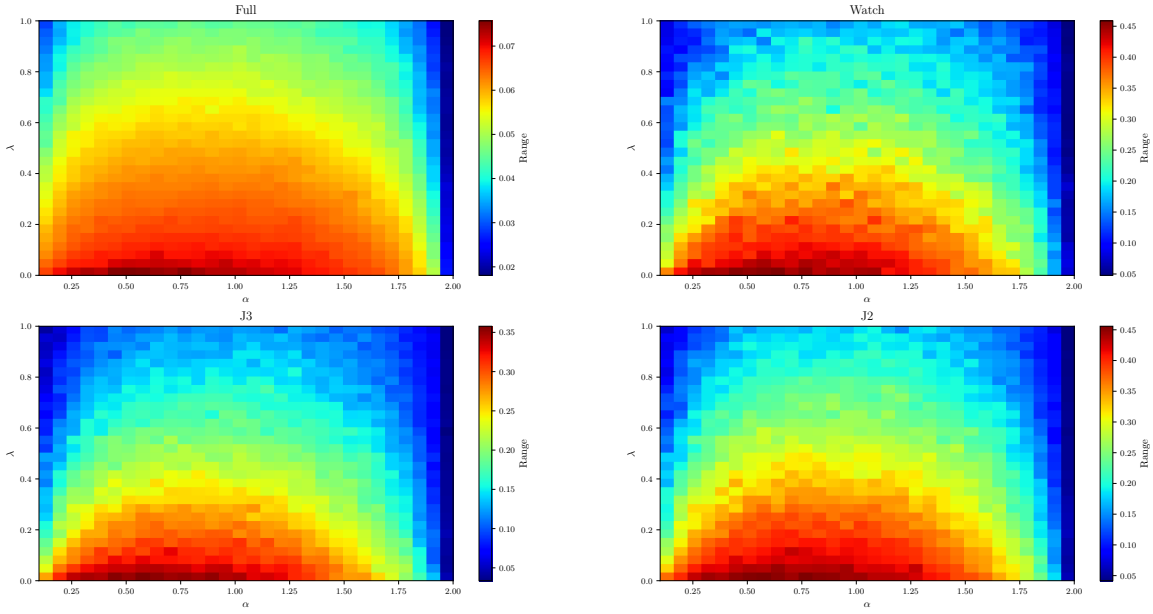


FIGURE 7.12: Range of order parameter values for the various groups (Full/J3/J2/Watch) in the crisis scenario using tempered stable noise with $\sigma_G = 0.6$.

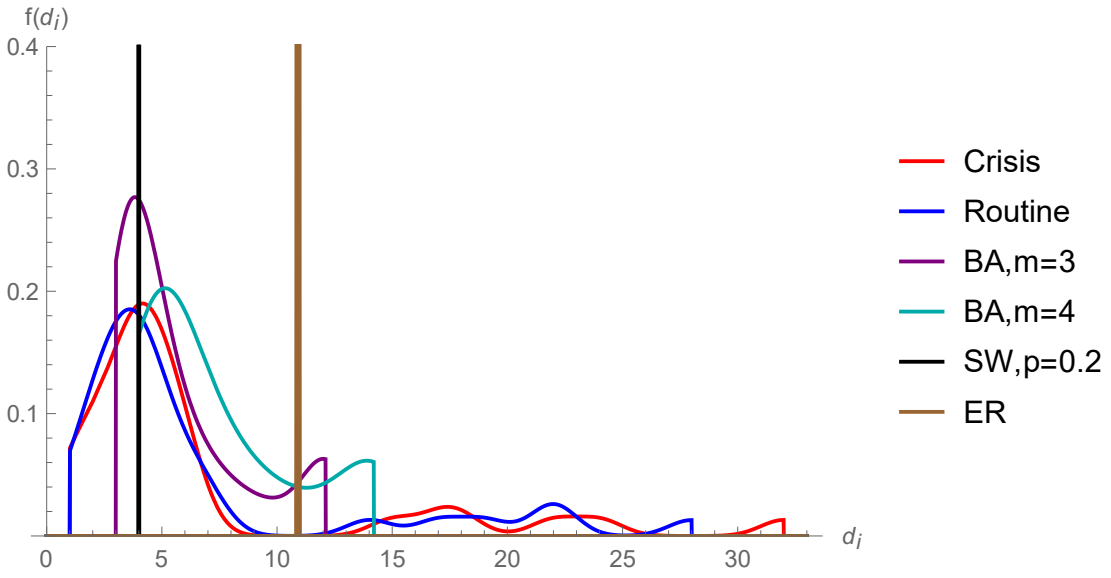


FIGURE 7.13: Degree distributions for the routine and crisis scenario networks compared to Barabasi-Alberts (BA), Small World (SW) and Erdos-Renyi (ER) cases.

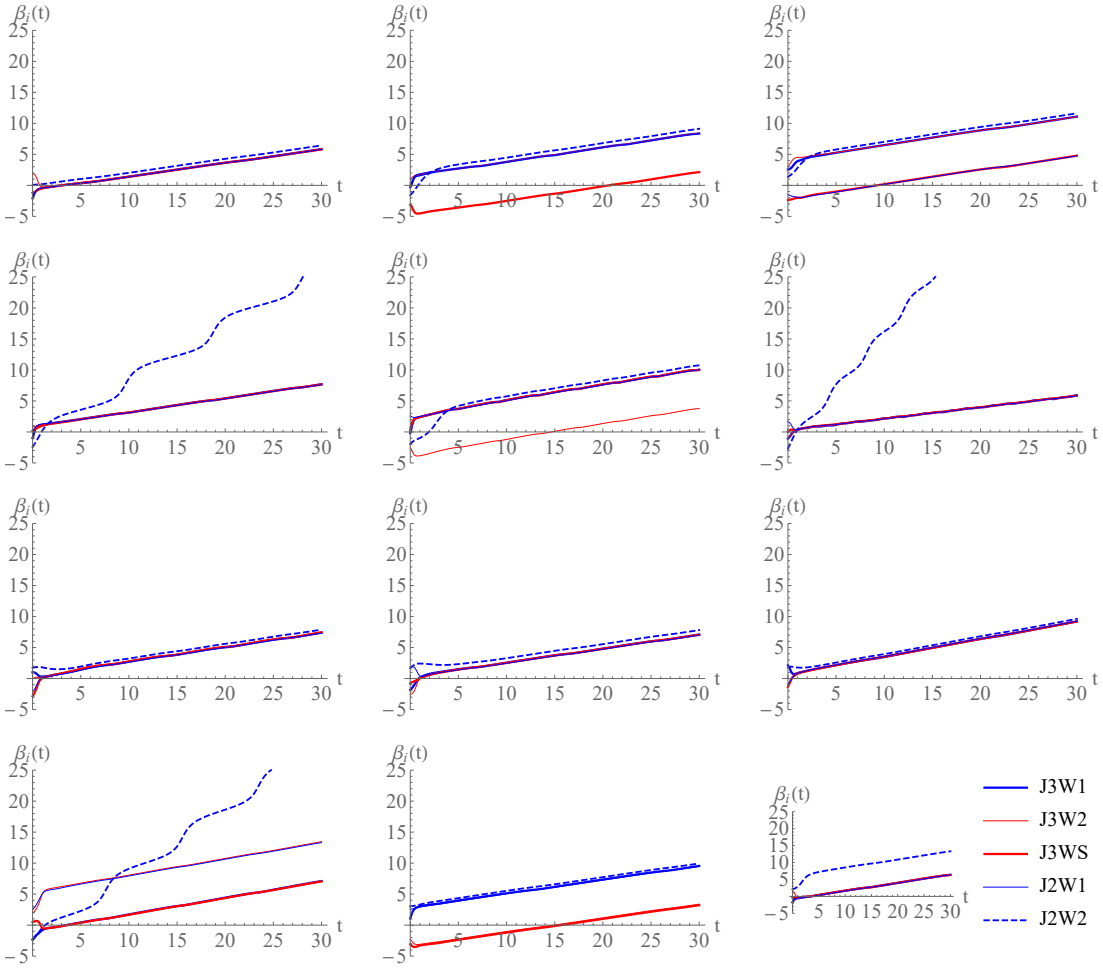


FIGURE 7.14: Individual paths for the stable noise case; note that the degree of J3WS is di = 2.

Chapter 8

Pairs Trading and Portfolio Optimisation

In this chapter we consider how to introduce the concept of synchronisation into a model for trading a portfolio of securities. This model aims to construct a market-neutral trading strategy that matches a long position with a short position in a pair of highly correlated securities. Our goal in this section is to extend this pairs scenario to the multiple asset setting in which the modelled financial system has an underlying topology which can be described by a graph and the dynamics of the drifts of the assets are driven by the Kuramoto equation. Once the model is established, we then consider the construction of the optimal portfolio.

8.1 An overview of pairs trading and synchronisation in financial markets

The aim of this work is to extend a commonly used pairs-trading model to a networked setting and devise an optimal trading strategy based upon the resulting synchronisation model. Pairs trading typically involves the construction of a portfolio of two closely related assets and implementing a strategy based upon the predictability of the joint behaviour of the assets. This strategy generally consists of a long position in one security and a short position in another based upon some predetermined ratio. Typically this portfolio is modelled as a mean-reverting process and the long and short positions taken are based upon wherever the observed values are larger or smaller than the respective predicted values. In the literature, trading algorithms are typically developed to take advantage of the structural dependency between assets by co-integrating the drifts of the

assets together in the named cointegration pairs model [19–21]. An example of the joint behaviour of two assets is provided on Figure 8.1 where we plot the midprice relative to the mean midprice of JPMorgan Chase (JPM) and Goldman Sachs (GS) over the course of a single trading day.

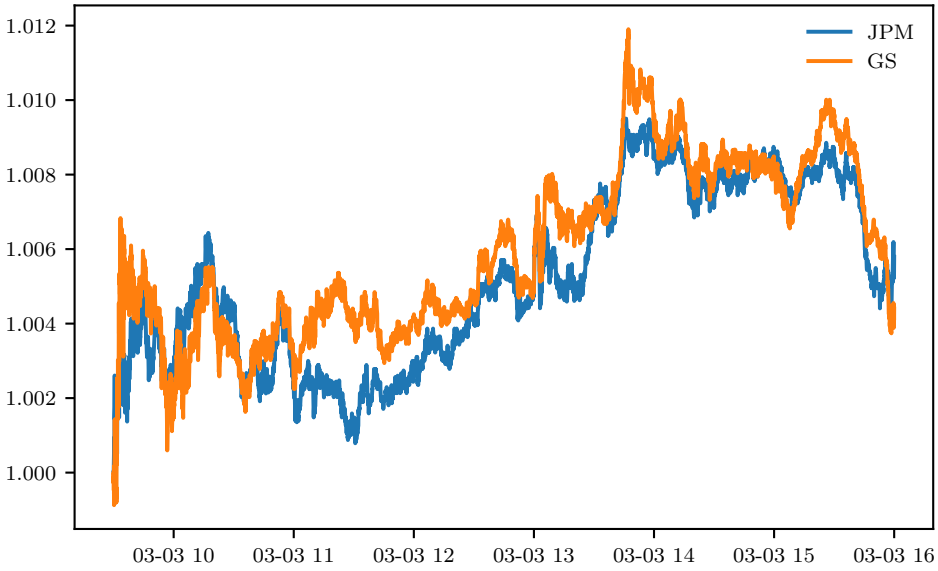


FIGURE 8.1: Midprice relative to mean midprice of JPM and GS on March 3, 2017 for the whole day of trading.

Structured behaviour in financial markets has been studied by Mantegna in [176] wherein the topology of the financial market is described through the means a connecting graph derived from the degree of similarity between the time evolution of stock prices. This degree of similarity is quantified through the correlation coefficient between the time series of two stocks i and j and is given by

$$\rho_{i,j} = \frac{\langle Y_i Y_j \rangle - \langle Y_i \rangle \langle Y_j \rangle}{\sqrt{\left(\langle Y_i^2 \rangle - \langle Y_i \rangle^2\right)\left(\left\langle Y_j^2 \right\rangle - \langle Y_j \rangle^2\right)}},$$

where Y_i is the return of stock i given by $Y_i = \log P_i(t) - \log P_i(t-1)$ and $P_i(t)$ is the closure at day t. The distance between two stocks i and j is then given by

$$d(i,j) = \sqrt{2(i - \rho_{ij})}.$$

The distances between stocks forms the $N \times N$ weighted symmetric matrix $\mathbf{D}$ which then represents the topology of the financial market. From the distance matrix, it also possible to represent the topology of the financial market through the construction of

a minimal spanning tree [177] which connects the n stocks in the market through a 1-connected graph. In the study of synchronisation dynamics in financial market networks, the Kuramoto model can then be employed in the following system [178]

$$\dot{\theta}_i = \omega_i + \sigma \sum_{j}^{N} e^{-\alpha d_{ij}} \sin\left(\theta_j - \theta_i, \right), \qquad i = 1, \ldots, N.$$

where α is a weighting parameter and d_{ij} are the elements of the distance matrix $\mathbf{D}$.

Our aim in this section is to take the existing cointegration model [19, 20] and using the derivation of the optimal portfolio as discussed in [21] as a framework, extend it by introducing a nonlinearity in the cointegration term such that the cointegration factor is now a synchronisation model with an underlying network structure.

8.2 The synchronised cointegration model

The setup of our synchronised cointegration model begins by considering a collection of risky assets $Y = \left(Y_1^t, \ldots, Y_t^n\right)_{0 \leq t \leq T}$ that satisfy the system of SDEs

$$\frac{dY_t^k}{Y_t^k} = \alpha_t^k dt + \sum_{i=1}^{n} \sigma_{ki} dW_t^i, \tag{8.1}$$

where the nonlinear cointegration factor that incorporates the underlying market structure is given by

$$\alpha_t^k = a_0^k + \sum_{i=1}^{n} A_{ki} F\left(\log Y_t^i - \log Y_t^k\right), \tag{8.2}$$

and $\mathbf{W}_t = \left(W_t^1, \ldots, W_t^n\right)_{0 \leq t \leq T}$ is a vector of independent Brownian motions, F is some known function $F : \mathbb{R} \to \mathbb{R}$ and $\boldsymbol{A}$ is a weighted adjacency matrix that encodes the network topology. The instantaneous covariance between the assets Y_t^i and Y_k^t is given by

$$\Omega_{ij} = \sum_{k=1}^{n} \sigma_{ik} \sigma_{kj},$$

and the matrix $\boldsymbol{\sigma}$ is the Cholesky decomposition of the variance-covariance matrix $\boldsymbol{\Omega} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T$.

It should be noted that in this newly proposed model, the dynamics of the original cointegration model as in [19, 20] can be retrieved through a linearisation of the new cointegration factor in Eq. (8.2). To see this, we first use Ito's theorem to show that the

log-prices satisfy the SDEs

$$d\log Y_t^k = (\alpha_t^k - \frac{1}{2}\Omega_{kk})dt + \sum_{i=1}^{n} \sigma_{ki} dW_t^i.$$

We now note that by substituting in α_t^k and linearising the function F around 0, the dynamics of $\log Y_t^k$ satisfy that of an Ornstein-Uhlenbeck process

$$\begin{aligned} d\log Y_t^k &= \left(a_0^k + \sum_{i=1}^{n} A_{ki} F\left(\log Y_t^i - \log Y_t^k\right) - \frac{1}{2}\Omega_{kk} \right) dt + \sum_{i=1}^{n} \sigma_{ki} dW_t^i, \\ &\approx \left(\omega^k - F'(0) \sum_{i=1}^{n} \mathcal{L}_{kj} \log Y_t^k \right) dt + \sum_{i=1}^{n} \sigma_{ki} dW_t^i \end{aligned}$$

where $\omega^k = a_0^k - \frac{1}{2}\Omega_{kk} + F(0)d_k$, $\mathcal{L}_{kj} = \begin{cases} A_{kj}, & k \neq j \\ -d_k, & k = j \end{cases}$ and $d_k = \sum_j A_{kj}$.

Now setting $Z_t^k = \log Y_t^k$ we arrive at the Ornstein-Uhlenbeck process

$$dZ = (\boldsymbol{\omega} - \boldsymbol{B} Z_t)dt + \boldsymbol{\sigma} d\boldsymbol{W}_t, \tag{8.3}$$

where $\boldsymbol{\omega} = (\omega^1, \ldots, \omega^n)$ and $\boldsymbol{B} = \text{diag}(H'(0), \ldots, H'(0))\mathcal{L}$. Thus we have shown that through linearisation, our model has the same mean-reverting process that is obtained in the original version of the cointegration model.

8.3 The optimal portfolio construction

We now turn our attention to the agent's optimisation problem for the model described in Eq. (8.1) in a similar manner as constructed by Cartea in [21]. Let $\pi = (\pi_t^0, \pi_t^1, \ldots, \pi_t^n)_{0 \leq t \leq T}$ denote the value of the amount invested in the riskless asset π_t^0 and risky assets $(\pi_t^1, \ldots, \pi_t^n)$ and denote the agents wealth process at time t by

$$X_t^\pi = \sum_{k=0}^{n} \pi_t^k = \sum m_t^k Y_t^k, \tag{8.4}$$

where m_k is the number of units invested in the the asset k. The dynamics of the agents wealth process X_t^π are then given by

$$\begin{aligned} dX_t^\pi &= \sum_{k=0}^{n} d(m_t^k Y_t^k) = \sum_{k=0}^{n} m_t^k dY_t^k = \sum_{k=1}^{n} \pi_t^k \frac{dY_t^k}{Y_t^k}, \\ &= \sum_{k=1}^{n} \left[\alpha_t^k dt + \sum_{i=1}^{n} \sigma_{ki} dW_t^i \right], \end{aligned}$$

where we have assumed that the interest rates corresponding to the riskless asset are zero over our time horizon, thus $dY_t^0 = 0$.

Using $u(x) = -\exp(-\gamma x)$ as the agents utility function, the performance criteria is given by

$$H^\pi(t, x, \boldsymbol{y}) = \mathbb{E}_{t,x,\boldsymbol{y}} \left[-\exp\left(-\gamma X_T^\pi\right) \right],$$

where $\mathbb{E}_{t,x,\boldsymbol{y}}$ denotes the expectation conditional on $X_t^\pi = x$ and $\boldsymbol{Y}_t = \boldsymbol{y}$. The agents value function is then given by

$$H(t, x, \boldsymbol{Y}) = \sup_{\pi \in \mathcal{A}} H^\pi(t, x, \boldsymbol{y}),$$

where the set of admissible strategies $\mathcal{A}$ obey the finite condition that

$$\mathbb{E}\left[\sum_{k=0}^{n} \int_0^T (\pi_u^k)^2 du \right] < \infty.$$

Before we construct the dynamic programming equation (DPE), we require the quadratic variation and cross-variations of the processes X and Y. These are given by

$$\begin{aligned} d[Y^i, Y^j]_t &= \sum_{k,l=1}^{n} Y_t^i Y_t^j \sigma_{ki} \sigma_{lj} d[W^i, W^j]_t = \sum_{k,l=1}^{n} Y_t^i Y_t^j \sigma_{ki} \sigma_{lj} dt = \left(\boldsymbol{Y}_t^T \boldsymbol{\Omega} \boldsymbol{Y}_t \right)_{ij} dt, \\ d[X^\pi, X^\pi]_t &= \sum_{i,j,k,l=1}^{n} \pi_t^k \pi_t^l \sigma_{ki} \sigma_{lj} d\left[W^i, W^j\right]_t = \sum_{i,j,k,l=1}^{n} \pi_t^k \pi_t^l \sigma_{ki} \sigma_{lj} dt = \boldsymbol{\pi}_t^T \boldsymbol{\Omega} \boldsymbol{\pi}_t dt, \\ d[X^\pi, Y^k]_t &= \sum_{i,j,l=1}^{n} \pi_t^l \sigma_{li} \sigma_{kj} Y_t^k d[W^i, W^j]_t = \sum_{j,l=1}^{n} \pi_t^l \sigma_{lj} \sigma_{kj} Y_t^k dt = \left(\boldsymbol{\pi}_t^T \boldsymbol{\Omega} \boldsymbol{\pi} \right)_k. \end{aligned}$$

The dynamic programming principle then states that the value function H should satisfy the following DPE

$$\begin{aligned}\partial_t H + &\boldsymbol{\alpha}^{\boldsymbol{T}} \boldsymbol{\mathcal{D}}_y H + \frac{1}{2}\boldsymbol{\mathcal{D}}^{\boldsymbol{\Omega}}_{yy} H \\ &+ \sup_{\pi}\left\{\boldsymbol{\pi}^T \boldsymbol{\alpha}\partial_x H + \frac{1}{2}(\boldsymbol{\pi}^T \boldsymbol{\Omega}\boldsymbol{\pi})\partial_{xx} H + \boldsymbol{\pi}^T \boldsymbol{\Omega}\boldsymbol{\mathcal{D}}_{xy} H\right\} = 0,\end{aligned} \tag{8.5}$$

subject to the boundary condition $H(T, x, \boldsymbol{y}) = -\exp(-\gamma x)$ and $\boldsymbol{\alpha} = (\alpha_t^1, \ldots, \alpha_t^n)_{0\leq t\leq T}$. The linear differential operators given in Eq. (8.5) are defined as follows

$$\begin{aligned}\boldsymbol{\mathcal{D}}_y H &= \left(y^1 \partial_{y^1} H, \ldots, y^n \partial_{y^n} H\right)^T, \\ \boldsymbol{\mathcal{D}}^{\boldsymbol{\Omega}}_{yy} H &= \sum_{i,j=1}^{n} y^i \Omega_{ij} y^j \partial_{y^i y^j} H, \\ \boldsymbol{\mathcal{D}}_{xy} H &= \left(y^1 \partial_{xy^1} H, \ldots, y^n \partial_{xy^n} H\right)^T.\end{aligned} \tag{8.6}$$

Assuming $\partial_{xx} H \neq 0$, we write the supremum term as

$$\begin{aligned}\mathcal{M} =& \boldsymbol{\pi}^T \boldsymbol{\alpha}\partial_x H + \frac{1}{2}(\boldsymbol{\pi}^T \boldsymbol{\Omega}\boldsymbol{\pi})\partial_{xx} H + \boldsymbol{\pi}^T \boldsymbol{\Omega}\boldsymbol{\mathcal{D}}_{xy} H, \\ =& \frac{1}{2}\partial_{xx} H \left\{\boldsymbol{\pi}^T \boldsymbol{\Omega}\boldsymbol{\pi} + 2\boldsymbol{\pi}^T \frac{\boldsymbol{\alpha}\partial_x H + \boldsymbol{\Omega}\boldsymbol{\mathcal{D}}_{xy} H}{\partial_{xx} H}\right\}, \\ =& \frac{1}{2}\partial_{xx} H \left\{\left(\boldsymbol{\pi}_t + \frac{\boldsymbol{\Omega}^{-1}\boldsymbol{\mathcal{L}} H}{\partial_{xx} H}\right)^T \boldsymbol{\Omega}\left(\boldsymbol{\pi}_t + \frac{\boldsymbol{\Omega}^{-1}\boldsymbol{\mathcal{L}} H}{\partial_{xx} H}\right) - \frac{\boldsymbol{\mathcal{L}}^T H \boldsymbol{\Omega}^{-1}\boldsymbol{\mathcal{L}} H}{(\partial_{xx} H)^2}\right\},\end{aligned} \tag{8.7}$$

where the linear operator $\boldsymbol{\mathcal{L}}$ is defined as

$$\boldsymbol{\mathcal{L}} H = \boldsymbol{\alpha}\partial_x H + \boldsymbol{\Omega}\boldsymbol{\mathcal{D}}_{xy} H,$$

and we have used that fact that the variance-covariance matrix $\boldsymbol{\Omega}$ is symmetric.

From the above expression in Eq. (8.7), we can identify the optimal investment as

$$\begin{aligned}\boldsymbol{\pi}^* &= -\frac{\boldsymbol{\Omega}^{-1}\boldsymbol{\mathcal{L}} H}{\partial_{xx} H}, \\ &= -\frac{\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha}\partial_x H + \boldsymbol{\mathcal{D}}_{\boldsymbol{xy}} H}{\partial_{xx} H}.\end{aligned}$$

The maximum value of M obtained at $\boldsymbol{\pi}^*$ is thus given by the second term in Eq. (8.7)

$$\mathcal{M} = -\frac{1}{2}\frac{\boldsymbol{\mathcal{L}}^T H \boldsymbol{\Omega}^{-1}\boldsymbol{\mathcal{L}} H}{\partial_{xx} H},$$

We thus arrive at the following form of the non-linear PDE for the value function

$$\partial_t H + \boldsymbol{\alpha}^T \boldsymbol{\mathcal{D}}_y H + \frac{1}{2}\boldsymbol{\mathcal{D}}^{\boldsymbol{\Omega}}_{yy} H - \frac{1}{2}\frac{\boldsymbol{\mathcal{L}}^T H \boldsymbol{\Omega}^{-1} \boldsymbol{\mathcal{L}} H}{\partial_{xx} H} = 0.$$

In the classical Merton problem, H takes the form $-\exp(-\gamma(x + h(t)))$ for some deterministic function $h(t)$. With the addition of the nonlinear co-integration factor, we expect that the function h additionally depends on $\boldsymbol{\alpha}$ and thus our ansatz for the form of H is

$$H(t, x, \boldsymbol{y}) = -e^{-\gamma(x+h(t,\boldsymbol{\alpha}))},$$

subject to $h(T, \boldsymbol{\alpha}) = 0$. Since $\boldsymbol{\alpha}$ is a function of $\boldsymbol{y}$, we can choose an alternative ansatz for H in terms of the function $\tilde{h}(t, \boldsymbol{y})$ where

$$H(t, x, \boldsymbol{y}) = -e^{-\gamma\left(x+\tilde{h}(t,\boldsymbol{y})\right)},$$

and $\tilde{h}(T, \boldsymbol{y}) = 0$.

Differentiation on H with respect to y^k acts as

$$\partial_{y^k} H = -\gamma H \partial_{y^k} \tilde{h}(t, \boldsymbol{y}).$$

And for the second derivative

$$\begin{aligned}\partial_{y^j y^k} H =& -\gamma \partial_{y^j}\left(H \partial_{y^k} \tilde{h}(t, \boldsymbol{y})\right), \\ =& \gamma^2 H \partial_{y^j} \tilde{h}(t, \boldsymbol{y}) \partial_{y^k} \tilde{h}(t, \boldsymbol{y}) - \gamma H \partial_{y^j y^k} \tilde{h}(t, \boldsymbol{y}).\end{aligned}$$

Thus we can write the differential operators in Eq. (8.6)

$$\begin{aligned}\mathcal{D}_y H =& \gamma H \mathrm{Diag}(\boldsymbol{y}) \partial_{\boldsymbol{y}} \tilde{h}, \\ \mathcal{D}_{xy} H =& \gamma^2 H \mathrm{Diag}(\boldsymbol{y}) \partial_y \tilde{h}, \\ \mathcal{D}^{\boldsymbol{\Omega}}_{yy} =& \gamma H \left(\sum_{i,j} y^i \Omega_{i,j} y^j \left[\partial_{y^i} \tilde{h} \partial_{y^j} \tilde{h} - \gamma \partial_{y^i y^j} \tilde{h}\right]\right), \\ \mathcal{L} H =& -\gamma H \boldsymbol{\alpha} + \gamma^2 H \boldsymbol{\Omega} \mathrm{Diag}(\boldsymbol{y}) \partial_y \tilde{h}.\end{aligned}$$

Using the above differential expressions from our ansatz, we can now write out the DPE in Eq. (8.5) as

$$\begin{aligned}0 =& -\gamma H\partial_t H - \gamma H\boldsymbol{\alpha}^T \text{Diag}(\boldsymbol{y})\partial_{\boldsymbol{y}} h,\\ &+\frac{1}{2}\gamma H\left(\sum_{i,j} y^i\Omega_{ij}y^j\left[\partial_{y^i}\tilde{h}\partial_{y^j}\tilde{h} - \gamma\partial_{y^iy^j}\tilde{h}\right]\right),\\ &+\frac{1}{2}(\gamma^2 H)^{-1}\left((-\gamma H\boldsymbol{\alpha} + \gamma^2 H\boldsymbol{\Omega}\text{Diag}(\boldsymbol{y})\partial_y\tilde{h})^T\boldsymbol{\Omega}^{-1}(-\gamma H\boldsymbol{\alpha} + \gamma^2 H\boldsymbol{\Omega}\text{Diag}(\boldsymbol{y})\partial_y\tilde{h})\right),\end{aligned}$$

which simplifies to

$$\begin{aligned}0 =& -\partial_t H - \boldsymbol{\alpha}^T \text{Diag}(\boldsymbol{y})\partial_{\boldsymbol{y}} h,\\ &+\frac{1}{2}\left(\sum_{i,j} y^i\Omega_{ij}y^j\left[\partial_{y^i}\tilde{h}\partial_{y^j}\tilde{h} - \gamma\partial_{y^iy^j}\tilde{h}\right]\right),\\ &+\frac{1}{2\gamma}\left((-\boldsymbol{\alpha} + \gamma\boldsymbol{\Omega}\text{Diag}(\boldsymbol{y})\partial_y\tilde{h})^T\boldsymbol{\Omega}^{-1}(-\boldsymbol{\alpha} + \gamma\boldsymbol{\Omega}\text{Diag}(\boldsymbol{y})\partial_y\tilde{h})\right).\end{aligned}$$

Noting that we can write the last term as

$$\begin{aligned}&(-\boldsymbol{\alpha} + \gamma\boldsymbol{\Omega}\text{Diag}(\boldsymbol{y})\partial_y\tilde{h})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\alpha} + \gamma\boldsymbol{\Omega}\text{Diag}(\boldsymbol{y})\partial_y\tilde{h})\\ &\qquad = \boldsymbol{\alpha}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha} - 2\gamma\boldsymbol{\alpha}^T\text{Diag}(\boldsymbol{y})\partial_y\tilde{h} + \gamma^2\partial_y\tilde{h}^T\text{Diag}(\boldsymbol{y})\boldsymbol{\Omega}\text{Diag}(\boldsymbol{y})\partial_y\tilde{h}.\end{aligned}$$

Substituting this back into our expression for the DPE and noting the cancellation of the second and third term in the expansion above, we obtain the PDE

$$\partial_t\tilde{h} + \frac{1}{2}\sum_{i,j} y^i\Omega_{ij}y^j\partial_{y^iy^j}\tilde{h} + \frac{\boldsymbol{\alpha}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha}}{2\gamma} = 0, \tag{8.8}$$

with the terminal condition $\tilde{h}(T, \boldsymbol{y}) = 0$. Solving this PDE for h, we can then obtain the solution for the optimal portfolio which in terms of our ansatz for H is given by

$$\begin{aligned}\boldsymbol{\pi}^* =& -\frac{\boldsymbol{\Omega}^{-1}\boldsymbol{\mathcal{L}}H}{\partial_{xx}H},\\ =& -\frac{\boldsymbol{\Omega}^{-1}\gamma H\boldsymbol{\alpha} + \gamma^2 H\text{Diag}(\boldsymbol{y})\partial_y\tilde{h}}{\gamma^2 H},\\ =&\frac{1}{\gamma}\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha} - \text{Diag}(\boldsymbol{y})\partial_y\tilde{h},\\ =&\frac{1}{\gamma}\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha} - \boldsymbol{\Gamma}\partial_\alpha\tilde{h}.\end{aligned} \tag{8.9}$$

Where we have made the change of derivative

$$\partial_{y^k} h(t, \boldsymbol{\alpha}) = \frac{\partial \boldsymbol{\alpha}}{\partial y^k} \partial_\alpha h(t, \boldsymbol{\alpha}),$$

And thus we can calculate

$$\frac{\partial \boldsymbol{\alpha}}{\partial y^k} = \frac{\boldsymbol{\Gamma}^{k^T}}{y^k},$$

where $\boldsymbol{\Gamma}^k$ the k'th column of the matrix $\boldsymbol{\Gamma}$ is defined by

$$\boldsymbol{\Gamma}_{kj} = \begin{cases} -\sum_{i,i\neq k} A_{ji} F'(\log y_t^i - \log y_t^j), & j = k \\ A_{jk} F'(\log y_t^k - \log y_t^j), & j \neq k. \end{cases}$$

8.4 Numerical implementation

We first consider a numerical scheme to directly solve for $\tilde{h}$ in (8.8) for $t \in [0, T]$. The numerical approach that we shall use is a finite-difference approach that incorporates a predictor-corrector Alternating Direction Implicit (ADI) method. The scheme works by alternative the directions that are treated implicitly by the finite difference scheme, that is in the first step the direction x_1 is treated implicitly and all others explicitly, in the second step the direction x_2 is implicit and all other explicit and so on. Thus all spatial variable end up being treated in a semi-implicit manner in a similarly to a Crank-Nicolson scheme. For a more detailed discussed on the method and the finite difference implementation of the method, see [179, 180].

We recall that the predictor-corrector ADI method can be formulated in the following manner for PDEs of the form

$$\frac{\partial V}{\partial t} + \sum_{h=1}^{p=1} \mu_h(t, x) \frac{\partial V}{\partial x_h} + \frac{1}{2} \sum_{h=1}^{p} \sum_{l=1}^{p} s_{h,l}(t, x) \frac{\partial^2 V}{\partial x_h \partial x_l} - r(t, x) V = 0, \tag{8.10}$$

where $s_{h,h}(t, x) \geq 0$ and $s_{h,l}(t, x) = s_{l,h}(t, x)$ for $h, l = 1, \ldots, p$ and subject to $V(T, x) = g(x), g : \mathbb{R}^P \to \mathbb{R}$. Given the PDE in (8.10), we can write it in the form

$$\frac{\partial V}{\partial t} + \sum_{h=1}^{p} \mathcal{L}_h V + \sum_{h=1}^{p} \sum_{l=h+1}^{p} \mathcal{L}_{h,l} V = 0, \tag{8.11}$$

where the operators $\mathcal{L}_h$ and $\mathcal{L}_{h,l}$ are defined by

$$\mathcal{L}_h = \mu_h(t,x)\frac{\partial}{\partial x_h} + \frac{1}{2}s_{h,h}(t,x)\frac{\partial^2}{\partial x_h^2} - p^{-1}r(t,x),$$
$$\mathcal{L}_{h,l} = s_{h,l}(t,x)\frac{\partial^2}{\partial x_h \partial x_l}.$$

Given the operator representation of the PDE in (8.11), the predictor-corrector method is then given by the following sequences of equations where the set of equations corresponding to the predictor are given by

$$\left(1-\theta\Delta_t\hat{\mathcal{L}}_1\right)U^{(1)} = \Delta_t\left(\Delta_t^{-1} + (1-\theta)\hat{\mathcal{L}}_1 + \sum_{h=2}^{p}\hat{\mathcal{L}}_h + \sum_{h=1}^{p}\sum_{l=h+1}^{p}\hat{\mathcal{L}}_{h,l}\right)\hat{V}(t_{i+1}),$$
$$\left(1-\theta\Delta_t\hat{\mathcal{L}}_2\right)U^{(2)} = U^{(1)} - \theta\Delta_t\hat{\mathcal{L}}_2\hat{V}(t_{i+1}),$$
$$\vdots$$
$$\left(1-\theta\Delta_t\hat{\mathcal{L}}_p\right)U^{(p)} = U^{(p-1)} - \theta\Delta_t\hat{\mathcal{L}}_p\hat{V}(t_{i+1}),$$

and the set of equations corresponding to the corrector are given by

$$\left(1-\theta\Delta_t\hat{\mathcal{L}}_1\right)Z^{(1)} = \Delta_t\left(\Delta_t^{-1} + (1-\theta)\hat{\mathcal{L}}_1 + \sum_{h=2}^{p}\hat{\mathcal{L}}_h\right.$$
$$\left.+(1-\lambda)\sum_{h=1}^{p}\sum_{l=h+1}^{p}\hat{\mathcal{L}}_{h,l}\right)\hat{V}(t_{i+1}) + \lambda\Delta_t\sum_{h=1}^{p}\sum_{l=h+1}^{p}\hat{\mathcal{L}}_{h,l}U^{(p)},$$
$$\left(1-\theta\Delta_t\hat{\mathcal{L}}_2\right)Z^{(2)} = Z^{(1)} - \theta\Delta_t\hat{\mathcal{L}}_2\hat{V}(t_{i+1}),$$
$$\vdots$$
$$\left(1-\theta\Delta_t\hat{\mathcal{L}}_p\right)\hat{V}(t_i) = Z^{(p-1)} - \theta\Delta_t\hat{\mathcal{L}}_p\hat{V}(t_{i+1}).$$

With m_p points in the x_h-direction for $h = 1,\ldots,p$, the computational cost the this scheme is $\mathcal{O}(\Pi_{h=1}^{p}m_h)$. For $p \leq 3$, the stability conditions of this method are $\theta \geq \frac{1}{2}$ and $\frac{1}{2} \leq \lambda \leq \theta$. For $p \geq 4$, the conditions are $\theta \geq \frac{1}{2}$ and

$$\frac{1}{2} \leq \lambda \leq \frac{p^{p-1}}{(p-1)^p}\theta.$$

Since the boundary conditions are not explicitly defined for this problem, we implement the Vetzal approximation [181] at the boundaries which approximates the partial derivatives at the boundaries in the following manner under a finite difference method

$$\begin{aligned}
\frac{\partial V}{\partial x}(i,0) &= \frac{-V(i,2)+4V(i,1)-3V(i,0)}{2\Delta x},\\
\frac{\partial^2 V}{\partial x^2}(i,0) &= \frac{V(i,2)-2V(i,1)+V(i,0)}{(\Delta x)^2},\\
\frac{\partial V}{\partial x}(i,N) &= \frac{3V(i,N)-4V(i,N-1)+3V(i,N-2)}{2\Delta x},\\
\frac{\partial^2 V}{\partial x^2}(i,) &= \frac{V(i,N)-2V(i,N-1)+V(i,N-2)}{(\Delta x)^2}.
\end{aligned} \tag{8.12}$$

where Δx is the step size in the finite difference mesh $x \in \{x_j\}_{j=1}^{N}$.

8.5 Results

As a starting point in the numerical evaluation of the pairs trading model, we look at the effect of the choice of the nonlinearity in the cointergration factor Eq. (8.2) upon the structure of the solution space of $\tilde{h}(0,\mathbf{y})$ obtained by solving the PDE in Eq. (8.8) with the Vetzal boundary approximations from Eq. (8.12) and terminal condition $\tilde{h}(T,\mathbf{y}) = 0$ for $T = 1$. In this numerical investigation, we consider two securities with the following parameterisation

$$\boldsymbol{\Omega} = \begin{pmatrix} 0.02 & 0.01 \\ 0.01 & 0.01 \end{pmatrix}, \qquad \boldsymbol{A} = \begin{pmatrix} 1. & 0.5 \\ 0.5 & 1. \end{pmatrix}, \qquad \boldsymbol{a}_0 = \mathbf{0}.$$

For the choice of the non-linearity in the Eq. (8.2), we investigate four different choices $F(x) \in \{x, |x|, \tanh(x), \sin(x)\}$ where the choice $F(x) = \sin(x)$ corresponds to the dynamics of the co-integration factor being governed by the Kuramoto model and $F(x) = x$ corresponds to the standard formulation of the co-integration factor in the literature. The results of the numerical investigation for these four different functions are provided upon Figure 8.2. Visually observing the structure of the function $\tilde{h}$ in the numerical output, we notice that all four functions exhibit a common structure and set of values along the line $x_1 = x_2$, i.e. when $F = 0$. This has been confirmed analytically in the linearisation of F around 0 in Eq. (8.3). The effect of the non-linearity is only noticeable when there is a sizeable deviation between the prices of the two securities thus the choice of the F specifically impacts the dynamics of re-synchronisation between the two securities. For the choice of $F(x) = x$ and to a somewhat lesser extent $F(x) = |x|$, the 'elastic pullback' that exists when the stocks deviate can be seen to be large relative to the other two choices which would imply a rapid 'snap back' to synchronisation between the two securities. The $\sin(x)$ and $\tanh(x)$ functions induce a more gradual return to synchronisation in such a scenario however the $\sin(x)$ function has the feature that after

a given level the 'elastic pullback' becomes weaker. It should be noted that in most cases where we are dealing with relative prices in the implementation of these models, that these large deviations between security prices are unlikely to be observed outside of extremely large volatilities. However it should be noted that these observations are only true for the selections of functions used in this analysis and should we wish for the non-linearity to be significant upon small deviations, an appropriate function F can be chosen for this purpose.

As an experimental implementation of the model Eq. (8.1) where we use $F(x) = \sin(x)$, i.e. the Kuramoto model, in the co-integration factor in Eq. (8.2), we simulate three securities with the following parameterisation

$$\boldsymbol{\Omega} = \begin{pmatrix} 0.05 & 0.01 & 0.005 \\ 0.01 & 0.025 & 0.001 \\ 0.005 & 0.001 & 0.01 \end{pmatrix}, \qquad \boldsymbol{A} = \begin{pmatrix} 1. & 0.1 & 1. \\ 0.1 & 1. & 0. \\ 1. & 0. & 1. \end{pmatrix}, \qquad \boldsymbol{a}_0 = \mathbf{0}.$$

and initial conditions

$$\boldsymbol{Y}_t^T = [11.1, 12.0, 11.0]\,.$$

The results of this simulation are provided on Fig 8.3 in which we plot the prices of the three assets Y_t^i given by Eq. (8.1), the co-integration factor α_t^i from Eq. (8.2), the wealth process X_t^π from Eq. (8.4) and the position π_t^i/Y_t^i where π_t^i is the dollar value invested in the asset Y_t^i and is given by the optimal portfolio in Eq. (8.9). Based upon the matrix A, we should expect that the dynamics assets Y_t^1 and Y_t^3 are strongly coupled together whereas this is less so for the dynamics of the asset Y_t^2 and this is visually the case in the simulated plot of the asset prices. Due to high coupling between Y_t^1 and Y_t^3, shifts in positions are most prevalent in these two assets whereas a position is very rarely taken in Y_t^2 as its impact on the dynamics of the other assets is relatively small. We also note that the wealth process is generally increasing as desired from an optimal portfolio.

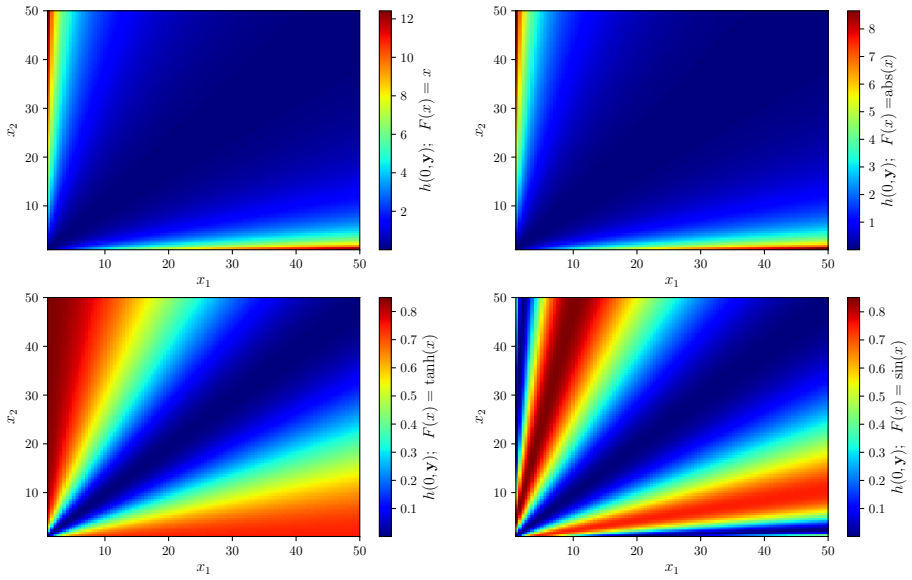


FIGURE 8.2: Numerical implementation of the ADI method upon the PDE in Eq. (8.8) using the Vetzal approximation at the boundaries as in Eq. (8.12) and terminal condition $\tilde{h}(T, \mathbf{y}) = 0$ with $T = 1$ and $\Delta T = 0.1$. Four different simulations have been run for four different choices of the non-linearity $F(x) \in \{x, |x|, \tanh(x), \sin(x)\}$ where the choice $F = \sin(x)$ in the bottom right plot represents the Kuramoto model.

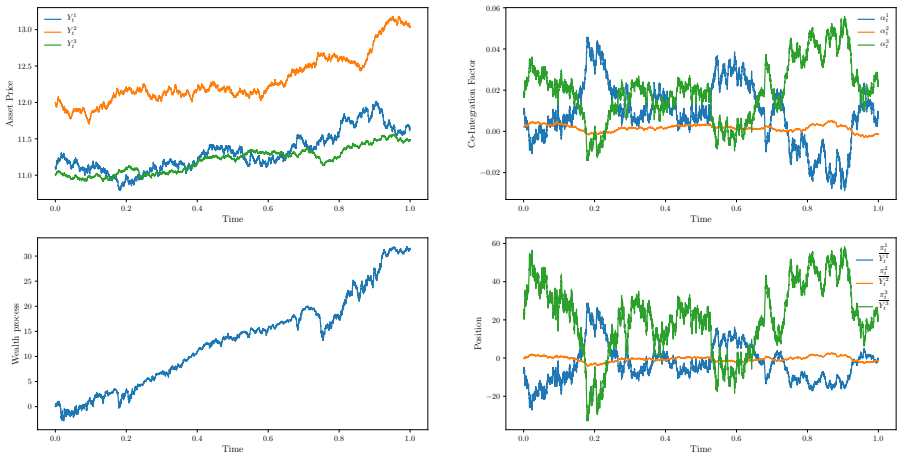


FIGURE 8.3: A single sample path of the co-integrated asset prices and the resulting optimal strategy. Top left: The asset prices. Top Right: The co-integration factors. Bottom left: The wealth process. Bottom right: Asset position.

Chapter 9

Conclusion and Future Work

To finalise this thesis, we now present a brief summary of the results presented and discuss future works that may come about as a result of this research.

We now summarise the novel results of this thesis. In Chapter 3 we studied the Blue vs Red model from an analytical and numerical point of view to understand how the phase lag values and the shape of the noise distribution can lead to steady or chaotic behaviour. This work has been published in the paper titled 'Two Network Kuramoto-Sakaguchi Model Under Tempered Stable Lévy Noise' in [22] and extends upon the previously studied deterministic [45] and Gaussian stochastic [6] versions of the Blue-Red system. Here our analytic study showed that the bulk behaviour can be described by dynamics in the presence of tilted ratchet potentials manifesting metastability in regimes that would otherwise have been expected to be chaotic and this was consist with the results from our numerical study of the full nonlinear system.
The Blue vs Red model was then extended in Chapter 4 to include spatial dynamics through the swarming framework as established in [12]. This Blue vs Red swarming model served as the basis for two proposed attrition models that could potentially be used to model a wargame between the Blue and Red networks based upon the Lanchester equations [13].
In Chapter 5 we have proposed new simulation methods for the simulation of conditioned Lévy processes based upon the WHMC algorithm [5]. We explored these WHMC extensions in the context of Lévy bridges, Lévy processes conditioned to stay positive and Lévy processes conditioned on their extrema. Although work has been done upon the theory of conditioned Lévy processes [15–17], discussion on the practical simulation of these conditioned processes is somewhat lacking in the literature and thus one of the aims of this chapter was to contribute towards this discussion.
The methods for simulating conditioned processes was then applied in Chapter 6 where

we considered a standard synchronised Kuramoto system [1] driven by stochastic noise upon an arbitrary network in which a random node was conditioned to escape from the group synchronised state. Here we established that the degree of the conditioned node is a good predictor for the dynamics of the system, in addition the closeness centrality of the conditioned node and the clustering coefficient were found to be potentially useful predictors.

In Chapter 7 we explored the implementation of stochastic noise in a Kuramoto model to capture the ability of agents to randomly 'jump' ahead in their decision state in time-stressed conditions where intuition becomes significant. This study based upon the data studied in [18], although the aforementioned study was outside of the context of the synchronisation of decision making. This chapter has been submitted for publication in the paper titled 'Modelling militarywatch staff using stochastic network synchronisation' in [23]. Here we observed a high probability that synchronisation may be lost as a consequence of the high information demand of some agents in the crisis scenario.

Finally in Chapter 8 we derived a stochastic synchronisation model in the context of pairs trading, derived the optimal portfolio for the model and numerically evaluated the model when the sychronisation dynamics are driven by the Kuramoto model. Similar pairs trading model have been studied in the literature [19–21] and our goal in this chapter was to provide a discussion on an extension to the sychronisation setting.

There are a number of extensions and prospects for future research that can come about as a result of the work presented in this thesis. From our work on the simulation of Lévy processes, we note that we can extend the idea of utilising the idea of the stochastic grid and increments based upon Wiener-Hopf factors to a variety of different conditioning problems. Some potential conditioning problems can can be explored are: conditioning both on the extrema and on the terminal value, conditioning on the extrema and to stay positive and/or negative and similarly for various combinations of conditionings. In terms of the application of our work on conditioned processes, there is potential for research in applications of the Lévy processes conditioned to stay positive in financial models, particularly in the context of the Bachelier model which has the drawback of negative prices that would be averted in this context. The use of a conditioned Lévy process in this Bachelier setting would be ideally suited to the modelling of penny stock dynamics of which little work is available in the current literature. A well studied area in stochastic processes is the calibration of such processes to real world data however it is less clear how to proceed in the case of effectively calibrating conditioned processes. In the context of our work on sychronisation modelling, we note that many of the models considered can be further extended through the application of β-family of Lévy processes as noise. The work in the stochastic synchronisation model in the pairs trading setting can be extended to formally consider real world data, i.e., calibrating the model

against the JPM and GS stock prices for a variety of different nonlinearities. The resulting optimal portfolios from these mode can than be compared to the existing pairs trading models in the literature. Finally, our proposed swarming model with Lanchester-type attrition has set the basis for future work which employs spatial dependency into attrition modelling.

Appendix A

Numerical Methods

A.1 Bayesian optimisation

Bayesian optimisation is different to other kinds of optimisation in that it constructs a probabilistic model for seeking the maximum of a function $f(x)$ on some bounded set $\mathcal{X} \subset \mathbb{R}^d$ and then exploits this model to make decisions about where in $\mathcal{X}$ to evaluate next. It uses all of the information available from previous evaluations of $f(x)$ and not simply local gradient and Hessian approximations. This gives an algorithm that finds the maximum of non-convex functions with relatively few evaluations at the cost of performing more computation to determine the next point to try. This approach is particularly well suited when the evaluation of $f(x)$ is expensive to perform. For a full review of Bayesian optimisation, see for e.g [7, 8].

The two major aspects of the implementation of Bayesian optimistion are the selection of the prior that expresses the assumptions about the function being optimised and the selection of an acquisition function which is used as a utility function from the model posterior which allows for the next point to evaluate.

A common choice of a prior in Bayesian optimisation due to its flexibility and tractability is the Gaussian process prior [107]. The Gaussian process is defined as the continuous time stochastic process of the form $f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$ for some mean function $m : \mathcal{X} \to \mathbb{R}$ and covariance function $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ with kernel k. Our observations are then taken to be of the form $\{\mathbf{x}_n, y_n\}_{n=1}^N$ where $y_n \sim \mathcal{N}(f(\mathbf{x}), \epsilon)$ where ϵ is the variance of noise introduced into the function observations. The flexibility of the Gaussian process comes into play in the choice of the covariance function.

A common choice is that of the squared exponential function

$$k_{SE}(\mathbf{x}, \mathbf{x}') = \theta_0 \exp\left\{-\frac{1}{2}\sum_{n=1}^{N}\left(\frac{x_n - x'_n}{\theta_n^2}\right)^2\right\}$$

where $\theta = [\theta_0, \theta_1, \ldots, \theta_n]$ are the so-called hyperparameters. Selection of the hyperparameters is typically done through the maximisation of the marginal likelihood under the Gaussian process given θ from an initial random seeding [182, 183].

One downside of the square exponential function is that the sample functions obtained from this covariance function are too smooth for practical optimisation use. A recommended alternative is that of the Matérn kernel [184] which is given by

$$k_M(\mathbf{x}, \mathbf{x}') = \frac{2^{1-\nu}}{\Gamma(\nu)}\left(2\sqrt{\nu}\left\|\mathbf{x} - \mathbf{x}'\right\|\right)^{\nu} K_\nu\left(2\sqrt{\nu}\left\|\mathbf{x} - \mathbf{x}'\right\|\right),$$

where $\Gamma(\cdot)$ and $K_\nu(\cdot)$ are the Gamma function and Bessel function of order ν respectively.

In the application of a Bayesian optimisation algorithm, selection of the next point $\mathbf{x}_{t+1}$ given the observations $\{\mathbf{x}_{1:t}, f(\mathbf{x}_{1:t})\}$ is obtained through the predictive distribution $P(f_{T+1}|\mathcal{D}_{1:t}, \mathbf{x}_{t+1})$. Through use of the Sherman-Morrison-Woodbury formula [185], the predictive distribution can be shown to be given by

$$P(f_{T+1}|\mathcal{D}_{1:t}, \mathbf{x}_{t+1}) \sim \mathcal{N}(\mu_t(\mathbf{x}_{t+1}), \sigma_t^2(\mathbf{x}_{t+1})).$$

The mean and variance functions in the predictive distribution are given by

$$\begin{aligned}\mu_t(\mathbf{x}_{t+1}) &= \mathbf{k}^T\mathbf{K}^{-1}\mathbf{f}_{1:t}\\ \sigma_t^2(\mathbf{x}_{t+1}) &= k(\mathbf{x}_{t+1}, \mathbf{x}_{t+1}) - \mathbf{k}^T\mathbf{K}^{-1}\mathbf{k},\end{aligned}$$

where $\mathbf{f}_{1:t} = f(\mathbf{x}_{1:t})$, the kernel matrix is given by

$$\mathbf{K} = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \ldots & k(\mathbf{x}_1, \mathbf{x}_t), \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_t, \mathbf{x}_1) & \ldots & k(\mathbf{x}_1, \mathbf{t}_t) \end{bmatrix},$$

and $\mathbf{k} = [k(\mathbf{x}_t, \mathbf{x}_1), k(\mathbf{x}_t, \mathbf{x}_2), \ldots, k(\mathbf{x}_t, \mathbf{x}_t)]$.

The second part of Bayesian optimisation after having constructed and updated the prior is the acquisition function which guide the algorithm in its search for the optimal value. The acquisition function, denoted by $a : \mathcal{X} \to \mathbb{R}^+$, seeks for this next value through a balancing of the principles of exploration, which seeks places with a high variance, and

exploitation, which seeks places with a low mean. The acquisition function balances these for our proxy optimisation to determine the next evaluation at $\text{argmax}_{\mathbf{x}} a(\mathbf{x})$.

We first denote the best current value as $\mathbf{x}_{\text{best}} = \text{argmin}_{\mathbf{x}_n} f(\mathbf{x}_n)$ and

$$\gamma(\mathbf{x}) = \frac{f(\mathbf{x}_{\text{best}}) - \mu(\mathbf{x}) - \xi}{\sigma(\mathbf{x})},$$

where ξ is a trade-off parameter that balances exploration against exploration. In practise it is recommended that ξ is initially high to encourage exploration and decreased towards 0 as the algorithm continues [186]. Some acquisition functions that are commonly used are the probability of Improvement [186] which looks to maximise the probability of improving over the best current value and is given by

$$a_{PI}(\mathbf{x}) = \Phi(\gamma(\mathbf{x})),$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal. The Expected Improvement [187] looks to maximise the expected improvement over the current best and is given by

$$a_{EI}(\mathbf{x}) = \sigma(\mathbf{x})\phi(\gamma(\mathbf{x})) + \sigma(\mathbf{x})\gamma(\mathbf{x})\Phi(\gamma(\mathbf{x})),$$

where $\phi(\gamma(\mathbf{x}))$ is the PDF of the standard normal distribution. And finally the GP Upper Confidence Bound [188] which seeks to exploit the lower (or upper under maximisation) confidence bounds to construct acquisition functions that minimise regret over the course of their optimisation and is given by

$$a_{LCB}(\mathbf{x}) = \mu(\mathbf{x}) - \kappa\sigma(\mathbf{x}),$$

where κ is a parameter that balances exploitation against exploration.

In summary, Bayesian optimisation can be summed in the following steps, finding x_t by optimising the acquisition function over the Gaussian process, sampling the objective function and finally augmenting the data and updating the Gaussian process. These steps are presented below in Algorithm 17:

Algorithm 17 Bayesian Optimisation

1: **for** $t = 1, 2 \ldots$ **do**
2: $\quad \mathbf{x}_t = \text{argmax}_{\mathbf{x}} a(\mathbf{x}|\mathcal{D}_{1:t-1} := \{\mathbf{x}_{1:t-1}.y_{1:t-1}\})$
3: $\quad y_t = f(\mathbf{x}_t) + \epsilon_t$
4: $\quad \mathcal{D}_{1:t} = \{\mathcal{D}_{1:t-1}, (\mathbf{x}_t.y_t)\}$

A.2 The 0-1 test

The 0-1 test [9, 10] provides a means for determining the presence of chaos in time series data that is a computationally efficient alternative to computing a Lyapunov exponent [189]. Utilising the 0-1 test compared with calculating the largest Lyapunov exponent has two clear advantages: the computational costs are greatly reduced and the complexity of the underlying dynamical system is irrelevant to the test as only the time-series data is used. The disadvantage of the simplicity of the 0-1 test is the loss of being able to quantify the chaos by observing the magnitude of the largest Lyapunov exponent. In additional to identifying the existence of chaos in deterministic dynamical system, the 0-1 tests singular reliance upon time series data allows it to be effectively used in the analysis of experimental data. To apply the 0-1 test upon the time series $O(t)$, we select an observable from the solution $\psi(O)$ and some parameter $c > 0$. The method used in this test is independent of the form of the function ψ. We now define

$$\begin{aligned} \theta(t) &= ct + \int_0^t \psi(O(s))ds, \\ p(t) &= \int_0^t \psi(O(s))\cos(\theta(s))ds. \end{aligned} \tag{A.1}$$

If the underlying dynamics are non-chaotic then $p(t)$ is bounded otherwise $p(t)$ behaves asymptotically like Brownian motion. To identify the growth of $p(t)$, the mean-square displacement is used:

$$M(t) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \left(p(t+\tau) - p(\tau)\right)^2 d\tau.$$

If the behaviour $p(t)$ is Brownian, then $M(t)$ will grow linearly in time; otherwise if $p(t)$ is bounded, then $M(t)$ will be bounded. To distinguish between these two properties, the asymptotic growth rate of the mean-square displacement is calculated:

$$K = \lim_{t\to\infty} \frac{\log(M(t))}{\log t}.$$

This can be calculated through a linear regression of $\log(M(t))$ and $\log t$ and will be near 1 if the underlying dynamics are chaotic and 0 otherwise.

The application of the 0-1 test relies on the validity of the claim of the asymptotic behaviour of $p(t)$ in (A.1) based on the chaotic or non-chaotic nature of the underlying system. For some $c \in (0, \pi)$, the translation variables are constructed through the

extended system given below

$$\begin{aligned} \dot{\theta} &= c, \\ \dot{p} &= \psi(O(t)) \cos\theta, \\ \dot{q} &= \psi(O(t) \sin\theta, \end{aligned} \tag{A.2}$$

where (θ, p, q) represent coordinates on the Euclidean group $\mathbf{E}(2)$ of rotations θ and translations (p, q) in the plane.

As an example of the regular and chaotic dynamics obtained through the system in (A.2), we consider the periodically forced self-excited oscillator known as the Duffing-Van der Pol oscillator, governed by the equation

$$\ddot{x} - \mu(1 - x^2)\dot{x} + \alpha x + \beta x^3 + \delta x^5 = f(t), \tag{A.3}$$

where $\mu > 0$, α,β and δ are constant parameters, and $f(t) = f\cos\omega t$. A plot of the translation variables given in (A.2) for the Duffing-Van der Pol oscillator (A.3) is provided on Figure A.1.

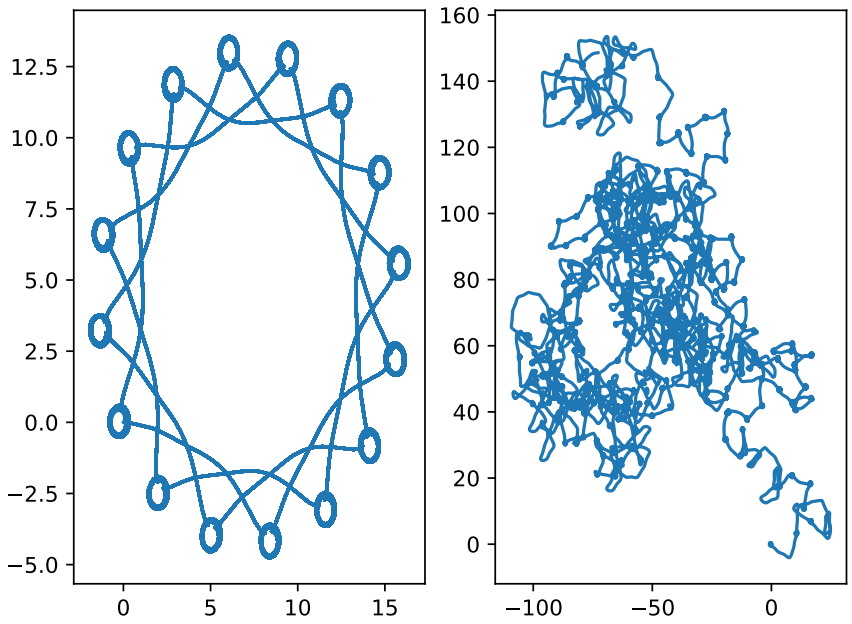


FIGURE A.1: Dynamics of the translation components (p, q) of the $\mathbf{E}(2)$-extension for the Duffing-van der Pol system $t \in [0, 1000]$ with $\alpha = 0.22$, $\beta = -0.58$, $\delta = 0.16$, $\mu = 0.68$ and $f = 1.8$. The trajectory in the right plot is bounded as the van der Pol system is regular when $\omega = 1.6$. The left plot shows the case when the trajectory undergoes a Brownian motion due to the chaotic dynamics of the Duffing van der Pol system for $\omega = 2$.

Although the $\mathbb{R}$-extension can be used as a detector for chaos, by the ergodic theorem the growth rate of $p(t)$ will be linear and the rate will be equal to that of the space average of $\psi(O(t))$. The result of this linear growth is that $K = 2$ regardless of regular or chaotic dynamics in the underlying system. This leads to the necessity of removing the linear term in $p(t)$ to differentiate between different dynamics. The construction of the skew product system in (A.2) is designed to remove the linear linear term. One of the key results that allow for the justification of the 0-1 test arises from the following theorem which summarises the major results from the work done in [190]:

Theorem A.1 *[Euclidean extensions of dynamical systems [190].] Given a dynamical system X with chaotic dynamics in the sense that periodic points are dense in X, then regardless of the ergodic properties of the transformation on X, generic Γ-extensions have trajectories that are unbounded when $\Gamma = \mathbb{R}^n$ for all n and $\Gamma = \mathbf{SE}(n)$ (for $n = 2$ or odd n) where $\Gamma = \mathbf{SE}(n) = \mathbf{SO}(n) \times \mathbb{R}^n$ is the special Euclidean group of rotations $\mathbf{SO}(n)$ and translations $\mathbb{R}^n$ in n-dimensional space. Furthermore for $\Gamma = \mathbf{SE}(n)$ (for $n = 2$ or odd n) this growth is sublinear.*

From the theorem in Theorem A.1 and the results in [191], the dynamics of the p and q components are shown to behave like a Brownian motion under the condition that the chaotic attractor is uniformly hyperbolic. There exists a weak condition upon the magnitude of the chaos of the underlying dynamic that guarantees the growth rate of $K = 1$ for the mean-square displacement. This weak condition states that if there are constants $C > 0$ and $\alpha > 2$ such that

$$\left| \int \psi(O(t)) \cos(\theta(t)) \psi(O_0) \cos\theta_0 dO_0 d\theta_0 \right| \leq Ct^{-\alpha}$$

for all $t > 0$ then K = 1 [192]. More simply, this weak condition states that if autocorrelation function of $\psi(O)\cos(\theta)$ decays at rate greater than quadratic, then $K = 1$ as desired.

In the practical implementation of the 0-1 test as discussed in [10], for some $c \in (0, \pi)$, the translation variables from (A.2) are constructed through the observable $\psi(O)$ as

$$p_c(n) = \sum_{j=1}^{n} \psi(j) \cos(jc),$$
$$q_c(n) = \sum_{j=1}^{n} \psi(j) \sin(jc),$$

where $n \in 1, 2, \ldots N$. The diffusive behaviour of p_c and q_c are then analysed through the mean square displacement which is calculated through

$$M_c(n) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} [p_c(j+n) - p_c(j)]^2 + [q_c(j+n) - q_c(j)]^2 \tag{A.4}$$

Using this definition requires $n << N$. A sample approximation to the limit can be obtained by using a value of $n_{cut} = N/10$ as a stopping point. An alternative approach as developed and proven in [193] uses a modified mean square displacement $D_c(n)$ which has the same asymptotic growth as $M_c(n)$ but has better convergence properties. For an observation $\psi(j)$, the autocorrelation function is given by

$$\rho(k) = \mathbb{E}(\psi(1)\psi(k+1)) - (\mathbb{E}(\psi))^2, \quad k = 0, 1, 2, \ldots$$

Provided the autocorrelation function is absolutely summable ($\sum_{k=0}^{\infty} |\rho(k)| < \infty$) then for each $c \in (0, \pi)$ we can write

$$M_c(n) = V(c)n + V_{osc}(c, n) + e(c, n),$$

where $e(c, n)/n \to 0$ as $n \to \infty$ uniformly in c and

$$V_{osc}(c, n) = (\mathbb{E}(\psi))^2 \frac{1 - \cos(nc)}{1 - \cos c}.$$

By subtracting the explicit term $V_{osc}(c, n)$ from the mean square displacement, we obtain an improved version in form

$$D_c(n) = M_c(n) - V_{osc}(c, n),$$

The removal of the oscillatory behaviour captured by the term $V_{osc}(n)$ allows for a better determination of the growth rate Kc as the error term $e(c, n)$ decays uniformly in $c \in (0, \pi)$.

A simple method to calculate the asymptotic growth rate K_c is done by taking a linear regression of the log-log plot of the mean-square displacement M_c. When using the modified mean-square displacement $D_c(n)$, some care must be taken as the subtraction of oscillatory term can result in negative values for $D_c(n)$. To address this issue we set

$$\tilde{D}_c(n) = D_c(n) - \min_{n=1,\ldots,n_{cut}} D_c(n),$$

and then obtain the asymptotic growth rate

$$K_c = \lim_{n\to\infty} \frac{\log \tilde{D}_c(n)}{\log n},$$

which can be approximated numerically by regression on the graph of $\log \tilde{D}_c(n)$ versus $\log n$. An alternative method is to calculate the correlation coefficient

$$K_c = \mathrm{corr}(\xi, \Delta) = \frac{\mathrm{cov}(\xi, \Delta)}{\sqrt{\mathrm{var}(\xi)\mathrm{var}(\Delta)}},$$

for the vectors $\xi = (1, 2, \ldots, n_{cut})$ and $\Delta = (D_c(1), D_c(2), \ldots, D_c(n_{cut}))$ In practice the correlation method proves to be a more precise method then that of the regression method.

In most applications of the 0-1 test, the random selection of c results in either $K_c \approx 0$ or $K_c \approx 1$ for regular or chaotic dynamics respectively. However there exist isolated values of c which result in a resonance effect at which $K_c \approx 2$ at these points regardless of the dynamics of the underlying system. To understand how these resonances arise, we must observe how the translation dynamics are structured in (A.4). If the Fourier decomposition of the observation ψ contains a term proportional to $\exp(-i\omega k)$ then there is a resonance as the $\cos(cj)$ term has a resonant frequency at $c = \omega$. At this frequency the translation variable is approximately $p_c \propto (n)$ and hence $M_c(n) \propto n^2$ regardless of the dynamics. The existence of resonant frequencies justifies the use of the median value of K_c rather than that of the mean, as outliers produced by such resonances do not affect the final result in the former case. To avoid these resonances that distorted the statistics, a restriction is placed on the range of c to a subset of the interval $[0, \pi]$. An example of a good interval range for use in computations is $c \in (\pi/4, 4\pi/5)$. A recommended sampling value N_c of c is 100 [10].

In Algorithm 18 we provide a summary of the practical implementation of the 0-1 test which incorporated all of the aforementioned details.

Algorithm 18 The 0-1 test given the observable $\psi(j)$ with $j \in 1, \ldots, N$

$n_{cut} \leftarrow N/10$
$N_c \leftarrow 100$
for $i \in 1, \ldots, N_c$ **do**
    $t \leftarrow 1, \ldots, n_{cut}$
    sample $c \sim \mathrm{U}(\pi/5, 4\pi/5)$
    $p \leftarrow \mathrm{cumsum}(\psi \cdot \cos(j \cdot c))$
    $q \leftarrow \mathrm{cumsum}(\psi \cdot \sin(j \cdot c))$
    for $n \in 1, \ldots, n_{cut}$ **do**
        $M[n] \leftarrow \mathrm{mean}\left((p[n : N_c] - p[1 : N_c - n])^2 + (q[n : N_c] - q[1 : N_c - n])^2\right)$
        $D[n] \leftarrow M[n] - (\mathrm{mean}(\psi))^2 \cdot (1 - \cos(n \cdot c))(1 - \cos(c))$
    $\tilde{D} \leftarrow D - \min(D)$
    $kcorr \leftarrow corr(t, \tilde{D})$
return $\mathrm{median}(kcorr)$

As an example of the implementation of the 0-1 test, we provide comparison of the 0-1 test and the maximum Lyapunov exponent upon the time series data simulated from the dynamical system in (A.3) where we vary the frequency of the driving force ω across the interval $\omega \in [1, 2.5]$. Within this range, the system is known to have chaotic behaviour in various parameter windows after $\omega = 1.141$ and for a more sustained parameter window after $\omega = 1.917$.

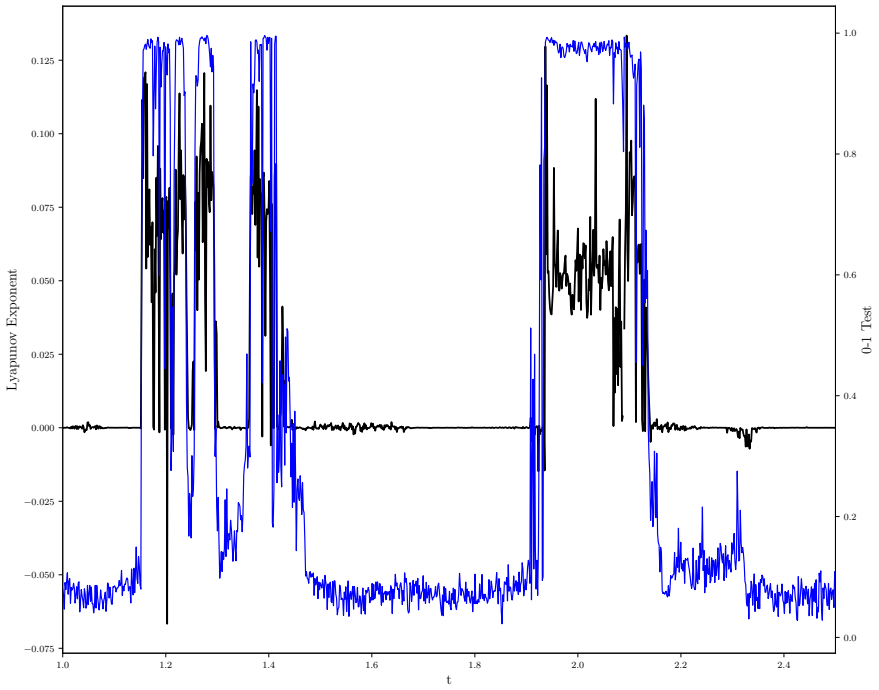


FIGURE A.2: Comparison of the 0-1 test and the maximum Lyapunov exponent for the Duffing-van der Pol system $t \in [0, 1000]$ with $\alpha = 0.22$, $\beta = -0.58$, $\delta = 0.16$, $\mu = 0.68$ and $f = 1.8$. The black line represents the value of the maximum Lyapunov exponent and the blue line represents the value of the 0-1 test.

A.3 The Feynman-Kac method

We shall now consider the probability measure change induced by the market price of risk vector $\boldsymbol{\lambda}_t$ which through the Radon-Nikodym derivative defined by

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = e^{\frac{1}{2}\int_0^T \|\boldsymbol{\lambda}_s\|^2 ds - \int_0^T \boldsymbol{\lambda}_s' d\boldsymbol{W}_s}$$

induces a new measure $\mathbb{P}^*$ so that Girsanov's theorem implies that

$$\boldsymbol{W}_t^* = -\int_0^t \boldsymbol{\lambda}_s ds + \boldsymbol{W}_t$$

are independent $\mathbb{P}^*$ Brownian motions. We shall now choose λ_t such that $\mathbb{P}^*$ is the risk-neutral measure, which implies that traded assets Y_t are martingales. Since

$$\begin{aligned} dY_t^k =& Y_t^K \alpha_t^k dt + \sum_{i=1}^n \sigma_{ki} Y_t^k dW_t^i \\ =& (\alpha_t^k + \sum_{i=1}^n \sigma_{ki}\lambda_t^i) Y_i^k dt + \sum_{i=1}^n \sigma_{ki} Y_t^k dW_t^{*i}, \end{aligned}$$

the martingale conditions implies that the processes are driftless, we require

$$\alpha_t^k = \sum_{i=1}^n \sigma_{ki}\lambda_t^i,$$

and hence, after applying the invertibility of $\boldsymbol{\sigma}$, we obtain the expression for the market price of risk vector

$$\boldsymbol{\lambda}_t = -\boldsymbol{\sigma}^{-1}\boldsymbol{\alpha},$$

and the driftless process

$$dY_t^k = \sum_{i=1}^n \sigma_{ki} Y_t^k dW_t^{*i}.$$

Using the probability measure $\mathbb{P}^*$, we can write the PDE in (8.8) as

$$\left(\partial_t + \mathcal{L}^{*,y}\right)\tilde{h} + \frac{\boldsymbol{\alpha}^T \Omega^{-1} \boldsymbol{\alpha}}{2\gamma} = 0,$$

with the terminal condition $\tilde{h}(T, \mathbf{y})$ and where $\mathcal{L}^{*,\mathbf{y}}$ is the $\mathbb{P}^*$-infinitesimal generator of $\mathbf{y}$. Using the Feynman-Kac formula [194], we can then write $h(t, \mathbf{y})$ in terms of the following expectation:

$$h(t, \mathbf{y}) = \mathbb{E}_{t,\mathbf{y}}^* \left[\int_t^T \frac{\boldsymbol{\alpha}_s^T \boldsymbol{\Omega}^{-1} \boldsymbol{\alpha}_s}{2\gamma} ds \right], \tag{A.5}$$

where $\mathbb{E}^*_{t,\mathbf{y}}$ denotes the $\mathbb{P}^*$-expectation given $\boldsymbol{y}_t = \boldsymbol{y}$. Taking into account the expression of the optimal portfolio in (8.9), we can then use the expectation in (A.5) to write

$$\boldsymbol{\pi}^* = \frac{1}{\gamma}\left(\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha} - \boldsymbol{\Gamma}\cdot\boldsymbol{\Omega}^{-1}\mathbb{E}^*_{t,\mathbf{y}}\left[\int_t^T \boldsymbol{\alpha}_s ds\right]\right),$$

which can then be numerically solved for through a Monte-Carlo procedure.

A.4 Multi-level Monte Carlo

Many of the WHMC algorithms described in this thesis can be extended in the Multi-level Monte Carlo setting. This standard Monte Carlo technique can be further extended to the multilevel Monte Carlo setting as discussed in detail by Ferreiro-Castilla et. al in [195]. The idea behind the multi-level Monte-Carlo setting relies upon a reduction of the variance of the traditional MC method and is summarised as follows. Fix $n_0, L \in \mathbb{N}$ and set $n_l = 2^l n_0$ for $l = 0, \ldots, L$ and define

$$F^{n_l} := (V(n_l, n_l/t), J(n_l, n_l/t)),$$

Thus the larger n becomes, the finer the approximation F^{n_l} becomes due to the convergence to $F(L_t, \overline{L}_t)$ in the limit $n_l \to \infty$. Since the finest approximation F^{n_L} can be written in terms of the coarser approximations as

$$\mathbb{E}[F^{n_L}] = \mathbb{E}[F^{n_0}] + \sum_{l=1}^{L}\mathbb{E}[F^{n_l} - F^{n_{l-1}}].$$

an estimate of F^n can be obtained by independently computing each of the expectations of the telescoping sum by a standard Monte-Carlo method:

$$\mathbb{E}[F(L_t, \overline{L}_t)] \approx \hat{F}_{ML}^{\mathcal{M}(n_0,L)} := \frac{1}{M_0}\sum_{m=1}^{M_0} F^{n_0,m} + \sum_{l=1}^{L}\frac{1}{M_l}\sum_{i=1}^{M_l}(F^{n_l,m} - F^{n_{l-1},m}),$$

where $\mathcal{M}(n_0, L) := \{M_l\}_{l=0}^L$ are the Monte Carlo trails in each level and $F^{n_l,m}$ denotes the m-th sample from F^{n_l}.

To see that this approach does indeed result in variance reduction, we first note that we can express the mean-square error between the WHMC estimate denoted as $\hat{F}^{n,M}$ and

the true quantity $\mathbb{E}\left[F(L_t, \bar{L}_t)\right]$ as follows

$$\begin{aligned}
\text{MSE}(\hat{F}^{n,M}) =& \mathbb{E}\left[\left(\hat{F}^{n,M} - \mathbb{E}\left[F(L_t, \bar{L}_t)\right]\right)^2\right], \\
=& \mathbb{E}\left[\left(\hat{F}^{n,M} - \mathbb{E}\left[\hat{F}^{n,M}\right]\right)^2\right] + \left(\mathbb{E}\left[\hat{F}^{n,M}\right] - \mathbb{E}\left[F(L_t, \bar{L}_t)\right]\right)^2, \\
=& \text{Var}\left[\hat{F}^{n,M}\right] + \left(\mathbb{E}\left[\hat{F}^{n,M}\right] - \mathbb{E}\left[F(L_t, \bar{L}_t)\right]\right)^2, \\
=& \frac{1}{M}\text{Var}\left[F^n\right] + \left(\mathbb{E}\left[F^n - F(L_t, \bar{L}_t)\right]\right)^2, \qquad \text{(A.6)}
\end{aligned}$$

where we have used the facts that $\mathbb{E}\left[\hat{F}^{n,M}\right] = \mathbb{E}\left[F^n\right]$ and $\text{Var}\left[\hat{F}^{n,M}\right] = \frac{1}{M}\text{Var}\left[F^n\right]$. Note that the first term corresponds to the variance of the estimator and the second term corresponds to the bias. The mean square error of the estimator $\hat{F}_{ML}^{\mathcal{M}(n_0,L)}$ from the multilevel Monte Carlo method can be expanded in a similar manner as in (A.6) to obtain

$$\begin{aligned}
\text{MSE}(\hat{F}_{ML}^{\mathcal{M}(n_0,L)}) =& \mathbb{E}\left[\left(\hat{F}_{ML}^{\mathcal{M}(n_0,L)} - \mathbb{E}\left[F(L_t, \bar{L}_t)\right]\right)^2\right], \\
=& \text{Var}\left[\hat{F}_{ML}^{\mathcal{M}(n_0,L)}\right] + \left(\mathbb{E}\left[\hat{F}_{ML}^{\mathcal{M}(n_0,L)}\right] - \mathbb{E}\left[F(L_t, \bar{L}_t)\right]\right)^2, \\
=& \frac{1}{M_0}\text{Var}\left[F^{n_0}\right] + \sum_{l=1}^{L}\frac{1}{M_l}\text{Var}\left[F^{n_l} - F^{n_{l-1}}\right] + \left(\mathbb{E}\left[F^n - F(L_t, \bar{L}_t)\right]\right)^2, \qquad \text{(A.7)}
\end{aligned}$$

where we have used the fact that $\mathbb{E}\left[\hat{F}_{ML}^{\mathcal{M}(n_0,L)}\right] = \mathbb{E}\left[F^n\right]$, $\text{Var}\left[\hat{F}_{ML}^{n_0,M}\right] = \frac{1}{M}\text{Var}\left[F^{n_0}\right]$ and $\text{Var}\left[\frac{1}{M_l}\sum_{i=1}^{M_l}(F^{n_l,m} - F^{n_{l-1},m})\right] = \text{Var}\left[\frac{1}{M_l}\sum_{i=1}^{M_l}(F^{n_l} - F^{n_{l-1}})\right]$.

Comparing the MSE in (A.7) to the MSE in (A.6) we note that the bias remains the same but the variance is reduced due to the convergence of the variance term $\text{Var}\left[F^{n_l} - F^{n_{l-1}}\right] \to 0$ as $n_l \to \infty$. In the WHMC setting, given the sequences $\left\{S_{n_l/t}^{(j)} : j \geq 1\right\}$ and $\left\{I_{n_l/t}^{(j)} : j \geq 1\right\}$, it can be shown that sampling from $F^{n_{l-1}}$ can be obtained through the sequences $\left\{S_{n_{l-1}/t}^{(i)} : i \geq 1\right\}$ and $\left\{I_{n_{l-1}/t}^{(i)} : i \geq 1\right\}$ as follows

$$\begin{aligned}
S_{n_{l-1}/t}^{(i)} =& \max\left\{\sum_{j=1}^{k-1}\left(S_{n_l/t}^{(\kappa_{i-1}+j)} + I_{n_l/t}^{(\kappa_{i-1}+j)}\right) + S_{n_l/t}^{(\kappa_i-\kappa_{i-1})}; k \in [1, \kappa_i - \kappa_{i-1}]\right\}, \\
I_{n_{l-1}/t}^{(i)} =& \sum_{j=\kappa_{i-1}+1}^{\kappa_i}\left(S_{n_l/t}^{(j)} + I_{n_l/t}^{(j)}\right) - S_{n_{l-1}/t}^{(i)},
\end{aligned}$$

where $\kappa_0 = 0$ and $\kappa_i - \kappa_{i-1}$ are a sequence of i.i.d geometrically distributed random variables with parameter 0.5. For a full discussion on how to then effectively sample from the coarser level $F^{n_{l-1}}$ given F^{n_l} we refer the reader to [195].

Appendix B

Topics in Stochastic Processes

B.1 The β-family of Lévy processes

The Lévy measure of the β-family as introduced by Kuznetsov in [70] is given by

$$\Pi(\mathrm{dx}) = c_1 \frac{e^{-\alpha_1\beta_1 x}}{(1-e^{-\beta_1 x})^{\lambda_1}} \mathbf{1}_{\{x>0\}} + c_2 \frac{e^{\alpha_2\beta_2 x}}{(1-e^{-\beta_2 x})^{\lambda_2}} \mathbf{1}_{\{x<0\}}\, \mathrm{dx},$$

allowing us to observe that this class of processes has both positive and negative jumps. From the application point-of-view, it is interesting to note that the processes in the β-family are similar to the class of tempered stable processes whose Lévy measure is given by

$$\Pi(\mathrm{dx}) = c_+ \frac{e^{-\alpha_+ x}}{x^{\lambda_+}} \mathbf{1}_{\{x>0\}} + c_- \frac{e^{\alpha_- x}}{|x|^{\lambda_-}} \mathbf{1}_{\{x<0\}}\, \mathrm{dx}\,.$$

In fact, the measure of the β-family converges to that of the tempered stable processes by setting $c_1 = c_+\beta^{\lambda_+}, c_2 = c_-\beta^{\lambda_-}, \alpha_1 = \alpha_+\beta^{-1}, \alpha_2 = \alpha_-\beta^{-1}, \beta_1 = \beta_2 = \beta$ and taking the limit $\beta \to 0$. The β-family is also similar to the CGMY family of processes [196] under the parameter restriction $c_1 = c_2$, $\lambda_1 = \lambda_2$ and $\beta_1 = \beta_2$.
The characteristic exponent of this class is given by

$$\begin{aligned}\Psi(u) =& ia\theta + \frac{1}{2}\sigma^2 u^2 + \frac{c_1}{\beta_1}\left\{B(\alpha_1, 1-\lambda_1) - B(\alpha_1 - \frac{iu}{\beta_1}, 1-\lambda_1)\right\} \\ &+ \frac{c_2}{\beta_2}\left\{B(\alpha_2, 1-\lambda_2) - B(\alpha_2 + \frac{iu}{\beta_2}, 1-\lambda_2)\right\},\end{aligned}$$

where $a, \sigma \in \mathbb{R}, c_1, c_2, \alpha_1, \alpha_2, \beta_1, \beta_2 > 0$, $\lambda_1, \lambda_2 \in (0,3)\setminus\{1,2\}$ and $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ is the Beta function.

For $\lambda > 0$, all the roots $\lambda + \Psi(u) = 0$ are simple and real and, looking at the case of the negative roots $\{i\zeta_n^- : n \geq 0\}$, are located on the negative imaginary axis as

$$\begin{aligned}
&\zeta_0^+ \in (0, \beta_2\alpha_2), \qquad \zeta_0^- \in (-\beta_1\alpha_1, 0),\\
&\zeta_n^+ \in (\beta_2(\alpha_2 + n - 1), \beta_2(\alpha_2 + n)) \qquad for\, n \geq 1,\\
&\zeta_n^- \in (\beta_1(-\alpha_1 - n), \beta_1(-\alpha_1 - n + 1)) \qquad for\, n \geq 1.
\end{aligned}$$

The semi-explicit law for the supremum of the β-process is then obtained from:

$$\mathbb{P}(\overline{X}_{\mathbf{e}_1/\lambda} \in dx) = -\left(\sum_{k\geq 0} c_k^- \zeta_k^- e^{\zeta_k^- x}\right) dx \qquad x \geq 0, \tag{B.1}$$

where

$$c_0^- = \prod_{n\geq 1} \frac{1 + \zeta_0^-/(\beta_1(n - 1 + \alpha_1))}{1 - \zeta_0^-/\zeta_n^-}$$

and

$$c_k^- = \frac{1 + \zeta_k^-/(\beta_1(k - 1 + \alpha_1))}{1 - \zeta_0^-/\zeta_n^-} \prod_{n\geq 1, n\neq k} \frac{1 + \zeta_k^-/(\beta_1(n - 1 + \alpha_1))}{1 - \zeta_k^-/\zeta_n^-}.$$

A similar expression holds for $\mathbb{P}(-\underline{X}_{\mathbf{e}_1/\lambda} \in dx)$ with the role of $\{\zeta_n^- : n \geq 0\}$ being played by $\{-\zeta_n^+ : n \geq 0\}$ and α_1, β_1 replaced by α_2, β_2.

Using Eq. (B.1), and its equivalent for the infimum process, one can sample efficiently from the law of $\overline{X}_{\mathbf{e}_1/\lambda}$ and $\underline{X}_{\mathbf{e}_1/\lambda}$. A suitable truncation of the infinite series is required to sample from these random variables.

B.2 Proof of Theorem 5.4

The proof of Theorem 5.4 is similar to that as given in [114].

To show that the first quantity in the 4-tuple, that is the first passage time, converges, we begin by introducing the auxiliary quantity

$$k_{L^\uparrow}^{(n)} := \inf\left\{k \in \{0, \ldots, n\} \,|\overline{L}^\uparrow_{\mathbf{g}(k,\lambda)} > u\right\}, \tag{B.2}$$

which yields the index of the stochastic grid upon which first passage time occurs. Since by the WHMC method, $\bar{L}^{\uparrow}_{\mathbf{g}(k,n/t)} \overset{d}{=} J(k,\lambda)$, it is sufficient to show that

$$\frac{t}{n}(\kappa_u^{(n)} \wedge n) \overset{d}{=} \frac{t}{n}(k_{L^{\uparrow}}^{(n)} \wedge n) \overset{\mathbb{P}}{\to} \tau_u \wedge t. \tag{B.3}$$

Due to the convergence result in (2.15) and by considering the event that the process does not cross the passage level u in a finite time, that is $t < \tau_u$ or $\left\{k_{L^{\uparrow}}^{(n)} = \infty\right\}$, we have that as $n \to \infty$

$$\mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)}=\infty\right\}}(t - \tau_u \wedge t) \leq \mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)}=\infty, \tau_u < t\right\}} t = \mathbf{1}_{\{\mathbf{g}(n,\lambda)<\tau_u, \tau_u<t\}} t \to 0 \quad \mathbb{Q}\text{-almost surely.}$$

In the case where the process does cross the passage level u, that is $t \geq \tau_u$ or $\left\{k_{L^{\uparrow}}^{(n)} < \infty\right\}$, we have that

$$\begin{aligned}\mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)}<\infty\right\}}(\frac{t}{n}k_{L^{\uparrow}}^{(n)} - \tau_u \wedge t) =& \mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)}<\infty, \tau_u \geq t\right\}}(\frac{t}{n}k_{L^{\uparrow}}^{(n)} - \tau_u \wedge t) + \mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)}<\infty, \tau<t\right\}}(\frac{t}{n}k_{L^{\uparrow}}^{(n)} - \tau_u \wedge t) \\ =& \mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)}<\infty, \tau_u \geq t\right\}}(\frac{t}{n}k_{L^{\uparrow}}^{(n)} - t) + \mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)}<\infty, \tau<t\right\}}(\frac{t}{n}k_{L^{\uparrow}}^{(n)} - \tau_u).\end{aligned} \tag{B.4}$$

By the SLLN, the first term in (B.4) converges in the limit $n \to \infty$ as follows:

$$\begin{aligned}\mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)}<\infty, \tau_u \geq t\right\}}(\frac{t}{n}k_{L^{\uparrow}}^{(n)} - t) \leq& \mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)}<\infty, \tau_u \geq t\right\}}\left|\frac{t}{n}k_{L^{\uparrow}}^{(n)} - t\right| \\ \leq& \mathbf{1}_{\{\mathbf{g}(n,\lambda), \tau_u \geq t\}} t \to 0 \quad \mathbb{Q}\text{-almost surely.}\end{aligned}$$

To show convergence in probability for the second term, we consider some $\epsilon > 0$ and some temporal partition $0 = t_0 < t_1 < \cdots < t_N = t$ where $t_i - t_{i-1} < \epsilon/2$ for all i. Thus we need to show that

$$\begin{aligned}\mathbb{P}\left(\mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)}<\infty, \tau<t\right\}}\left|\frac{t}{n}k_{L^{\uparrow}}^{(n)} - \tau_u)\right| > \epsilon\right) =& \sum_{i=1}^{N} \mathbb{P}\left(\mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)}<\infty, \tau \in [t_{i-1}, t_i)\right\}}\left|\frac{t}{n}k_{L^{\uparrow}}^{(n)} - \tau_u\right| > \epsilon\right) \\ \to& 0.\end{aligned} \tag{B.5}$$

Noting that we can produce an upper bound for each term of the summation as

$$\begin{aligned}\mathbb{P}\left(\mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)}<\infty, \tau \in [t_{i-1}, t_i)\right\}}\left|\frac{t}{n}k_{L^{\uparrow}}^{(n)} - \tau_u\right| > \epsilon\right) \leq& \mathbb{P}\left(t_{i-1} - \frac{t}{n}\inf\{k \geq 0 | \mathbf{g}(k,\lambda) > t_{i-1}\} > \frac{\epsilon}{2}\right) \\ &+ \mathbb{P}\left(\frac{t}{n}\inf\{k \geq 0 | \mathbf{g}(k,\lambda) > t_i\} - t_i > \frac{\epsilon}{2}\right),\end{aligned}$$

and the fact that each of the terms above converge in the limit as $n \to \infty$

$$\begin{aligned}\mathbb{P}\left(t_{i-1} - \frac{t}{n}\inf\{k \geq 0 | \mathbf{g}(k,\lambda) > t_{i-1}\} > \frac{\epsilon}{2}\right) &= \mathbb{P}\left(\frac{n(t_{i-1} - \frac{\epsilon}{2})}{t} > \inf\{k \geq 0 | \mathbf{g}(k,\lambda) > t_{i-1}\}\right) \\ &= \mathbb{P}\left(t_{i-1} \leq \mathbf{g}(\frac{n(t_{i-1} - \frac{\epsilon}{2})}{t}, \lambda)\right) \\ &\to \mathbb{P}\left(t_{i-1} \leq t_{i-1} - \frac{\epsilon}{2}, \lambda)\right) \quad \mathbb{Q}\text{-almost surely} \\ &= 0,\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}\left(\frac{t}{n}\inf\{k \geq 0 | \mathbf{g}(k,\lambda) > t_i\} - t_i > \frac{\epsilon}{2}\right) &= \mathbb{P}\left(\inf\{k \geq 0 | \mathbf{g}(k,\lambda) > t_i\} > \frac{n(t_i + \frac{\epsilon}{2})}{t}\right) \\ &= \mathbb{P}\left(\mathbf{g}(\frac{n(t_i + \frac{\epsilon}{2})}{t}, \lambda) \leq t_i\right) \\ &\to \mathbb{P}\left(t_i + \frac{\epsilon}{2}, \lambda) \leq t_i\right) \quad \mathbb{Q}\text{-almost surely} \\ &= 0.\end{aligned}$$

Thus we have shown that (B.5) holds and can conclude that the first quantity in 4-tuple has the desired result from (B.3). For the second component, that is the overshoot of the process, we introduce the two auxiliary quantities

$$\begin{aligned}k_{\mathbf{g}}^{(n)} &:= \inf\{\mathbf{g}(h,\lambda) | \mathbf{g}(k,\lambda) > \tau_u\}, \\ \sigma_+^{(n)} &:= \mathbf{1}_{\left\{k_{L^\uparrow}^{(n)} < \infty\right\}} \mathbf{g}(k_{L^\uparrow}^{(n)}, \lambda) + \mathbf{1}_{\left\{k_{L^\uparrow}^{(n)} = \infty\right\}} \mathbf{g}(n,\lambda),\end{aligned}$$

where $k_{L^\uparrow}^{(n)}$ is given in (B.2) and the two auxiliary quantities $k_{\mathbf{g}}^{(n)}$ and $\sigma_+^{(n)}$ give the first passage time of the process in terms of the stochastic grid and are equal when the process passes the level u in the given time t however the latter term is equal to t in the event that the first passage time is not hit. (That is $\mathbf{1}_{\left\{k_{L^\uparrow}^{(n)} < \infty\right\}} k_{\mathbf{g}}^{(n)} = \mathbf{1}_{\left\{k_{L^\uparrow}^{(n)} < \infty\right\}} \sigma_+^{(n)}$). In terms of this auxiliary quantity, we wish to show that

$$V^\uparrow(\kappa_u^{(n)} \wedge n, \lambda) - u \overset{d}{=} L^\uparrow_{\sigma_+^{(n)}} - u \overset{\mathbb{P}}{\to} L^\uparrow_{\tau_u \wedge t} - u.$$

We first note that in the event $\left\{k_{L^\uparrow}^{(n)} = \infty\right\}$, we use the SLLN to write

$$\begin{aligned}\mathbf{1}_{\left\{k_{L^\uparrow}^{(n)} = \infty\right\}} L^\uparrow_{\sigma_+^{(n)}} = \mathbf{1}_{\left\{\bar{L}^\uparrow_{\mathbf{g}(n,\lambda) \leq u}\right\}} L^\uparrow_{\mathbf{g}(n,\lambda)} &\to \mathbf{1}_{\left\{\bar{L}_t^\uparrow \leq u\right\}} L_t^\uparrow \quad \mathbb{Q}\text{-almost surely} \\ &= \mathbf{1}_{\{\tau_u \geq t\}} L_t^\uparrow.\end{aligned}$$

For the event $\left\{k_{L^\uparrow}^{(n)} < \infty\right\}$, we first note that we can write

$$\begin{aligned}\mathbf{1}_{\left\{k_{L^\uparrow}^{(n)}<\infty\right\}} = \mathbf{1}_{\left\{\bar{L}^\uparrow_{\mathbf{g}(n,\lambda)>u}\right\}} &\to \mathbf{1}_{\left\{\bar{L}^\uparrow_t>u\right\}} \quad \mathbb{Q}\text{-almost surely} \\ &= \mathbf{1}_{\{\tau_u<t\}}.\end{aligned}$$

Thus to show convergence for the second quantity in the 4-tuple, all that remains is to show that in the limit as $n \to \infty$

$$\mathbf{1}_{\left\{k_{L^\uparrow}^{(n)}<\infty,\tau_u\le t\right\}}\left(L^\uparrow_{\sigma_+^{(n)}} - L^\uparrow_{\tau_u}\right) \xrightarrow{\mathbb{P}} 0.$$

Fixing some $\epsilon > 0$ and letting $\epsilon' > 0$, since $\sigma_+^{(n)} = k_g^{(n)} > \tau_u$ we can obtain the following upper bound on our probability through

$$\begin{aligned}&\mathbb{P}\left(\mathbf{1}_{\left\{k_{L^\uparrow}^{(n)}<\infty,\tau_u\le t\right\}}\left|L^\uparrow_{\sigma_+^{(n)}} - L^\uparrow_{\tau_u}\right| > \epsilon\right) \\ =\mathbb{P}\left(\mathbf{1}_{\left\{k_{L^\uparrow}^{(n)}<\infty,\tau_u\le t\right\}}\left(\mathbf{1}_{\left\{\sigma_+^{(n)}\in[\tau_u,\tau_u+\epsilon']\right\}} + \mathbf{1}_{\left\{\sigma_+^{(n)}>\tau_u+\epsilon'\right\}}\right)\left|L^\uparrow_{\sigma_+^{(n)}} - L^\uparrow_{\tau_u}\right| > \epsilon\right) \\ \le \mathbb{P}\left(\sup_{s\in[0,\epsilon']}\left|L^\uparrow_{\tau_u+s} - L^\uparrow_{\tau_u}\right| > \epsilon\right) + \mathbb{P}\left(\mathbf{1}_{\left\{k_{L^\uparrow}^{(n)}<\infty,\tau_u\le t\right\}}\left(\sigma_+^{(n)} - \tau_u\right) > \epsilon'\right),\end{aligned}$$

by right continuity and path regularity, we can make the first term arbitrary small through the choice of ϵ'. For the second term, we consider the temporal partition $0 = t_0 < t_1 < \cdots < t_N = t$ such that $t_i - t_{i-1} < \epsilon'/2$ for all i and in the limit $n \to \infty$ write

$$\begin{aligned}\mathbb{P}\left(\mathbf{1}_{\left\{k_{L^\uparrow}^{(n)}<\infty,\tau_u\le t\right\}}\left(\sigma_+^{(n)} - \tau_u\right) > \epsilon'\right) &= \sum_{i=1}^N \mathbb{P}\left(\mathbf{1}_{\left\{k_{L^\uparrow}^{(n)}<\infty,\tau_u\in[t_{i-1},t_i)\right\}}\left(\sigma_+^{(n)} - \tau_u\right) > \epsilon'\right) \\ &\le \sum_{i=1}^N \mathbb{P}\left(\inf\{\mathbf{g}(k,\lambda)|\mathbf{g}(k,\lambda) > t_i\} - t_{i-1} > \epsilon'\right) \\ &= \sum_{i=1}^N \mathbb{P}\left(\inf\{\mathbf{g}(k,\lambda)|\mathbf{g}(k,\lambda) > t_i\} - t_i > \frac{\epsilon'}{2}\right) \\ &\to \sum_{i=1}^N \mathbb{P}(t_i - t_i > \frac{\epsilon'}{2}) \quad \mathbb{Q}\text{-almost surely} \\ &= 0.\end{aligned}$$

For the third component of the 4-tuple, we introduce the auxiliary quantity

$$\sigma_-^{(n)} := \mathbf{1}_{\left\{k_{L^\uparrow}^{(n)}<\infty\right\}}\mathbf{g}(k_{L^\uparrow}^{(n)} - 1, \lambda) + \mathbf{1}_{\left\{k_{L^\uparrow}^{(n)}=\infty\right\}}\mathbf{g}(n,\lambda),$$

which corresponds to the last time before first passage in terms of the stochastic grid when the process passes the level u in the given time t and t otherwise. In terms of this auxiliary quantity, we wish to show that

$$u - V^{\uparrow}((\kappa_u^{(n)} - 1) \wedge n, \lambda) \stackrel{d}{=} u - L^{\uparrow}_{\sigma_-^{(n)}} \stackrel{\mathbf{P}}{\to} u - L^{\uparrow}_{\tau_u \wedge t}.$$

Similarly as in the argument for the second quantity in the 4-tuple, it is sufficient to show that in the limit as $n \to \infty$

$$\mathbf{1}_{\left\{k^{(n)}_{L^{\uparrow}} < \infty, \tau_u < t\right\}} \left(L^{\uparrow}_{\sigma_-^{(n)}} - L^{\uparrow}_{\tau_u^-} \right) \stackrel{\mathbf{P}}{\to} 0.$$

By construction, we have that

$$\mathbf{1}_{\left\{k^{(n)}_{L^{\uparrow}} < \infty\right\}} \sigma_-^{(n)} < \mathbf{1}_{\left\{k^{(n)}_{L^{\uparrow}} < \infty\right\}} \tau_u,$$

thus by fixing some $\epsilon > 0$ and letting $\epsilon' > 0$ we can obtain, in a similar manner as before, the following upper bound on our probability as

$$\begin{aligned}
&\mathbb{P}\left(\mathbf{1}_{\left\{k^{(n)}_{L^{\uparrow}} < \infty, \tau_u \leq t\right\}} \left| L^{\uparrow}_{\sigma_-^{(n)}} - L^{\uparrow}_{\tau_u^-} \right| > \epsilon \right) \\
&= \mathbb{P}\left(\mathbf{1}_{\left\{k^{(n)}_{L^{\uparrow}} < \infty, \tau_u \leq t\right\}} \left(\mathbf{1}_{\left\{\sigma_-^{(n)} \in [\tau_u - \epsilon, \tau_u)\right\}} + \mathbf{1}_{\left\{\sigma_-^{(n)} < \tau_u - \epsilon'\right\}} \right) \left| L^{\uparrow}_{\sigma_-^{(n)}} - L^{\uparrow}_{\tau_u^-} \right| > \epsilon \right) \\
&\leq \mathbb{P}\left(\sup_{s \in [0, \epsilon']} \left| L^{\uparrow}_{\tau_u^- - s} - L^{\uparrow}_{\tau_u^-} \right| > \epsilon \right) + \mathbb{P}\left(\mathbf{1}_{\left\{k^{(n)}_{L^{\uparrow}} < \infty, \tau_u \leq t\right\}} \left(\tau_u - \sigma_-^{(n)} \right) > \epsilon' \right),
\end{aligned}$$

by left continuity and path regularity, we can make the first term arbitrary small through the choice of ϵ'. For the second term, we consider the temporal partition $0 = t_0 < t_1 < \cdots < t_N = t$ such that $t_i - t_{i-1} < \epsilon'/2$ for all i and in the limit $n \to \infty$ write

$$\begin{aligned}
\mathbb{P}\left(\mathbf{1}_{\left\{k^{(n)}_{L^{\uparrow}} < \infty, \tau_u \leq t\right\}} \left(\tau_u - \sigma_-^{(n)} \right) > \epsilon' \right) &= \sum_{i=1}^{N} \mathbb{P}\left(\mathbf{1}_{\left\{k^{(n)}_{L^{\uparrow}} < \infty, \tau_u \in [t_{i-1}, t_i)\right\}} \left(\tau_u - \sigma_-^{(n)} \right) > \epsilon' \right) \\
&\leq \sum_{i=1}^{N} \mathbb{P}\left(t_{i-1} - \sup\left\{ \mathbf{g}(k, \lambda) | \mathbf{g}(k, \lambda) < t_i \right\} > \epsilon \right) \\
&= \sum_{i=1}^{N} \mathbb{P}\left(t_i - \sup\left\{ \mathbf{g}(k, \lambda) | \mathbf{g}(k, \lambda) < t_i \right\} > \frac{\epsilon}{2} \right) \\
&\to \sum_{i=1}^{N} \mathbb{P}(t_i - t_i > \frac{\epsilon}{2}) \quad \mathbb{Q}\text{-almost surely} \\
&= 0.
\end{aligned}$$

For the final quantity in the 4-tuple, we wish to show that

$$u - J^{\uparrow}((\kappa_u^{(n)-1}) \wedge n, \lambda) \stackrel{d}{=} u - \overline{L}^{\uparrow}_{\sigma_-^{(n)}} \stackrel{\mathbb{P}}{\to} u - \overline{L}^{\uparrow}_{\tau_u \wedge t}.$$

Similarly as before, by fixing some $\epsilon > 0$ and letting $\epsilon' > 0$ we can obtain the following upper bound on our probability as

$$\begin{aligned}
&\mathbb{P}\left(\mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)} < \infty, \tau_u \leq t\right\}} \left| \overline{L}^{\uparrow}_{\sigma_-^{(n)}} - \overline{L}^{\uparrow}_{\tau_u^-} \right| > \epsilon \right) \\
&= \mathbb{P}\left(\mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)} < \infty, \tau_u \leq t\right\}} \left(\mathbf{1}_{\left\{\sigma_-^{(n)} \in [\tau_u - \epsilon, \tau_u)\right\}} + \mathbf{1}_{\left\{\sigma_-^{(n)} < \tau_u - \epsilon'\right\}}\right) \left| \overline{L}^{\uparrow}_{\sigma_-^{(n)}} - \overline{L}^{\uparrow}_{\tau_u^-} \right| > \epsilon \right) \\
&\leq \mathbb{P}\left(\sup_{s \in [0, \epsilon']} \left| \overline{L}^{\uparrow}_{\tau_u^- - s} - \overline{L}^{\uparrow}_{\tau_u^-} \right| > \epsilon \right) + \mathbb{P}\left(\mathbf{1}_{\left\{k_{L^{\uparrow}}^{(n)} < \infty, \tau_u \leq t\right\}} \left(\tau_u - \sigma_-^{(n)}\right) > \epsilon' \right),
\end{aligned}$$

by left continuity and path regularity, we can make the first term arbitrary small through the choice of ϵ'. The desired convergence then follows by repeating the arguments for the third component.

B.3 Markovian projection

The simplification of a complex process by a simpler process was first introduced in [197] and discussed further in [198] is given by the following result

Theorem B.1 *[Markovian projection.] Let $\boldsymbol{X}(t)$ be a p-dimensional stochastic process given by the strong solution of the SDE*

$$d\boldsymbol{X}(t) = \boldsymbol{\mu}(t)dt + \boldsymbol{\lambda}(t)^T d\boldsymbol{W}(t),$$

where $\boldsymbol{\lambda}(t)$ is a $(d \times p)$-matrix valued adapted process, $\boldsymbol{W}(t)$ is a d-dimensional Brownian motion and

$$\mathbb{E}\left(\int_0^t \left\| \boldsymbol{\lambda}(s)^T \boldsymbol{\lambda}(s) \right\| \right) < \infty$$

for all $t \geq 0$. Then there exists a $(d \times p)$-matrix valued function b and a p-dimensional vector valued function a such that

$$\boldsymbol{b}(t,x)^T \boldsymbol{b}(t,x) = \mathbb{E}\left(\boldsymbol{\lambda}(t)^T \boldsymbol{\lambda}(t) | \boldsymbol{X}(t) = x\right), \qquad \boldsymbol{a}(t,x) = \mathbb{E}\left(\boldsymbol{\mu}(t) | \boldsymbol{X}(t) = x\right)$$

for any $t \geq 0$. The SDE given by

$$d\boldsymbol{Y}(t) = \boldsymbol{a}(t, \boldsymbol{Y}(t))dt + \boldsymbol{b}(t, \boldsymbol{Y}(t))^T d\boldsymbol{W}(t)$$

then admits a weak solution and furthermore the process $\boldsymbol{Y}(t)$ *has the same distribution as the process* $\boldsymbol{X}(t)$ *for* $t \geq 0$.

Thus through Markovian projection, we are able to construct a mimicking process Y that has the same marginal distribution as the original process X.

Bibliography

[1] Y. Kuramoto, *Chemical oscillations, waves, and turbulence*, vol. 19. Springer Science & Business Media, 2012.

[2] P. Protter, D. Talay, *et al.*, "The Euler scheme for Lévy driven stochastic differential equations." *The Annals of Probability*, vol. 25, no. 1, pp. 393–423, 1997.

[3] R. Kawai and H. Masuda, "On simulation of tempered stable random variates." *Journal of Computational and Applied Mathematics*, vol. 235, no. 8, pp. 2873–2887, 2011.

[4] J. M. Chambers, C. L. Mallows, and B. W. Stuck, "A method for simulating stable random variables." *Journal of the American Statistical Association*, vol. 71, no. 354, pp. 340–344, 1976.

[5] A. Kuznetsov, A. E. Kyprianou, J. C. Pardo, K. van Schaik, *et al.*, "A Wiener-Hopf Monte Carlo simulation technique for Lévy processes." *The Annals of Applied Probability*, vol. 21, no. 6, pp. 2171–2190, 2011.

[6] A. B. Holder, M. L. Zuparic, and A. C. Kalloniatis, "Gaussian noise and the two-network frustrated Kuramoto model." *Physica D: Nonlinear Phenomena*, vol. 341, pp. 10–32, 2017.

[7] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms." in *Advances in neural information processing systems*, pp. 2951–2959, 2012.

[8] E. Brochu, V. M. Cora, and N. De Freitas, "A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning." *arXiv preprint arXiv:1012.2599*, 2010.

[9] G. A. Gottwald and I. Melbourne, "A new test for chaos in deterministic systems." in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 460, pp. 603–611, The Royal Society, 2004.

[10] G. A. Gottwald and I. Melbourne, "On the implementation of the 0–1 test for chaos." *SIAM Journal on Applied Dynamical Systems*, vol. 8, no. 1, pp. 129–145, 2009.

[11] A. Kalloniatis, "Concepts for complexity enabled warfare." MCW SRI Symposium 2, 2018.

[12] K. P. O'Keeffe, H. Hong, and S. H. Strogatz, "Oscillators that sync and swarm." *Nature Communications*, vol. 8, no. 1, p. 1504, 2017.

[13] J. G. Taylor, "Lanchester-Type Models of Warfare, Volume II." tech. rep., 1980.

[14] M. Bladt, M. Sørensen, *et al.*, "Simple simulation of diffusion bridges with application to likelihood inference for diffusions." *Bernoulli*, vol. 20, no. 2, pp. 645–675, 2014.

[15] L. Chaumont, "Sur certains processus de Lévy conditionnés à rester positifs." *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 47, no. 1-2, pp. 1–20, 1994.

[16] L. Chaumont, R. Doney, *et al.*, "On Lévy processes conditioned to stay positive.." *Electronic Journal of Probability*, vol. 10, pp. 948–961, 2005.

[17] J. Bertoin, *Lévy processes*, vol. 121. Cambridge University Press, 1998.

[18] A. Kalloniatis, I. Ali, T. Neville, P. La, I. Macleod, M. Zuparic, and E. Kohn, "The situation awareness weighted network (SAWN) model and method: theory and application." *Applied ergonomics*, vol. 61, pp. 178–196, 2017.

[19] P. C. B. Phillips, "Error correction and long-run equilibrium in continuous time." *Econometrica: Journal of the Econometric Society*, pp. 967–980, 1991.

[20] J.-C. Duan and S. R. Pliska, "Option valuation with co-integrated asset prices." *Journal of Economic Dynamics and Control*, vol. 28, no. 4, pp. 727–754, 2004.

[21] Á. Cartea, S. Jaimungal, and J. Penalva, *Algorithmic and high-frequency trading*. Cambridge University Press, 2015.

[22] A. C. Kalloniatis, T. A. McLennan-Smith, D. O. Roberts, and M. L. Zuparic, "Two-network Kuramoto-Sakaguchi model under tempered stable Lévy noise." *Physical Review E*, vol. 99, no. 1, p. 012205, 2019.

[23] A. C. Kalloniatis, T. A. McLennan-Smith, and D. O. Roberts, "Modelling military watch staff using stochastic network synchronisation." *Submitted to the Journal of the Operational Research Society*, 2018.

[24] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical processes on complex networks*. Cambridge university press, 2008.

[25] J. L. Gross, J. Yellen, and P. Zhang, *Handbook of graph theory*. Chapman and Hall/CRC, 2013.

[26] B. Mohar, Y. Alavi, G. Chartrand, and O. R. Oellermann, "The Laplacian spectrum of graphs." *Graph theory, Combinatorics, and Applications*, vol. 2, no. 871-898, p. 12, 1991.

[27] B. Mohar, "Eigenvalues, diameter, and mean distance in graphs." *Graphs and Combinatorics*, vol. 7, no. 1, pp. 53–64, 1991.

[28] P. Eröds and A. Rényi, "On random graphs I." *Publ. Math. Debrecen*, vol. 6, pp. 290–297, 1959.

[29] J. A. Storer, *An introduction to data structures and algorithms*. Springer Science & Business Media, 2012.

[30] P. Erdos and A. Rényi, "On the evolution of random graphs." *Publ. Math. Inst. Hung. Acad. Sci*, vol. 5, no. 1, pp. 17–60, 1960.

[31] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, "Random graphs with arbitrary degree distributions and their applications." *Physical Review E*, vol. 64, no. 2, p. 026118, 2001.

[32] B. Bollobás, "Degree sequences of random graphs." *Discrete Mathematics*, vol. 33, no. 1, pp. 1–19, 1981.

[33] B. Bollobás, *Modern graph theory*, vol. 184. Springer Science & Business Media, 2013.

[34] R. Diestel, *Graph theory*. Springer Publishing Company, Incorporated, 2018.

[35] S. H. Strogatz, "From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators." *Physica D: Nonlinear Phenomena*, vol. 143, no. 1-4, pp. 1–20, 2000.

[36] J. A. Acebrón, L. L. Bonilla, C. J. P. Vicente, F. Ritort, and R. Spigler, "The Kuramoto model: A simple paradigm for synchronization phenomena." *Reviews of modern physics*, vol. 77, no. 1, p. 137, 2005.

[37] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, "Critical phenomena in complex networks." *Reviews of Modern Physics*, vol. 80, no. 4, p. 1275, 2008.

[38] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, "Synchronization in complex networks." *Physics reports*, vol. 469, no. 3, pp. 93–153, 2008.

[39] F. Dörfler and F. Bullo, "Synchronization in complex networks of phase oscillators: A survey." *Automatica*, vol. 50, no. 6, pp. 1539–1564, 2014.

[40] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, and M. Zanin, "The structure and dynamics of multilayer networks." *Physics Reports*, vol. 544, no. 1, pp. 1–122, 2014.

[41] E. Montbrió, J. Kurths, and B. Blasius, "Synchronization of two interacting populations of oscillators." *Physical Review E*, vol. 70, no. 5, p. 056125, 2004.

[42] E. Barreto, B. Hunt, E. Ott, and P. So, "Synchronization in networks of networks: The onset of coherent collective behavior in systems of interacting populations of heterogeneous oscillators." *Physical Review E*, vol. 77, no. 3, p. 036107, 2008.

[43] Y. Kawamura, H. Nakao, K. Arai, H. Kori, and Y. Kuramoto, "Phase synchronization between collective rhythms of globally coupled oscillator groups: Noiseless nonidentical case." *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 20, no. 4, p. 043110, 2010.

[44] P. S. Skardal and J. G. Restrepo, "Synchronization of Kuramoto oscillators in networks of networks." *arXiv preprint arXiv:1206.3822*, 2012.

[45] A. C. Kalloniatis and M. L. Zuparic, "Fixed points and stability in the two-network frustrated Kuramoto model." *Physica A: Statistical Mechanics and its Applications*, vol. 447, pp. 21–35, 2016.

[46] B. C. Bag, K. Petrosyan, and C.-K. Hu, "Influence of noise on the synchronization of the stochastic Kuramoto model." *Physical Review E*, vol. 76, no. 5, p. 056210, 2007.

[47] H. Khoshbakht, F. Shahbazi, and K. A. Samani, "Phase synchronization on scale-free and random networks in the presence of noise." *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10020, 2008.

[48] R. Tönjes, "Synchronization transition in the Kuramoto model with colored noise." *Physical Review E*, vol. 81, no. 5, p. 055201, 2010.

[49] R. K. Esfahani, F. Shahbazi, and K. A. Samani, "Noise-induced synchronization in small world networks of phase oscillators." *Physical Review E*, vol. 86, no. 3, p. 036204, 2012.

[50] L. DeVille, "Transitions amongst synchronous solutions in the stochastic Kuramoto model." *Nonlinearity*, vol. 25, no. 5, p. 1473, 2012.

[51] D. Traxl, N. Boers, and J. Kurths, "General scaling of maximum degree of synchronization in noisy complex networks." *New Journal of Physics*, vol. 16, no. 11, p. 115009, 2014.

[52] A. C. Kalloniatis and D. O. Roberts, "Synchronisation of networked Kuramoto oscillators under stable Lévy noise." *Physica A: Statistical Mechanics and its Applications*, vol. 466, pp. 476–491, 2017.

[53] D. Roberts and A. C. Kalloniatis, "Synchronisation under shocks: The Lévy Kuramoto model." *Physica D: Nonlinear Phenomena*, 2017.

[54] R. Metzler, "R. Metzler and J. Klafter, Phys. Rep. 339, 1 (2000).." *Phys. Rep.*, vol. 339, p. 1, 2000.

[55] R. Metzler and J. Klafter, "The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics." *Journal of Physics A: Mathematical and General*, vol. 37, no. 31, p. R161, 2004.

[56] Y. Kuramoto, "Self-entrainment of a population of coupled non-linear oscillators." in *International symposium on mathematical problems in theoretical physics*, pp. 420–422, Springer, 1975.

[57] D. Applebaum, *Lévy processes and stochastic calculus*. Cambridge University Press, 2009.

[58] I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*, vol. 113. Springer Science & Business Media, 2012.

[59] B. Øksendal, "Stochastic differential equations." in *Stochastic differential equations*, pp. 65–84, Springer, 2003.

[60] S. M. Iacus, *Simulation and inference for stochastic differential equations: with R examples*. Springer Science & Business Media, 2009.

[61] D. Talay, *Numerical solution of stochastic differential equations*. Taylor & Francis, 1994.

[62] K.-i. Sato, *Lévy processes and infinitely divisible distributions*. Cambridge University Press, 1999.

[63] P. Tankov, *Financial modelling with jump processes*, vol. 2. CRC press, 2004.

[64] A. E. Kyprianou, *Introductory lectures on fluctuations of Lévy processes with applications*. Springer Science & Business Media, 2006.

[65] C. Klüppelberg, A. E. Kyprianou, R. A. Maller, *et al.*, "Ruin probabilities and overshoots for general Lévy insurance risk processes." *The Annals of Applied Probability*, vol. 14, no. 4, pp. 1766–1801, 2004.

[66] B. Mandelbrot, "The Pareto-Levy law and the distribution of income." *International Economic Review*, vol. 1, no. 2, pp. 79–106, 1960.

[67] J. Nolan, *Stable distributions: models for heavy-tailed data*. Birkhauser New York, 2003.

[68] J. Rosiński, "Tempering stable processes." *Stochastic processes and their applications*, vol. 117, no. 6, pp. 677–707, 2007.

[69] U. Küchler and S. Tappe, "Tempered stable distributions and processes." *Stochastic Processes and their Applications*, vol. 123, no. 12, pp. 4256–4293, 2013.

[70] A. Kuznetsov *et al.*, "Wiener–Hopf factorization and distribution of extrema for a family of Lévy processes." *The Annals of Applied Probability*, vol. 20, no. 5, pp. 1801–1830, 2010.

[71] A. Kuznetsov, A. E. Kyprianou, J. C. Pardo, *et al.*, "Meromorphic Lévy processes and their fluctuation identities." *The Annals of Applied Probability*, vol. 22, no. 3, pp. 1101–1135, 2012.

[72] A. Ferreiro-Castilla and W. Schoutens, "The β-Meixner model." *Journal of Computational and Applied Mathematics*, vol. 236, no. 9, pp. 2466–2476, 2012.

[73] S. Asmussen and J. Rosiński, "Approximations of small jumps of Lévy processes with a view towards simulation." *Journal of Applied Probability*, vol. 38, no. 2, pp. 482–493, 2001.

[74] S. Cohen and J. Rosiński, "Gaussian approximation of multivariate Lévy processes with applications to simulation of tempered stable processes." *Bernoulli*, pp. 195–210, 2007.

[75] J. Gil-Pelaez, "Note on the inversion theorem." *Biometrika*, vol. 38, no. 3-4, pp. 481–482, 1951.

[76] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series." *Mathematics of computation*, vol. 19, no. 90, pp. 297–301, 1965.

[77] V. M. Zolotarev, *One-dimensional stable distributions*, vol. 65. American Mathematical Soc., 1986.

[78] M. C. K. Tweedie, "An index which distinguishes between some important exponential families." in *Statistics: Applications and new directions: Proc. Indian statistical institute golden Jubilee International conference*, vol. 579, p. 6o4, 1984.

[79] G. Casella, C. P. Robert, and M. T. Wells, "Generalized accept-reject sampling schemes." *Lecture Notes-Monograph Series*, pp. 342–347, 2004.

[80] C. P. Robert, *Monte carlo methods*. Wiley Online Library, 2004.

[81] B. Baeumer and M. M. Meerschaert, "Tempered stable Lévy motion and transient super-diffusion." *Journal of Computational and Applied Mathematics*, vol. 233, no. 10, pp. 2438–2448, 2010.

[82] L. Devroye, "Random variate generation for exponentially and polynomially tilted stable distributions." *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 19, no. 4, p. 18, 2009.

[83] A. Ferreiro-Castilla, A. E. Kyprianou, and R. Scheichl, "An Euler–Poisson scheme for Lévy driven stochastic differential equations." *Journal of Applied Probability*, vol. 53, no. 1, pp. 262–278, 2016.

[84] P. Carr, "Randomization and the American put." *The Review of Financial Studies*, vol. 11, no. 3, pp. 597–626, 1998.

[85] M. E. Caballero and L. Chaumont, "Conditioned stable Lévy processes and the Lamperti representation." *Journal of Applied Probability*, vol. 43, no. 4, pp. 967–983, 2006.

[86] H. Sakaguchi and Y. Kuramoto, "A soluble active rotater model showing phase transitions via mutual entertainment." *Progress of Theoretical Physics*, vol. 76, no. 3, pp. 576–581, 1986.

[87] A. C. C. Coolen and C. Pérez-Vicente, "Partially and fully frustrated coupled oscillators with random pinning fields." *Journal of Physics A: Mathematical and General*, vol. 36, no. 16, p. 4477, 2003.

[88] V. Nicosia, M. Valencia, M. Chavez, A. Díaz-Guilera, and V. Latora, "Remote synchronization reveals network symmetries and functional modules." *Physical review letters*, vol. 110, no. 17, p. 174102, 2013.

[89] S. Kirkland and S. Severini, "α-Kuramoto partitions from the frustrated Kuramoto model generalise equitable partitions." *Applicable Analysis and Discrete Mathematics*, pp. 29–38, 2015.

[90] A. C. Kalloniatis, "From incoherence to synchronicity in the network Kuramoto model." *Physical Review E*, vol. 82, no. 6, p. 066202, 2010.

[91] R. N. Mantegna and H. E. Stanley, "Scaling behaviour in the dynamics of an economic index." *Nature*, vol. 376, no. 6535, p. 46, 1995.

[92] I. Koponen, "Analytic approach to the problem of convergence of truncated Lévy flights towards the Gaussian stochastic process." *Physical Review E*, vol. 52, no. 1, p. 1197, 1995.

[93] A. Kullberg and D. del Castillo-Negrete, "Transport in the spatially tempered, fractional Fokker–Planck equation." *Journal of Physics A: Mathematical and Theoretical*, vol. 45, no. 25, p. 255101, 2012.

[94] D. del Castillo-Negrete, "Anomalous transport in the presence of truncated Lévy flights." in *Fractional Dynamics: Recent Advances*, pp. 129–157, World Scientific, 2012.

[95] I. Podlubny, *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*, vol. 198. Elsevier, 1998.

[96] S. G. Samko, A. A. Kilbas, O. I. Marichev, *et al.*, "Fractional integrals and derivatives." *Theory and Applications, Gordon and Breach, Yverdon*, vol. 1993, p. 44, 1993.

[97] L. M. Pecora and T. L. Carroll, "Master stability functions for synchronized coupled systems." *Physical review letters*, vol. 80, no. 10, p. 2109, 1998.

[98] B. Lindner, M. Kostur, and L. Schimansky-Geier, "Optimal diffusive transport in a tilted periodic potential." *Fluctuation and Noise Letters*, vol. 1, no. 01, pp. R25–R39, 2001.

[99] A. K. Lee, J.-R. Lee, and K. Lee, "Asymmetric step-like characteristics in a tilted rocking ratchet potential." *Physica B: Condensed Matter*, vol. 407, no. 21, pp. 4298–4302, 2012.

[100] K. J. Challis and M. W. Jack, "Tight-binding approach to overdamped Brownian motion on a multidimensional tilted periodic potential." *Physical Review E*, vol. 87, no. 5, p. 052102, 2013.

[101] P. Reimann, "Brownian motors: noisy transport far from equilibrium." *Physics reports*, vol. 361, no. 2-4, pp. 57–265, 2002.

[102] M. L. Zuparic and A. C. Kalloniatis, "Analytic solution to space-fractional Fokker–Planck equations for tempered-stable Lévy distributions with spatially linear, time-dependent drift." *Journal of Physics A: Mathematical and Theoretical*, vol. 51, no. 3, p. 035101, 2017.

[103] Á. Cartea and D. del Castillo-Negrete, "Fluid limit of the continuous-time random walk with general Lévy jump distribution functions." *Physical Review E*, vol. 76, no. 4, p. 041105, 2007.

[104] A. D. Polyanin and V. E. Nazaikinskii, *Handbook of linear partial differential equations for engineers and scientists*. CRC press, 2015.

[105] M. L. Zuparic, A. C. Kalloniatis, and D. O. Roberts, "Noise-driven current reversal and stabilization in the tilted ratchet potential subject to tempered stable Lévy noise." *Physical Review E*, vol. 96, no. 5, p. 052116, 2017.

[106] B. D. McKay *et al.*, "Practical graph isomorphism." 1981.

[107] C. E. Rasmussen, "Gaussian processes in machine learning." in *Advanced lectures on machine learning*, pp. 63–71, Springer, 2004.

[108] W. P. Hughes Jr, "A salvo model of warships in missile combat used to evaluate their staying power." *Naval Research Logistics (NRL)*, vol. 42, no. 2, pp. 267–289, 1995.

[109] C. McLemore, D. Gaver, and P. Jacobs, "A model for geographically distributed combat interactions of swarming naval and air forces." *Naval Research Logistics (NRL)*, vol. 63, no. 7, pp. 562–576, 2016.

[110] V. Protopopescu, R. Santoro, and J. Dockery, "Combat modeling with partial differential equations." *European Journal of Operational Research*, vol. 38, no. 2, pp. 178–183, 1989.

[111] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles." *Physical Review Letters*, vol. 75, no. 6, p. 1226, 1995.

[112] K. Itô, P. Henry Jr, *et al.*, *Diffusion processes and their sample paths*. Springer Science & Business Media, 2012.

[113] R. Doney, "Fluctuation theory for Lévy processes." in *Lévy processes*, pp. 57–66, Springer, 2001.

[114] A. Ferreiro-Castilla and K. Van Schaik, "Applying the Wiener-Hopf Monte Carlo simulation technique for Lévy processes to path functionals." *Journal of Applied Probability*, vol. 52, no. 1, pp. 129–148, 2015.

[115] S. Asmussen and H. Albrecher, *Ruin probabilities*. World Scientific Publishing Co Pte Ltd, 2010.

[116] R. A. Doney and A. E. Kyprianou, "Overshoots and undershoots of Lévy processes." *The Annals of Applied Probability*, pp. 91–106, 2006.

[117] E. Hoyle, L. P. Hughston, and A. Macrina, "Lévy random bridges and the modelling of financial information." *Stochastic Processes and their Applications*, vol. 121, no. 4, pp. 856–884, 2011.

[118] F. Caravenna, L. Chaumont, *et al.*, "An invariance principle for random walk bridges conditioned to stay positive." *Electronic Journal of Probability*, vol. 18, 2013.

[119] J. E. Figueroa-López, P. Tankov, *et al.*, "Small-time asymptotics of stopped Lévy bridges and simulation schemes with controlled bias." *Bernoulli*, vol. 20, no. 3, pp. 1126–1164, 2014.

[120] K.-K. Kim and S. Kim, "Simulation of tempered stable Lévy bridges and its applications." *Operations Research*, vol. 64, no. 2, pp. 495–509, 2016.

[121] H. E. Daniels, "Saddlepoint approximations in statistics." *The Annals of Mathematical Statistics*, pp. 631–650, 1954.

[122] O. Barndorff-Nielsen and D. R. Cox, "Edgeworth and saddle-point approximations with statistical applications." *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 279–312, 1979.

[123] H. Daniels, "Discussion of" The regression analysis of binary sequences" by DR Cox." *J. Roy. Statist. Soc. Ser. B*, vol. 20, pp. 236–238, 1958.

[124] C. Ribeiro and N. Webber, "A Monte Carlo method for the normal inverse Gaussian option valuation model using an inverse Gaussian bridge." *Preprint, City University*, 2003.

[125] C. Ribeiro and N. Webber, "Valuing path-dependent options in the variance-gamma model by Monte Carlo with a gamma bridge." *Journal of Computational Finance*, vol. 7, no. 2, pp. 81–100, 2004.

[126] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks." *Reviews of Modern Physics*, vol. 87, no. 3, p. 925, 2015.

[127] J. Hindes and I. B. Schwartz, "Epidemic extinction paths in complex networks." *Physical Review E*, vol. 95, no. 5, p. 052317, 2017.

[128] P. Holme and L. Tupikina, "Epidemic extinction in networks: Insights from the 12,110 smallest graphs." *arXiv preprint arXiv:1802.08849*, 2018.

[129] F. M. Atay, T. Biyikoglu, and J. Jost, "Synchronization of networks with prescribed degree distributions." *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 1, pp. 92–98, 2006.

[130] F. Harary, "The maximum connectivity of a graph." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 48, no. 7, p. 1142, 1962.

[131] R. S. Aguilar-Saven, "Business process modelling: Review and framework." *International Journal of production economics*, vol. 90, no. 2, pp. 129–149, 2004.

[132] M. J. Ashworth and K. M. Carley, "Can tools help unify organization theory? Perspectives on the state of computational modeling." *Computational and Mathematical Organization Theory*, vol. 13, no. 1, pp. 89–111, 2007.

[133] D. Schreiber, "Validating agent-based models: From metaphysics to applications." in *Annual Conference of the Midwestern Political Science Association, Chicago, IL: April*, 2002.

[134] G. K. Bharathy and B. Silverman, "Validating agent based social systems models." in *Simulation Conference (WSC), Proceedings of the 2010 Winter*, pp. 441–453, IEEE, 2010.

[135] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks." *Reviews of Modern Physics*, vol. 74, no. 1, p. 47, 2002.

[136] T. Ichinomiya, "Frequency synchronization in a random oscillator network." *Physical Review E*, vol. 70, no. 2, p. 026116, 2004.

[137] E. Oh, D.-S. Lee, B. Kahng, and D. Kim, "Synchronization transition of heterogeneously coupled oscillators on scale-free networks." *Physical Review E*, vol. 75, no. 1, p. 011104, 2007.

[138] J. Gómez-Gardeñes, Y. Moreno, and A. Arenas, "Synchronizability determined by coupling strengths and topology on complex networks." *Physical Review E*, vol. 75, no. 6, p. 066106, 2007.

[139] A. Dekker, "Studying organisational topology with simple computational models." *Journal of Artificial Societies and Social Simulation*, vol. 10, no. 4, p. 6, 2007.

[140] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world'networks." *nature*, vol. 393, no. 6684, p. 440, 1998.

[141] H. Hong, M.-Y. Choi, and B. J. Kim, "Synchronization on small-world networks." *Physical Review E*, vol. 65, no. 2, p. 026139, 2002.

[142] K. A. Penson and K. Górska, "Exact and explicit probability densities for one-sided Lévy stable distributions." *Physical review letters*, vol. 105, no. 21, p. 210604, 2010.

[143] D. Roberts and A. C. Kalloniatis, "Synchronisation under shocks: The Lévy Kuramoto model." *Physica D: Nonlinear Phenomena*, vol. 368, pp. 10–21, 2018.

[144] A. Kalloniatis and D. E. Fairbairn, "A new paradigm for dynamical modelling of networked C2 processes." *13TH ICCRTS*, 2008.

[145] U. Neisser, "Cognition and reality. Principles and implication of cognitive psychology." *San Francisko: WH Freeman and Company*, 1976.

[146] C. A. Miller, "The Risks of Discretization: What Is Lost in (Even Good) Levels-of-Automation Schemes." *Journal of Cognitive Engineering and Decision Making*, vol. 12, no. 1, pp. 74–76, 2018.

[147] M. R. Endsley and D. Garland, "Theoretical underpinnings of situation awareness: A critical review." *Situation awareness analysis and measurement*, vol. 1, p. 24, 2000.

[148] C. Wickens, "Automation stages & levels, 20 years after." *Journal of Cognitive Engineering and Decision Making*, vol. 12, no. 1, pp. 35–41, 2018.

[149] J. R. Boyd, "Organic design for command and control." *A discourse on winning and losing*, 1987.

[150] T. Moon, E. Kruzins, and G. Calbert, "Analyzing the OODA cycle." *Phalanx*, vol. 35, no. 2, pp. 9–35, 2002.

[151] J. Hasík, "Beyond the Briefing: Theoretical and Practical Problems in the Works and Legacy of John Boyd." *Contemporary Security Policy*, vol. 34, no. 3, pp. 583–599, 2013.

[152] F. Osinga, "Getting'A Discourse on Winning and Losing: A Primer on Boyd's 'Theory of Intellectual Evolution." *Contemporary Security Policy*, vol. 34, no. 3, pp. 603–624, 2013.

[153] K. E. Weick, "Educational organizations as loosely coupled systems." *Administrative science quarterly*, pp. 1–19, 1976.

[154] R. I. Beekun and W. H. Glick, "Development and test of a contingency framework of coupling: Assessing the covariation between structure and culture." *The Journal of Applied Behavioral Science*, vol. 37, no. 4, pp. 385–407, 2001.

[155] C. Perrow, *Normal Accidents: Living with High Risk Technologies-Updated Edition*. Princeton University Press, 2011.

[156] R. Ratcliff and G. McKoon, "The diffusion decision model: theory and data for two-choice decision tasks." *Neural computation*, vol. 20, no. 4, pp. 873–922, 2008.

[157] H. W. J. Rittel and M. M. Webber, "Dilemmas in a general theory of planning." *Policy sciences*, vol. 4, no. 2, pp. 155–169, 1973.

[158] G. A. Klein, *Sources of power: How people make decisions*. MIT press, 2017.

[159] C. K. Morewedge and D. Kahneman, "Associative processes in intuitive judgment." *Trends in Cognitive Sciences*, vol. 14, no. 10, pp. 435–440, 2010.

[160] M. M. Meerschaert, P. Roy, and Q. Shao, "Parameter estimation for exponentially tempered power law distributions." *Communications in Statistics-Theory and Methods*, vol. 41, no. 10, pp. 1839–1856, 2012.

[161] M. Grabchak, "Does value-at-risk encourage diversification when losses follow tempered stable or more general Lévy processes?." *Annals of Finance*, vol. 10, no. 4, pp. 553–568, 2014.

[162] P. S. Griffin, R. A. Maller, and D. Roberts, "Finite time ruin probabilities for tempered stable insurance risk processes." *Insurance: Mathematics and Economics*, vol. 53, no. 2, pp. 478–489, 2013.

[163] D. Hainaut and P. Devolder, "Mortality modelling with Lévy processes." *Insurance: Mathematics and Economics*, vol. 42, no. 1, pp. 409–418, 2008.

[164] K. J. Palmer, M. S. Ridout, and B. J. T. Morgan, "Modelling cell generation times by using the tempered stable distribution." *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 57, no. 4, pp. 379–397, 2008.

[165] A. Wyłomańska, G. Żak, P. Kruczek, and R. Zimroz, "Application of tempered stable distribution for selection of optimal frequency band in gearbox local damage detection." *Applied Acoustics*, vol. 128, pp. 14–22, 2017.

[166] A. C. Kalloniatis, "21 st International Command and Control Research and Technology Symposium 2016 C2 in a Complex Connected Battlespace."

[167] F. Manual, "101-5 Staff Organization and Operations." *Washington, DC: Headquarters, Department of the Army*, p. G1, 1997.

[168] G. C. Kamena, "The Dying Art of Battle-Rhythm." *Combat Training Center Quarterly Bulletin*, pp. 99–3, 1999.

[169] S. Wasserman and K. Faust, *Social network analysis: Methods and applications*, vol. 8. Cambridge University Press, 1994.

[170] R. G. Sargent, "A tutorial on verification and validation of simulation models." in *Proceedings of the 16th conference on Winter simulation*, pp. 114–121, IEEE Press, 1984.

[171] E. J. Rykiel Jr, "Testing ecological models: the meaning of validation." *Ecological Modelling*, vol. 90, no. 3, pp. 229–244, 1996.

[172] C. R. Berger and R. J. Calabrese, "Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication." *Human communication research*, vol. 1, no. 2, pp. 99–112, 1974.

[173] W. A. Afifi and J. L. Weiner, "Introduction. Information Seeking Across Contexts." *Human Communication Research*, vol. 28, no. 2, pp. 207–212, 2002.

[174] L. Donaldson, *The contingency theory of organizations*. Sage, 2001.

[175] T. Bruns and G. M. Stalker, "The management of innovation." *Tavistock, London*, 1961.

[176] R. N. Mantegna, "Hierarchical structure in financial markets." *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 11, no. 1, pp. 193–197, 1999.

[177] C. H. Papadimitriou and K. Steiglitz, *Combinatorial optimization: algorithms and complexity*. Courier Corporation, 1998.

[178] T. K. D. Peron and F. A. Rodrigues, "Collective behavior in financial markets." *EPL (Europhysics Letters)*, vol. 96, no. 4, p. 48004, 2011.

[179] I. J. D. Craig and A. D. Sneyd, "An alternating-direction implicit scheme for parabolic equations with mixed derivatives." *Computers & Mathematics with Applications*, vol. 16, no. 4, pp. 341–350, 1988.

[180] L. B. G. Andersen and V. V. Piterbarg, *Interest Rate Modeling*. Atlantic Financial Press, 2010.

[181] K. R. Vetzal, "An improved finite difference approach to fitting the initial term structure." *Journal of Fixed Income*, vol. 7, pp. 62–82, 1998.

[182] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions." *Journal of Global optimization*, vol. 13, no. 4, pp. 455–492, 1998.

[183] M. J. Sasena, *Flexibility and efficiency enhancements for constrained global design optimization with kriging approximations*. PhD thesis, University of Michigan Ann Arbor, 2002.

[184] B. Matérn, *Spatial variation*, vol. 36. Springer Science & Business Media, 2013.

[185] W. H. Press, *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge University Press, 2007.

[186] H. J. Kushner, "A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise." *Journal of Basic Engineering*, vol. 86, no. 1, pp. 97–106, 1964.

[187] J. Mockus, *Bayesian approach to global optimization: theory and applications*, vol. 37. Springer Science & Business Media, 2012.

[188] N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger, "Gaussian process optimization in the bandit setting: No regret and experimental design." *arXiv preprint arXiv:0912.3995*, 2009.

[189] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, "Determining Lyapunov exponents from a time series." *Physica D: Nonlinear Phenomena*, vol. 16, no. 3, pp. 285–317, 1985.

[190] M. Nicol, I. Melbourne, and P. Ashwin, "Euclidean extensions of dynamical systems." *Nonlinearity*, vol. 14, no. 2, p. 275, 2001.

[191] M. Field, I. Melbourne, and A. Török, "Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions." *Ergodic Theory and Dynamical Systems*, vol. 23, no. 1, p. 87, 2003.

[192] V. N. Biktashev and A. V. Holden, "Deterministic Brownian motion in the hypermeander of spiral waves." *Physica D: Nonlinear Phenomena*, vol. 116, no. 3-4, pp. 342–354, 1998.

[193] G. A. Gottwald and I. Melbourne, "On the validity of the 0–1 test for chaos." *Nonlinearity*, vol. 22, no. 6, p. 1367, 2009.

[194] H. Pham, *Continuous-time stochastic control and optimization with financial applications*, vol. 61. Springer Science & Business Media, 2009.

[195] A. Ferreiro-Castilla, A. E. Kyprianou, R. Scheichl, and G. Suryanarayana, "Multilevel Monte Carlo simulation for Lévy processes based on the Wiener–Hopf factorisation." *Stochastic Processes and their Applications*, vol. 124, no. 2, pp. 985–1010, 2014.

[196] P. Carr, H. Geman, D. B. Madan, and M. Yor, "The fine structure of asset returns: An empirical investigation." *The Journal of Business*, vol. 75, no. 2, pp. 305–332, 2002.

[197] I. Gyöngy, "Mimicking the one-dimensional marginal distributions of processes having an Itô differential." *Probability theory and related fields*, vol. 71, no. 4, pp. 501–516, 1986.

[198] G. Brunick, S. Shreve, *et al.*, "Mimicking an Itô process by a solution of a stochastic differential equation." *The Annals of Applied Probability*, vol. 23, no. 4, pp. 1584–1628, 2013.