sedimentary and hard rock regions in the shallow crust. Fig. 14 shows a simplified broad-scale geological map of the Australian continent with terranes marked according to age and dominant exposure type. The correlation between higher velocities and the presence of exposed crystalline basement, and lower velocities and the presence of sedimentary basins of substantial thickness, is quite distinct. The Archean terranes of central and western Australia, comprising the Gawler, Yilgarn and Pilbara cratons (Fig. 14), are clearly revealed as regions of elevated velocity, although the full extent of the Pilbara craton is somewhat masked by the presence of overlying sediments. Although sampling much deeper in the lithosphere, the shear wave speed images obtained from regional surface wave tomography (Fishwick et al. 2005) of the Australian continent show a very similar pattern of high velocities beneath the Archean cratons.

Although less distinct, there is also a relationship between elevated velocities and the presence of Proterozoic and Palaeozoic basement. For example, the Kimberley Craton and Lachlan Orogen appear to correspond to regions of higher velocity; in the latter case, this region includes much of the southern Great Dividing Range, which contains extensive exposure of igneous rocks and metamorphosed sediments (Foster & Gray 2000). Sedimentary basins, which cover vast tracts of the Australian continent, are generally responsible for most of the low velocity features present in Fig. 13. For example, the Great Artesian Basin in central eastern Australia, which encompasses the Eromanga, Surat and Carpentaria basins (Fig. 14) contains large regions in which the sediment thickness exceeds 2 km (Laske & Masters 1997). The presence of such thick sediments would be responsible for the distinct low velocity zone observed in central and central eastern Australia (Fig. 13).

The Canning Basin in northwest Australia also hosts extremely thick sedimentary sequences—up to 6 km in some places (Laske & Masters 1997); this would explain the presence of the low velocity zone dividing the Kimberley and Pilbara Cratons.

In southeast Australia, the vast intracratonic Murray Basin is not very clearly defined in the velocity images (Fig. 13), but this is probably due to the sediment layer being relatively thin (Knight et al. 1995). However, the small scale length low velocity lineations that can be observed, due to the multiscale nature of the tomographic technique used to recover structure, probably represent the presence of pre-Tertiary infra-basins that underlie the Murray Basin (Arroucau et al. 2010). One remarkable feature of the tomography results is that all three Cainozoic basins in Bass Strait—the Bass, Otway and Gippsland basins (Fig. 14) appear as distinct low velocity zones, even though they are largely resolved by the continent-wide data set alone. In Tasmania, the divide between the Proterozoic West Tasmania Terrane and Phanerozoic East Tasmania Terrane is represented by a lower velocity transition zone, which is consistent with the recent results of (Young et al. 2011).

The posterior information on hyperparameters \((n, \lambda, a, b)\) is shown in Fig. 15. Note that the collected velocity models in the ensemble solution have an average of 1200 cells. Each cell is defined by three parameters (2-D location of nodes + velocity) which makes the dimension of the model space around 3600. The Monte Carlo integration is feasible because it was implemented on parallel computing architecture. To give the reader an idea of the computational cost of such an inversion, each ‘outer-loop’ iteration of the algorithm (as shown in Fig. 5) requires approximately five days,
so each panel of Fig. 13 represents about 15 days of computation time. The inferred information on the level of noise in travel times indicates that the uncertainties provided for the WOMBAT arrays have been rescaled to around $\lambda = 0.28$. The posterior value on hyperparameter $a$ indicates that the data noise for the large scale is expected to increase 0.85 s each time the interstation distance increases by 1 degree with an expected data error of 0.6 s at 0° degrees (see lower panels in Fig. 15).

Fig. 16(a) shows the error map for the $L_1$ solution obtained with transdimensional tomography. This is constructed by taking the standard deviation of the ensemble of sampled Voronoi models at each point of the velocity field. This locally shows how well the solution model in 13(b) is constrained. As expected, well sampled areas in Western Australia, South East Australia and Tasmania show a lower velocity uncertainty.

Because the underlying parametrization is mobile, it is also interesting to look at the spatial density of Voronoi nodes across the ensemble of models collected. To do this we discretized the region into cells of $0.5 \times 0.5^\circ$, and calculated the average number of Voronoi nodes per cell over the posterior ensemble (see Fig. 16b). This map displays the average size of Voronoi cells at each point of the velocity field. A large number of small Voronoi cells are concentrated within WOMBAT arrays with larger cells elsewhere, thereby demonstrating the adaptive character of the transdimensional parametrization.

It is interesting to see that the estimated error on the model is not necessarily correlated with the density of cells. The Archaean cratons in Western Australia are well constrained without need of small cells (This area shows low values for model uncertainty in 16(a) with the lowest density of cells in 16b). There is good ray coverage in Western Australia, and one might expect the algorithm to introduce a lot of small cells to provide a high level of detail. However, the velocity field seems to be quite homogeneous and there is no need to introduce high levels of complexity in this region. This example shows the parsimonious nature of the algorithm and indicates that the transdimensional parametrization not only adapts to the density of rays but also to the character of the velocity structure itself.

### 6.3 Comparison with the Subspace Inversion

Finally, we compare our results with maps obtained with a standard fixed grid optimization approach. As described in Section 5.5, the three data sets are simultaneously inverted with the subspace method (Kennett et al. 1988; Rawlinson et al. 2006, 2008). Here the inverse problem is regularized using ‘ad-hoc’ damping and smoothing, and the level of data noise is not accounted for. Fig. 17 shows two solutions obtained with the two grid sizes showed in Fig. 9. For each grid size, regularization parameters were successively ‘tuned’ with an iterative L-curve method as in Rawlinson et al. (2006).

As with the synthetic experiments, the coarse grid solution in Fig. 17(a) misses the details and small scale features under WOMBAT arrays that have been recovered with transdimensional tomography. Although the node spacing seems appropriate for the large scale data set, it is too coarse and cannot map information present under Southeast Australia and Tasmania. By using a finer grid, in Fig. 17(b) the velocity field in Southeast Australia presents more details but the structure under Tasmania is still missing. In this case, small artefacts resulting from data noise are introduced such as under Yilgarn craton, and overall the amplitudes are worse due to damping.

By visual inspection of the two images on Fig. 17, one can observe a striking feature of fixed grid inversion approaches: the scale of recovered velocity heterogeneities is spatially constant over the velocity field. Here the various scales of structural heterogeneities in the Earth, as well as inhomogeneities in data coverage, are not accounted for.

Although results are generally similar to transdimensional tomography for both choices of grid size, information appears to be lost in the well sampled areas compared to the transdimensional solution. Furthermore, here there is no way to measure which of these two solutions best describes the underlying seismic structure, which can result in misinterpretation. These are best fitting models and there is no information available on the level of uncertainty on the recovered velocity field.

### 7 Conclusion and Future Directions

We have shown here that transdimensional tomography is particularly suited for inversion of multiple data sets that sample the Earth at different scales. Synthetic and real data examples have illustrated the adaptive character of the parametrization which enables us to image small scale features in well sampled areas without introducing spurious artefacts elsewhere. The level of smoothing is spatially

---

**Figure 15.** Posterior probability distribution on hyperparameters for the $L_1$ misfit solution shown in Fig. 13(b). The three inferred noise parameters $\lambda$, $a$ and $b$ define the level of data noise in data sets as given by (7) and (8).
variable and is naturally determined by the data. Contrary to other multiscale tomography methods, recovered structure is not only regulated by the density of rays, but also by the inferred data noise and by the structure of the underlying velocity field.

As the complexity of the model is variable, the estimated level of data noise takes an important role in the inversion as it directly determines the number of model parameters needed to fit the data to the required level. We have shown that an extended Bayesian

Figure 16. Top panel: error map (km s$^{-1}$) associated with the L1 norm solution in 13(b). Bottom panel: density of Voronoi nuclei across the ensemble of sampled models. The colour scale represents the expected number of nuclei per pixel.
formulation called Hierarchical Bayes can take into account the lack of knowledge on data uncertainty. When assessment of measurement errors is difficult to achieve \textit{a priori} (as in ambient noise tomography), this procedure treats the standard deviation of data noise as an unknown and makes a joint posterior inference on both model complexity and data uncertainty. The Hierarchical Bayes procedure turns out to be particularly useful when dealing with multiple data types having different unknown levels of noise. With scant prior knowledge on data errors, the algorithm is able to infer the level of information brought by each data type and to naturally adjust the fit to different data sets.

In our ambient noise tomography application, the data noise from WOMBAT arrays was naturally rescaled while the noise for the large scale data set was parametrized as a linear function of the interstation...
distance. The inversion resulted in a parsimonious velocity map with a spatial resolution adapted to the quantity of information present in the data.

Uncertainty assessment on apparent traveltimes from ambient noise cross-correlation is an active area of research and Hierarchical Bayes could be used as a tool in the future to quantify the behaviour of noise with different parameters like interstation distance, azimuthal source distribution, or recording time.

It will be soon possible to incorporate new data from on-going deployments at different scales in Australia. Saygin & Kennett (2012) recently processed additional traveltimes that sample the crust at the continental scale. There are also supplementary ambient noise data for Tasmania available (Young et al. 2011). Furthermore, 67 short-period seismometers have recently been positioned across the Gawler and Curnamona Cratons in South Australia (Salmon & Arroucau 2010). Station spacing was approximately 60 km and covers the area from the Streaky Bay in the west to the New South Wales border in the east. Stations recorded continuous three component data for a period of 6–8 months and ambient noise traveltime are currently being processed.

Another possibility is the inclusion of azimuthal anisotropy in the inversion. It is possible to observe an azimuthal dependence on the path-averaged velocities extracted from WOMBAT arrays. Therefore, instead of inverting for a single velocity value within each cell, one could invert for three anisotropic parameters per cell (a maximum velocity, a minimum velocity and a direction).

Other possible extensions include combining ambient noise recordings with receiver functions, earthquake surface wave dispersion measurements, regional and teleseismic traveltimes or SKS splitting measurements. The Hierarchical Bayes procedure may prove to be practically useful for joint inversion, because it is able to naturally weight the contribution of different data types in the misfit function, thus removing the user driven selection of weighting factors.

ACKNOWLEDGMENTS

We thank Kerry Gallagher for discussions on Hierarchical Bayes, Alberto Malinverno, Andrew Curtis and Stan Dosso for useful feedback, and Erdinc Saygin for assisting with seismic data. This project was supported by French-Australian Science and Technology travel grant (FR090051) under the International Science Linkages program from the Department of Innovation, Industry, Science and Research, and Australian Research Council Discovery project DP110102098. Calculations were performed on the TerraWulf II, a computational facility supported through the AuScope inversion laboratory. AuScope Ltd is funded under the National Collaborative Research Infrastructure Strategy (NCRIS) and the Education Infrastructure Fund (EIF), both Australian Commonwealth Government Programmes.

REFERENCES


