# Parsimonious Seismic Tomography with Poisson Voronoi Projections: Methodology and Validation

Hongjian Fang*[1], Robert D. van der Hilst[1], Maarten V. de Hoop[2], Konik Kothari[3], Sidharth Gupta[3], and Ivan Dokmanić[3]

## Abstract

**Ill-posed seismic inverse problems are often solved using Tikhonov-type regularization, that is, incorporation of damping and smoothing to obtain stable results. This typically results in overly smooth models, poor amplitude resolution, and a difficult choice between plausible models. Recognizing that the average of parameters can be better constrained than individual parameters, we propose a seismic tomography method that stabilizes the inverse problem by projecting the original high-dimension model space onto random low-dimension subspaces and then infers the high-dimensional solution from combinations of such subspaces. The subspaces are formed by functions constant in Poisson Voronoi cells, which can be viewed as the mean of parameters near a certain location. The low-dimensional problems are better constrained, and image reconstruction of the subspaces does not require explicit regularization. Moreover, the low-dimension subspaces can be recovered by subsets of the whole dataset, which increases efficiency and offers opportunities to mitigate uneven sampling of the model space. The final (high-dimension) model is then obtained from the low-dimension images in different subspaces either by solving another normal equation or simply by averaging the low-dimension images. Importantly, model uncertainty can be obtained directly from images in different subspaces. Synthetic tests show that our method outperforms conventional methods both in terms of geometry and amplitude recovery. The application to southern California plate boundary region also validates the robustness of our method by imaging geologically consistent features as well as strong along-strike variations of San Jacinto fault that are not clearly seen using conventional methods.**

**Cite this article as** Fang, H., R. D. van der Hilst, M. V. de Hoop, K. Kothari, S. Gupta, and I. Dokmanić (2019). Parsimonious Seismic Tomography with Poisson Voronoi Projections: Methodology and Validation, *Seismol. Res. Lett.* **91**, 343–355, doi: 10.1785/0220190141.

**Supplemental Material**

## Introduction

Seismic tomography has been one of the most important tools for imaging 3D subsurface structures since the pioneering local, regional, and global studies by Aki *et al.* (1977), Romanowicz (1979), and Dziewonski *et al.* (1977), respectively. However, the uneven distribution of earthquakes and stations usually results in ill-posed tomographic inverse problems, calling for regularization to obtain a physically plausible result. Depending on the *a priori* information available, regularization can be included explicitly in the misfit functions used to solve the inverse problem or implicitly controlled by the user.

Among different explicit regularization techniques, the most widely used ones involve the use of so-called damping or smoothing terms, which force the result to be smooth and close to the *a priori* model (Aster *et al.*, 2011). Other types of regularization include sparsity, which exploits the fact that a model can be represented with only a fraction of coefficients in certain domains (Chiao and Kuo, 2001; Chiao and Liang, 2003; Loris *et al.*, 2007, 2010; Simons *et al.*, 2011; Fang and Zhang, 2014) and compactness constraints in cases where the target structure is not smooth (Ajo-Franklin *et al.*, 2007). These methods can produce high-resolution images, but the computational cost is high and the performance often case dependent.

Besides explicit regularization that is incorporated directly into the misfit function, the choice of model parameterization and how to weigh data can also affect the results (Kárason and Van Der Hilst, 2000). Discretization choices include cell- or grid-based representation in local, regional, and global tomography, with cells or grids regularly distributed or irregularly

1. Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.; 2. Department of Computational and Applied Mathematics, Rice University, Houston, Texas, U.S.A.; 3. Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.

*Corresponding author: hfang@mit.edu

adapted to the density of data sampling (Bijwaard *et al.*, 1998; Kárason and Van Der Hilst, 2000). In global problems, harmonic bases have been adopted to decrease the dimension of the model space and allow direct inversion (Dziewonski, 1984). Large grid cells (or low harmony orders) lead to high model stability but might not represent the small-scale structure. In contrast, small grid cells (or coefficients of high-order harmonics) can represent different scales of features, but the available data might not be able to constrain them well. Thus, the trade-off between model complexity and model stability should be taken into consideration when discretizing models. On the data side, different weighting schemes can be used to reduce artifacts due to preferential sampling caused by spatial clustering of events and stations (Bijwaard *et al.*, 1998). Combined, these subjective factors produce different models that each fits the data according to some misfit criterion, leading to a difficult choice of which is more reliable than the others.

Here, we take a different approach toward inversion, which addresses some of the issues mentioned earlier and is also (computationally) efficient. Instead of solving the ill-posed full inverse problem with regularization, we choose to stabilize the inversion by projecting the original high-dimension model space onto low-dimension subspaces and infer the final solution from them. The subspaces are formed by constant functions, that is, constant wavespeed, in randomly distributed Voronoi (Poisson Voronoi [PV]) cells (Sambridge *et al.*, 1995; Okabe *et al.*, 2000). In each subspace, we obtain the low-dimension image by simply solving a normal equation without using any explicit regularization. We can do this many times, each with a random distribution of Voronoi cells. The final result is then retrieved by combining all low-dimension images, either by solving another normal equation or simply by averaging the low-dimension images in different subspaces. Explicit regularization such as damping and smoothing is not necessary because the inverse problem in the projected subspace is better constrained (Gupta *et al.*, 2018). We note the subspace used here is different from the subspace inversion of Kennett *et al.* (1988), who proposed to solve inverse problems with misfit functions dependent on multiple types of parameters, but both approaches aim for model dimension reduction. Gupta *et al.* (2018) used a similar technique to obtain high-resolution images with extremely sparse data coverage, but instead of solving normal equations they trained a deep neural net with artificial data to represent the generalized matrix inverse. In terms of data weighting, our method allows us to use subsets of the whole dataset for different realizations to make data coverage more homogeneous while still benefiting from data redundancy. The strategy that we consider bears some similarity to the Backus–Gilbert inversion (Backus and Gilbert, 1968), which focuses on the mean instead of individual parameters.

In the following sections, we first revisit briefly conventional seismic travel-time tomography and then introduce PV-based inversion. We validate our method with synthetic tests and field data in the southern California plate boundary region. Lastly, we discuss how to choose the number of Voronoi cells and subspaces, the effects of using a subset of the whole dataset, and the possibility of representing uncertainty based on low-resolution images in different subspaces.

## Methodology

### Seismic travel-time tomography based on Tikhonov-type regularization

Linearized seismic travel-time tomography can be formulated as minimizing the quadratic objective function that characterizes data fitting as

$$\Delta\hat{\mathbf{m}} = \text{argmin}\|\mathbf{G}\Delta\mathbf{m} - \Delta\mathbf{d}\|^2, \quad (1)$$

in which $\mathbf{G}$ is the sensitivity matrix, which can be obtained, for instance, by numerically solving the Eikonal equation (e.g., Rawlinson and Sambridge, 2004), 3D raytracing (e.g., Um and Thurber, 1987), or finite-frequency theory (e.g., Dahlen *et al.*, 2000); $\Delta\mathbf{m}$ and $\Delta\mathbf{d}$ are the vectors of model update and data residual between calculated and observed travel times, respectively. Because of the ill-conditioning of the $\mathbf{G}$ matrix, which is caused by the uneven sampling of the model space (due to ray bending and irregular source and receiver distribution) and a particular model discretization, regularization terms must be incorporated to obtain stable results.

The most widely used regularization terms are damping and smoothing, which prevent the model from changing too much from the initial model and make model perturbations smooth. Then, the seismic tomography problem can be recast as

$$\Delta\hat{\mathbf{m}} = \text{argmin}\|\mathbf{G}\Delta\mathbf{m} - \Delta\mathbf{d}\|^2 + \lambda\|\mathbf{L}\Delta\mathbf{m}\|^2 + \mu\|\mathbf{I}\Delta\mathbf{m}\|^2, \quad (2)$$

in which $\mathbf{L}$ is a smoothing matrix, which is usually chosen to be the first- or second-order Tikhonov matrix (Aster *et al.*, 2011), $\mathbf{I}$ is the identity matrix used to obtain a minimum norm solution, and $\lambda$ and $\mu$ are the regularization parameters and are often chosen based on an L-curve (Hansen, 2001; Aster *et al.*, 2011) or generalized cross validation (Golub *et al.*, 1979; MacCarthy *et al.*, 2011). Then, the optimization problem can be solved with singular value decomposition, subspace inversion (Kennett *et al.*, 1988), least squares (LSQR; Paige and Saunders, 1982), or other iterative solvers. Other regularization approaches, for example, using sparsity of the model in certain domains, claim to have better resolution and be data adaptive (Loris *et al.*, 2007; Simons *et al.*, 2011; Fang and Zhang, 2014), but they still depend heavily on the *a priori* information. Apart from the explicit regularization terms, how to discretize the model can be treated as an implicit regularization, which is usually determined based on the user's preferences.

### PV-cells-based seismic tomography

In PV-based tomography, we first discretize the model on a fine grid to enable resolving structures over different scales

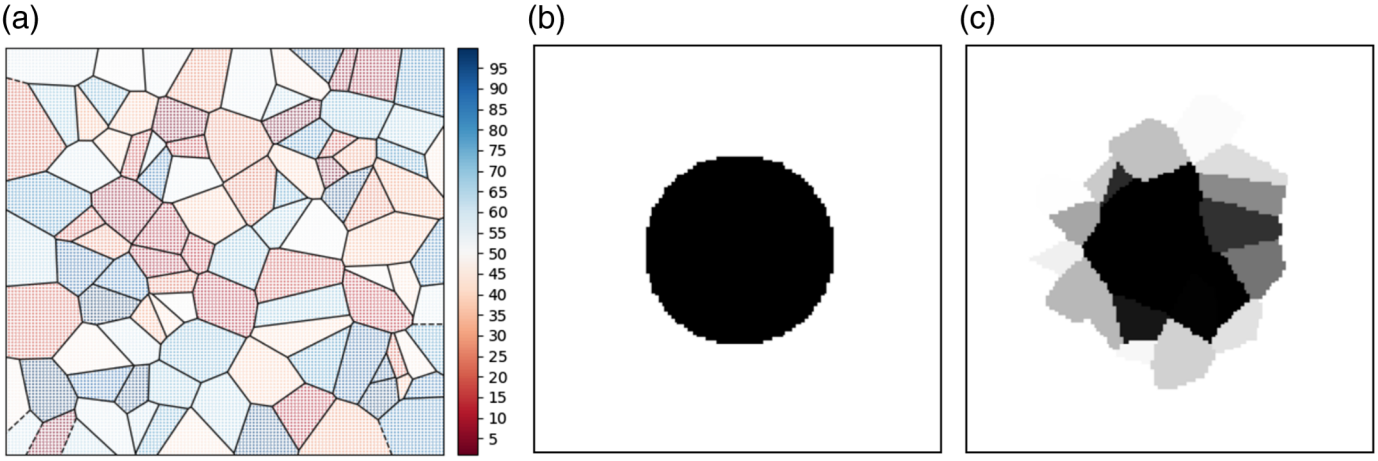



**Figure 1.** Schematic diagram of a high-dimension 2D model and its projection in a subspace formed by Poisson Voronoi (PV) cells. (a) 100 randomly distributed Voronoi cells. The original fine discretization (128 by 128 grid, i.e., 16,384 model parameters) is shown as dots with different colors. Color bar shows the indexes of different cells. (b) Original high-dimension model with the dimension of 128 by 128. (c) Projected image of (b) to a subspace formed by (a) 100 Voronoi cells. The color version of this figure is available only in the electronic edition.

using the available data. This results in a model dimension $m$ that could be larger than the quantity of observed data $n$, and thus an under(mixed)-determined sensitivity matrix $\mathbf{G}$, even with homogeneous data sampling. Instead of solving the high-dimension inverse problem directly, with regularization, we project the original sensitivity matrix into a series $j$ ($j = 1 \ldots N$) of subspaces with dimension $l$, with $l \ll m$

$$\tilde{\mathbf{G}}_j = \mathbf{G}\mathbf{P}_j, \tag{3}$$

in which $\mathbf{P}_j \in \mathbb{R}^{ml}$ is $j$th projection matrix. Figure 1 shows an example of a 2D projection using Voronoi cells, with $m = 16,384$ (128 by 128) and $l = 100$. The projected sensitivity matrix $\tilde{\mathbf{G}}$ is more well conditioned than the original one, that is, closer to full rank (farther from singularity). The model for the $j$th projected low-dimension subspace $\Delta\hat{\mathbf{m}}_j^s$ can be obtained by solving

$$\Delta\hat{\mathbf{m}}_j^s = \text{argmin}\|\tilde{\mathbf{G}}_j\Delta\tilde{\mathbf{m}}_j - \Delta\mathbf{d}\|^2, \tag{4}$$

without any regularization. After solving equation (4) individually for $N$ different subspaces formed by PV (i.e., random) cells, the high-dimension model can be obtained by solving

$$\Delta\hat{\mathbf{m}} = \text{argmin}\|\mathbf{P}\Delta\mathbf{m} - \Delta\tilde{\mathbf{m}}\|^2, \tag{5}$$

in which $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_N]^T$, the superscript $T$ stands for the matrix transpose, and $\Delta\tilde{\mathbf{m}} = [\Delta\hat{\mathbf{m}}_1^s, \Delta\hat{\mathbf{m}}_2^s, \ldots, \Delta\hat{\mathbf{m}}_N^s]^T$. The LSQR solution to equation (5) is

$$\Delta\hat{\mathbf{m}} = \mathbf{P}^g\Delta\tilde{\mathbf{m}}, \tag{6}$$

in which $\mathbf{P}^g = (\mathbf{P}^T\mathbf{P})^{-1}\mathbf{P}^T$ if the inverse of $\mathbf{P}^T\mathbf{P}$ exists and the superscript $g$ stands for the generalized inverse. Alternatively, the solution can be approximated with

$$\Delta\hat{\mathbf{m}} \approx \frac{1}{N}\mathbf{P}^T\Delta\tilde{\mathbf{m}}, \tag{7}$$

which is a simple average over the different projected subspaces. This is, in effect, a filtered version of the exact solution to equation (5), because we approximated $(\mathbf{P}^T\mathbf{P})^{-1}$ as $\frac{1}{N}\mathbf{I}$. We note that this approximation has been widely applied to interface and source imaging, for example, the backprojection technique to image earthquake rupture using high-frequency teleseismic data (Ishii *et al.*, 2005). The usage of PV cells makes equation (7) converge extremely fast. In addition to the average model, equation (7) yields the variance of the solutions from different realizations, at a certain location, which can be used as a measure of uncertainty. Similar random-number-based techniques have also been used for estimating the diagonal entries of large matrices (MacCarthy *et al.*, 2011) and in full-waveform inversion to extract the point spread function in the Hessian matrix for resolution assessment (Trampert *et al.*, 2012; Fichtner and Leeuwen, 2015).

## Method Validation

### A 2D synthetic test

We first illustrate and validate our method using a simple 2D synthetic test. The input model with a 5% low-wavespeed anomaly in the center is shown in Figure 2a, and the station distribution used for this test is taken from Zigone *et al.* (2015) in the southern California plate boundary region. We then calculate the travel times between different pairs of stations (Fig. 2a) using straight rays. This would be similar to, for instance, the first step in obtaining phase or group velocity maps in surface-wave or ambient-noise tomography. We note that one could use more accurate kernels but straight rays are adequate for illustration purposes. After adding Gaussian noise

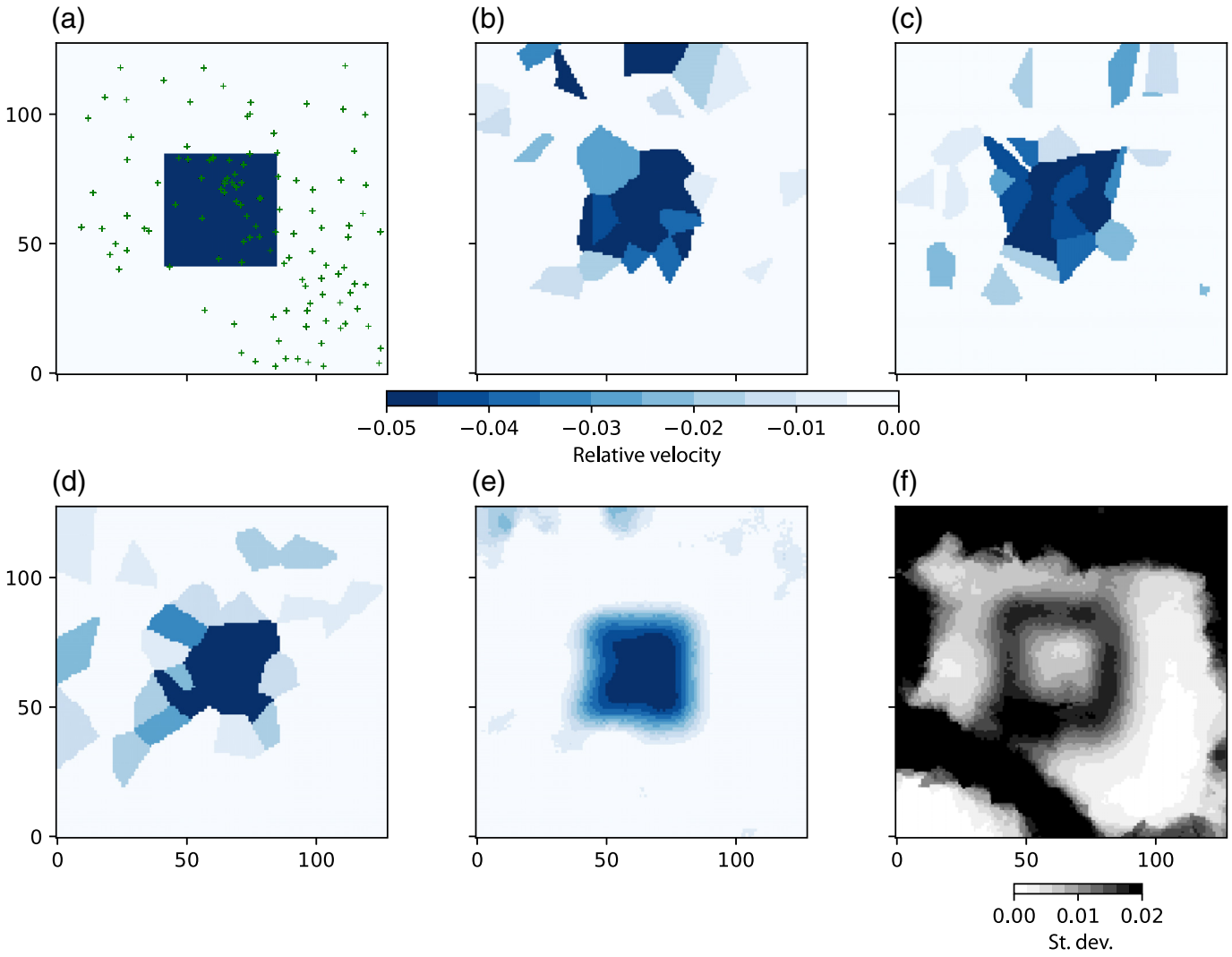



**Figure 2.** A 2D synthetic test. (a) The input model with a 5% low-wavespeed anomaly in the center. Green crosses show the station distribution. (b–d) Inverted results from three different realizations, each with 100 Voronoi cells. (e) Recovered image by averaging 500 realizations. (f) Standard deviation estimated from 500 subspaces. The color version of this figure is available only in the electronic edition.

with a standard deviation of 0.15 s, we apply the PV-based method to (randomly selected) 1000 synthetic travel times.

We discretize the original model using a fine grid consisting of 128 by 128 nodes in the $X$ and $Y$ directions (Fig. 1a), which leads to 16,384 model parameters. The number of Voronoi cells used should be related to the noise in the data and the number of measurements, which will be discussed in detail later. After several tests with different numbers of subspaces formed by PV cells, we found that 50 cells already lead to a stable and good reconstruction, but for this illustration we use 100 cells. We project the high-dimension model onto 500 randomly permuted subspace distributions formed by 100 random Voronoi cells (Fig. 1a). The number of realizations ($N = 500$) is excessive but guarantees stability. We then solve equation (4) for each realization independently without explicit regularization.

Figure 2b–d shows the results for three different realizations, with each one recovering the input model to a certain degree with low-wavespeed anomalies in the center. As expected, artifacts exist in regions where the data sampling is poor and anomalies are distorted by the shapes of the Voronoi cells. The noisy features in the recovered images in the projected subspaces show that the projections are still sensitive to noise in the data, but the projection matrices are less ill-conditioned than the original ones, with which recovery without regularization would not yield plausible results. We obtain the final model by averaging all 500 different realizations based on equation (7). The image obtained from averaging (Fig. 2e) suppresses the artifacts in individual realizations and the effect caused by different shapes of Voronoi cells. Besides, the standard deviation estimation (Fig. 2f) obtained from the variance of low-dimension images in different realizations captures the large uncertainties in regions with poor data coverage as well as the bimodal distribution near sharp boundaries.

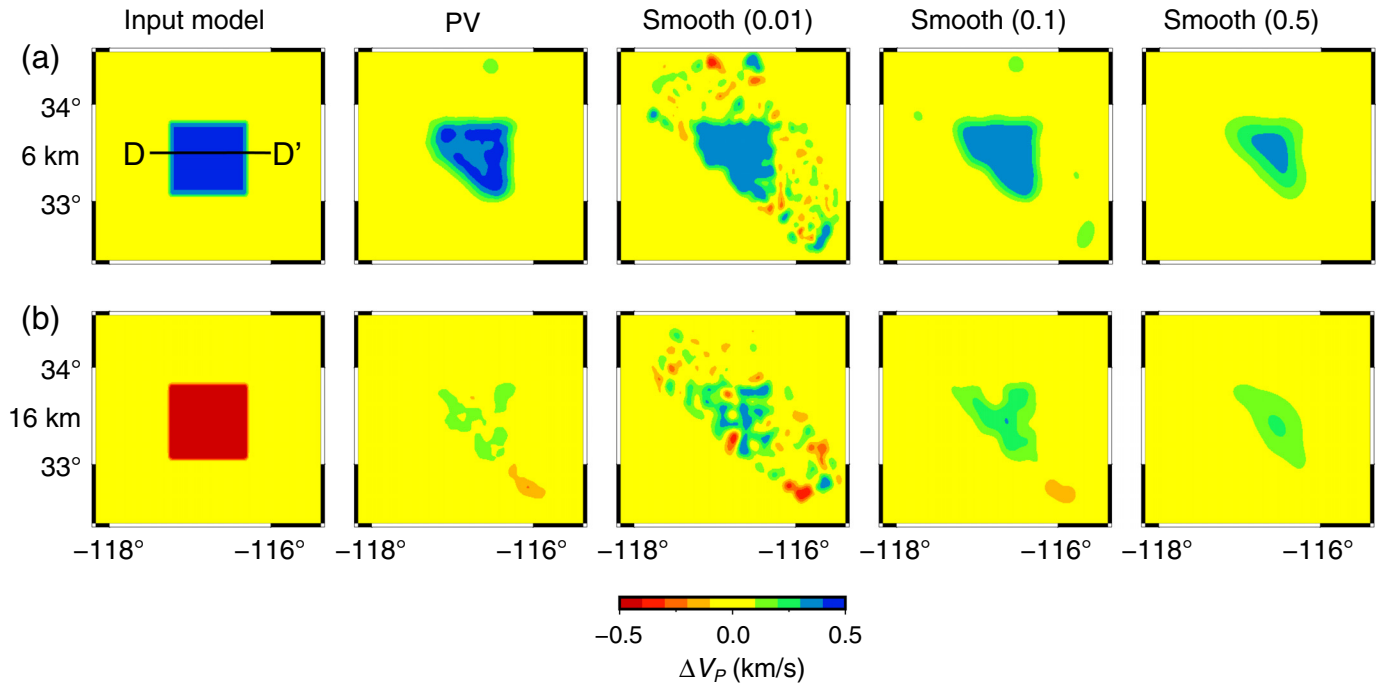



**Figure 3.** Horizontal slices of $\Delta V_P$ in the 3D synthetic test. The first column shows the input model, followed by obtained model from PV-based inversion, conventional inversion with smoothing parameters of 0.01, 0.1, and 0.5 at depth of (a) 6 and (b) 16 km. The thick black line in (a) shows the position of cross section for Figure 4. The color version of this figure is available only in the electronic edition.

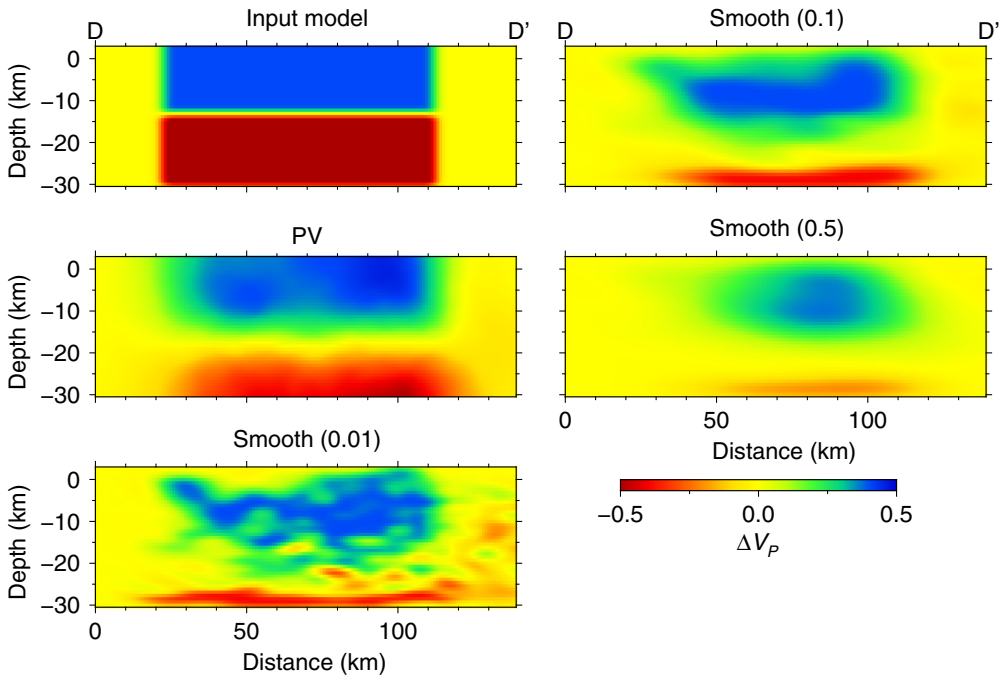



**Figure 4.** Vertical slice of $\Delta V_P$ for the input model, PV-based inversion, and conventional inversion with different smoothing parameter. The position is shown in Figure 3a. The color version of this figure is available only in the electronic edition.

## Validation with a 3D synthetic test

To validate our method further, we designed a 3D toy model with high $V_P$ (0.5 km/s larger than the reference) in the upper 11 km and low $V_P$ (0.5 km/s smaller than the reference) at greater depths (Figs. 3a and 4a). A fast-marching method (de Kool *et al.*, 2006) was adopted to calculate the travel times and ray paths with the same earthquake and station distribution as the field data (Fig. 5a). After adding Gaussian noise with a standard deviation of 0.06 s to the synthetic data, we applied the two methods mentioned in the Methodology section to invert for $V_P$.

The model obtained from PV-based inversion shows more accurate amplitude and fewer

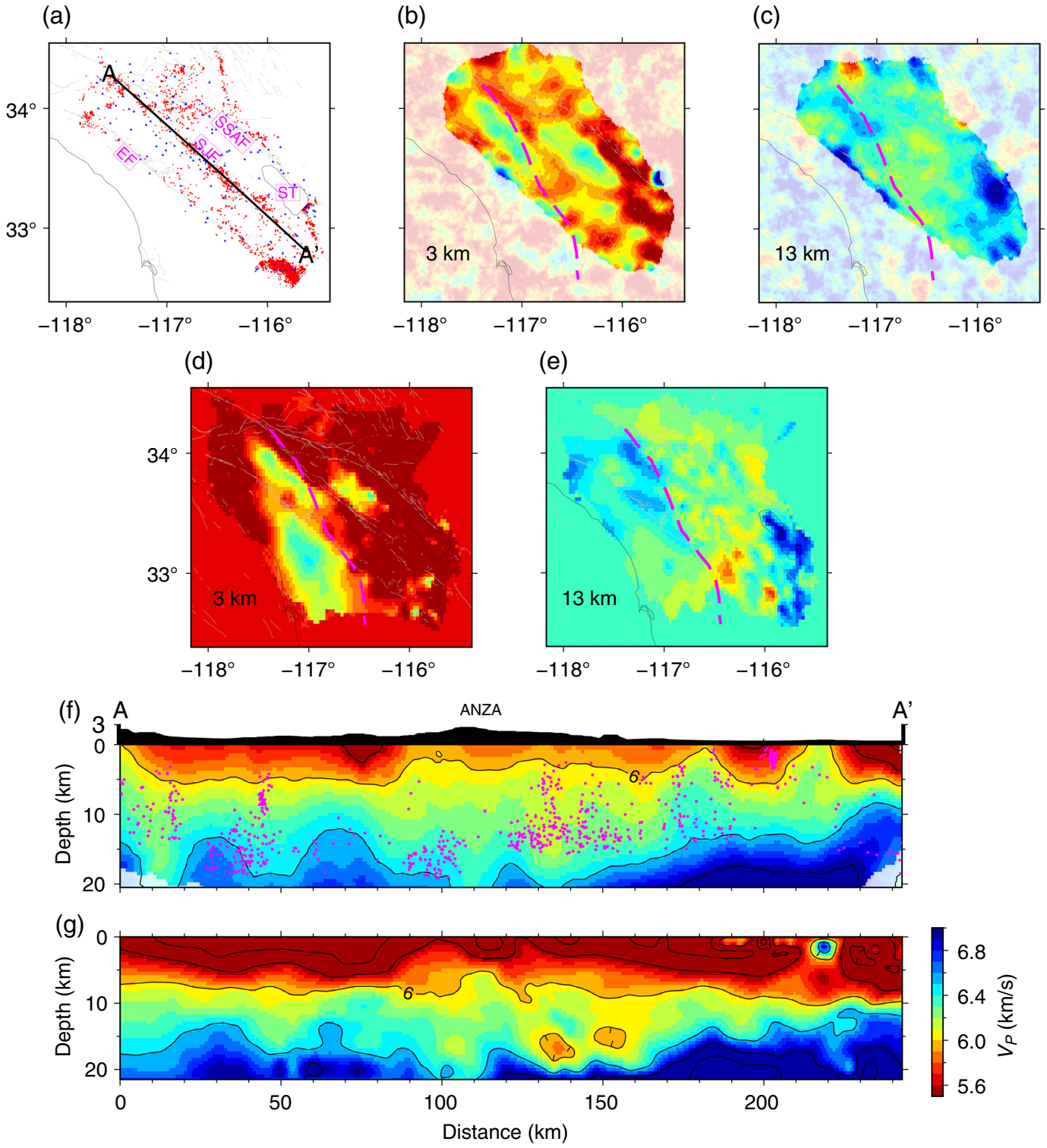



**Figure 5.** Application to field data in southern California plate boundary region. (a) Distribution of earthquakes (red dots) and stations (blue triangles). Gray lines show the fault traces. EF, Elsinore fault; SAF, San Andres fault; SJF, San Jacinto fault; ST, Salton trough. (b,c) Horizontal slice of $V_P$ at 3 and 13 km. (d,e) Same as (b,c) but for the model from Fang *et al.* (2016). A vertical slice of $V_P$ along the SJF from (f) PV-based method and (g) Fang *et al.* (2016). The position is shown as a black thick line in (a). Pink dots in (f) show the relocated earthquakes within 5 km from the cross section. Topographic variation along A–A′ is plotted on top of (f). Magenta lines in horizontal slices show the Peninsular Range compositional boundary. The shaded area in (b,c,f) corresponds to regions with large standard deviation (>0.3 km/s in Fig. 7). The color version of this figure is available only in the electronic edition.

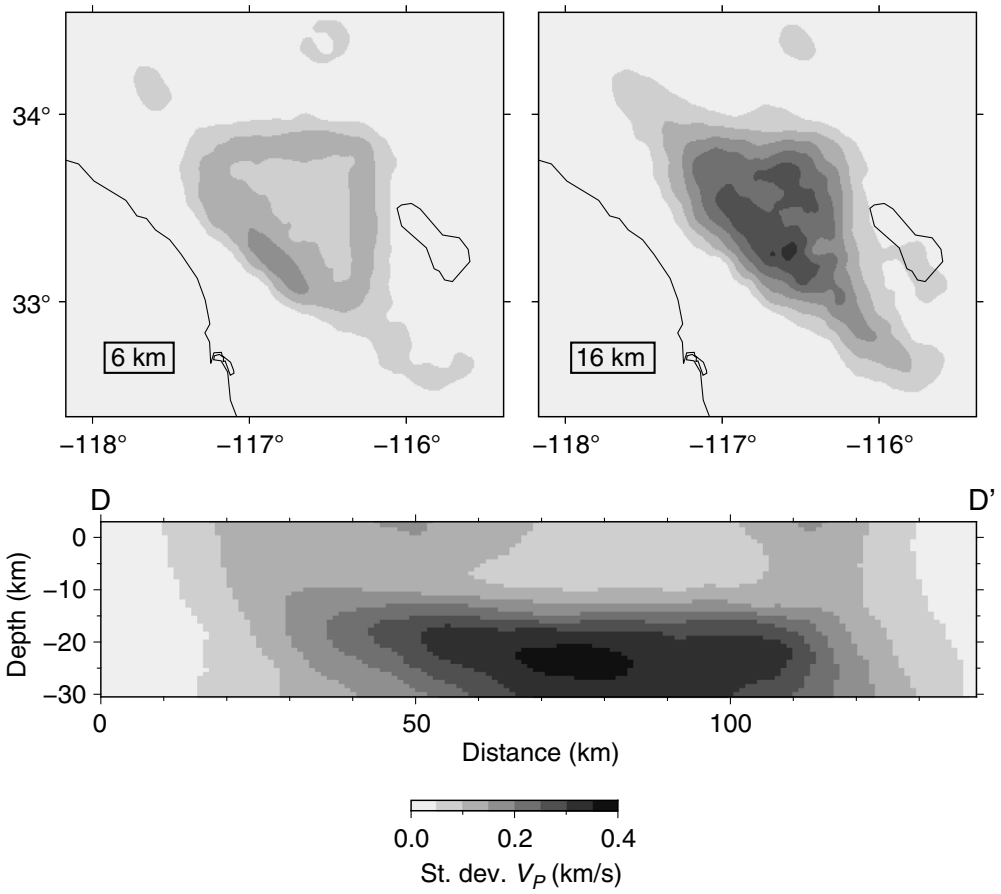



**Figure 6.** Standard deviation estimated from the 3D synthetic test at 6 and 16 km and a vertical slice shown in Figure 3a.

artifacts compared to conventional tomography with different smoothing parameter ($\lambda$) (Figs. 3 and 4). The conventional tomography method works fine with appropriate regularization parameter, with $\lambda$ in equation (2) equal to 0.1 for this particular case. Smaller values lead to artifacts, whereas larger value results in diminished amplitude and smeared boundaries (Fig. 3). In contrast, the PV-based inversion method reconstructs the low $V_P$ anomaly at greater depth (Fig. 4), whereas the conventional method tends to underestimate this anomaly because of the larger trade-off between different model parameters in high-dimension space. Both methods fail to recover the low-wavespeed anomaly at 16 km, but the PV-based method shows less smearing compared to the conventional method (Fig. 3), and the large uncertainty also indicates the data do not constrain this part of the model space well (Fig. 6).

## Application to the southern California plate boundary using *P*-wave travel-time data

We applied the previous method to *P*-wave travel-time data in the southern California plate boundary region (Allam and Ben-Zion, 2012). The dataset includes about 200,000 *P*-wave arrivals recorded by 139 stations from almost 5500 earthquakes (see Data and Resources). We use the same initial 1D model as in Allam and Ben-Zion (2012), and the model is discretized with a grid interval of 0.01° in the horizontal directions and 0.5 km in the vertical direction, which results in 3,797,996 model parameters. The travel times and ray paths are calculated using a fast marching method (de Kool *et al.*, 2006). After a similar trade-off analysis as the 2D synthetic test, we choose 100 subspace realizations, each formed by 500 Voronoi cells. We only use a selected subset of the complete datasets for each subspace; not only do they reduce computation cost but also reduce effects of preferential sampling due to event clustering. The choice of subdatasets for each realization will be discussed in detail later.

The recovered model shows features that are similar to a recent model from joint inversion of body- and surface-wave data Fang *et al.* (2016), such as low $V_P$ along the southern San Andres fault in the upper 5 km related to broad fault damage; low $V_P$ in the upper 7 km and high $V_P$ at greater depth in the Salton trough area, which corresponds to crustal thinning due to the rifting in the Gulf of Mexico; as well as a velocity contrast along the northern part of the San Jacinto fault (SJF; Fig. 5b,c). The central part of the SJF is relatively slow, which corresponds well with high topography and Moho uplifting to maintain isostatic balance in the mantle (Lewis *et al.*, 2000). In the southeast part of the SJF, the depth of seismicity is shallower than that of the northwest part (Fig. 5f), correlating well with the velocity model and high heat flow data (Enescu *et al.*, 2009), which may indicate a shallower brittle-to-ductile transition depth in this region. The strong along-strike variation along the SJF, which is not clearly seen in Allam and Ben-Zion (2012), seems well constrained in both the new model and that of Fang *et al.* (2016), but the new model seems simpler with fewer features that look like artifacts (Fig. 5f,g). Moreover, the velocity contrast aligns better with the Peninsular Range compositional boundary and the fault trace in the northwest part of SJF, compared to the model of Fang *et al.* (2016) (Fig. 5b,d). Further model validation using waveform simulation will be left to future work. The estimated

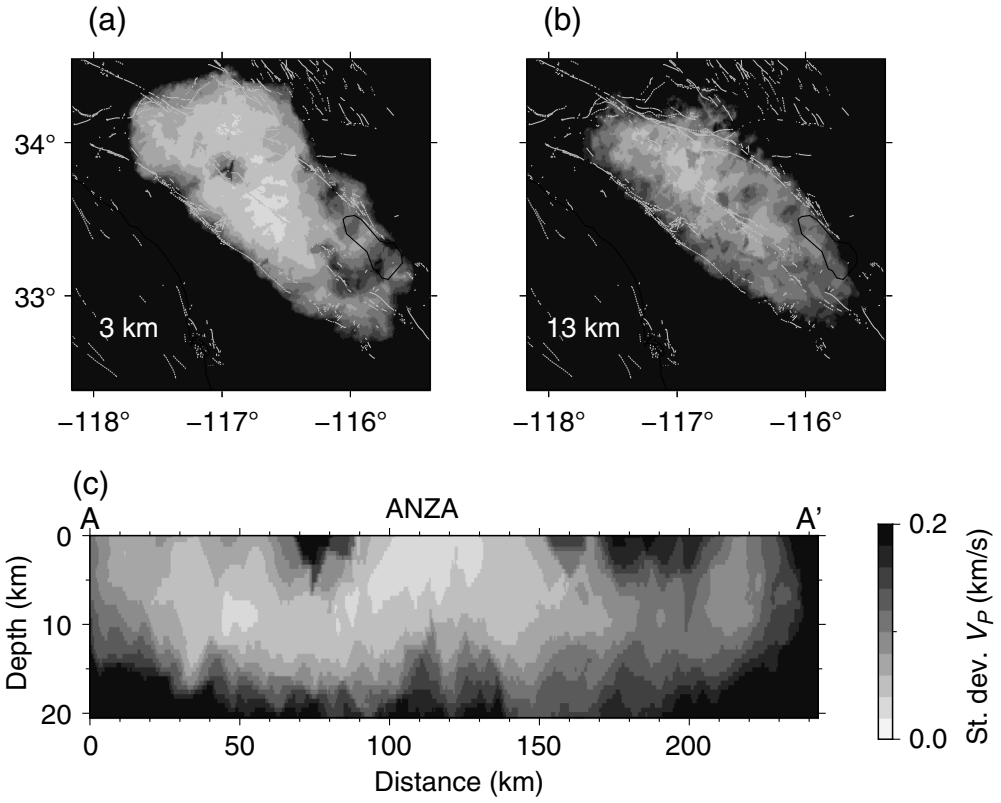



**Figure 7.** Standard deviation estimated from field data application at (a) 3 and (b) 13 km, and (c) a vertical slice along the SJF. The position of the vertical slice is shown in Figure 5a.

standard deviation (Fig. 7) is consistent with data coverage, with regions of good data sampling show smaller values than regions with sparse data sampling. In addition, large standard deviation in the northwest part of SJF (Fig. 7a) may help indicate large velocity contrast.

## Discussions

### Determining the number of Voronoi cells and realizations

Because we do not use explicit regularization, the only subjective parameters in our inversion system are the number of Voronoi cells, that is, the dimension $l$ of each subspace, and the number $N$ of realizations needed to obtain a robust solution. Similar to the trade-off between model complexity and stability in conventional tomography, a small number of cells in PV-based inversion tends to be stable for the LSQR solution, but will have difficulty fitting the data. In contrast, a large number of Voronoi cells improve the data fit, but they will also fit the noise and lead to ill-conditioning of the sensitivity matrix. Thus, the number of Voronoi cells should depend on the quantity and quality (i.e., noise level) of the data.

We tested different numbers of cells in the 2D synthetic model using $l = 20$, 50, 100, 200, and 300 Voronoi cells, each with $N = 500$ different realizations (Fig. 8 and Fig. S1, available in the supplemental material to this article). The data fit versus dimension $l$ (Fig. 8a) is similar to the shape of an L-curve (Hansen, 2001), indicating there is a trade-off, as expected, between data fit and the number of Voronoi cells. The result for $l = 20$ (Fig. 8b) shows good recovery of the low-wavespeed anomaly in the center, but with diminished amplitude and smeared boundaries. With the number of Voronoi cells increasing to 300 (Fig. 8f), the data fit improves, but there are more artifacts in regions with poor data sampling. The result from 100 cells seems best able to recover the true amplitude and the relatively sharp boundaries with good data fit and limited artifacts. As expected, the estimated standard deviation (Fig. S1) shows ring-like features along the boundaries and large values in regions with sparse data coverage when the cell number is larger than 50. Combined with the recovered models, the results indicate (as expected) a trade-off between data fit and cell number $l$, but not as strong as conventional tomography, in which only slightly different regularization parameters can change the model dramatically (Fig. 3).

To determine the optimal number of realizations, that is $N$, we plot the recovered models and the estimated uncertainty from different realizations. Figure 9 shows the results associated with their standard deviation using $N = 20$, 50, 100, and 300 realizations, respectively. The difference between 100 and 300 realizations is quite small. In fact, the result from 20 realizations already captures most of the features, albeit with some noise. The decrease of mean squared error between the recovered model and the input model becomes insignificant after about 50 realizations, indicating a small number of realizations are usually sufficient for obtaining stable results.

We stress that choosing the number of Voronoi cells is easier than choosing the regularization parameters in conventional tomography. Besides the data quantity and quality, the number of Voronoi cells is also a function of the complexity in the actual structure. A simple model with smooth structure may only need a small number of cells (Fig. S2). More cells are necessary to represent models with sharp boundaries or complex structures (Figs. S3 and S4), but the number usually ranges from several dozen to several hundred in local to regional scales using travel-time data, which are consistent with the

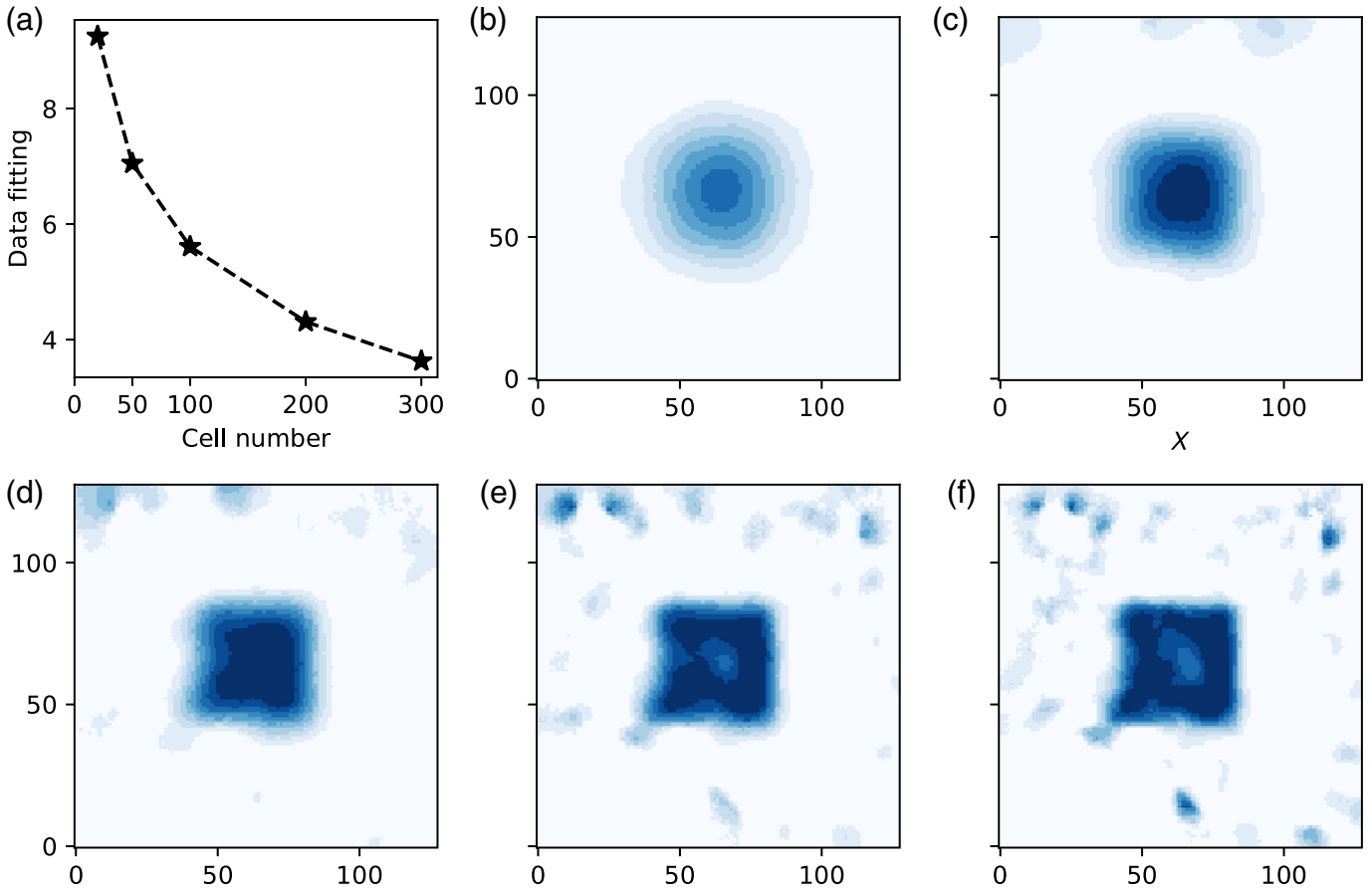


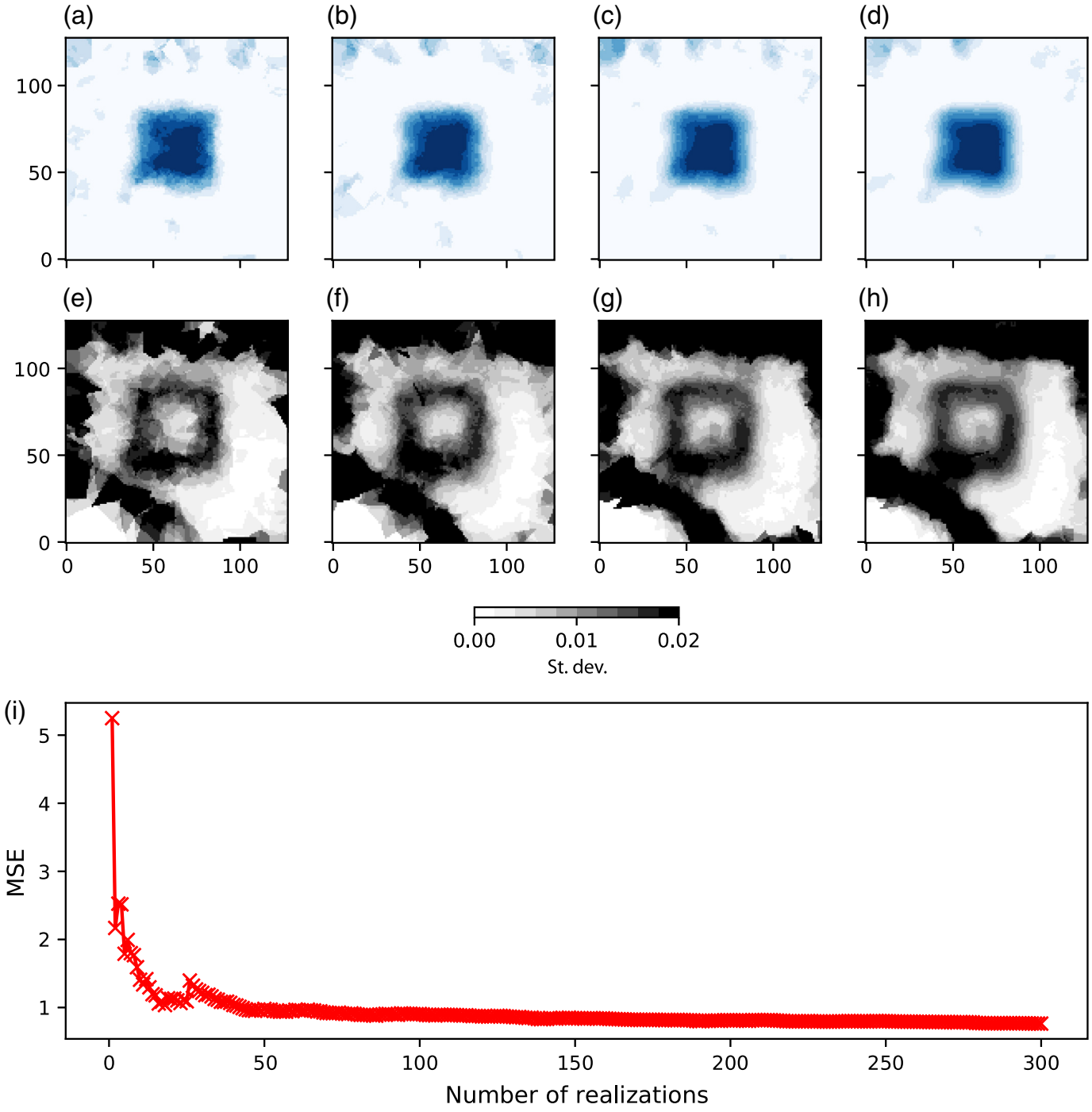

**Figure 8.** (a) The trade-off curve of data fitting versus number of Voronoi cells. (b–f) The result using 20, 50, 100, 200, and 300, respectively, each obtained from averaging of 500 realizations. The color bar is the same as Figure 2. The color version of this figure is available only in the electronic edition.

posterior distribution of cell number in nonlinear inversions (e.g., Bodin and Sambridge, 2009; Agostinetti *et al.*, 2015; Zhang *et al.*, 2018). Thus, it is possible to use the posterior distribution of cell numbers in nonlinear inversion as *a priori* information for choosing PV subspace dimension $l$. Furthermore, choosing $l$ is quite robust in that the solution changes little for choices around a reasonable value of $l$. In contrast, choosing the damping and smoothing parameters in conventional tomography requires test inversions with a broad spectrum of different values, and the solution may be more sensitive to this value than the choice of $l$.

## The standard deviation estimation

Model uncertainty is critical in seismic tomography, yet most linearized seismic tomography methods do not give direct information about it. The most commonly used technique to characterize the resolution is the checkerboard test, although it tells us little about the true resolution and has its intrinsic drawbacks (Lévêque *et al.*, 1993; Rawlinson and Spakman, 2016). The use of Markov chain Monte Carlo (MCMC) sampling to explore the model space has gained in popularity in the past decade (e.g., Bodin and Sambridge, 2009; Agostinetti *et al.*, 2015; Burdick and Lekić, 2017), but its applications are limited to small-scale problems where the dimension of the model space is less than a few hundred or large-scale problems where the forward calculation is efficient (Burdick and Lekić, 2017).

In our PV-based tomography, the projected models in different subspaces can fit the data equally well and all lie in or very close to the high-dimension model space. Thus, we could in principle characterize model uncertainty using the variance in the projected models, at least partly. Indeed, the standard deviations for the 2D and 3D synthetic tests (Figs. 6 and 9) look similar to those from a Monte Carlo sampler (e.g., Bodin and Sambridge, 2009; Galetti and Curtis, 2018), with the ring-like features around the sharp boundaries showing the bimodal distribution of model parameters in those regions. Moreover, large differences between the recovered model and the input model are also associated with large standard deviation (Fig. S5), indicating the estimated standard deviation could, at least partly, capture the real model uncertainty when the noise in the data is known *a priori*. For the field data application, the standard deviation is large both in regions with poor data sampling and in regions with sharp velocity contrasts such as the northwest part of SJF (Fig. 7). At this stage, we should mention that the uncertainty estimation is preliminary because it is still not clear to what degree the projected subspaces capture the whole model




space, but we argue that even with this caveat in mind it is more informative than conventional synthetic tests for characterizing uncertainty.

## Data selection for each realization

Earthquakes used in passive seismic tomography mostly occur in specific tectonically active regions. This leads to redundant sampling in certain regions and leaves others undersampled. Commonly used techniques to prevent the results from being dominated by clusters of earthquakes include selecting events to be optimally distributed (Evans *et al.*, 1994), clustering events for summary rays (Bijwaard *et al.*, 1998; Kárason and van der Hilst, 2001), and downweighing the data in regions within event clusters (e.g., Schaeffer and Lebedev, 2013).

**Figure 9.** Obtained results of the 2D synthetic test for (a) 20, (b) 50, (c) 100, and (d) 300 realizations as well as their corresponding standard deviation estimation (e–h). (i) The mean square error (MSE) between the input and obtained model versus the number of realizations. The color bar for (a–d) is the same as Figure 2. The color version of this figure is available only in the electronic edition.

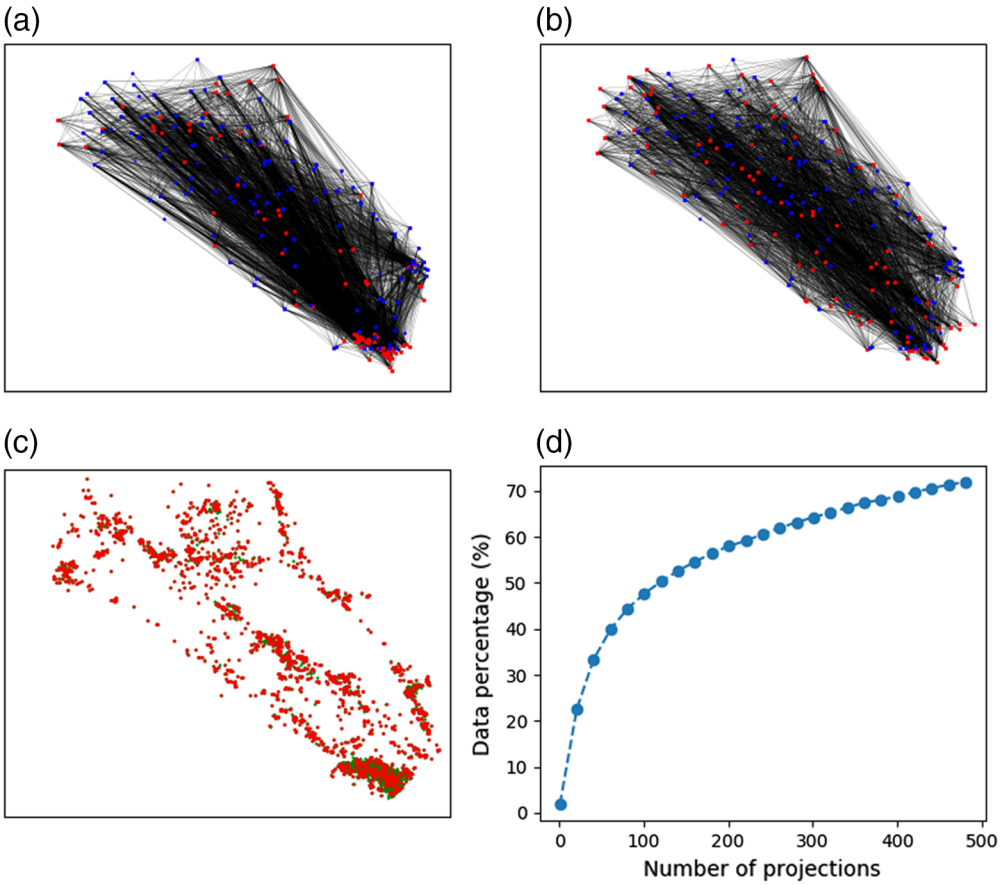



**Figure 10.** (a) Map view of the overall data distribution characterized by one-tenth of the rays of the complete dataset. (b) Map view of data sampling from a subset of data for a certain subspace. (c) Map view of the complete set of events (red dots) and events used in the PV inversion (green dots). (d) Data used in the PV-based inversion with the increasing number of realizations. The color version of this figure is available only in the electronic edition.

In our field data case, using the full dataset, which includes about 200,000 measurements, seems unnecessary for determining 500 unknowns in each subspace realization, and sampling is dominated by data from nearby events (Fig. 10). Both for efficiency and for suppressing the effects of uneven data coverage, for each realization we use a subset of all available data. To this effect, we first cluster earthquakes using blocks with similar size that span the source regions and randomly chose a subset of earthquakes in each block for each subspace reconstruction. We selected about 3000 measurements from 150 events for each subspace reconstruction to achieve more homogeneous sampling (Fig. 10b). There will be events in events cluster regions that may not be selected for inverting the model in the subspace. However, travel-time data from these events could serve as validation datasets to check if the obtained model could also fit those data. By choosing different subsets of the data for different subspaces in a more or less random way, we obtain results that are less affected by event clustering and still make use of nearly all data. Indeed, after 100 realizations, the event distribution used (Fig. 10c) in the field data case is similar to the full event distribution, and the subsets of data used in the inversion approach 50% of the whole dataset (Fig. 10d). More subspaces (or using more data per subspace) would increase the percentage of the subset data to nearly 80% for 500 realizations (for 3000 data per subspace), but the remaining data seem redundant and will unlikely add more information about the final results if the data noise from similar paths is correlated. Indeed, results (Fig. S6) using about 2000 measurements from 100 events show very similar results as the results using 3000 measurements.

## Conclusions

We developed a PV-based seismic tomography method, which bypasses the need for explicit regularization. Our method mitigates the ill-posedness of the inverse problem by first projecting the high-dimension model space onto a number of low-dimension subspaces formed by randomly distributed Voronoi cells and then reconstructing the final model from combinations of these low-dimensional subspaces. In addition to computation efficiency, the method is attractive as the choice of free parameters (i.e., the dimension $l$ and the number $N$ of subspace distributions) is more straightforward than the choice of regularization parameters in conventional tomography, and, the results are more robust, as they do not vary much between different choices of $l$ and $N$ around reasonable values. Furthermore, for each realization a subset of data can be selected to reduce the effects of uneven data coverage. Lastly, the method can be used to characterize model uncertainty, which is of great importance when interpreting the final results. Synthetic tests show that (for a certain type of sensitivity kernel) our method can recover anomaly amplitudes and relatively sharp boundaries better than conventional tomography methods. The application to the *P*-wave travel-time data in the southern California plate boundary region demonstrates its robustness by revealing geologically consistent features such as velocity contrasts across faults, fault damage related low-wavespeed anomalies, and known geologic boundaries. The low-dimension and independence of subspaces,

along with the use of subsets of the data, makes our method inherently parsimonious and highly efficient (and parallelizable). This PV approach is attractive for several geophysical inverse problems such as surface-wave tomography, earthquake slip inversion, and global scale tomography, where MCMC-based inversion and other full model space searches are still computationally expensive.

## Data and Resources

The earthquake catalog and *P*-wave travel-time data used in this study are mostly from Allam and Ben-Zion (2012) and can be obtained from https://github.com/HongjianFang/sjfz_data (last accessed May 2019). The inversion code is available now upon request and will be accessible soon on the author's Github site. Supplemental material for this article includes validation of Poisson Voronoi projections and model comparison.

## Acknowledgments

The article benefited from the insightful discussion with Malcolm Sambridge at Australian National University. The authors thank Cliff Thurber for his helpful comments on the early version of the manuscript. Constructive comments from Nick Rawlinson and an anonymous reviewer greatly improved the article. Some of the figures are created using Generic Mapping Tools (GMT) plotting software (Wessel and Smith, 1998).

## References

Agostinetti, N., G. Giacomuzzi, and A. Malinverno (2015). Local three-dimensional earthquake tomography by trans-dimensional Monte Carlo sampling, *Geophys. J. Int.* **201,** 1598–1617, doi: 10.1093/gji/ggv084.

Ajo-Franklin, J. B., B. J. Minsley, and T. M. Daley (2007). Applying compactness constraints to differential traveltime tomography, *Geophysics* **72,** no. 4, R67–R75.

Aki, K., A. Christoffersson, and E. S. Husebye (1977). Determination of the three-dimensional seismic structure of the lithosphere, *J. Geophys. Res.* **82,** 277–296, doi: 10.1029/jb082i002p00277.

Allam, A. A., and Y. Ben-Zion (2012). Seismic velocity structures in the southern California plate-boundary environment from double-difference tomography, *Geophys. J. Int.* **190,** 1181–1196, doi: 10.1111/j.1365-246X.2012.05544.x.

Aster, R. C., B. Borchers, and C. H. Thurber (2011). *Parameter Estimation and Inverse Problems*, Second Ed., Vol. 90, Academic Press, Cambridge, United Kingdom.

Backus, G., and F. Gilbert (1968). The resolving power of gross earth data, *Geophys. J. Roy. Astron. Soc.* **16,** 169–205, doi: 10.1111/j.1365-246x.1968.tb00216.x.

Bijwaard, H., W. Spakman, and E. R. Engdahl (1998). Closing the gap between regional and global travel time tomography, *J. Geophys. Res.* **103,** no. B12, 30,055–30,078.

Bodin, T., and M. Sambridge (2009). Seismic tomography with the reversible jump algorithm, *Geophys. J. Int.* **178,** 1411–1436, doi: 10.1111/j.1365-246X.2009.04226.x.

Burdick, S., and V. Lekić (2017). Velocity variations and uncertainty from transdimensional P-wave tomography of North America, *Geophys. J. Int.* **209,** no. 2, 1337–1351, doi: 10.1093/gji/ggx091.

Chiao, L., and B. Kuo (2001). Multiscale seismic tomography, *Geophys. J. Int.* **145,** 517–527, doi: 10.1046/j.0956-540x.2001.01403.x.

Chiao, L.-Y., and W.-T. Liang (2003). Multiresolution parameterization for geophysical inverse problems, *Geophysics* **68,** 199–209, doi: 10.1190/1.1543207.

Dahlen, F., S.-H. Hung, and G. Nolet (2000). Fréchet kernels for finite-frequency traveltimes—I. Theory, *Geophys. J. Int.* **141,** no. 1, 157–174.

de Kool, M., N. Rawlinson, and M. Sambridge (2006). A practical grid-based method for tracking multiple refraction and reflection phases in three-dimensional heterogeneous media, *Geophys. J. Int.* **167,** no. 1, 253–270.

Dziewonski, A. M. (1984). Mapping the lower mantle: Determination of lateral heterogeneity in *P* velocity up to degree and order 6, *J. Geophys. Res.* **89,** no. B7, 5929–5952.

Dziewonski, A. M., B. H. Hager, and R. J. O'Connell (1977). Large-scale heterogeneities in the lower mantle, *J. Geophys. Res.* **82,** 239–255, doi: 10.1029/jb082i002p00239.

Enescu, B., S. Hainzl, and Y. Ben-Zion (2009). Correlations of seismicity patterns in Southern California with surface heat flow data, *Bull. Seismol. Soc. Am.* **99,** no. 6, 3114–3123.

Evans, J. R., D. Eberhart-Phillips, and C. Thurber (1994). User's manual for SIMULPS12 for imaging Vp and Vp/Vs; a derivative of the "Thurber" tomographic inversion SIMUL3 for local earthquakes and explosions, *U.S. Geol. Surv. Open-File Rept. 94-431*, 101 pp.

Fang, H., and H. Zhang (2014). Wavelet-based double-difference seismic tomography with sparsity regularization, *Geophys. J. Int.* **199,** no. 2, 944–955.

Fang, H., H. Zhang, H. Yao, A. Allam, D. Zigone, Y. Ben-Zion, C. Thurber, and R. D. van der Hilst (2016). A new algorithm for three-dimensional joint inversion of body wave and surface wave data and its application to the Southern California plate boundary region, *J. Geophys. Res.* **121,** 3557–3569, doi: 10.1002/2015JB012702.

Fichtner, A., and T. v. Leeuwen (2015). Resolution analysis by random probing, *J. Geophys. Res.* **120,** no. 8, 5549–5573.

Galetti, E., and A. Curtis (2018). Transdimensional electrical resistivity tomography, *J. Geophys. Res.* **123,** no. 8, 6347–6377.

Golub, G. H., M. Heath, and G. Wahba (1979). Generalized cross-validation as a method for choosing a good ridge parameter, *Technometrics* **21,** no. 2, 215–223.

Gupta, S., K. Kothari, M. V. de Hoop, and I. Dokmanić (2018). *Deep Mesh Projectors for Inverse Problems*, available at http://arxiv.org/abs/1805.11718v3 (last accessed May 2019).

Hansen, P. (2001). The L-curve and its use in the numerical treatment of inverse problems. Computational inverse problems in electrocardiology, *Adv. Comput. Bioeng.* **5,** 119.

Ishii, M., P. M. Shearer, H. Houston, and J. E. Vidale (2005). Extent, duration and speed of the 2004 Sumatra-Andaman earthquake imaged by the Hi-Net array, *Nature* **435,** no. 7044, 933.

Kárason, H., and R. D. van der Hilst (2000). Constraints on mantle convection from seismic tomography, in *The History and Dynamics of Global Plate Motion*, M. A. Richards, R. Gordon, and R. D. van der Hilst (Editors), Geophys. Monogr. Ser., Vol. 121, AGU, Washington, D. C., 277–288.

Kárason, H., and R. D. van der Hilst (2001). Tomographic imaging of the lowermost mantle with differential times of refracted and

diffracted core phases (PKP, Pdiff), *J. Geophys. Res.* **106,** no. B4, 6569–6587.

Kennett, B., M. Sambridge, and P. Williamson (1988). Subspace methods for large inverse problems with multiple parameter classes, *Geophys. J. Int.* **94,** no. 2, 237–247.

Lévěque, J.-J., L. Rivera, and G. Wittlinger (1993). On the use of the checker-board test to assess the resolution of tomographic inversions, *Geophys. J. Int.* **115,** no. 1, 313–318.

Lewis, J. L., S. M. Day, H. Magistrale, J. Eakins, and F. Vernon (2000). Regional crustal thickness variations of the Peninsular Ranges, southern California, *Geology* **28,** no. 4, 303–306.

Loris, I., H. Douma, G. Nolet, I. Daubechies, and C. Regone (2010). Nonlinear regularization techniques for seismic tomography, *J. Comput. Phys.* **229,** 890–905, doi: 10.1016/j.jcp.2009.10.020.

Loris, I., G. Nolet, I. Daubechies, and F. Dahlen (2007). Tomographic inversion using $l_1$-norm regularization of wavelet coefficients, *Geophys. J. Int.* **170,** 359–370, doi: 10.1111/j.1365-246x.2007.03409.x.

MacCarthy, J., B. Borchers, and R. Aster (2011). Efficient stochastic estimation of the model resolution matrix diagonal and generalized cross-validation for large geophysical inverse problems, *J. Geophys. Res.* **116,** no. B10, doi: 10.1029/2011JB008234.

Okabe, A., B. Boots, K. Sugihara, and S. N. Chiu (2000). Spatial tessellations: Concepts and applications of voronoi diagrams, in *Wiley Series in Probability and Statistics*, Second Ed., John Wiley, Chichester, United Kingdom.

Paige, C. C., and M. A. Saunders (1982). LSQR: An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Software* **8,** no. 1, 43–71.

Rawlinson, N., and M. Sambridge (2004). Wave front evolution in strongly heterogeneous layered media using the fast marching method, *Geophys. J. Int.* **156,** no. 3, 631–647.

Rawlinson, N., and W. Spakman (2016). On the use of sensitivity tests in seismic tomography, *Geophys. J. Int.* **205,** no. 2, 1221–1243.

Romanowicz, B. A. (1979). Seismic structure of the upper mantle beneath the United States by three-dimensional inversion of body wave arrival times, *Geophys. J. Roy. Astron. Soc.* **57,** 479–506, doi: 10.1111/j.1365-246x.1979.tb04790.x.

Sambridge, M., J. Braun, and H. McQueen (1995). Geophysical parametrization and interpolation of irregular data using natural neighbours, *Geophys. J. Int.* **122,** 837–857, doi: 10.1111/j.1365-246x.1995.tb06841.x.

Schaeffer, A., and S. Lebedev (2013). Global shear speed structure of the upper mantle and transition zone, *Geophys. J. Int.* **194,** no. 1, 417–449.

Simons, F. J., I. Loris, G. Nolet, I. C. Daubechies, S. Voronin, J. Judd, P. Vetter, J. Charléty, and C. Vonesch (2011). Solving or resolving global tomographic models with spherical wavelets, and the scale and sparsity of seismic heterogeneity, *Geophys. J. Int.* **187,** 969–988, doi: 10.1111/j.1365-246x.2011.05190.x.

Trampert, J., A. Fichtner, and J. Ritsema (2012). Resolution tests revisited: The power of random numbers, *Geophys. J. Int.* **192,** no. 2, 676–680.

Um, J., and C. Thurber (1987). A fast algorithm for two-point seismic ray tracing, *Bull. Seismol. Soc. Am.* **77,** no. 3, 972–986.

Wessel, P., and W. H. Smith (1998). New, improved version of generic mapping tools released, *Eos Trans. AGU* **79,** no. 47, 579–579.

Zhang, X., A. Curtis, and E. Galetti (2018). 3-D Monte Carlo surface wave tomography, *Geophys. J.*, 1644–1658, doi: 10.1093/gji/ggy362.

Zigone, D., Y. Ben-Zion, M. Campillo, and P. Roux (2015). Seismic tomography of the Southern California plate boundary region from noise-based Rayleigh and Love waves, *Pure Appl. Geophys.* **172,** no. 5, 1007–1032.

**Manuscript received 10 June 2019**
**Published online 30 October 2019**