Transdimensional Markov Chain Monte Carlo Methods

Jesse Kolb, Vedran Lekić (Univ. of MD)

Supervisor: Kris Innanen

UNIVERSITY OF CALGARY

CREWES

Motivation for Different Inversion Technique

- Inversion techniques typically provide a single best-fit model while the fit of other models may be only slightly worse.
- Inversion algorithms often get stuck in local minima.
- Non-uniqueness is often uncharacterized.
- Dimensionality (e.g. number of layers) may be unknown prior to inversion.
- Uncertainty analysis is desired on individual features within a result.

Bayes' Theorem

Incorporates prior knowledge and data to obtain posterior probability density function (PDF):

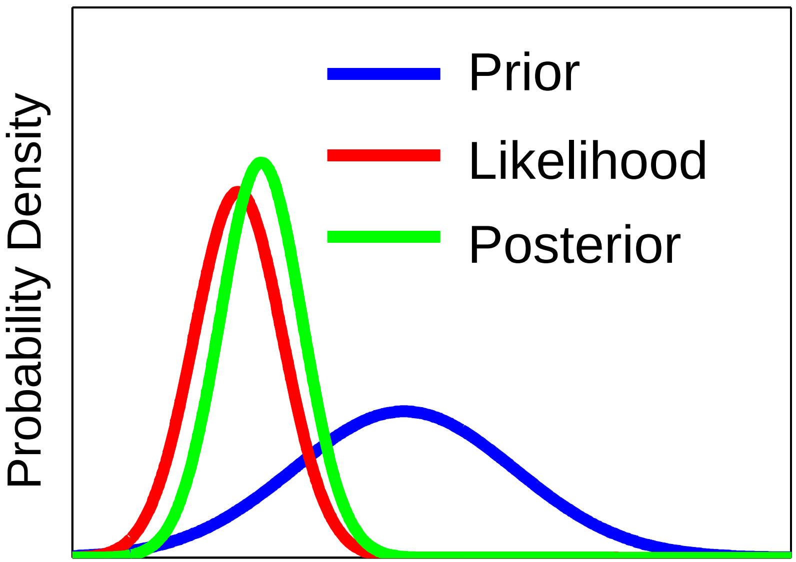
Likelihood PDF

Prior PDF

$$p(\boldsymbol{m}|\boldsymbol{d}) = \frac{p(\boldsymbol{d}|\boldsymbol{m})p(\boldsymbol{m})}{p(\boldsymbol{d})}$$

Posterior PDF

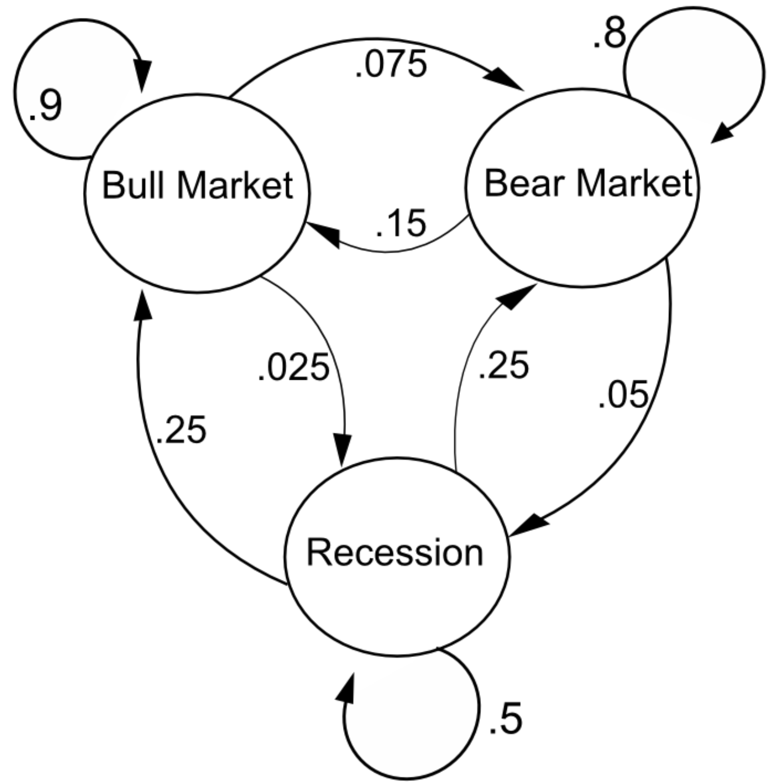
Evidence



What *is* Markov Chain Monte Carlo?

Markov chain: a process in which the next state only depends on the current state

Monte Carlo: Using random numbers to estimate properties of a solution



Example of 3-state Markov chain. Courtesy of Gareth Jones.

Metropolis-Hastings Algorithm

Tests new set of model parameters, $\boldsymbol{m}'$, in order to sample the posterior probability distribution, $p(\boldsymbol{m}|\boldsymbol{d})$.

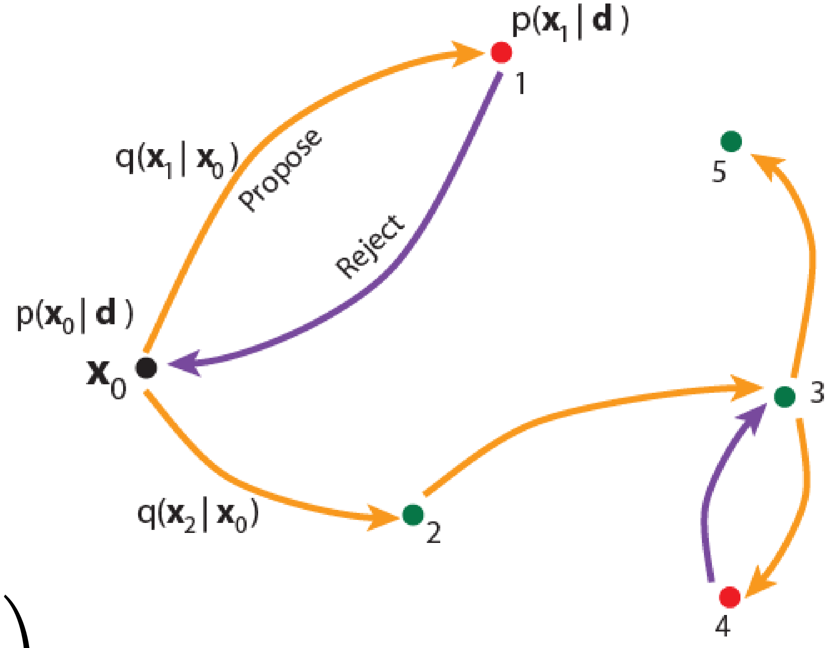
Proposal distribution for $\boldsymbol{m}'$

given current model, $\boldsymbol{m}$:

$$q(\boldsymbol{m}'|\boldsymbol{m})$$

Acceptance probability:

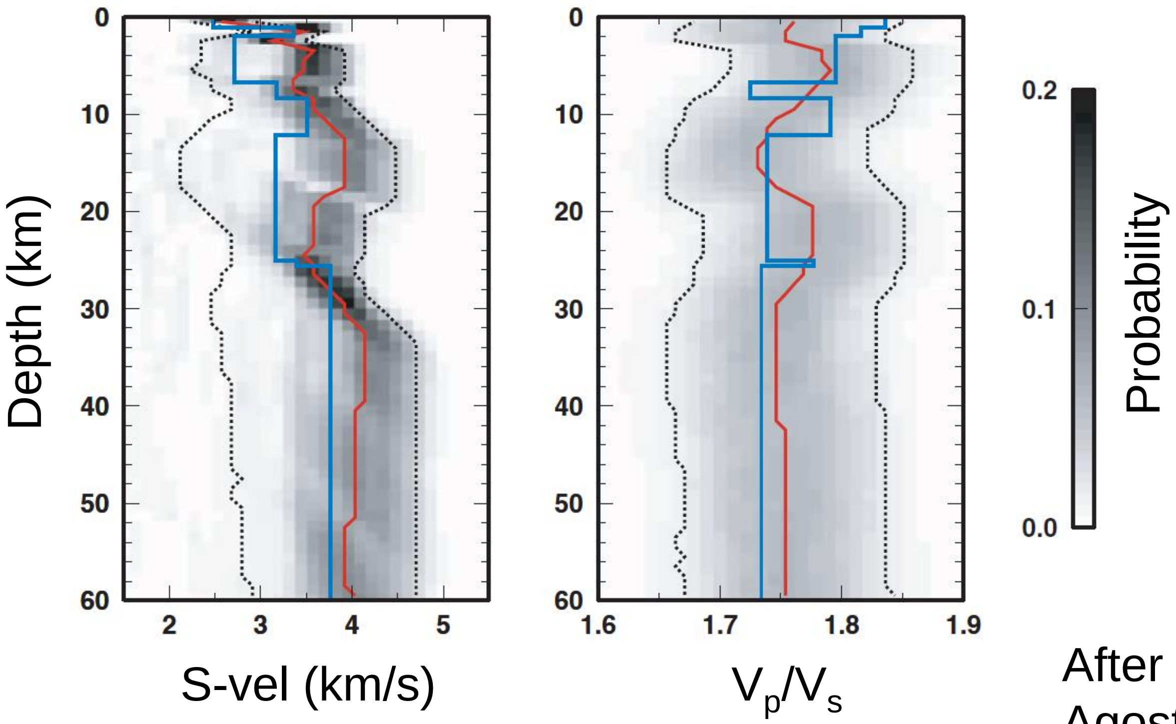
$$\alpha = min\left(1, \frac{p(\boldsymbol{m}')p(\boldsymbol{d}|\boldsymbol{m}')q(\boldsymbol{m}|\boldsymbol{m}')}{p(\boldsymbol{m})p(\boldsymbol{d}|\boldsymbol{m})q(\boldsymbol{m}'|\boldsymbol{m})}\right)$$



Example of progression of states. After Sambridge (2010).

MCMC Results

After Markov chain has converged, models are binned and the ratio of models in a given bin is proportional to posterior probability in that bin



After Piana Agostinetti and Malinverno (2010)

Reversible-Jump (Transdimensional) MCMC

Allows for jumps between parameter spaces of different dimensions

Acceptance probability changes from

$$\alpha = min\left(1, \frac{p(\boldsymbol{m'})p(\boldsymbol{d}|\boldsymbol{m'})q(\boldsymbol{m}|\boldsymbol{m'})}{p(\boldsymbol{m})p(\boldsymbol{d}|\boldsymbol{m})q(\boldsymbol{m'}|\boldsymbol{m})}\right)$$

To

$$\alpha = min\left(1, \frac{p(\boldsymbol{m'})p(\boldsymbol{d}|\boldsymbol{m'})q(\boldsymbol{m}|\boldsymbol{m'})}{p(\boldsymbol{m})p(\boldsymbol{d}|\boldsymbol{m})q(\boldsymbol{m'}|\boldsymbol{m})}|\mathbf{J}|\right)$$

Jumps between parameter spaces are commonly done using Birth-Death MCMC

RJMCMC for Receiver Function Deconvolution

- Receiver functions use P to S or S to P conversions from teleseismic waves to infer velocity structure
- Assumed relation between Parent (P) and Daughter (D) waveforms:

$$D = P * G + \varepsilon$$

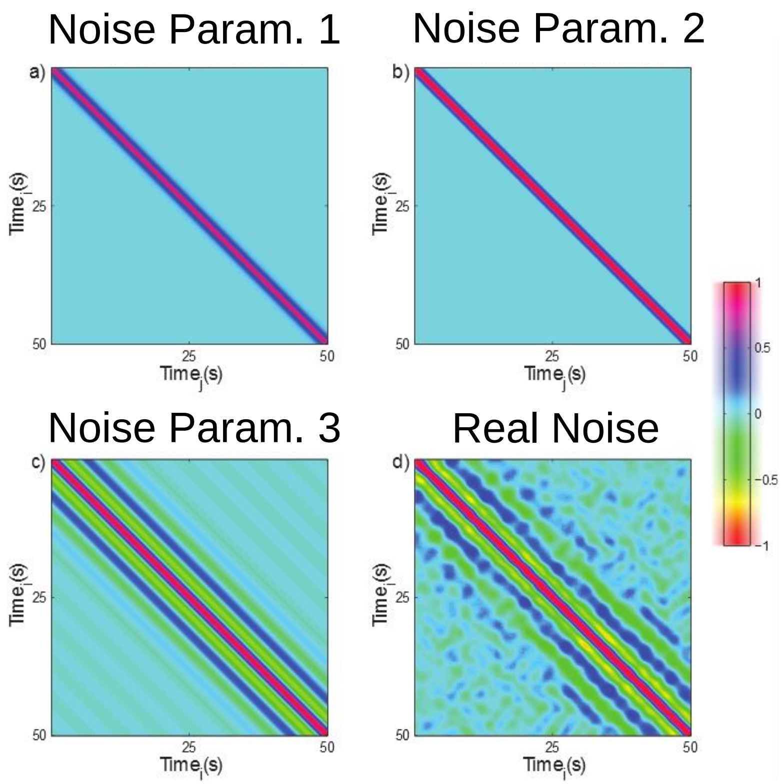
- Receiver function, G, is parameterized as an unknown number of Gaussians at unknown lag times with unknown widths and amplitudes.
- Likelihood of observed daughter waveform given G:

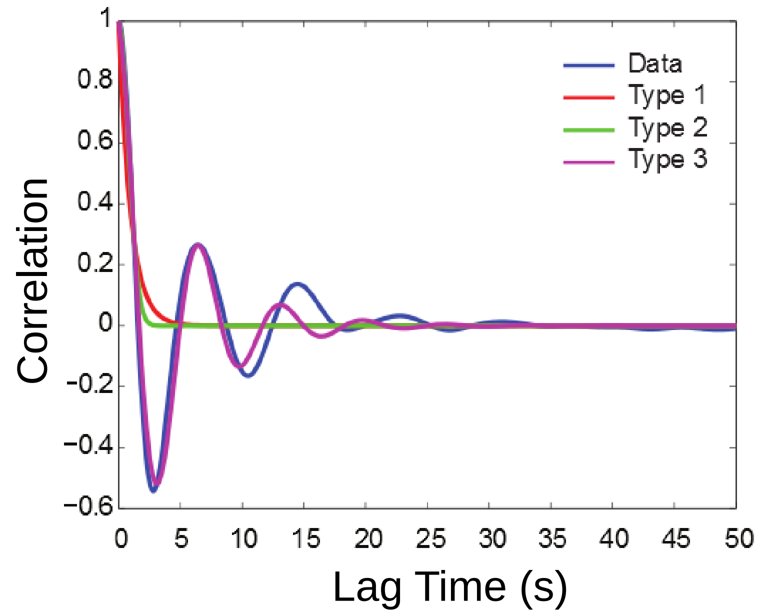
$$p(D|G) = \frac{1}{\sqrt{(2\pi)^n |C_D|}} e^{(P*G-D)^T {C_D}^{-1} (P*G-D)}$$

Noise Parameterization

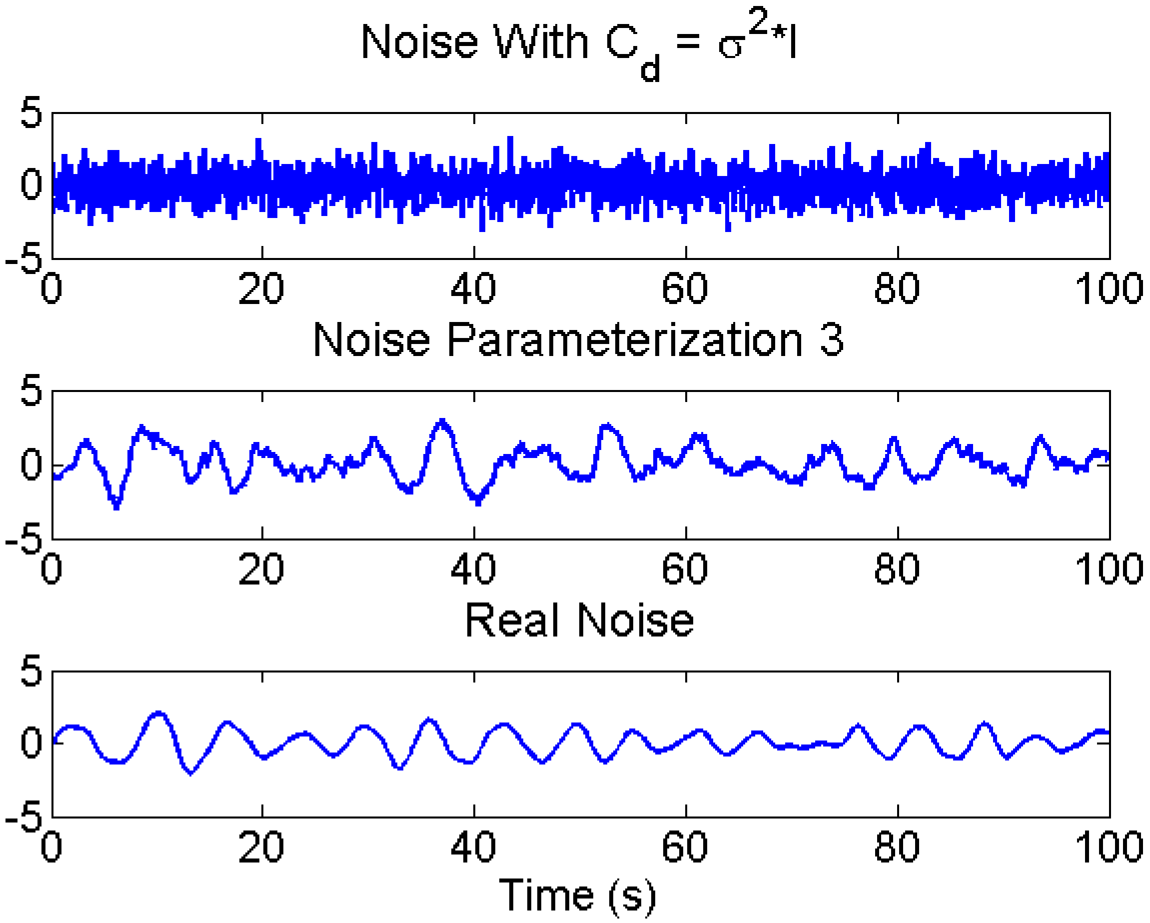
We parameterize the noise covariance matrix with hyperparameters σ, λ, and ω_0:

$$C_{D_{ij}} = \sigma^2 e^{-\lambda|t_j - t_i|} \cos(\omega_0 \lambda |t_j - t_i|)$$





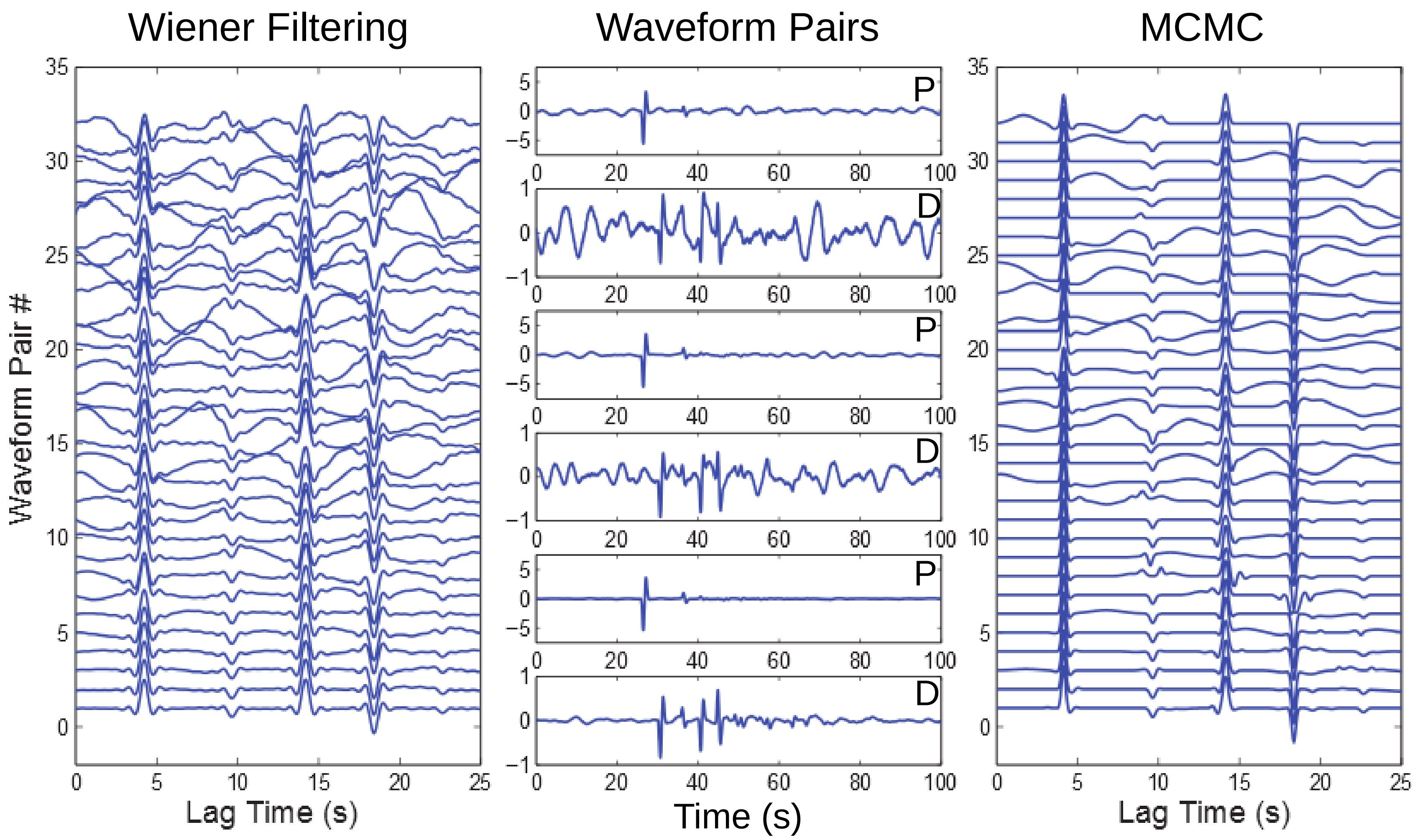
Why Correlation Matters



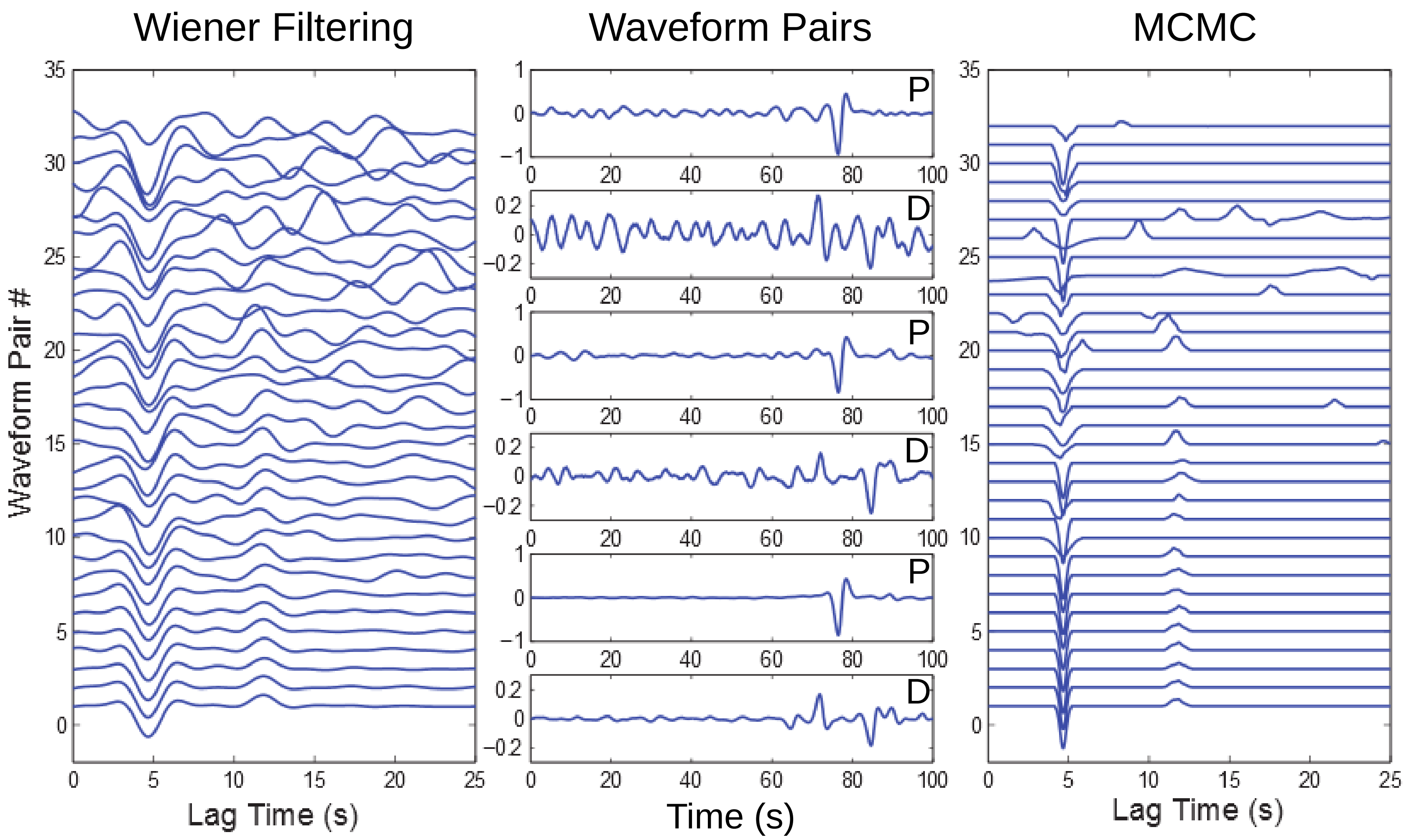
Process

- Start with initial model, G, of no Gaussians.
- Convolve P with G (forward model) and use to calculate likelihood.
- Propose new model by either “birthing” or “killing” a Gaussian, changing an existing Gaussian’s amplitude, width, or location, or changing a noise hyperparameter.
- Calculate forward model, likelihood, and acceptance probability of new model. If acceptance probability is greater than a random number from 0 to 1, accept the model. Otherwise, reject it.
- Repeat proposal and accept/reject process until convergence and then continue, saving models afterwards.

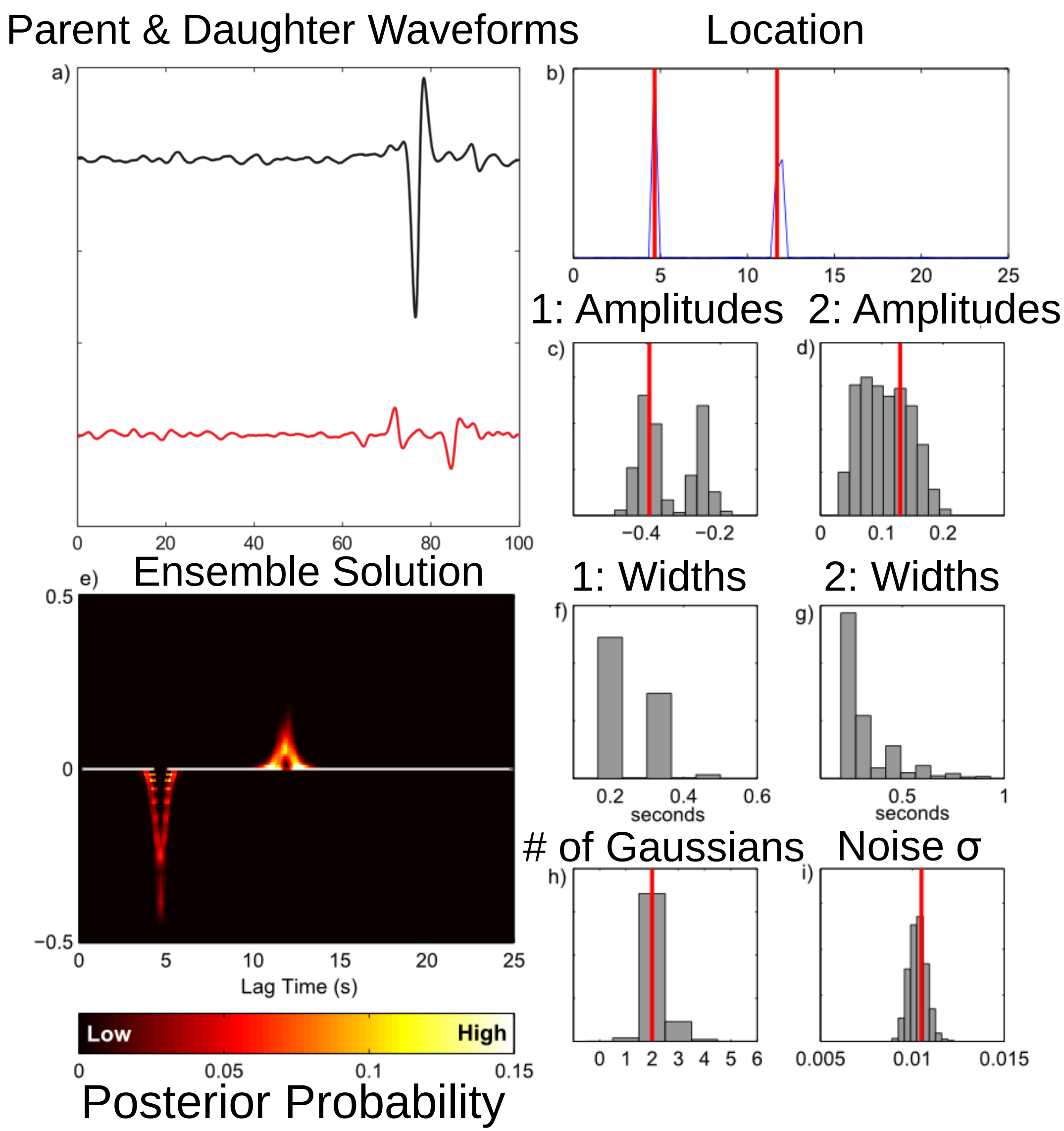
PS MCMC results



SP MCMC results



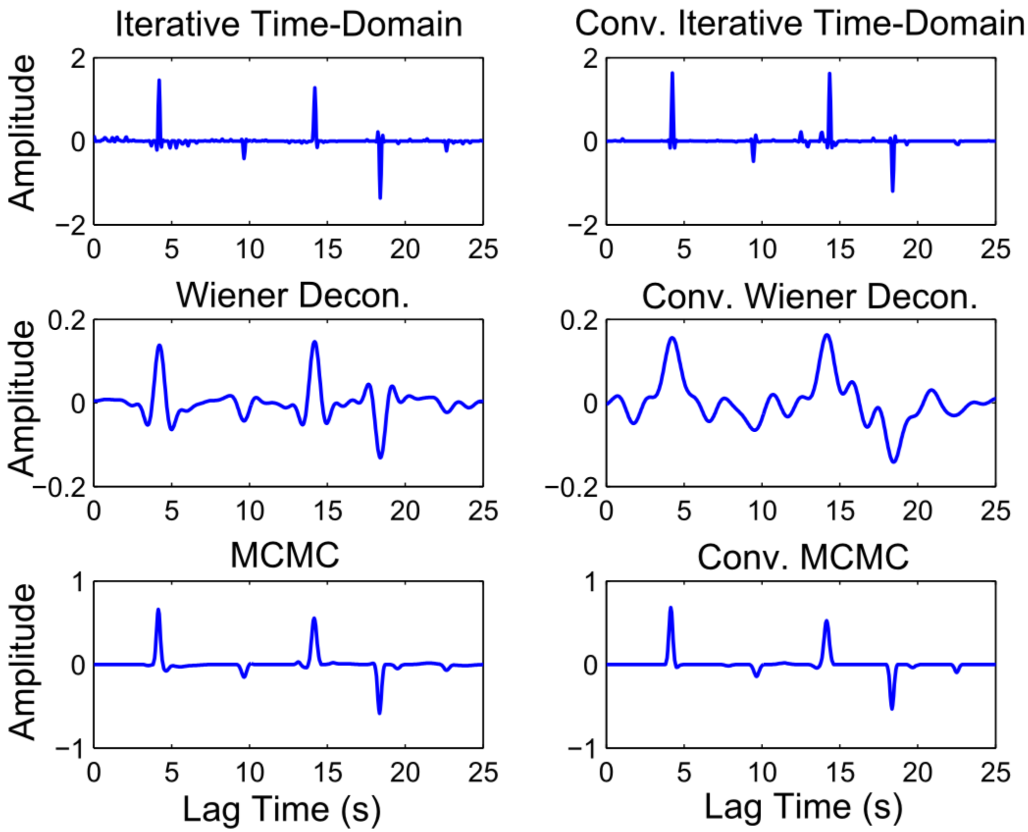
SP Result Analysis



Spectral Hole Test

Solve for G when using STF with spectral holes:

$$(D * STF) = (P * STF) * G + \varepsilon$$



Model Comparisons

Can compare hypotheses *H* which may be parameterizations or forward models:

$$p(\boldsymbol{m}|\boldsymbol{d}, H) = \frac{p(\boldsymbol{d}|\boldsymbol{m}, H)p(\boldsymbol{m}|H)}{p(\boldsymbol{d}|H)}$$

With evidence, $p(\boldsymbol{d}|H)$:

$$p(\boldsymbol{d}|H) = \int p(\boldsymbol{d}|\boldsymbol{m}, H)p(\boldsymbol{m}|H)d\boldsymbol{m}$$

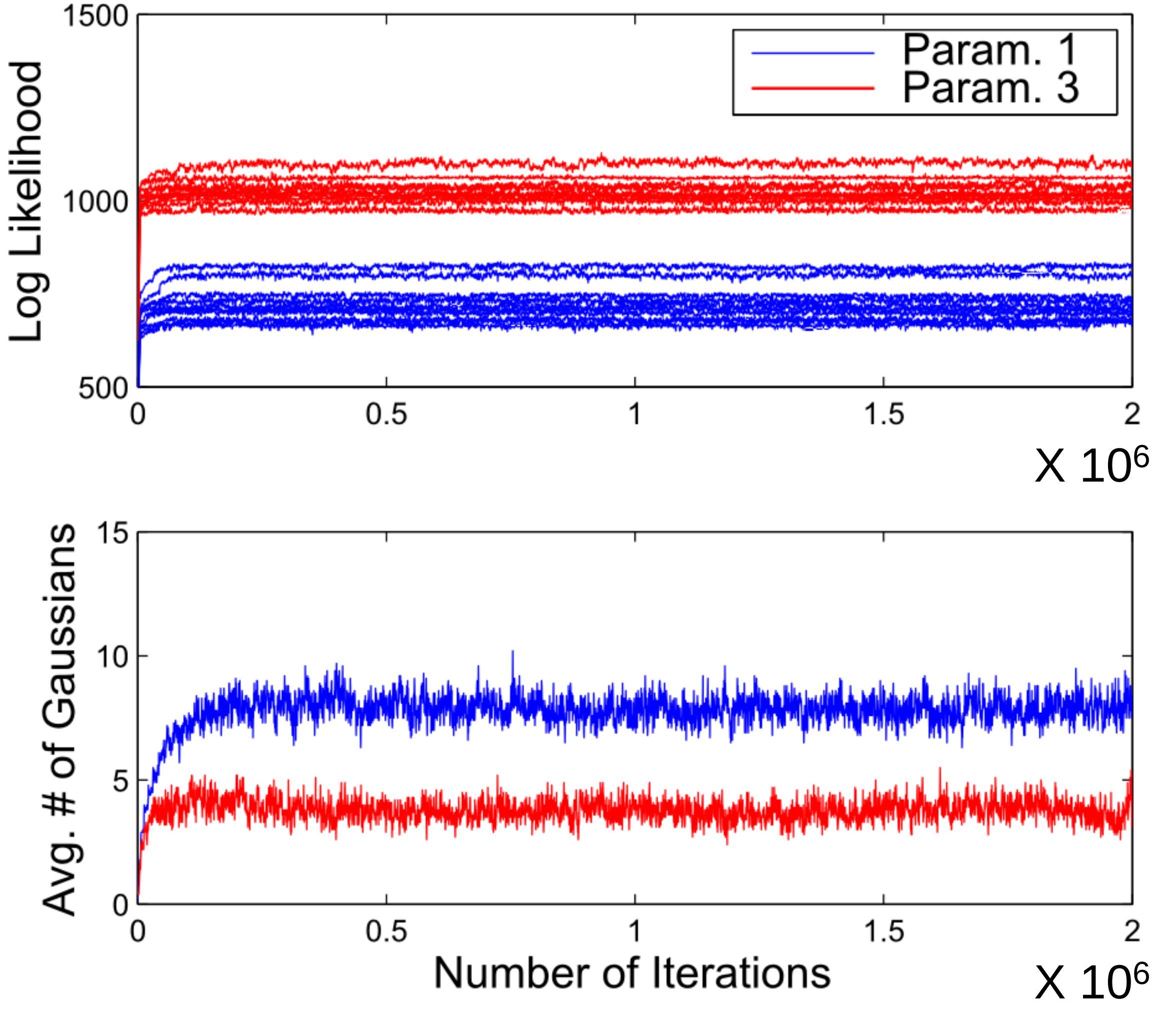
Point approximations using evidence:

$$AIC = -2\log[p(\boldsymbol{d}|\boldsymbol{m}_{mle})] + 2k$$
$$BIC = -2\log[p(\boldsymbol{d}|\boldsymbol{m}_{mle})] + k\log(n)$$
$$DIC = E^{\theta}[-2\log[p(\boldsymbol{d}|\boldsymbol{m})] + k_{eff}$$

Where $\boldsymbol{m}_{mle}$ is maximum likelihood estimate of the model, k is number of parameters, and n is sample size

Noise Parameterization 1 vs. 3

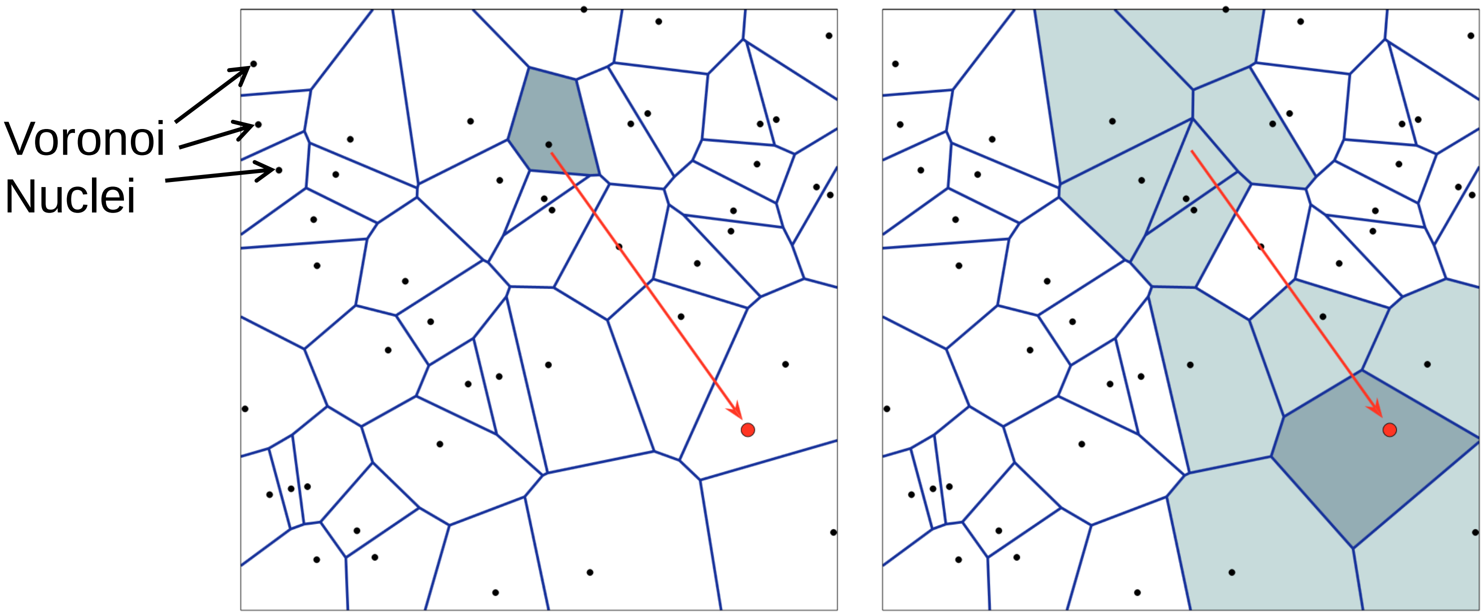
Parameterization 3 has higher likelihoods *and* uses fewer Gaussians.



Tomography

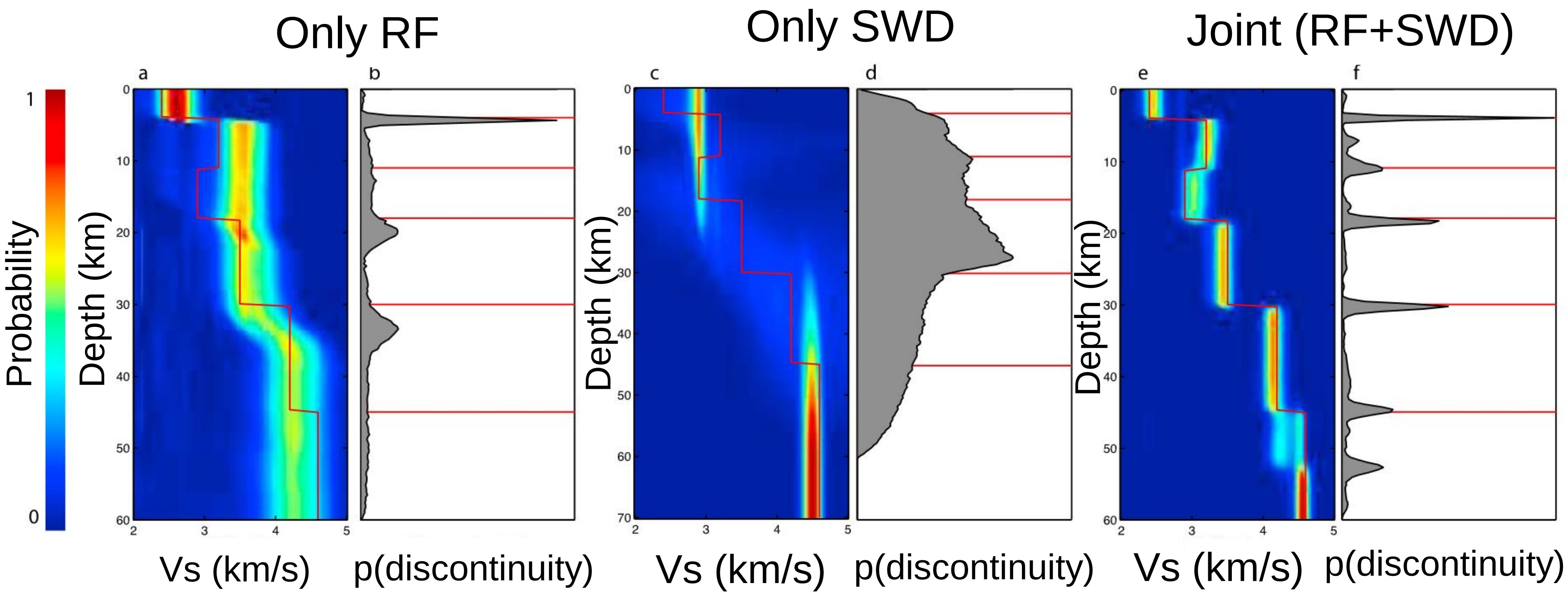
MCMC with parameterization of unknown # of Voronoi nuclei with unknown velocities and locations was proposed (Sambridge et al., 1995).

Allows for cells of varying shapes and sizes, and multiple parameterizations for the same structure.



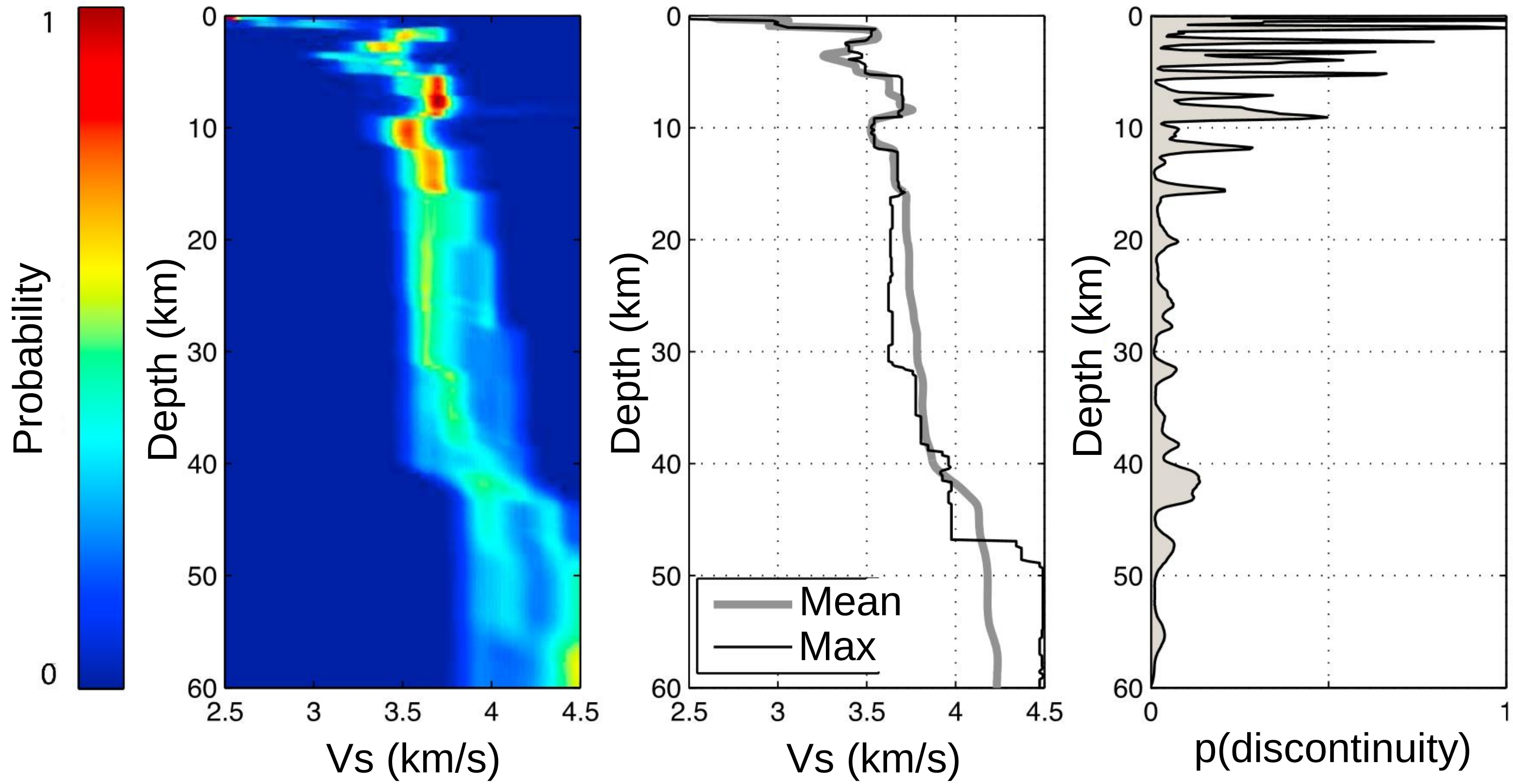
Movement of a Voronoi nucleus to a new location. The velocity structure in each of the teal cells has been changed. After Bodin et al. (2009).

Joint Inversion



After Bodin et al., (2012)

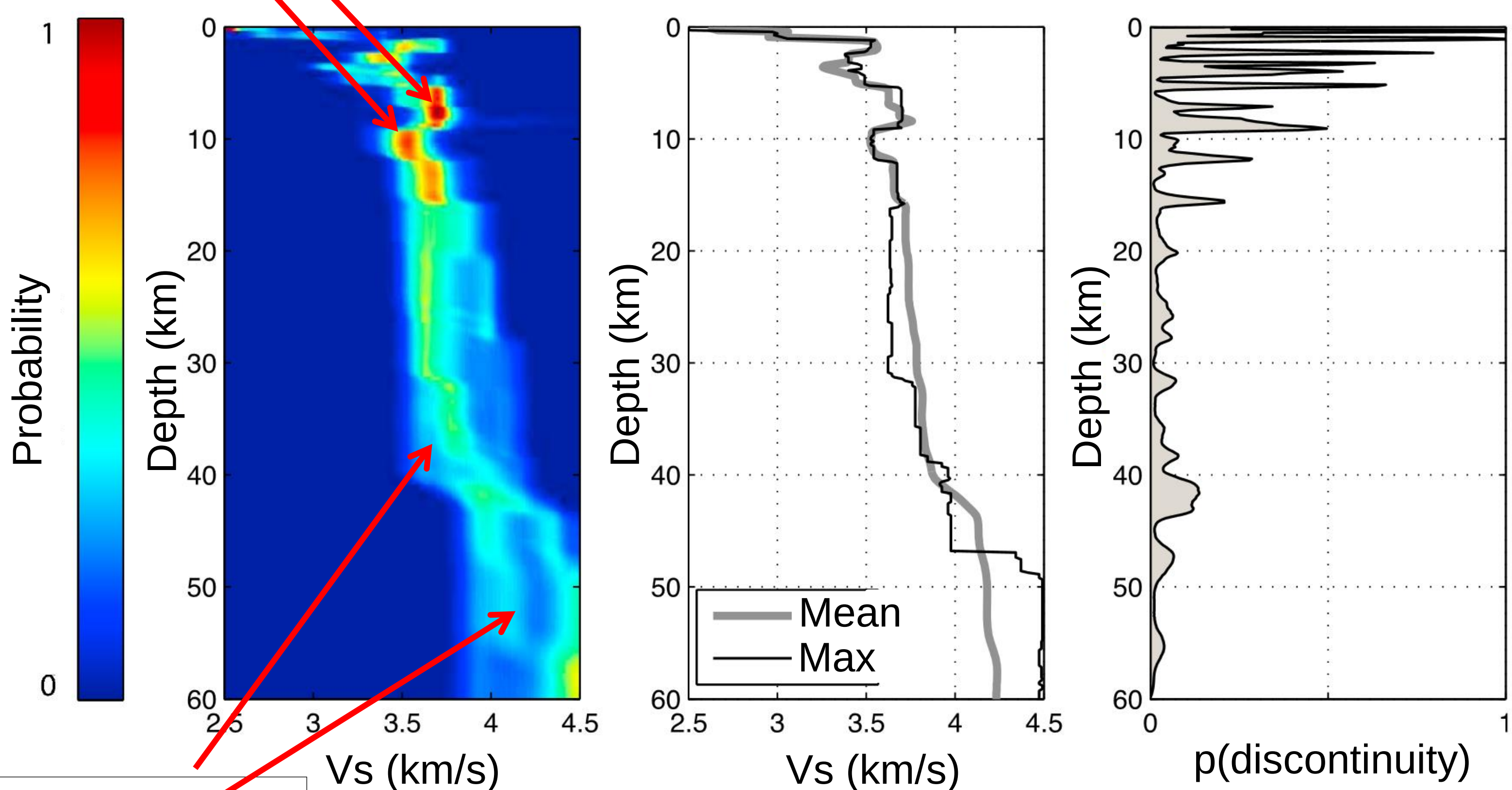
Joint Inversion



After Bodin et al., (2012)

Joint Inversion

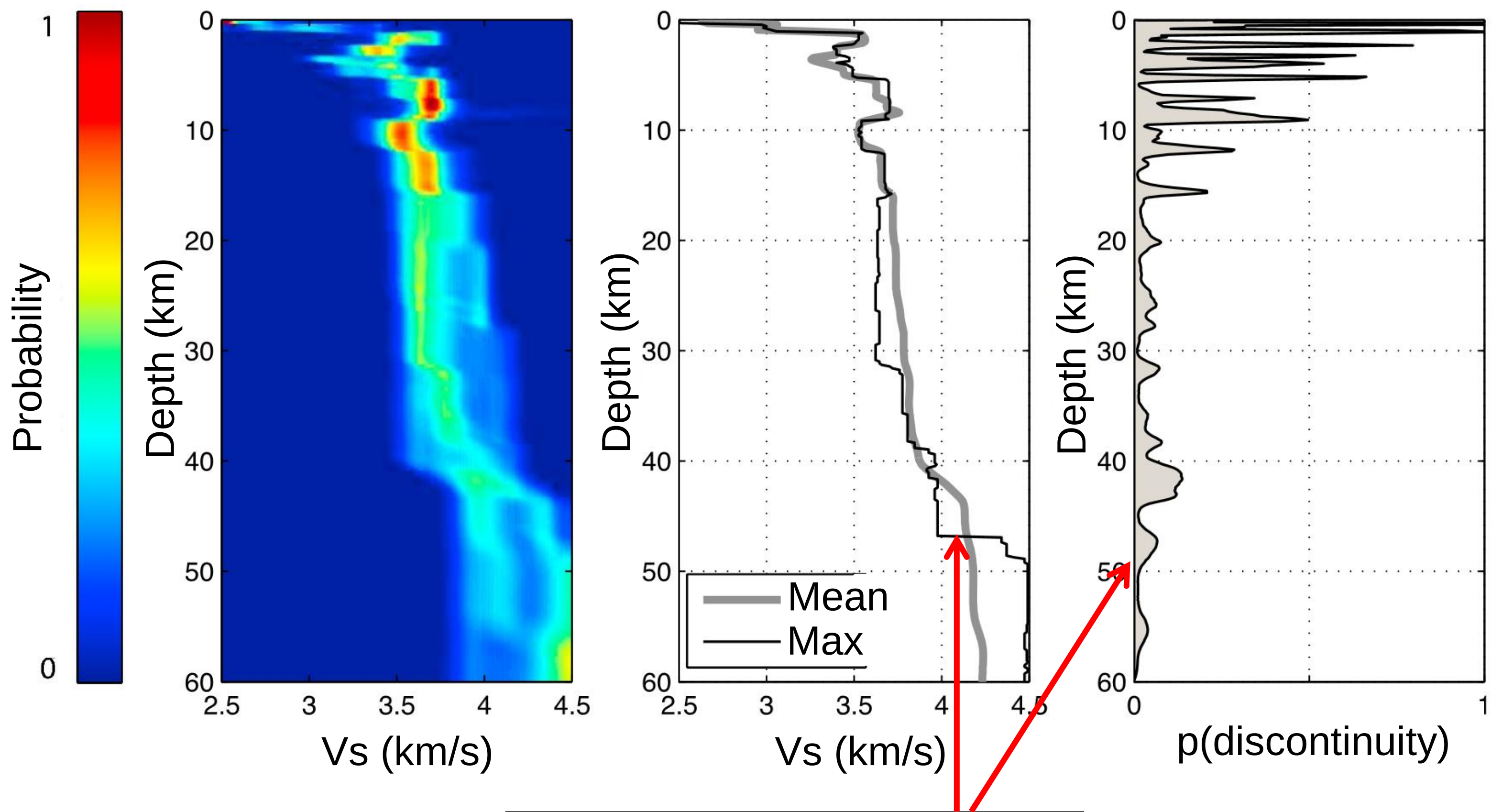
Well-Constrained Features



More uncertainty on these velocities

After Bodin et al., (2012)

Joint Inversion



Sharp increase in max. posterior model not necessarily a discontinuity

After Bodin et al., (2012)

Conclusions

Transdimensional MCMC:

- Is a data-driven approach allowing for dimensionality to be decided by the data;
- Results in an ensemble solution of models that can make uncertainty analysis simpler;
- Provides uncertainties on individual features of solutions;
- Can jump out of local minima and can sample non-unique solutions.

Acknowledgements

Vedran Lekic

Fellow “transdimensionalers”

CREWES

References

- Agostinetti, N. Piana, and A. Malinverno. "Receiver function inversion by trans-dimensional Monte Carlo sampling." *Geophysical Journal International* 181.2 (2010): 858-872.
- Bodin, Thomas, and Malcolm Sambridge. "Seismic tomography with the reversible jump algorithm." *Geophysical Journal International* 178.3 (2009): 1411-1436.
- Bodin, Thomas, et al. "Transdimensional inversion of receiver functions and surface wave dispersion." *Journal of Geophysical Research: Solid Earth (1978–2012)* 117.B2 (2012).
- Jones, Gareth. *Finance Markov Chain Example State Space*. Digital image. *Wikipedia.org*. Wikipedia, n.d. Web. <http://en.wikipedia.org/wiki/File:Finance_Markov_chain_example_state_space.svg>.
- Sambridge, Malcolm (2010). “Uncertainty in Transdimensional Inverse Problems.” PDF File. Lecture Slides.

Happy Valentine’s Day!