A Two-Stage Markov Chain Monte Carlo Method for Seismic Inversion

Susan E. Minkoff, Georgia K. Stuart, and Felipe Pereira

Department of Mathematical Sciences
The University of Texas at Dallas

ICERM Workshop
Recent Advances in Seismic Modeling and Inversion: From Analysis to Applications
Brown University, November 9, 2017

Outline

Goal is to estimate both the parameters describing the subsurface (e.g., layer depths and layer velocities) and to quantify the uncertainty in these estimates.

I. Background and Motivation behind Bayesian Inversion

II. Multilevel Markov chain Monte Carlo Simulation

   A. McMC Process and Bayes' Rule

   B. Operator Upscaling for the Acoustic Wave Equation

III. Numerical Experiments for Layer Depth and Velocity

   A. Flat Layers with Known Layer Positions and Unknown Velocities

   B. Unknown layer Positions and Known Velocities

   C. Unknown layer Positions and Velocities

IV. Conclusions and Future Work

Exploration Seismology

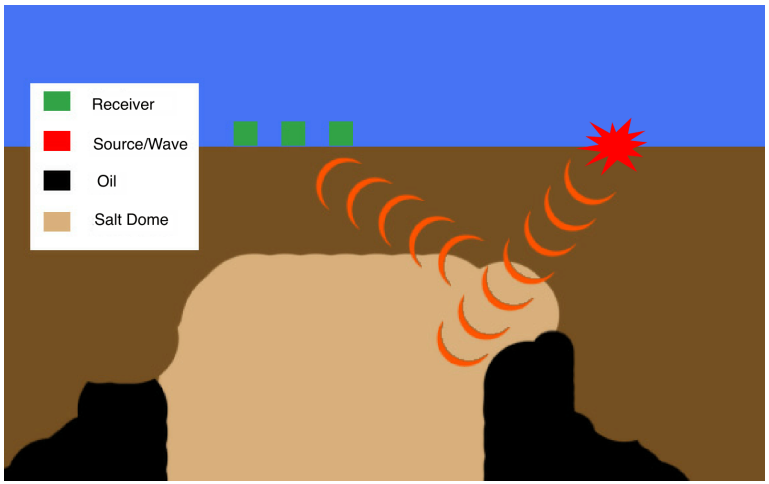


Figure: The reflection seismology process. Waves are generated at the source and reflect off the interfaces between different materials.

- In exploration seismology, seismic waves (acoustic or elastic) can be used to image the subsurface of the earth.
- In a typical seismic experiment:
  1. A source creates a disturbance in the form of a wave.
  2. This wave travels through the earth and reflects off of material property interfaces.
  3. Seismometers on the surface of the earth or in wells record the returning wave.
- This recorded seismic data can be used to image the earth's subsurface.
- In velocity inversion, the result is a map of wavespeed that can be used to determine lithology.

Motivation: Uncertainty Quantification of the Velocity Field

- A deterministic approach to waveform inversion results in a single model of the desired parameter without uncertainty information.
- A Bayesian approach allows us to characterize and quantify uncertainty.
- We focus on inversion for layer depth and layer velocities, but this approach can be extended to other material parameters and even (micro)seismic events.
- Gouveia and Scales[1] applied Bayesian methods to full waveform inversion for the velocity estimation problem. However, they make simplifying assumptions of the posterior distribution being Gaussian and do not use McMC.
- We follow Tarantola[2] and use a Markov chain Monte Carlo process to sample from the posterior distribution of the wavefield. This allows us to avoid assumptions of normality, which means a better characterization of the uncertainty.
- Hong and Sen[3] use a multiscale genetic algorithm McMC mthod with multiple chains at different coarse scales that inform the more expensive fine-grid chain.

[1] Gouveia, W. and Scales, J., 1998, "Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis," *J. Geophy. Res.*, **103**, pp. 2759–2779.

[2] Tarantola, A., 2005, *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM.

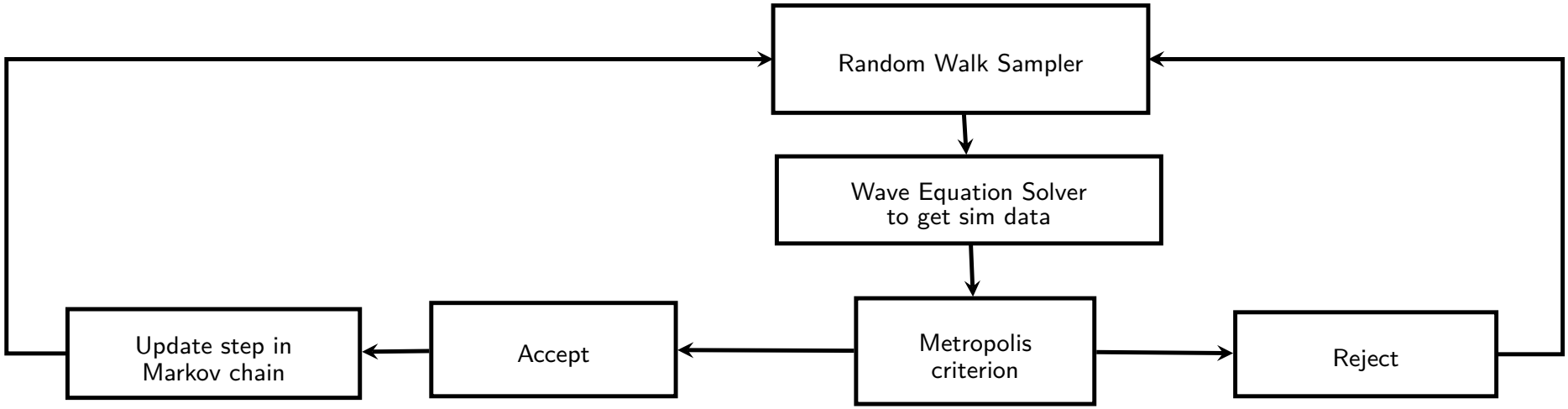
[3] Hong, T. and Sen, M. 2009, "A new McMC algorithm for seismic waveform inversion and corresponding uncertainty analysis," *Geophysical Journal International*, **177**, pp. 14–32.

Upscaling and McMC Velocity Inversion

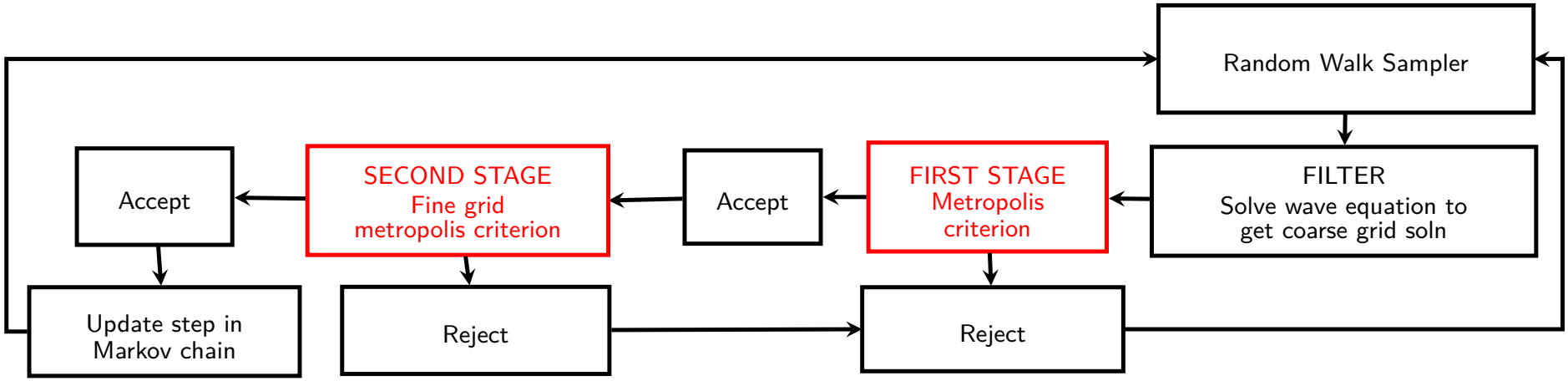
- Problem: McMC requires many samples (thousands) to converge to steady state, and each sample must be run through a forward simulator to see if it is acceptable for the characterization of the posterior distribution.
- Often 90% of samples are rejected!
- Proposed solution: use a coarse grid solution to quickly reject samples, then simulate on the full fine grid if upscaled sample is accepted.
- Multilevel McMC has been used for determination of parameters in fluid flow simulation. See, for example:
  1. Efendiev, Y., Datta-Gupta, A., Ginting, V., Ma, X., and Mallick, B., 2005, "An efficient two-stage Markov chain Monte Carlo method for dynamic data integration," *Water Resources Research*, **41**, W12423.
  2. Akbarabadi, M., Borges, M., Jan, A., Pereira, F., and Piri, M., 2015,"A Bayesian Framework for the Validation of Models for subsurface flows: synthetic experiments," **19**, pp. 1231–1250.
- We are first to use idea for seismic inversion.
  1. Stuart, G., Yang, W., Minkoff, S., and Pereira, F., 2016, "A two-stage Markov chain Monte Carlo method for velocity estimation and uncertainty quantification", *SEG Technical Program Expanded Abstracts*, pp. 3682–3687.

One-Stage vs. Two-Stage McMC

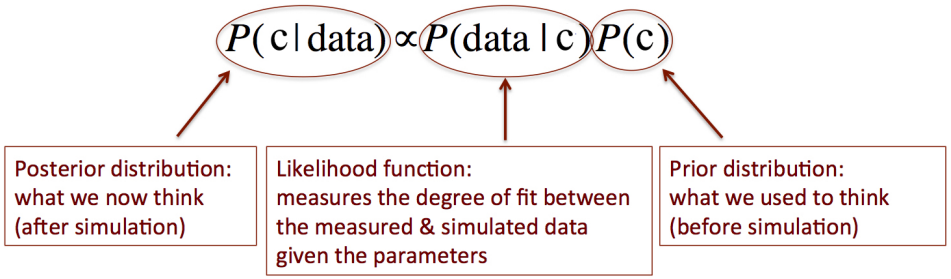
One Stage



Two Stage



Bayes' Rule



We assume the likelihood function has the form:

$$P(d_m|C) = \exp\left(-\frac{\|d_m - d_s\|^2}{2\sigma^2\|d_s\|^2}\right)$$

where σ is the precision parameter which we vary in later experiments and should reflect errors in measurements, modeling and the numerical approximation, d_m is the measured or observed data, and d_s is the simulated data from our new velocity sample.

After obtaining the simulated receiver data, we decide whether to accept or reject the proposed velocity field via the Metropolis Criterion:

Accept C with probability:

$$\rho(C_n, C) = \min\left\{1, \frac{P(C|d_m)}{P(C_n|d_m)}\right\}.$$

Where $P(C|d_m)$ is the posterior, C is the proposed velocity field, and C_n is the last accepted velocity field.

McMC Process for Layer Depth Experiments

We generate stochastic perturbations of the velocity field for the McMC process via a random walk sampler of the form:

$$\theta^{n+1} = \beta\theta^n + \sqrt{1-\beta^2}\hat{\theta} \tag{1}$$

where θ^{n+1} is the new random vector, θ^n is the previous value, β is the tuning parameter, and $\hat{\theta} \sim \mathcal{N}(0,1)$.

New layer depths z at each point are then generated via the formula

$$z_{new} = z_0 + \sigma_{step}\theta^{n+1} \tag{2}$$

and new velocities v at each point are generated from

$$v_{new} = v_0 + \sigma_{step}\theta^{n+1} \tag{3}$$

where z_0 and v_0 are the initial layer depth and initial velocity value respectively, and σ_{step} is the standard deviation of the (Gaussian) prior distribution (in these experiments it is taken to be 25 m for layer depth and 500 m/s for velocity).

Upscaling of the Acoustic Wave Equation

- We use the 2D constant-density acoustic wave equation to solve for pressure, p, given velocity, c.

$$\frac{1}{c^2(x,z)}\frac{\partial^2 p}{\partial t^2} - \Delta p = f$$

- Modeling the propagation of an acoustic wave can be computationally expensive.

- Operator upscaling decomposes the solution into two parts: the fine grid solution on independent subdomains solved in parallel and a small coarse grid problem over the whole domain solved in serial.

- The upscaling we use is embarrassingly parallel with near perfect speedup due to a simplifying assumption: homogeneous Neumann boundary conditions on each coarse block[4].

[1] Vdovina, T., Minkoff, S. E., and Korostyshevskaya, O., 2005, "Operator Upscaling for the Acoustic Wave Equation", *Multiscale Modeling & Simulation*, **4**, pp. 1305-1338.

Upscaling of the Acoustic Wave Equation

We rewrite the second-order acoustic wave equation as a first-order system in space by introducing acceleration, $\vec{v}$. The acoustic wave equation becomes

$$\vec{v} = -\nabla p,$$
$$\frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = -\nabla \cdot \vec{v} + f.$$

Step 1: solve fine grid problems over each coarse block

$$\langle \rho(\vec{v}^c + \delta\vec{v}), \delta\vec{u}\rangle = \langle p, \nabla \cdot \delta\vec{u}\rangle,$$
$$\left\langle \frac{1}{\rho c^2}\frac{\partial^2 p}{\partial t^2}, w \right\rangle = -\langle \nabla \cdot (\vec{v}^c + \delta\vec{v}), w\rangle + \langle f, w\rangle .$$

Step 2: solve the coarse grid problem over the whole domain

$$\langle \rho(\vec{v}^c + \delta\vec{v}), \vec{u}^c\rangle = \langle p, \nabla \cdot \vec{u}^c\rangle$$

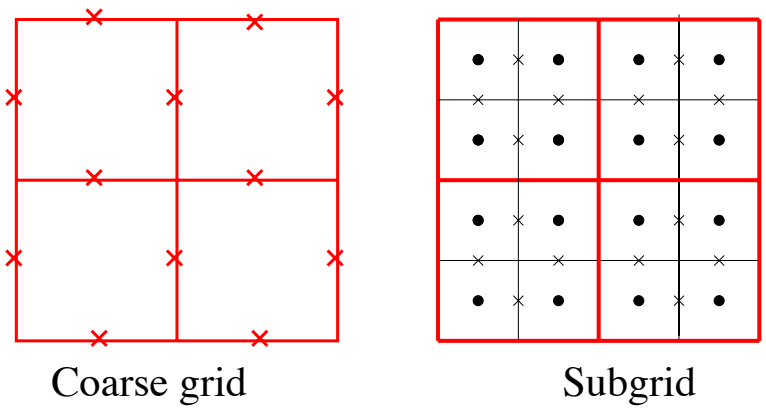


Figure: Diagram showing the location of pressures (dots) and accelerations (x's) on the fine and upscaled grid.

Numerical Experiments

All our numerical experiments have the following setup:

- Computational region is 960m x 960m with 1m spacing between grid points.
- For the two stage trials, each coarse block contains 16 fine blocks in each direction (x and z), for a total of 60 x 60 coarse blocks within the computational region.
- Simulation data is recorded on 480 receivers located at a depth of 90m.
- Each simulation took 10,000 time steps for a total time of 0.5 s.

Our numerical experiments perturb the velocity values and layer depths using perturbations drawn from a Gaussian distribution. For example,

$$c(\vec{x}) = M(z) + C(\vec{x})$$

where $c(\vec{x})$ is the modeled velocity field at location $\vec{x} = (x, z)$, $M(z)$ is the deterministic position of the velocity layer interfaces, and $C(\vec{x})$ is a stochastic perturbation of the model.

First Experiment: Flat Layers with Fixed Layer Positions and Varying Velocity

- Our first numerical experiment perturbs the velocity values with fixed layer positions.
- We invert for velocity values of the 4 middle layers.
- We compare one-stage and one-stage McMC runs. Tuning parameter is $\beta = 0.9$. Fine grid precision parameter $\sigma = 0.01$. For two-stage McMC $\sigma_c = 0.025$.

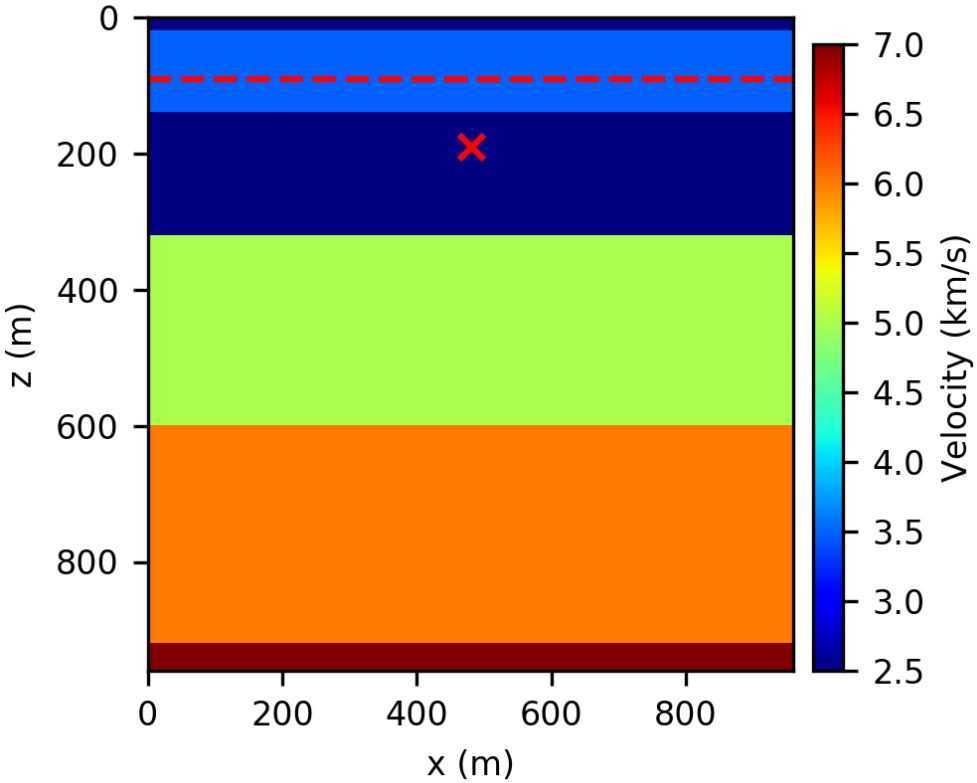


Figure: True velocity field for Experiment 1. Dashed red line shows the receiver positions. Red x gives the source position.

First Experiment: Filter vs. Fine Grid Residuals

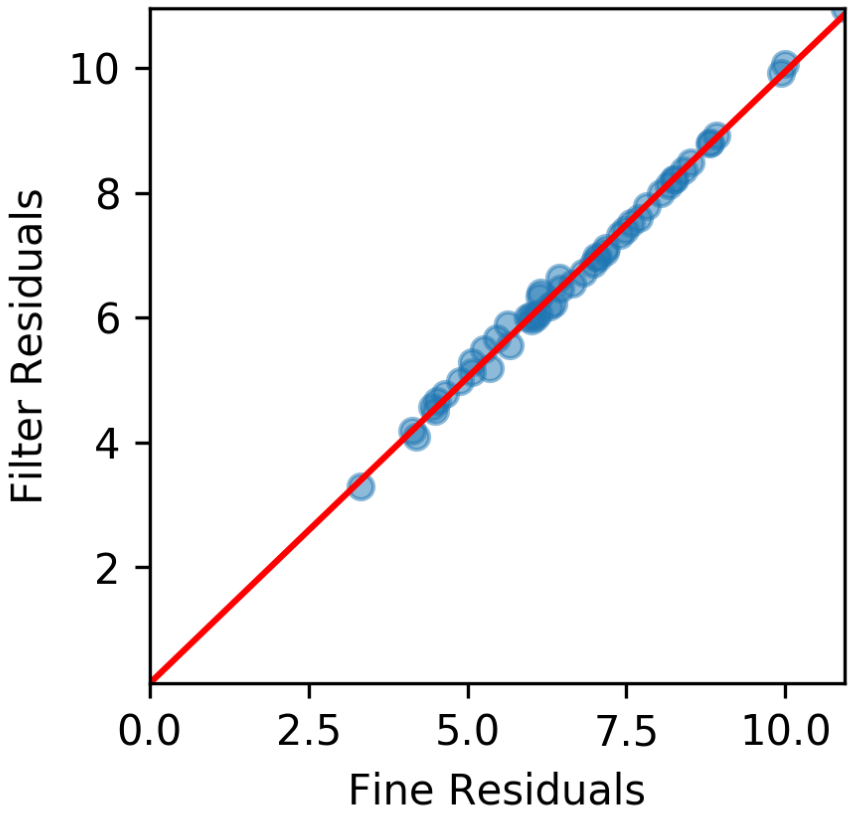
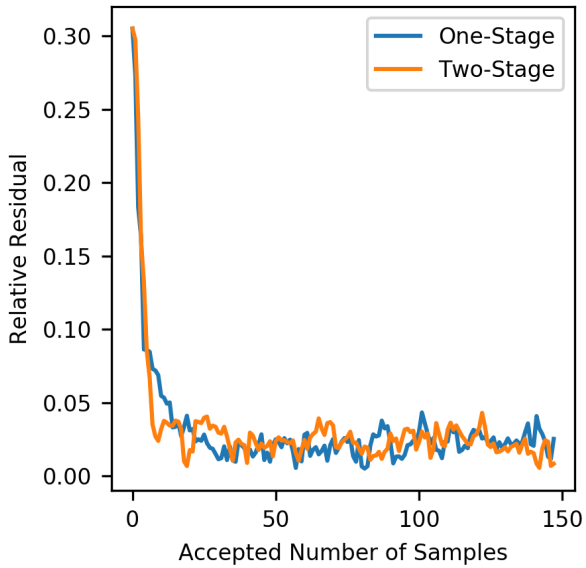


Figure: Comparison of fine grid residuals, $\|d_{\text{ref}} - d_{\text{fine}}\|$ vs. the upscaled residuals, $\|d_{\text{ref}} - d_{\text{filter}}\|$, for Experiment 1 (blue dots). The red line is the perfect correlation line.

- In order to use the upscaled wave solver as a filter for the McMC process, we must show that the residuals on the fine grid and residuals on the coarse grid are correlated.
- Each point is an ordered pair of fine residuals and filter residuals created from the same velocity perturbation.
- Here we see an excellent agreement between the filter and the fine grid residuals. This suggests the upscaling is a valid filter for this problem.

First Experiment: McMC Results

- For both one-stage and two-stage McMC we see a quick decrease in residuals. Both exited the burn-in period after about 25 initial acceptances. Both converged to a residual of about 0.01.
- Our acceptance ratio on the two-stage was nearly 10x the ratio on one-stage.



Run	Samples Tried	Filter Accepted	Fine Accepted	Acceptance Ratio
One-stage	4615	N/A	148	0.032
Two-stage	5202	603	148	0.245

First Experiment: McMC Results

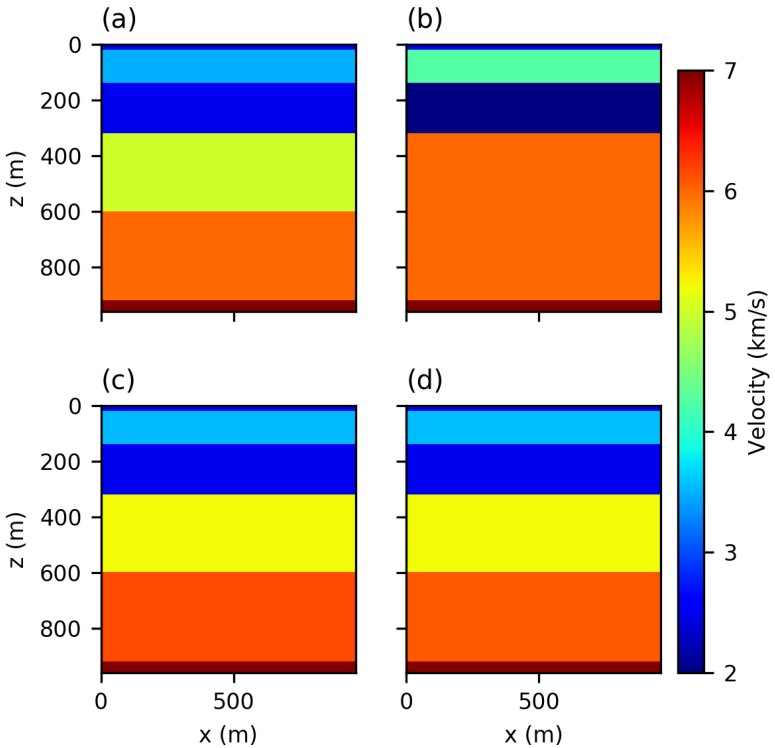


Figure: Comparison of velocity fields for Experiment 1. (a) True velocity field, (b) starting velocity field for both the one-stage and two-stage experiments, (c) average velocity field for the one-stage experiment, (d) average velocity field for the two-stage experiment.

First Experiment: McMC Results

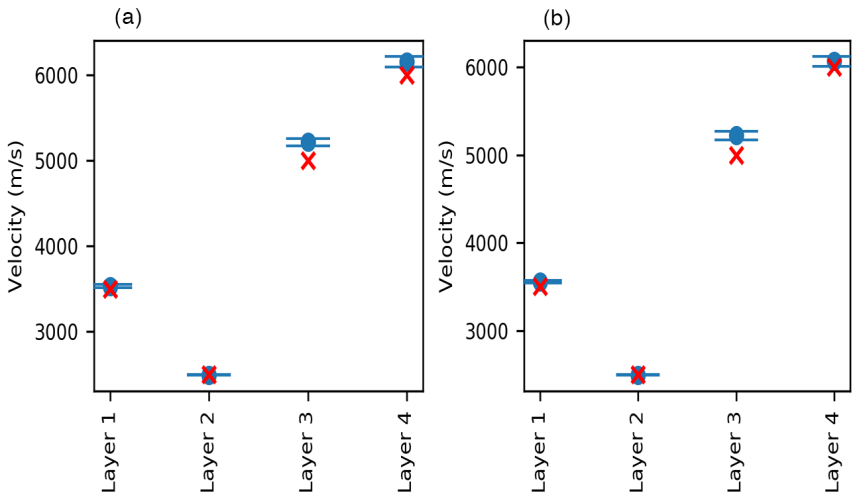
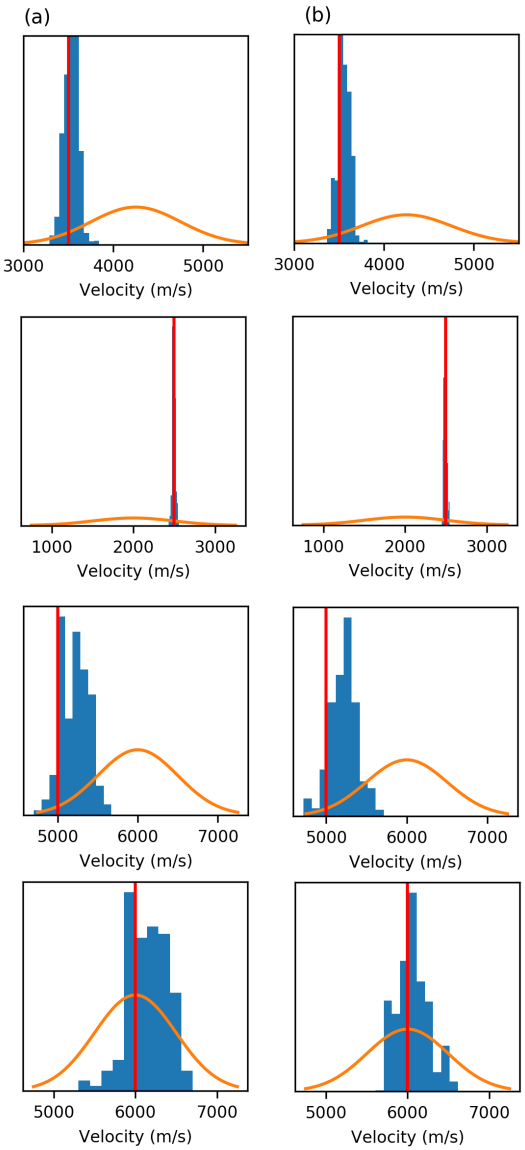


Figure: The 99 % confidence intervals for each layer's velocity for Experiment 1: (a) one-stage experiment, (b) two-stage experiment. Blue dot shows the average of the posterior. The red x shows the true velocity value.

Note: Confidence intervals for the posterior means do not capture true values for all layers, but the two-stage run replicates the confidence intervals produced by one-stage run.

First Experiment: McMC Results

Histograms showing the constructed posterior distribution for (a) one-stage McMC simulation and (b) two-stage McMC simulation in Experiment 1. Four pictures in each column correspond to the velocity of each of the four layers estimated, from the shallowest layer (top) to the deepest layer (bottom). Red line shows the true velocity value. Orange curve shows the prior distribution. Blue histogram shows the posterior distribution recovered from the McMC process.



Experiment 2: Varying Position of Layers Assuming Known Velocities

1. For the second experiment we varied the velocity layer depth, with no requirement of flat layers.
2. We allowed for 9 points per layer to be varied on each of 3 layers independently.
3. Only ran two-stage McMC simulation.
4. tuning parameter ($\beta = 0.95$) in the random walk sampler and the precision parameters for the fine and coarse grid likelihoods are ($\sigma = 0.01$ and $\sigma_c = 0.025$).

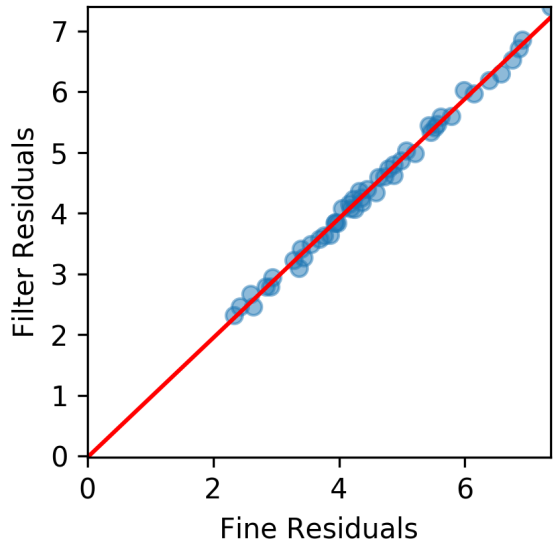


Figure: Comparison of fine grid residuals, $\|d_{\text{ref}} - d_{\text{fine}}\|$, vs. the upscaled residuals, $\|d_{\text{ref}} - d_{\text{filter}}\|$.

Experiment 2: Two-Stage McMC Results (Variable Layer Positions)

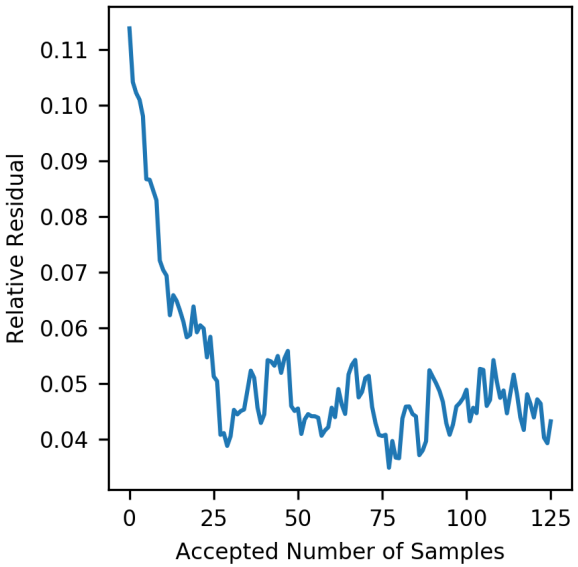


Figure: Relative residuals in the likelihood function for Experiment 2. Burn-in includes approximately first 25 velocity field acceptances. Residual converges to about 0.05.

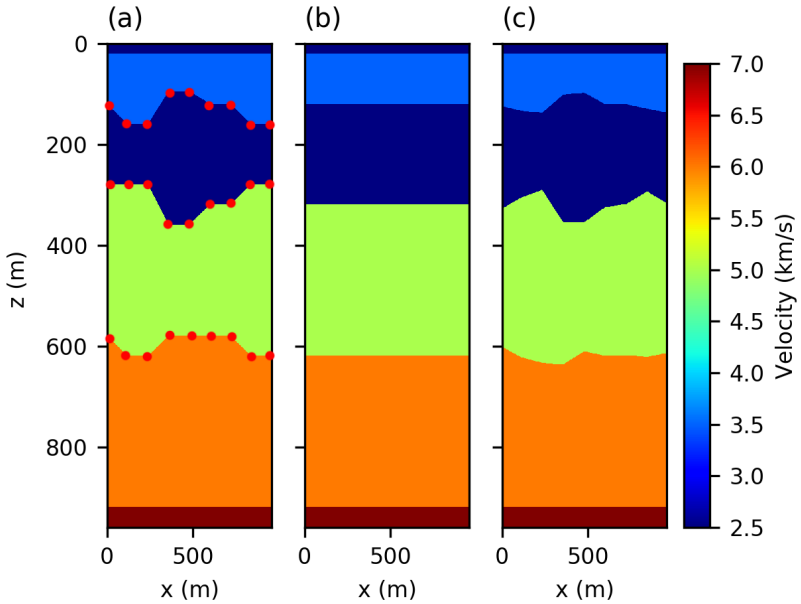
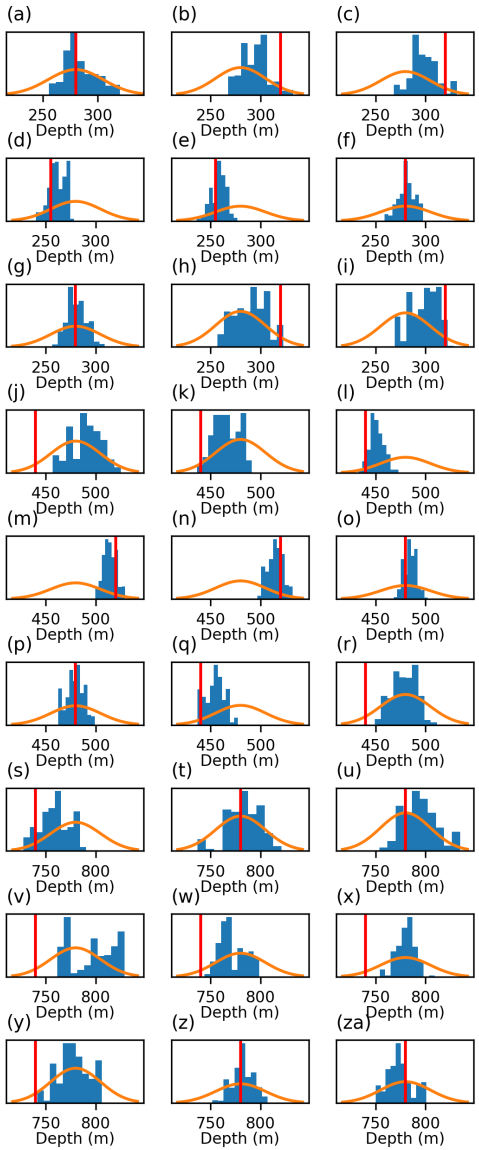


Figure: Comparison of velocity fields for Experiment 2. (a) True velocity field, (b) Initial velocity field, (c) Average velocity field at the end of the two-stage McMC simulation. Red dots in (a) show placement of the 9 nodes on each estimated layer.

Experiment 2: McMC Results (Variable Layer Positions)

Histograms showing the posterior distributions recovered from the two-stage McMC process for Experiment 2 (blue). Orange curves are prior distributions. True layer depth at each node is shown in red. Plots (a) through (i) are the node positions for the first variable layer going from left to right in the domain; (j) through (r) correspond to the second layer, and (s) through (za) correspond to the third layer.



Experiment 3: Flat Layered Experiment with Variable Layer Positions and Velocities

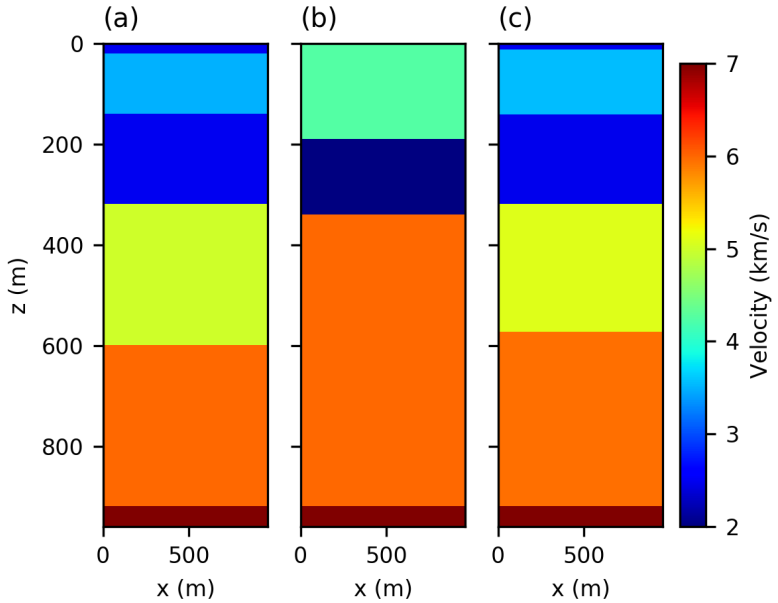


Figure: (a) True velocity field, (b) Initial velocity field, (c) Average velocity field after two-stage McMC simulation. Acceptance ratio is 0.07 with only 43 velocity fields accepted on fine grid.

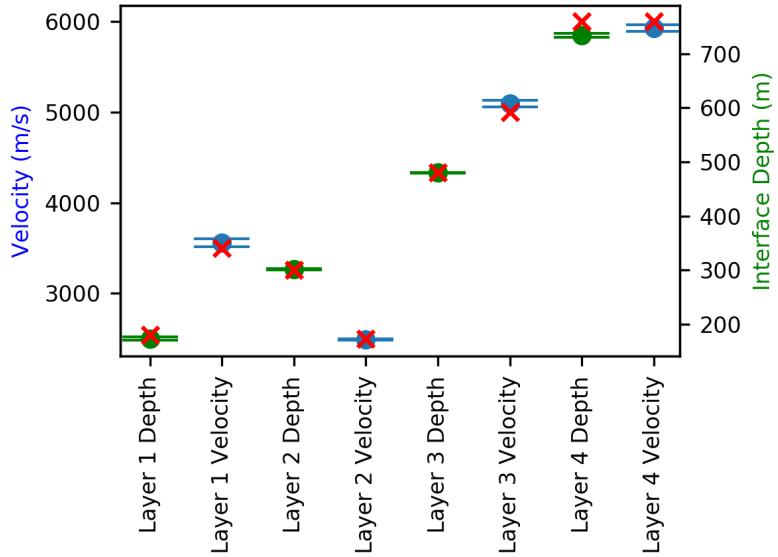


Figure: The 99% confidence intervals for Experiment 3. Green confidence intervals are layer position (right axis). Blue confidence intervals are velocity (left axis). Red crosses show the true value of the depth or velocity.

Experiment 3: Two-Stage McMC Results

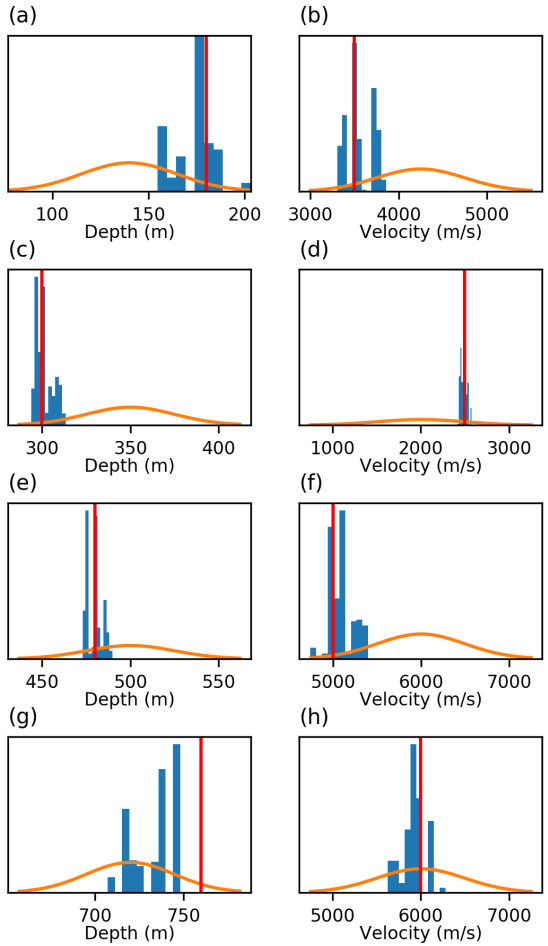


Figure: Histograms (blue) showing the posterior distribution for Experiment 3. Red lines are the true position or velocity. Orange curves show the prior distribution. First column corresponds to layer depths. Second column corresponds to layer velocities. Each row of pictures depicts one layer from shallowest to deepest.

Conclusions and Future Work

1. Two-stage McMC potentially allows us to determine both layer depths and velocities while quantifying uncertainty in Bayesian seismic inversion.
2. One-stage and two-stage McMC give very similar inversion results.
3. The acceptance rate for the two-stage McMC is considerably higher than for single-stage McMC. We save time by quickly rejecting unacceptable samples.
4. In future we plan to apply the method to more realistic problems and data.

Acknowledgements: This research was supported by the National Science Foundation's Enriched Doctoral Training Program, DMS grant #1514808. We thank Rob Meek and Matt McChesney of Pioneer Natural Resources and Weihua Yang (UTD) for their help with this project. The numerical simulations were performed on XSEDE which is supported by National Science Foundation grant number ACI-1548562. F. Pereira was also funded in part by a Science Without Borders/CNPq-Brazil grant #400169/2014-2 and UT Dallas.

PIONEER
NATURAL RESOURCES