

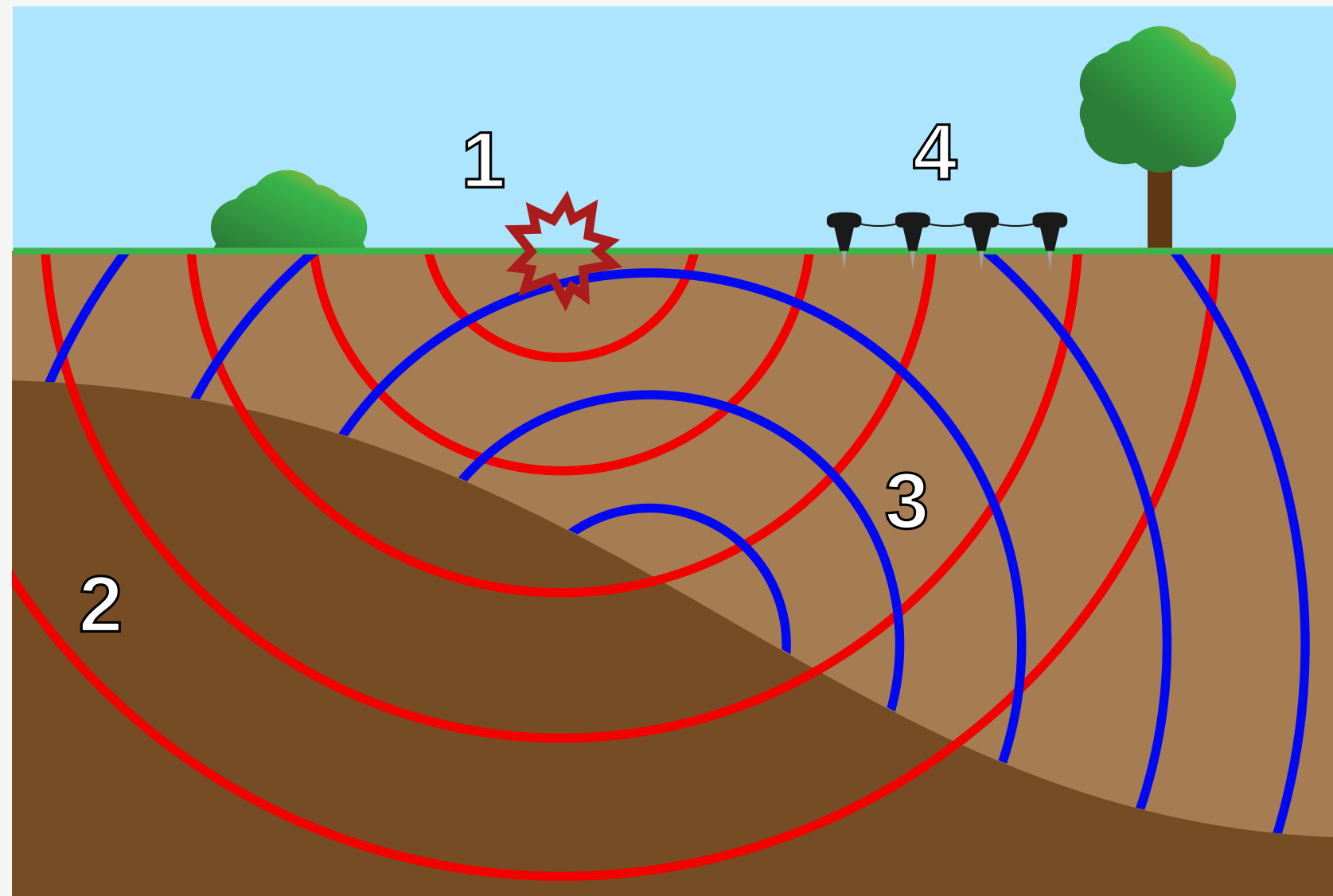
COMPUTATIONALLY EFFICIENT METHODS FOR UNCERTAINTY QUANTIFICATION IN SEISMIC INVERSION



Georgia K. Stuart

The Department of Mathematical Sciences
The University of Texas at Dallas

28 September 2020



1. A **seismic disturbance**.
2. **Seismic waves** propagating through the subsurface.
3. **Reflected seismic waves** created by the change in material.
4. **Geophones** that record the direct (red) and reflected (blue) waves.

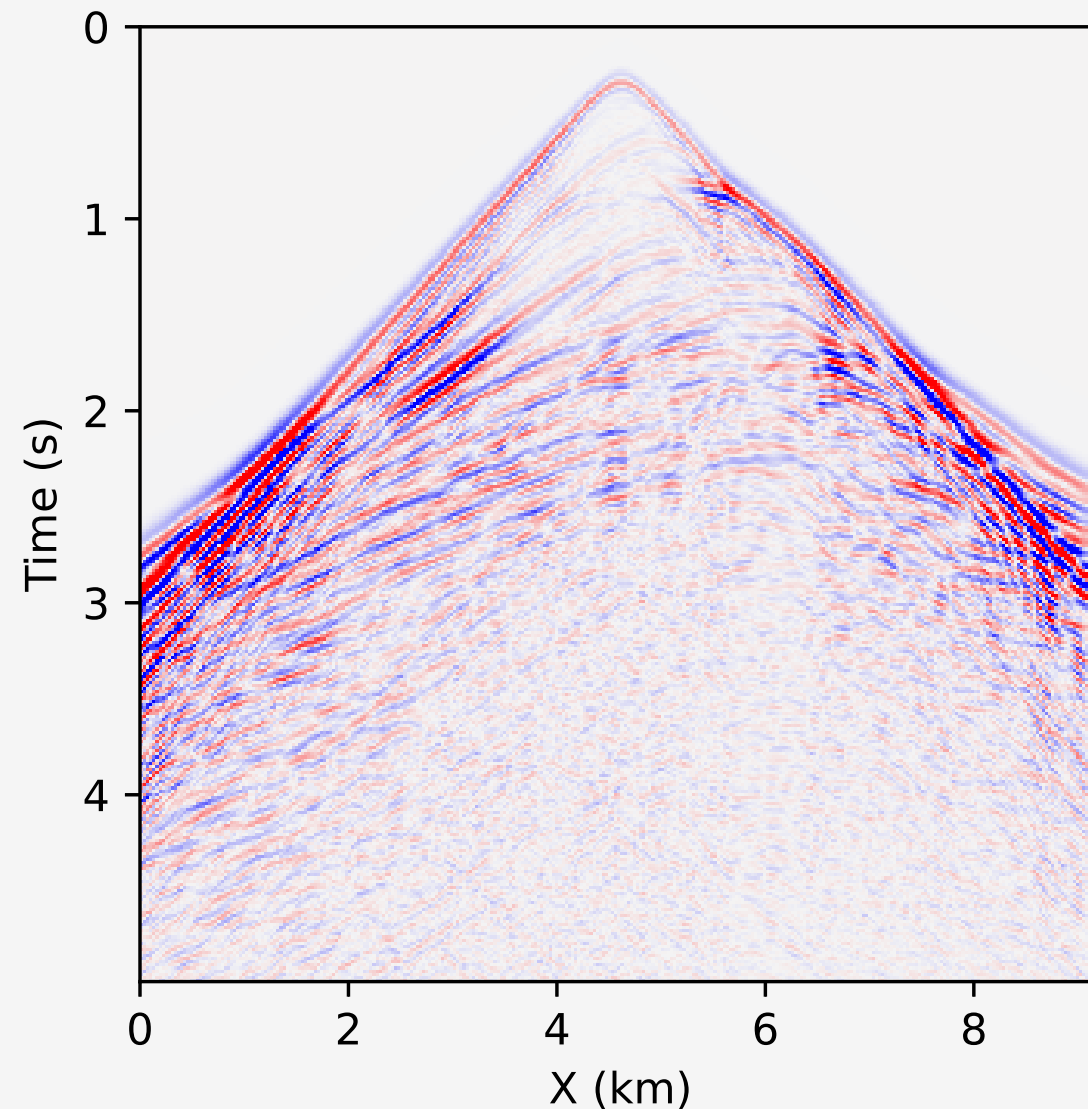


Figure: Synthetic receiver data with a t^2 gain applied to emphasize the later events.

✚ In **Full Waveform Inversion (FWI)**, the goal is to **minimize** an objective function, e.g.:

$$\frac{1}{2} \|F(\theta) - D\|^2$$

where θ describes the velocity field, D is the observed seismic data, and F is the wave solver.

✚ We use the **2D constant-density acoustic wave equation**:

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

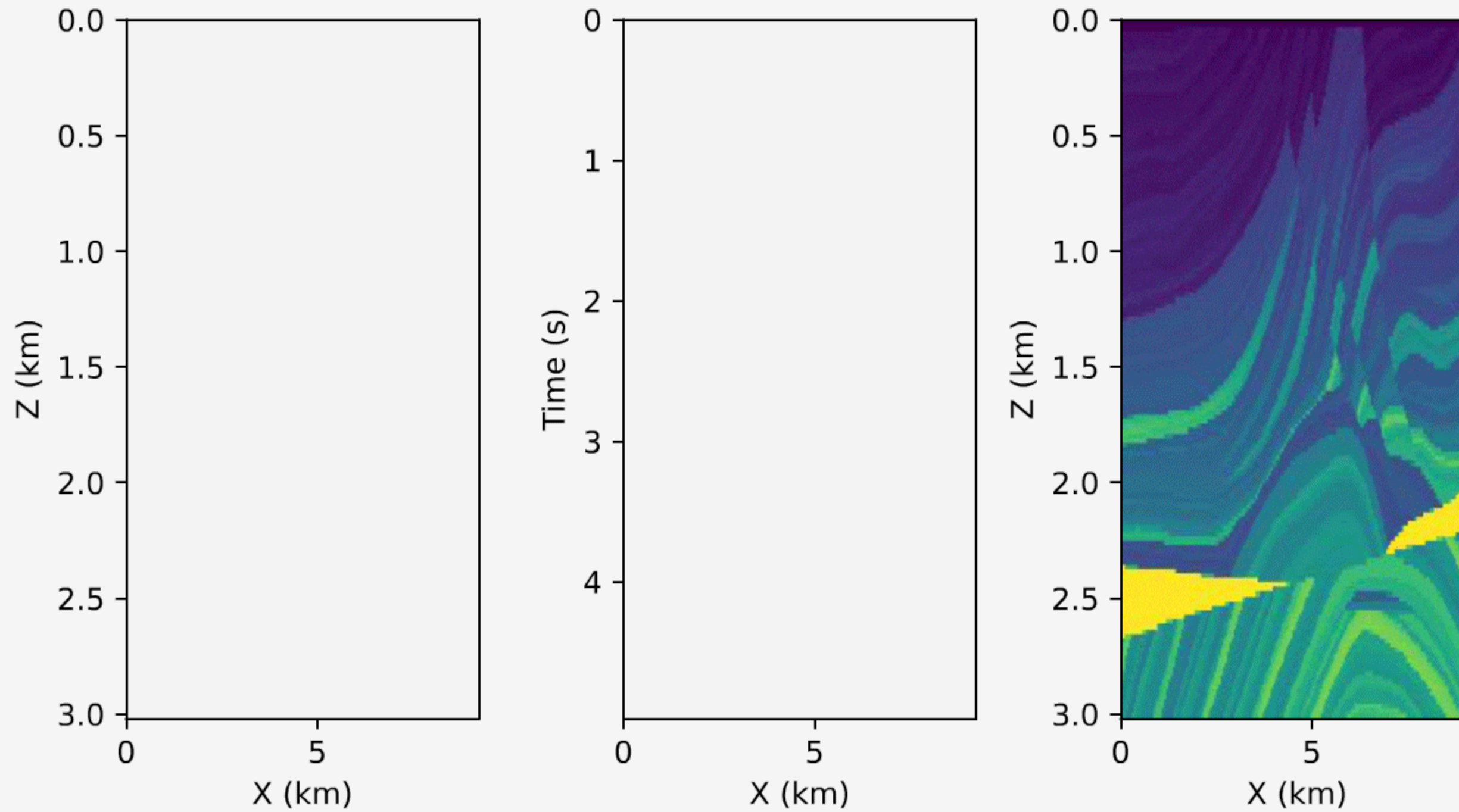
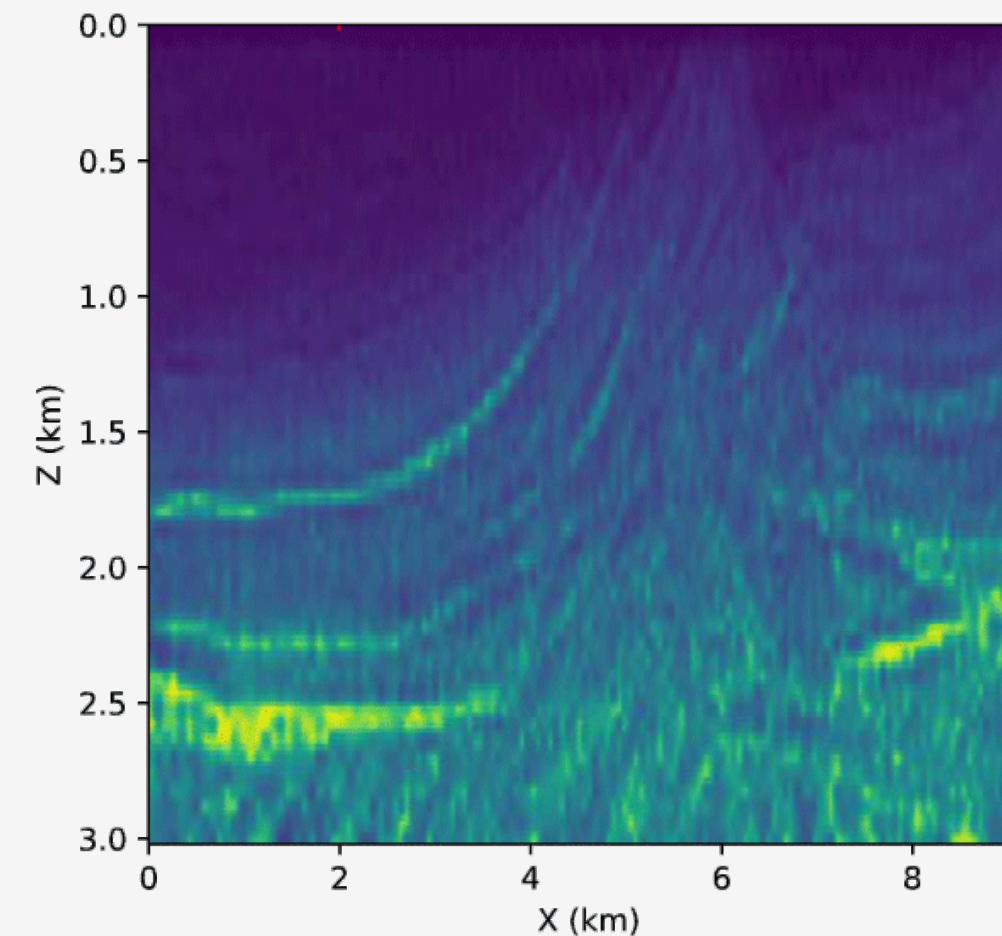


Figure: Left: an acoustic wave propagating through the velocity field on the right. Center: Receiver data by time. Right: the velocity profile.

- Traditional FWI methods result in a **single velocity field**.
- UQ methods result in **distributions of velocity fields**.
- UQ indicates where we have **more or less certainty** about the estimate from FWI.



BAYES' RULE



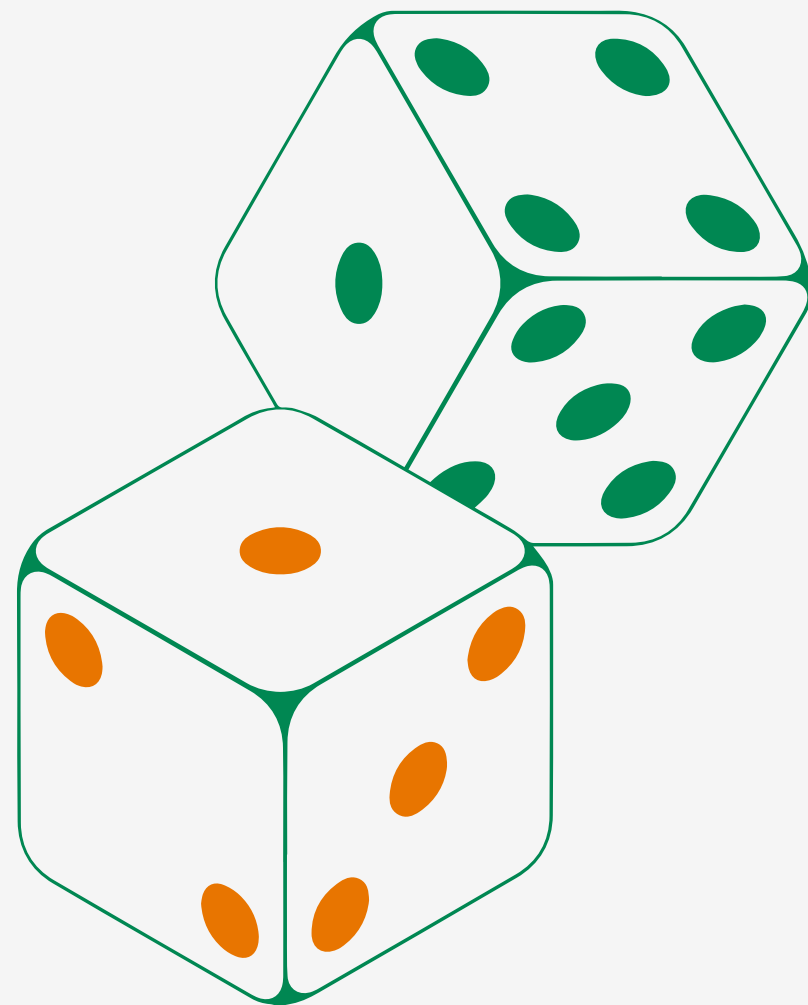
↯ We assume the **likelihood function** has the form

$$\pi(D|\theta) = \exp \left(-\frac{\|F(\theta) - D\|^2}{2\sigma^2} \right)$$

where $F(\theta)$ is the simulated data, D is the observed data, and σ is the precision parameter.

↯ The **prior distribution** can take many forms, e.g. uniform or Gaussian.

↯ However, the **posterior is not necessarily Gaussian**.



- ↯ **Deterministic** approaches to Bayesian FWI require **many assumptions**.
- ↯ **Stochastic** approaches require fewer assumptions.
- ↯ **Markov chain Monte Carlo** methods sample from the posterior distribution **without assumptions on the shape of the distribution**.

MARKOV CHAIN MONTE CARLO (MCMC)



- ✚ MCMC can take **many models (tens of thousands to millions)** to converge.
- ✚ Each model must be run through a forward simulator (**wave equation**).
- ✚ A single chain can take **a week or more** on a cluster.
- ✚ Often **80% to 90%** of the samples are rejected!

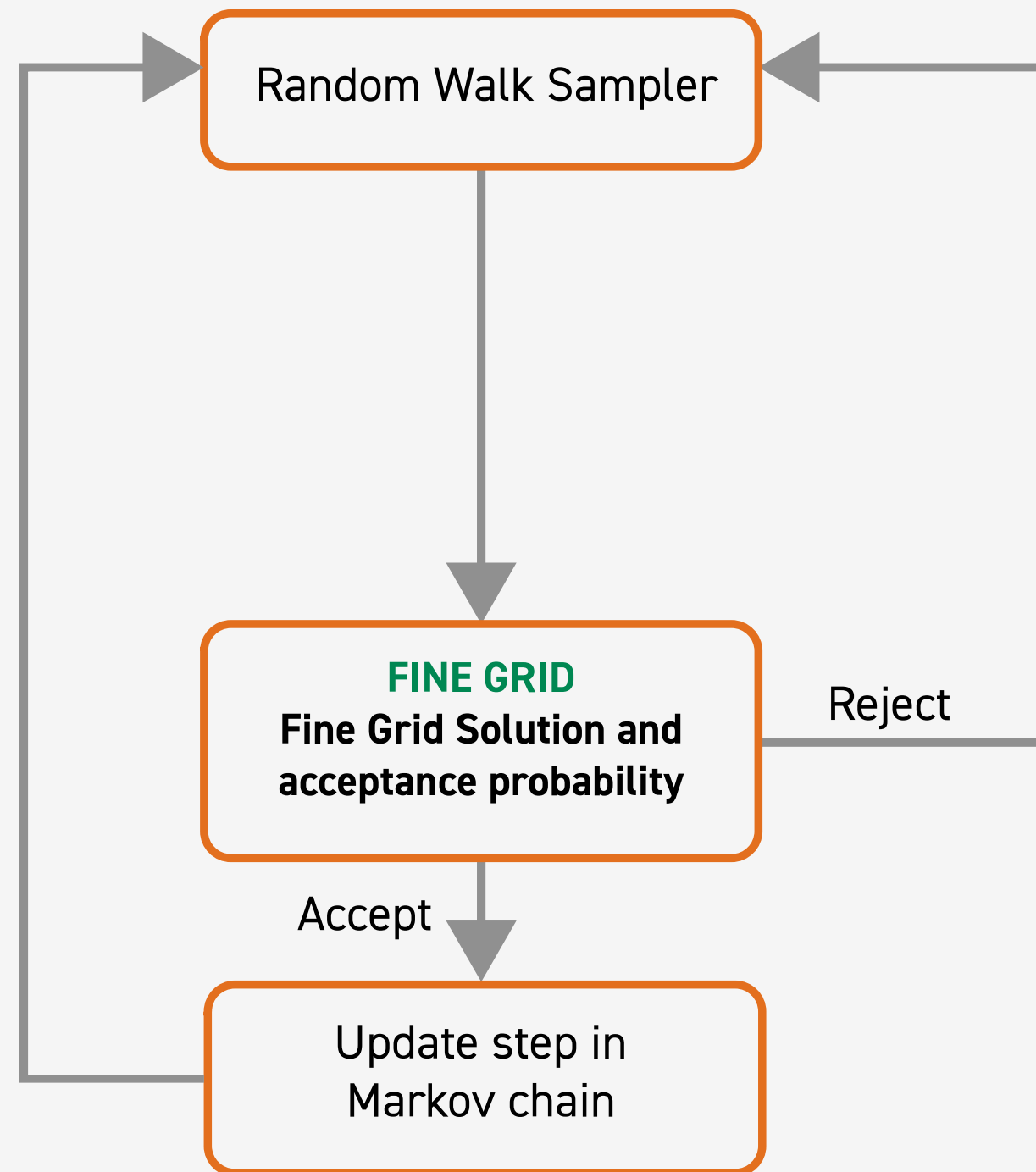


HOW CAN WE REDUCE THE
COMPUTATIONAL COST OF
MCMC METHODS FOR FWI?

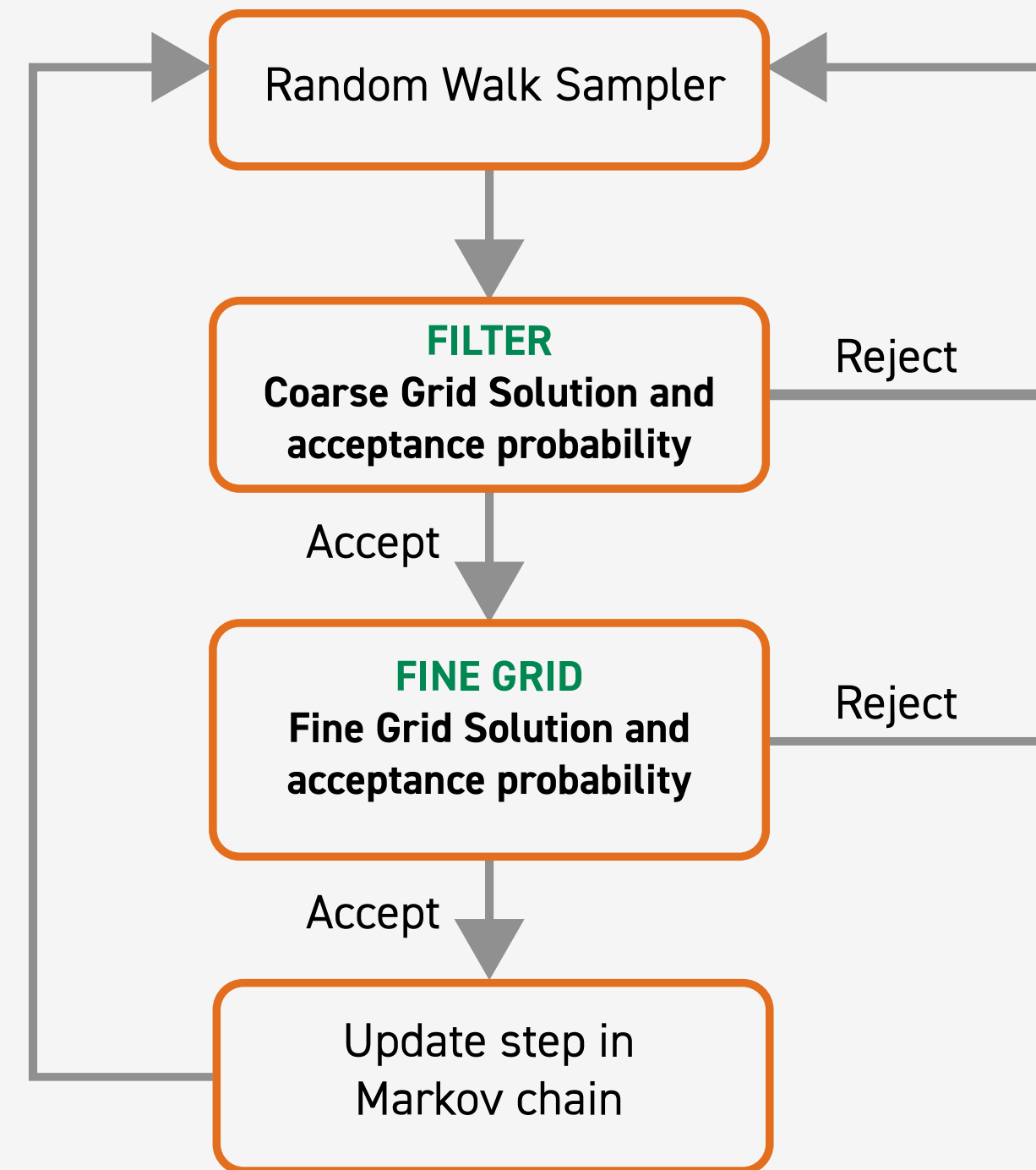
TWO-STAGE MARKOV CHAIN MONTE CARLO



ONE STAGE MCMC



TWO-STAGE MCMC



⚡ Modeling wave propagation can be **computationally expensive**.

⚡ **Operator upscaling**¹ decomposes the solution into two parts:

1. **Fine grid** problem on independent subdomains
2. Small **coarse grid** problem over the whole domain

⚡ In this upscaling technique **we do NOT upscale the model**.

(1) Vdovina et al. (2005), Korostyshevskaya and Minkoff (2006), Vdovina and Minkoff (2008)

1. Write the acoustic wave equation as a system in space by introducing acceleration, \vec{v}

$$\vec{v} = -\nabla p$$

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

2. Solve in parallel for fine grid pressure and acceleration over each independent coarse block. No communication is required at this stage.

3. Solve for coarse grid acceleration over the whole domain.

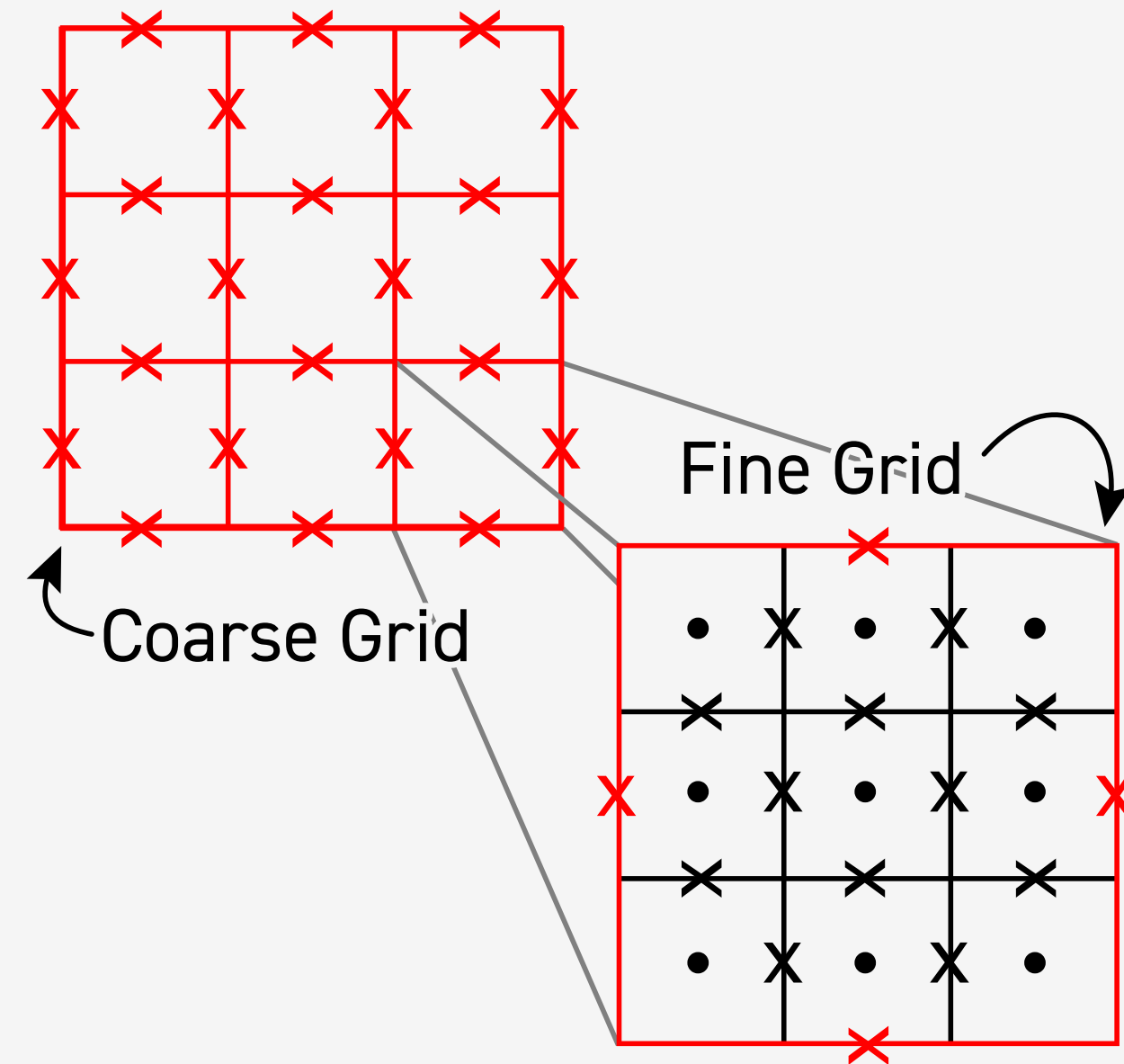
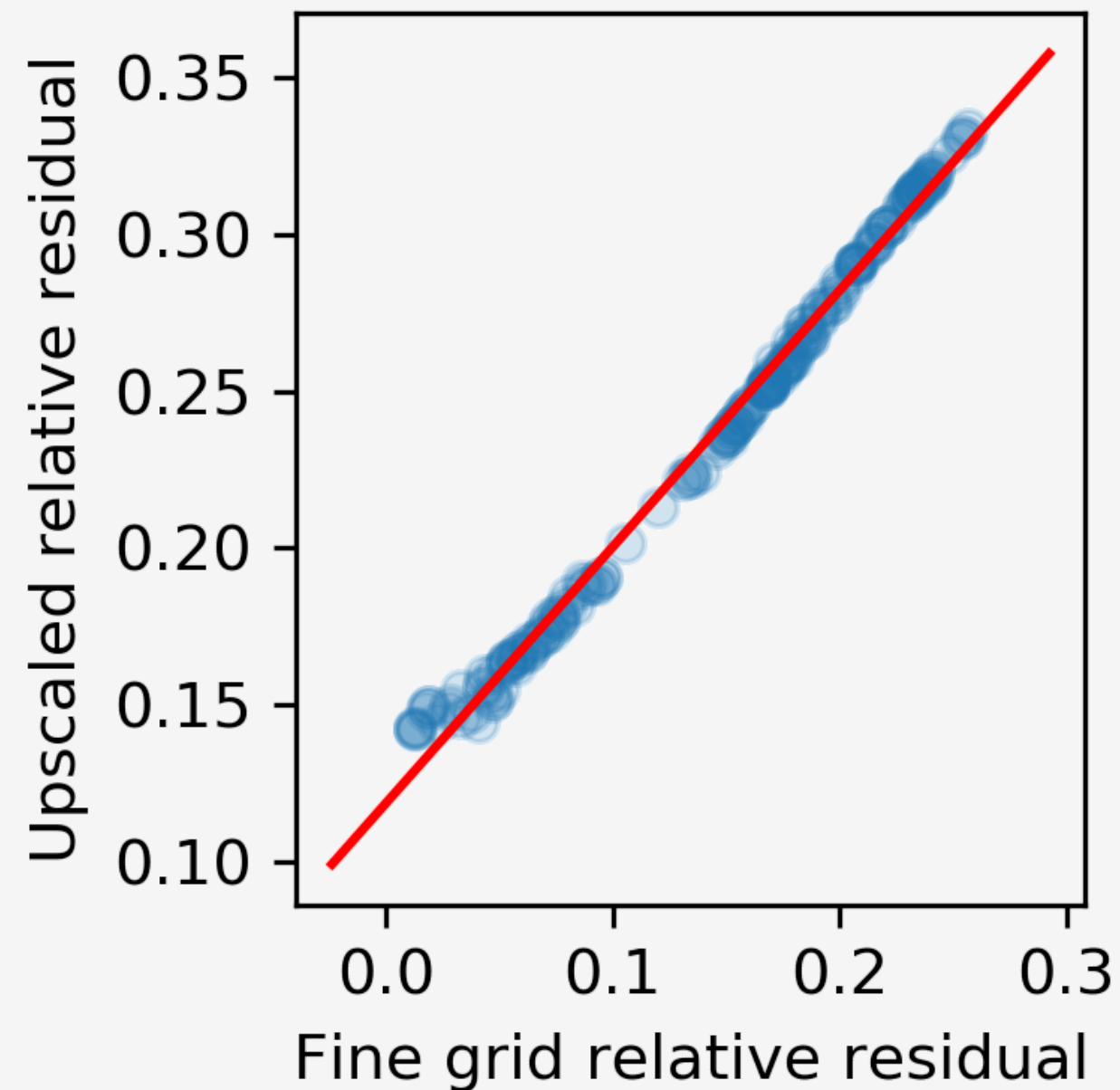


Figure: A picture of the coarse grid (red) and coarse grid acceleration (red X's) and the fine grid (black) and fine grid unknowns (black X's: acceleration, black dots: pressure)



✓ We see a **strong linear relationship** between the fine grid relative residuals and the upscaled relative residuals for a layered velocity model.

✓ This indicates that the upscaling filter is a **good surrogate** for the fine grid solver.

RESULTS: TWO-STAGE MCMC WITH UPSCALING

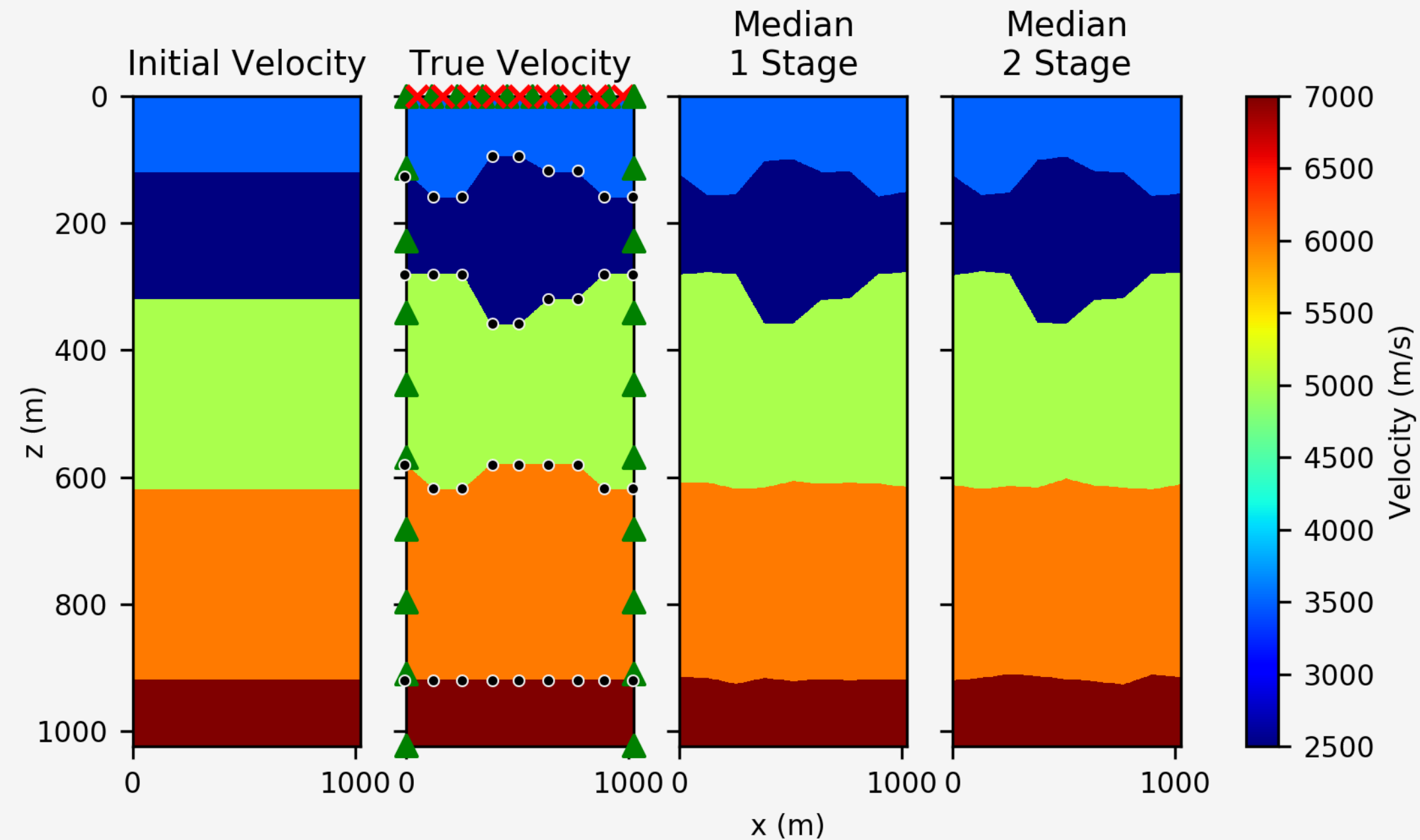


Figure: A comparison of the initial velocity, true velocity, median of the posterior from one-stage MCMC, and the median of the posterior distribution from two-stage MCMC. The true velocity shows the location of a line of sources (red X's), receivers (green triangles), and the unknown nodes that describe the interfaces (black dots). Published in Stuart (2019b).

RESULTS: TWO-STAGE MCMC WITH UPSCALING

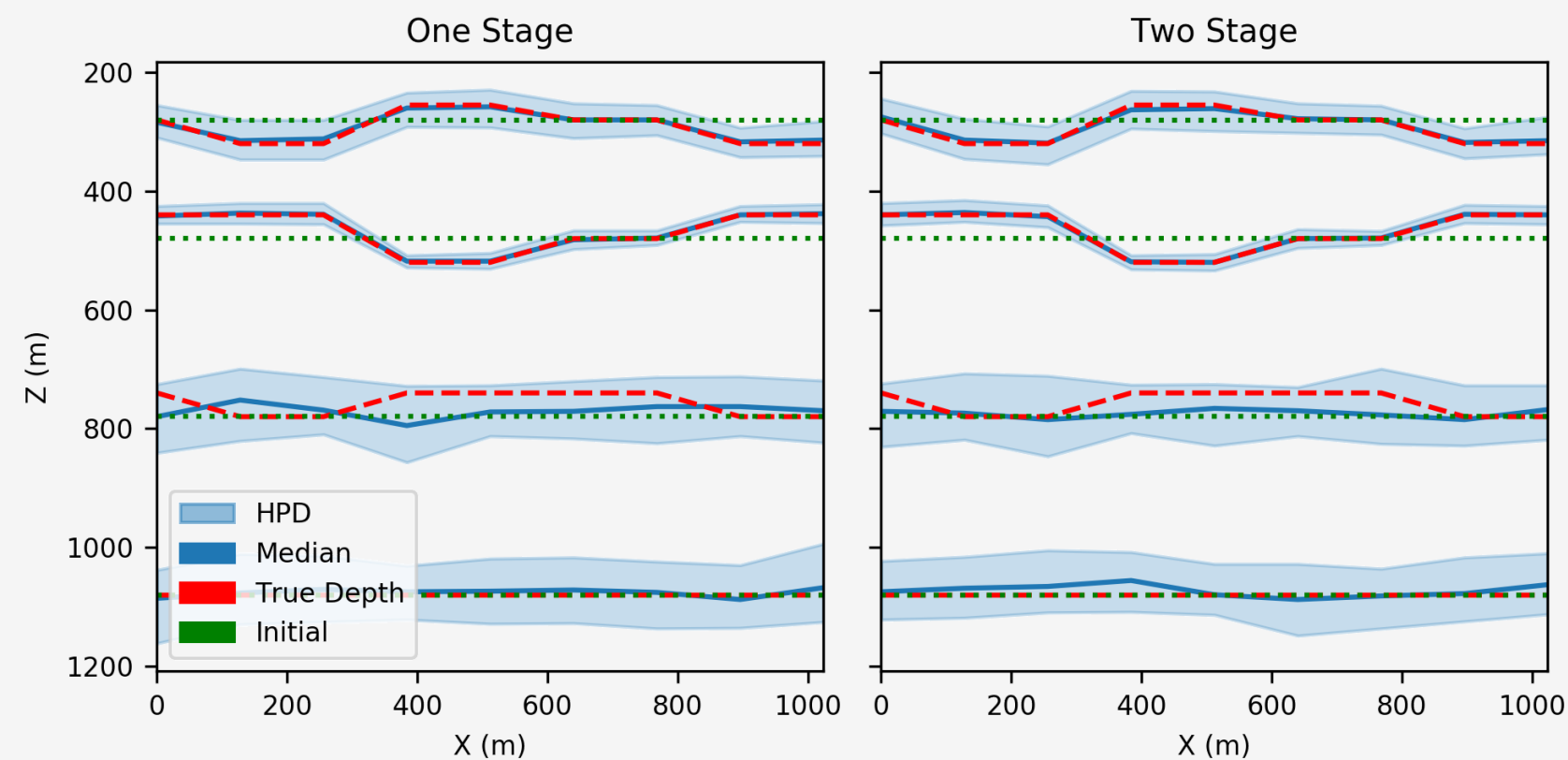


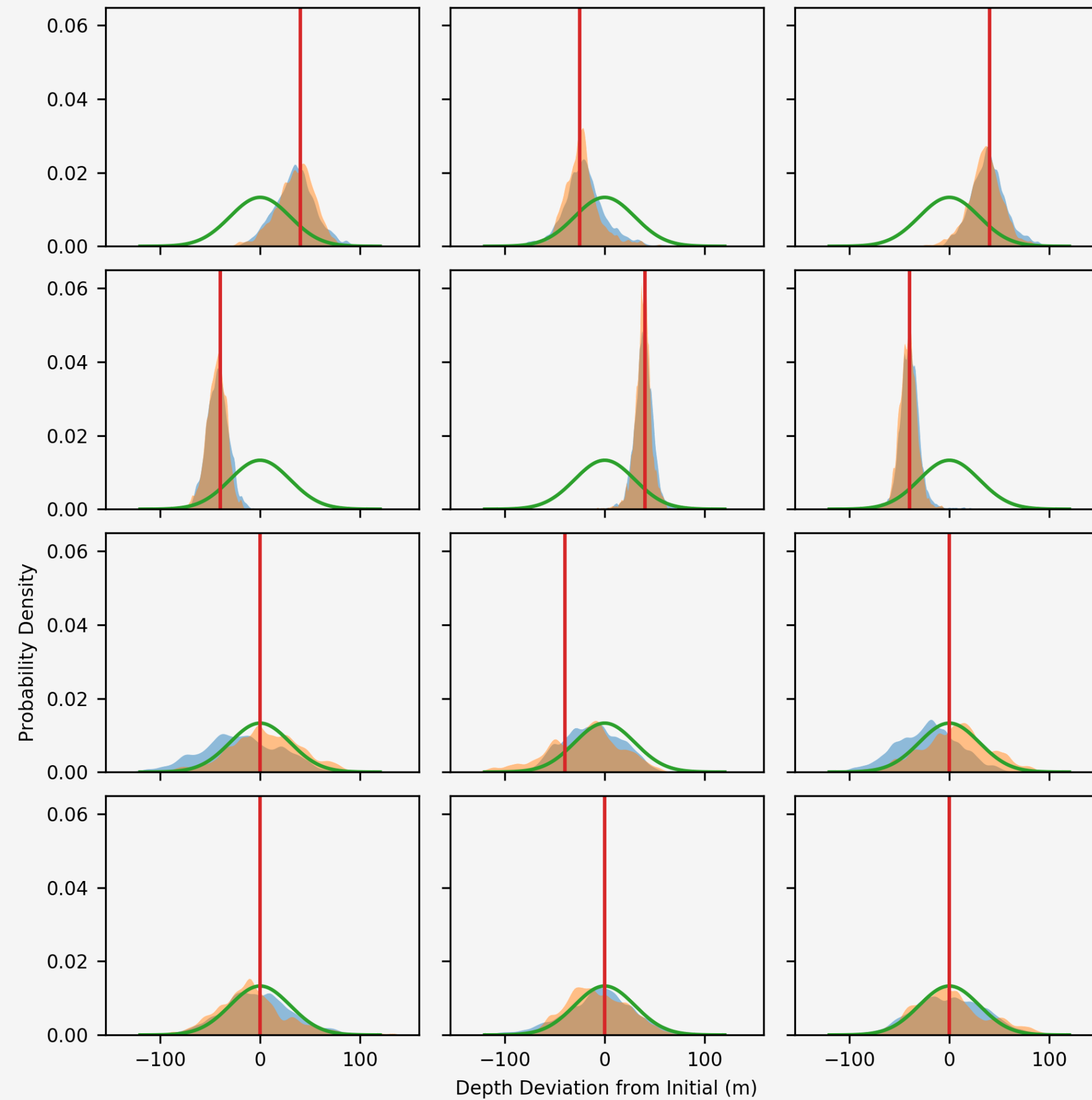
Figure: A comparison between one-stage MCMC highest posterior density (HPD) intervals and two-stage MCMC HPD intervals.

✓ Acceptance rate
increased from 10% to 40%.

✓ Time per sample
decreased by 22%
(40% in other experiments).

✓ Time per rejection
decreased by 33%.

RESULTS: TWO-STAGE MCMC WITH UPSCALING



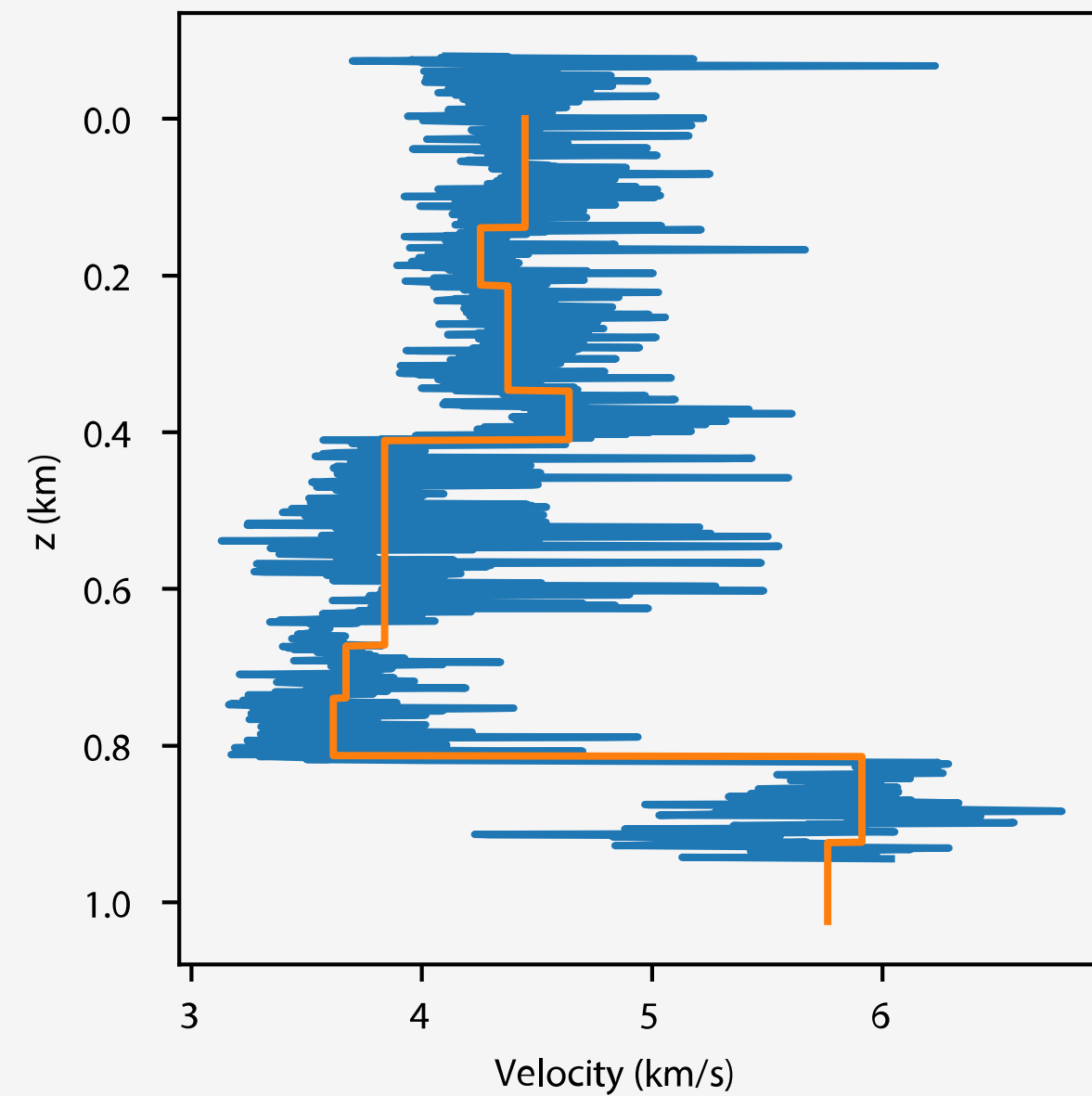


Figure: Well log from the Midland, TX basin (blue, courtesy of Pioneer Natural Resources and 9-layer block (orange).

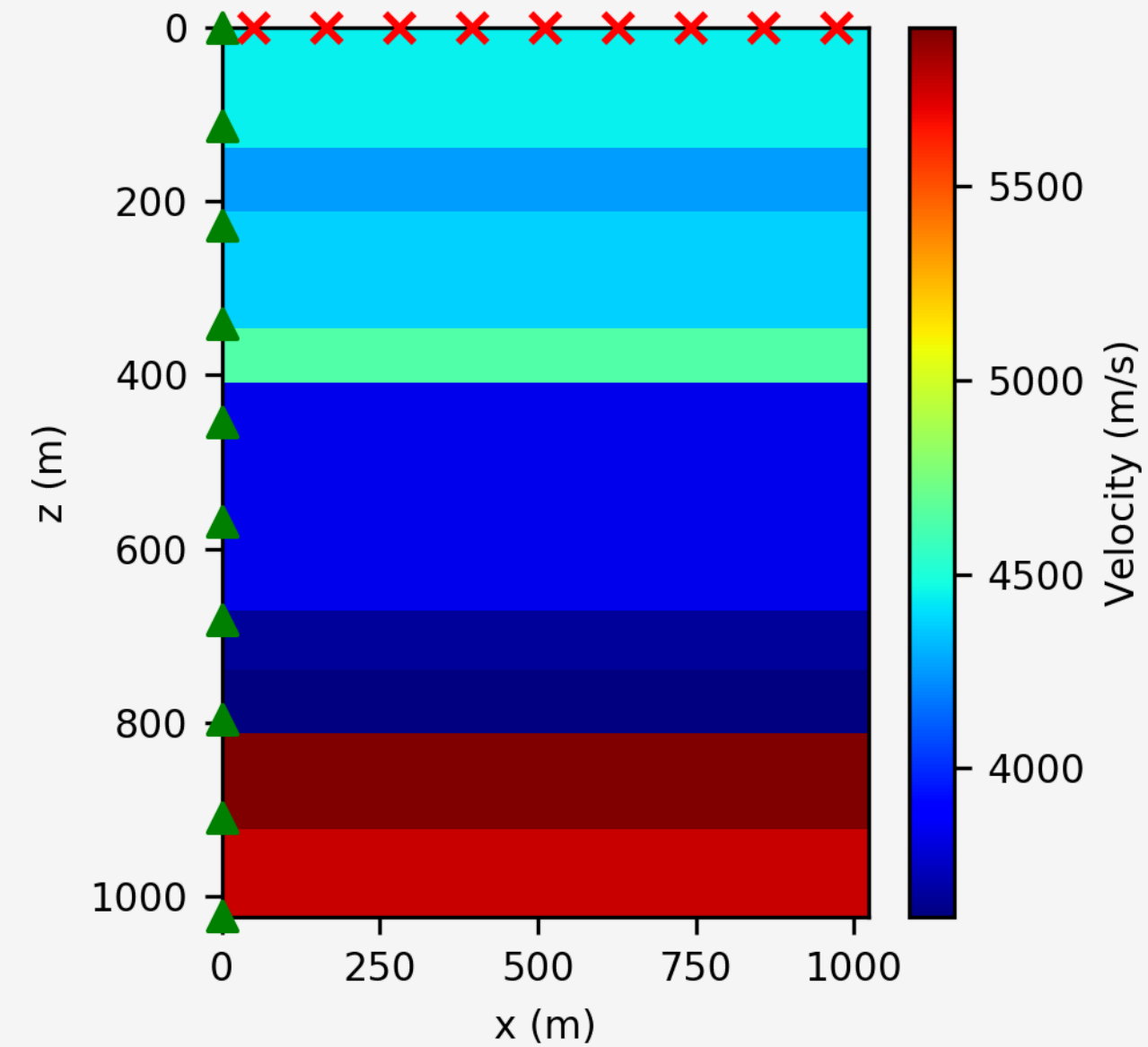


Figure: Flat layered experimental setup with nine unknowns (Stuart et al. 2019a)



Figure: The fine grid residual norm vs. neural network filter residual norm with continuous learning.

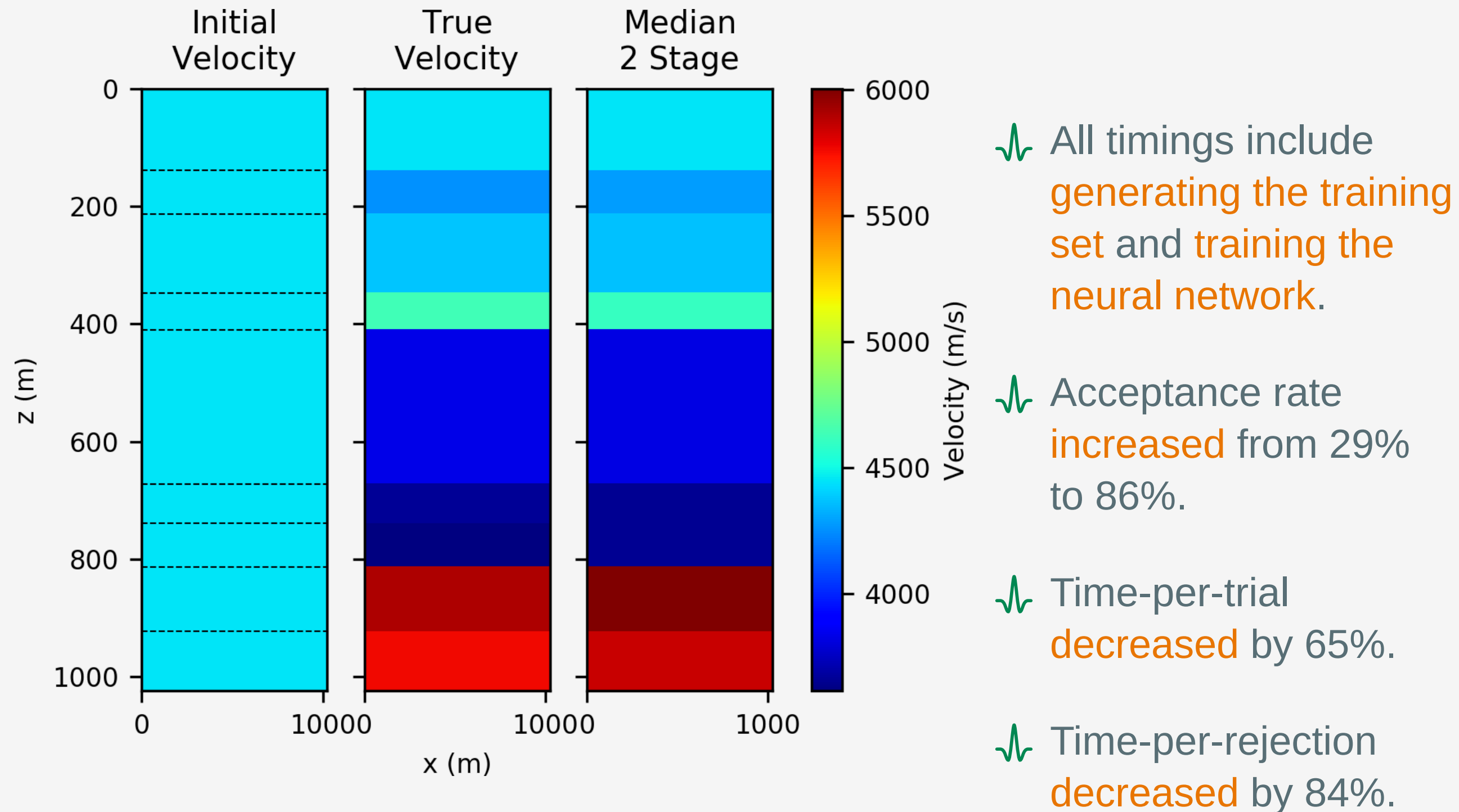


Figure: The initial, true, and median velocity fields for the neural net two-stage MCMC experiment. The dashed lines in the initial velocity field mark the positions of pre-set interfaces.

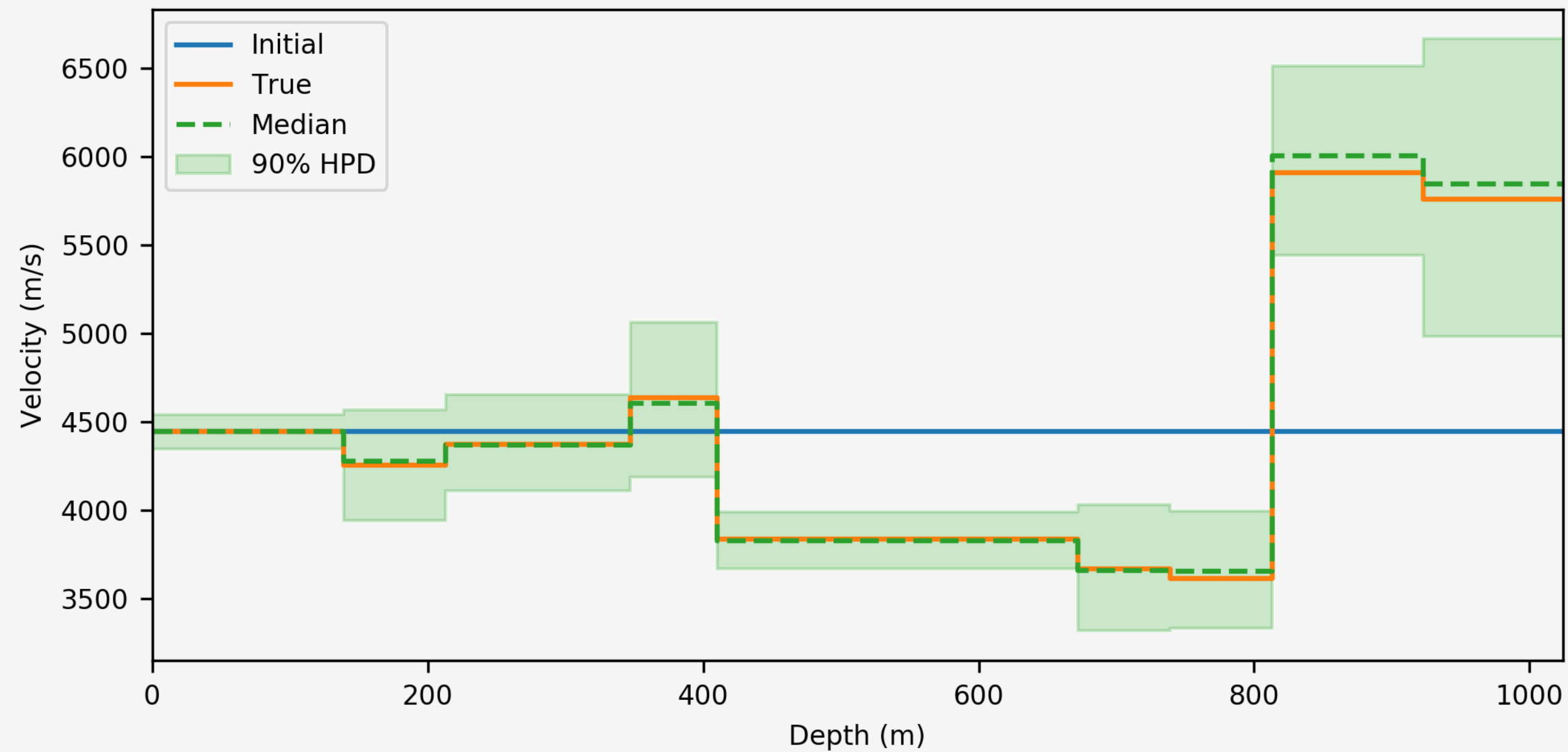


Figure: A vertical slice of the initial (blue), true (orange), and median (green dashed) velocity fields. With 90% highest posterior density intervals.

TROUBLE: THE RANDOM WALK SAMPLER



- ✓ In **theory**, MCMC will converge to the target distribution.
- ✓ In **practice**, methods based on random walk sampling (RWS) can handle a limited number of unknowns (**< 100 in our experience**)
- ✓ RWS produces samples that are highly **correlated**.¹

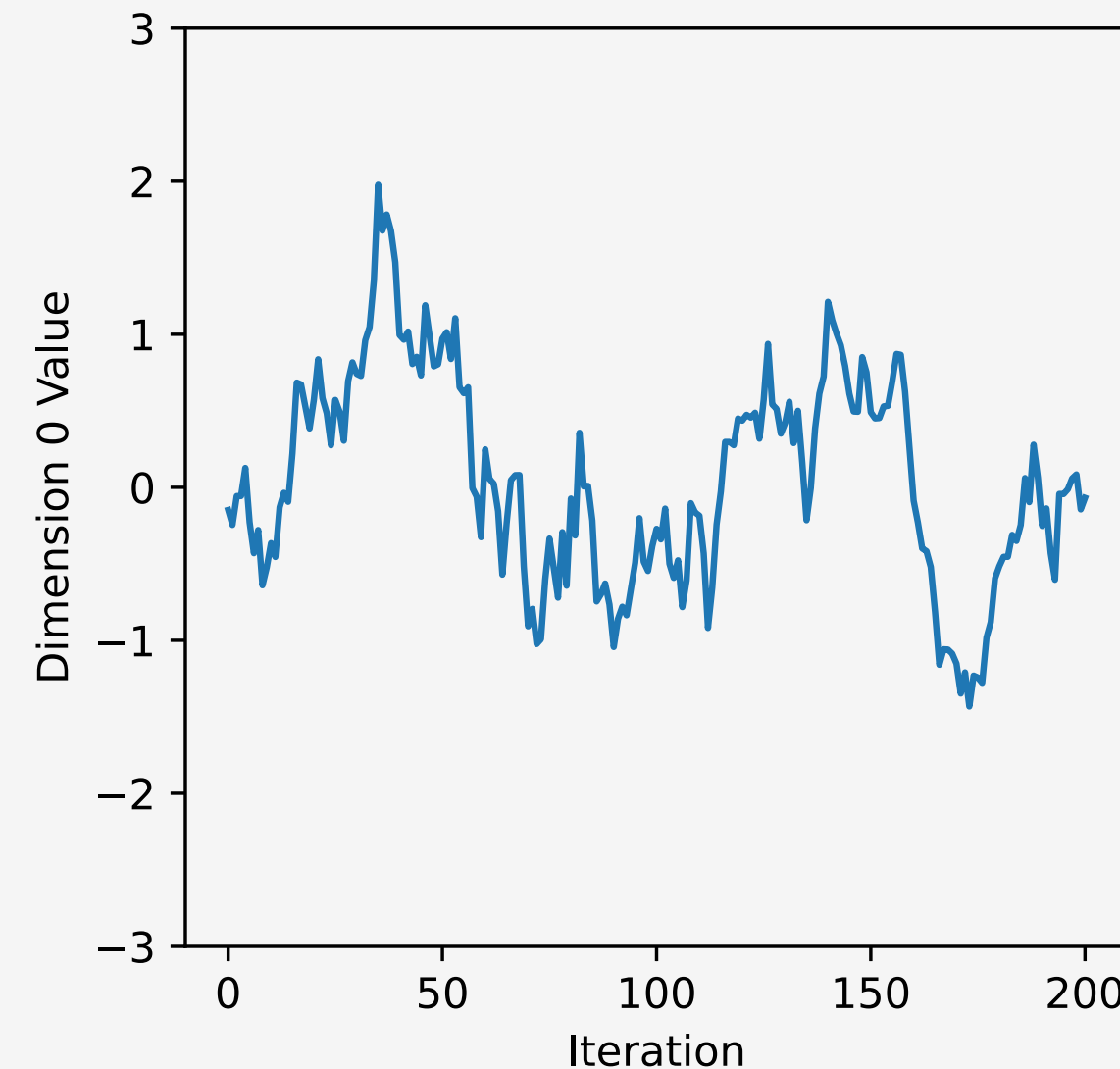


Figure: A view of one dimension of a 100-dimensional Gaussian sampled with Metropolis-Hastings MCMC.

THE RANDOM WALK SAMPLER
PRACTICALLY LIMITS THE
NUMBER OF UNKNOWNNS WE
CAN USE

- ✓ HMC uses **Hamiltonian Mechanics** and **gradient information** of the posterior distribution to draw samples that are less correlated.
- ✓ This results in an algorithm that can handle **higher dimensions** and **converges in fewer samples**.

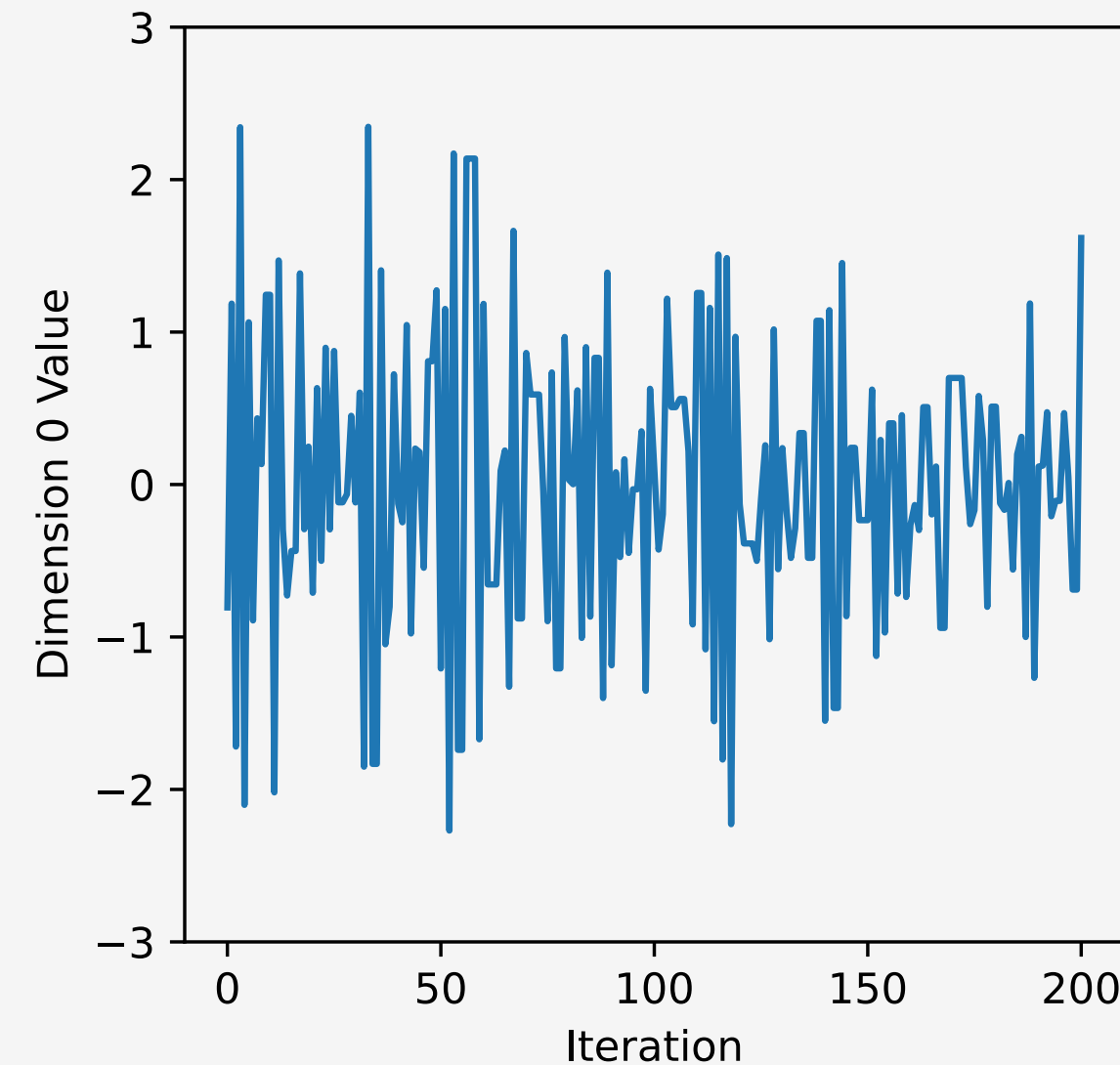


Figure: A view of one dimension of a 100-dimensional Gaussian sampled with HMC.

Potential
Energy

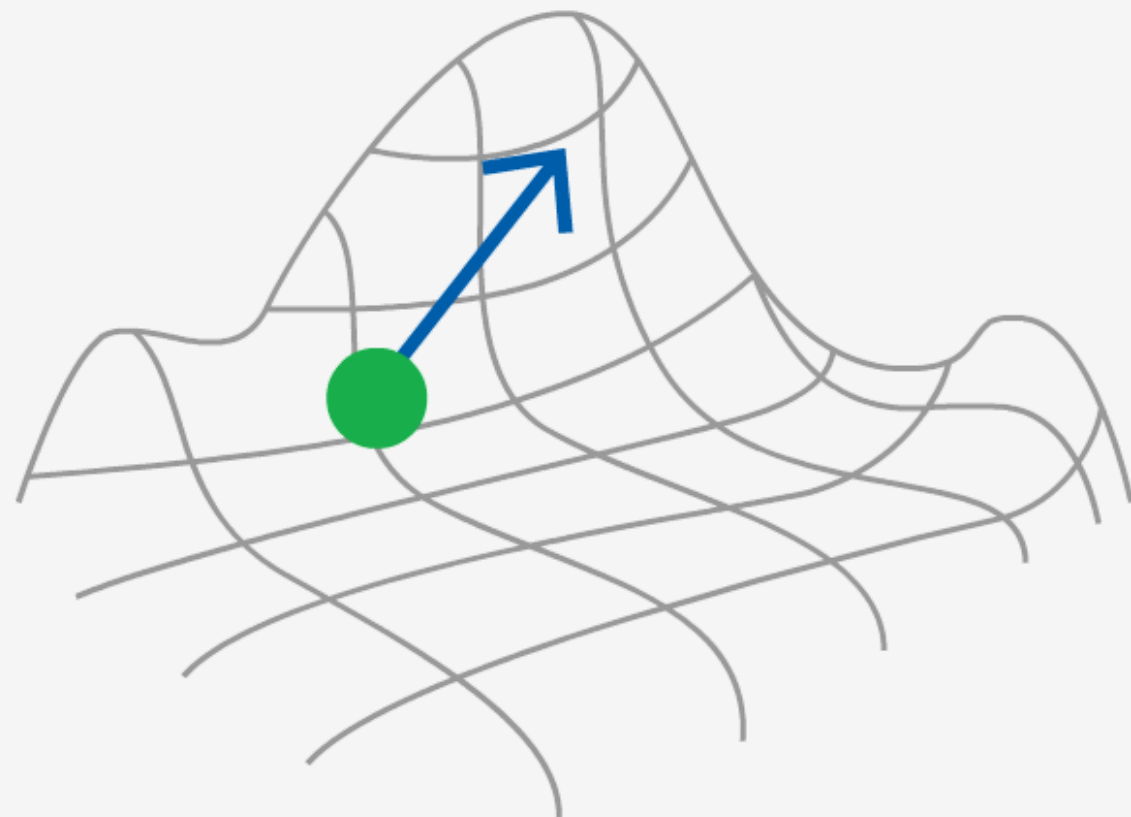
Kinetic Energy

$$H(q, p) = U(q) + K(p)$$

Hamiltonian

Position

Momentum



HAMILTON'S EQUATIONS AND THE POSTERIOR DISTRIBUTION



- ✚ The posterior distribution is embedded in the potential energy by way of the **canonical distribution**.

$$U(q) = -\log[\pi(q|D)] = -\log[\pi(q)\pi(D|q)]$$

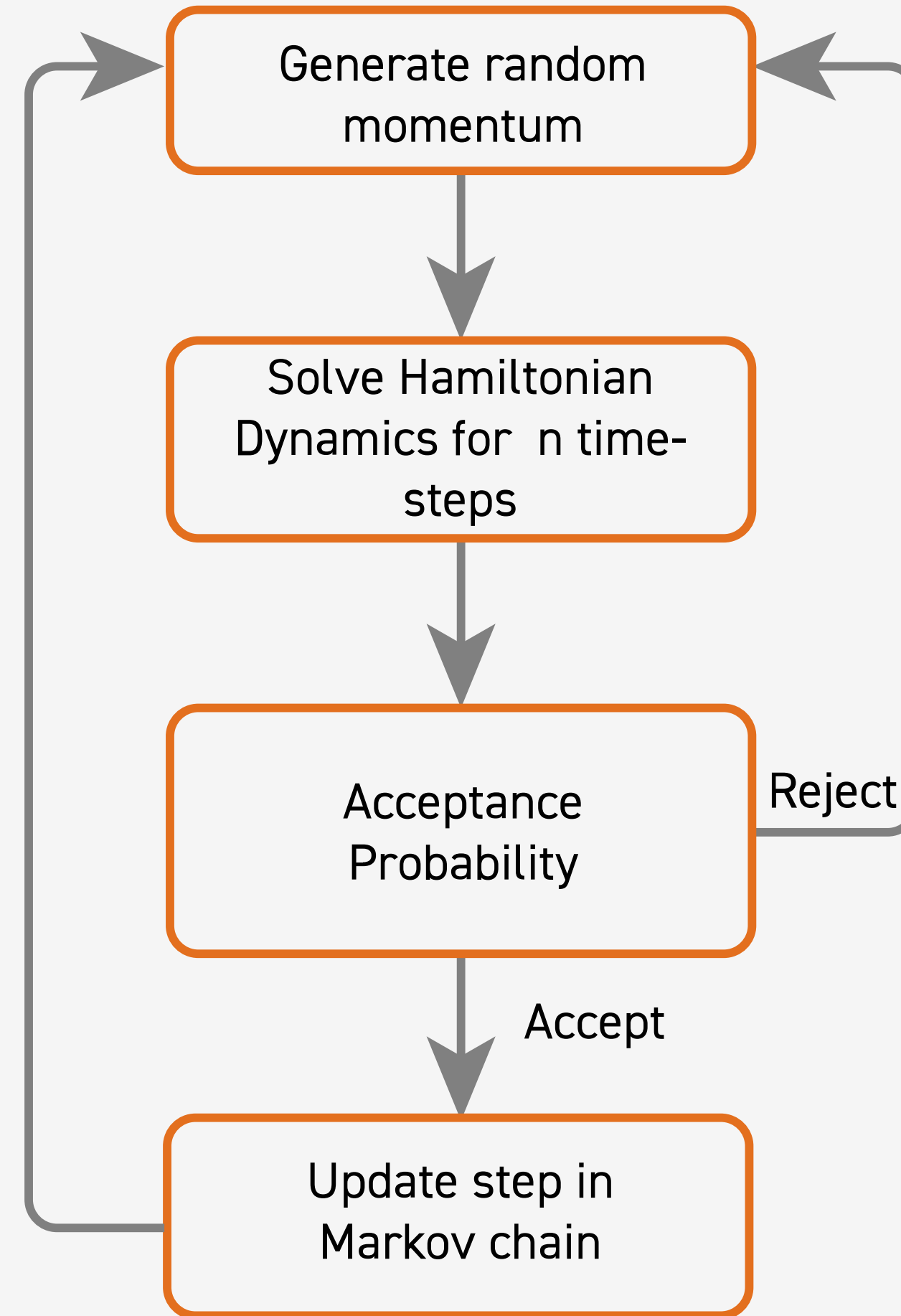
where $\pi(q|D)$ is the **posterior distribution**, $\pi(q)$ is the **prior distribution**, and $\pi(D|q)$ is the **likelihood function**.

- ✚ For Hamilton's equations, we need to take the **gradient of the log likelihood**.

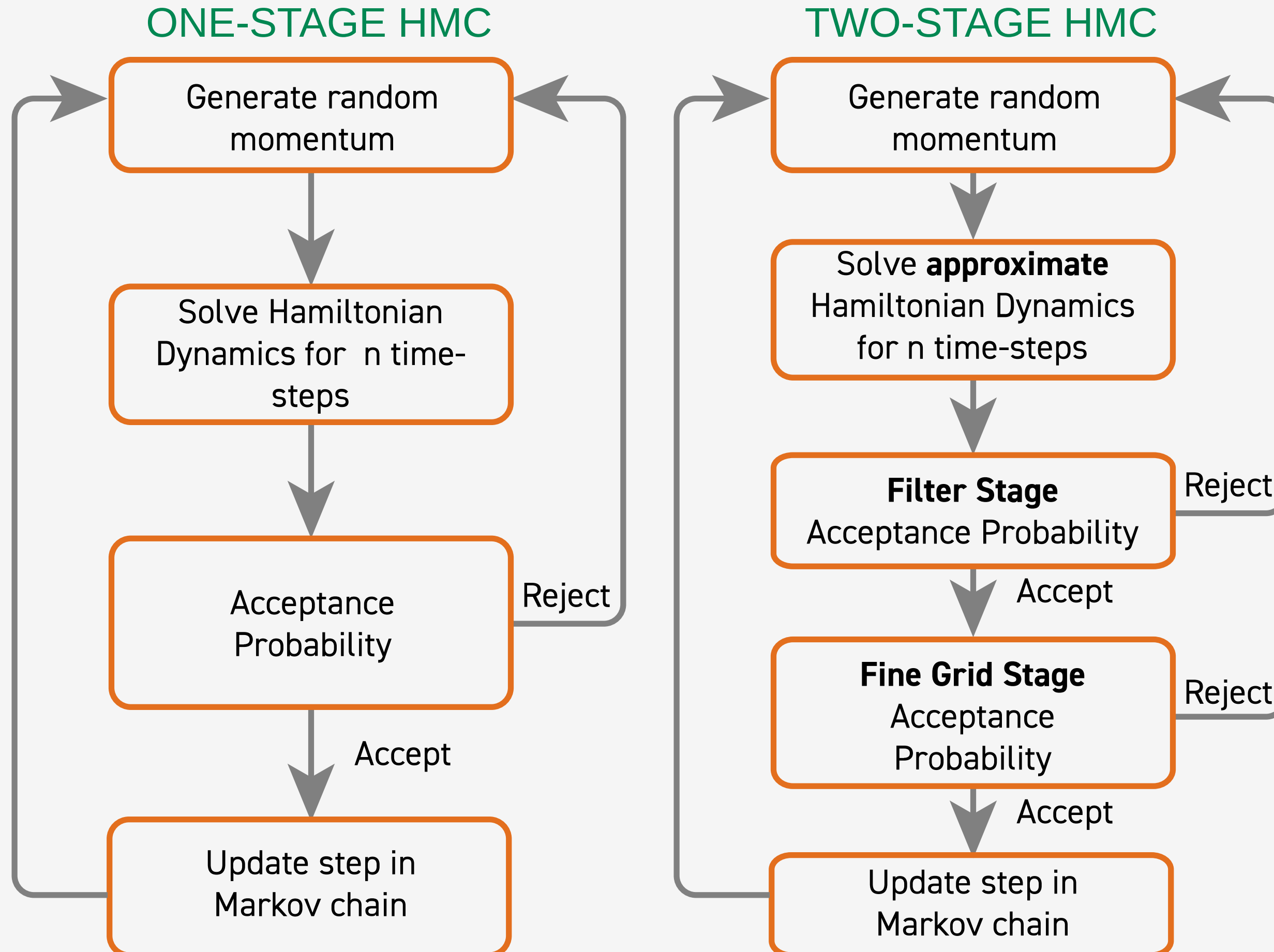
$$-\nabla \log[\pi(D|q)] = \nabla \frac{\|F(q) - D\|^2}{2\sigma^2}$$

THE LEAPFROG DISCRETIZATION



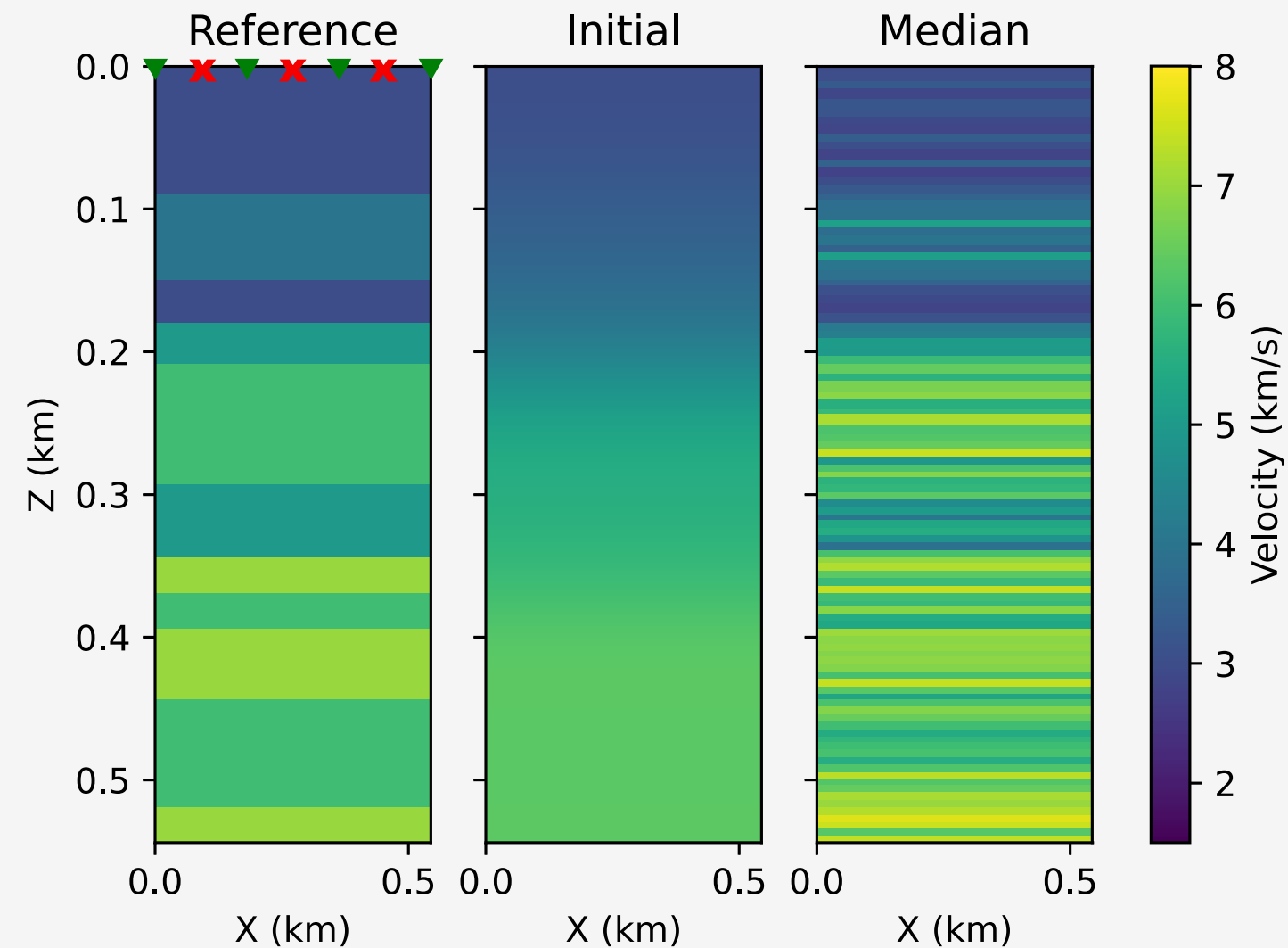


HAMILTONIAN MONTE CARLO
REQUIRES NUMEROUS
EXPENSIVE GRADIENT
CALCULATIONS TO PRODUCE
EACH SAMPLE.



NEURAL NETWORK-ENHANCED TWO-STAGE HMC (NNHMC)





⚡ All timings include
generating the training
set and training the
neural network.

⚡ Time-per-trial
decreased by 85%.

Figure: The initial (left), true (middle), and median (right) velocity fields for the 100-unknown neural net two-stage HMC experiment. On the left image, red x's represent a line of sources and green triangles represent a line of receivers.

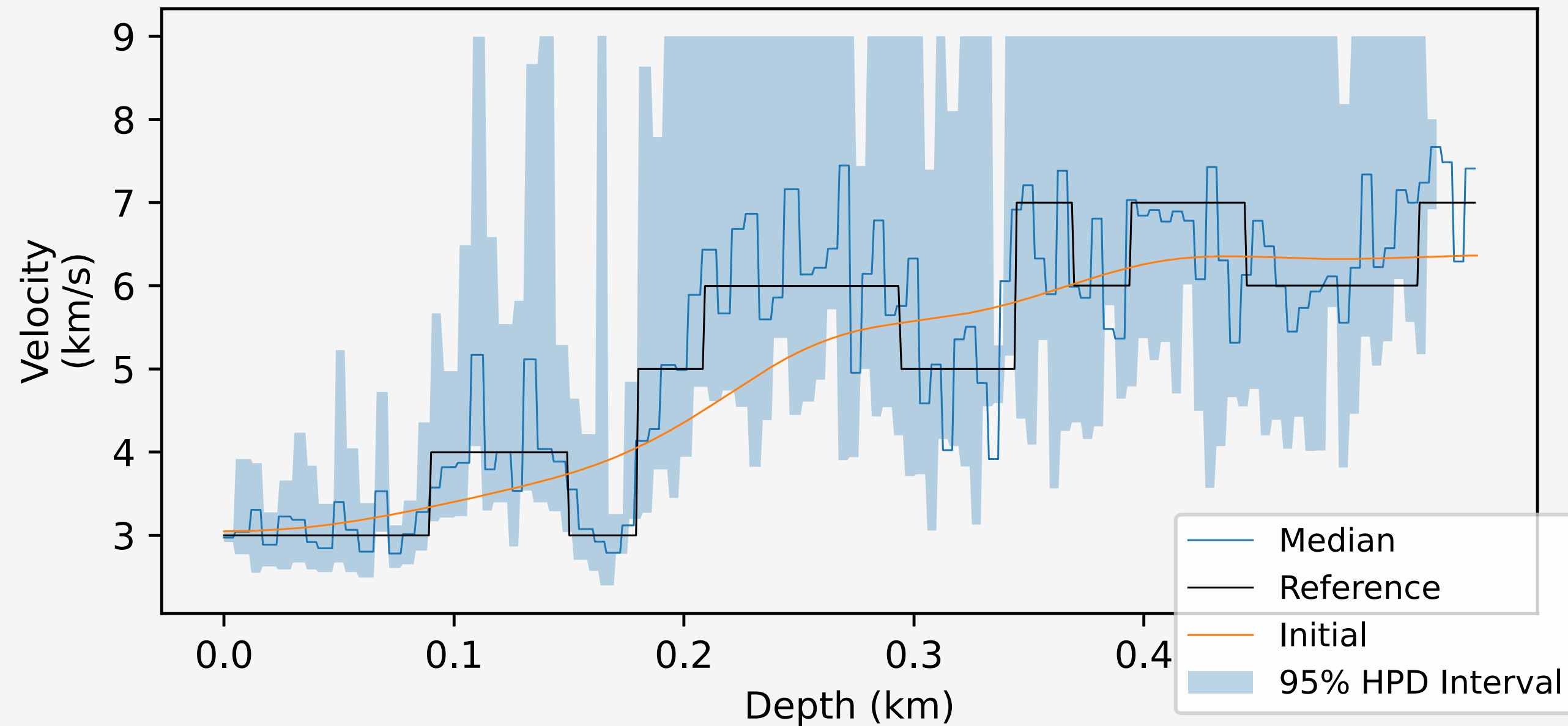


Figure: A one-dimensional slice of the median (blue), reference (black), and initial (orange) velocity fields, with 95% HPD interval shown with blue shading.

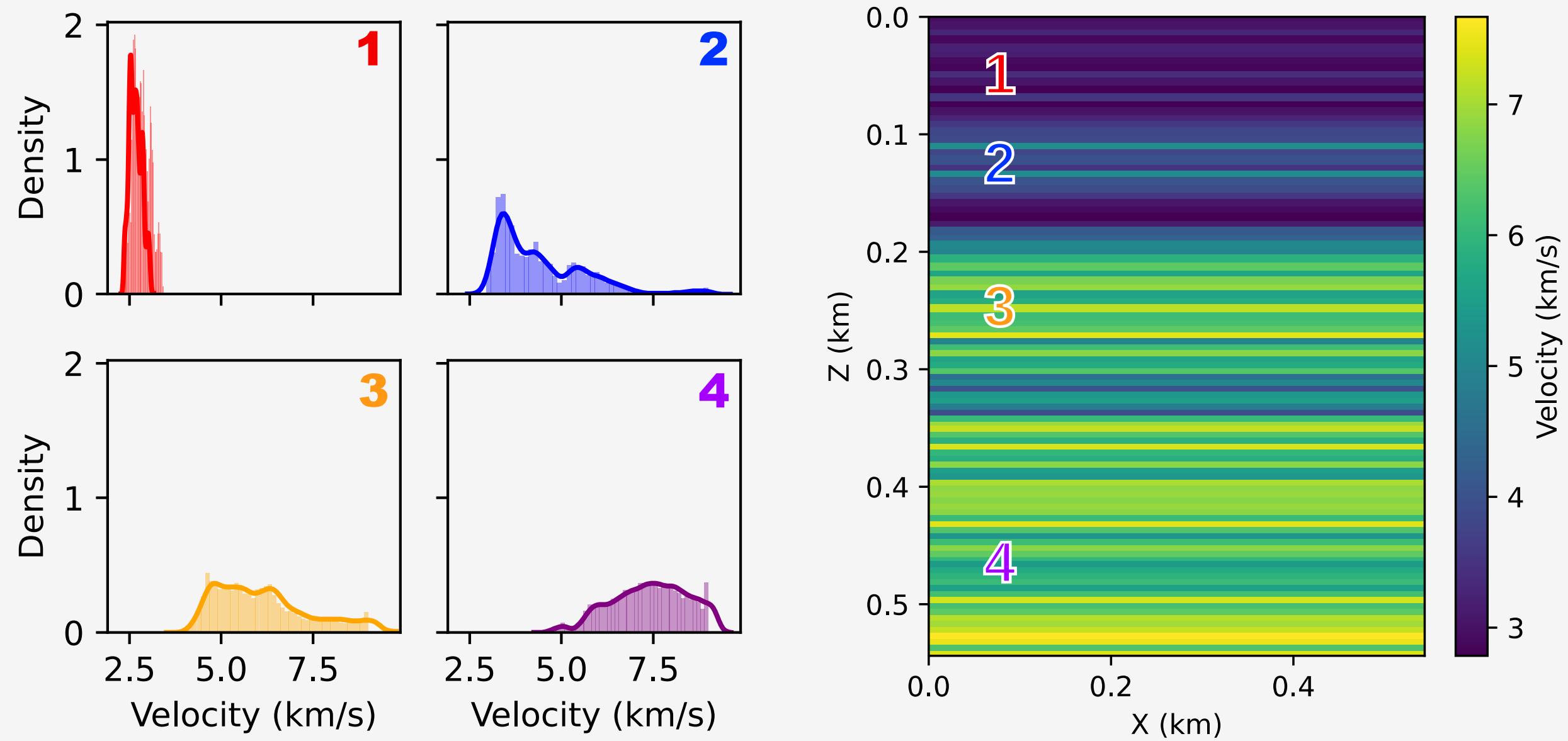


Figure: Four representative posterior distributions (left) from the marked locations (right).

HMC REQUIRES USER-
SPECIFIED PARAMETERS TO
DISCRETIZE THE HAMILTONIAN
DYNAMICS

- ✚ The **No-U-Turn Sampler (NUTS)** modifies HMC to have an adaptive trajectory length L .
- ✚ **Eliminates costly tuning runs** for the trajectory length in the leapfrog algorithm.
- ✚ From the starting point (blue dot), we **we randomly select a forward trajectory** (yellow dot) or a **backward trajectory** (green dots).
- ✚ The number of leapfrog steps **doubles** with each recursive NUTS iteration.



NUMERICAL EXPERIMENT: NUTS

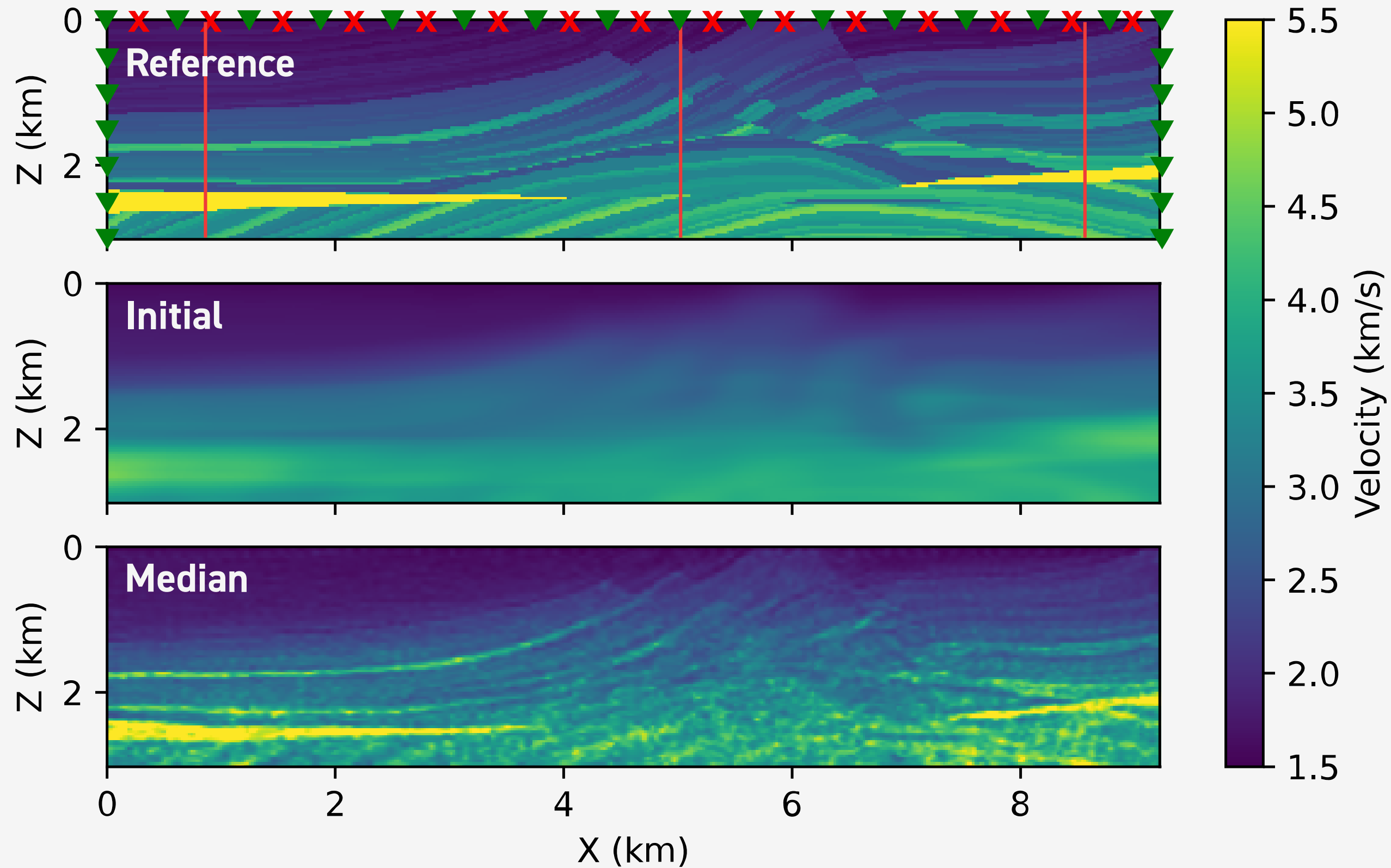


Figure: Top: the reference velocity field with markings for sources (red X's) and receivers (green triangles) and red lines to mark the location of the vertical slices (next slide). Middle: The initial velocity field. Bottom: the median velocity field.

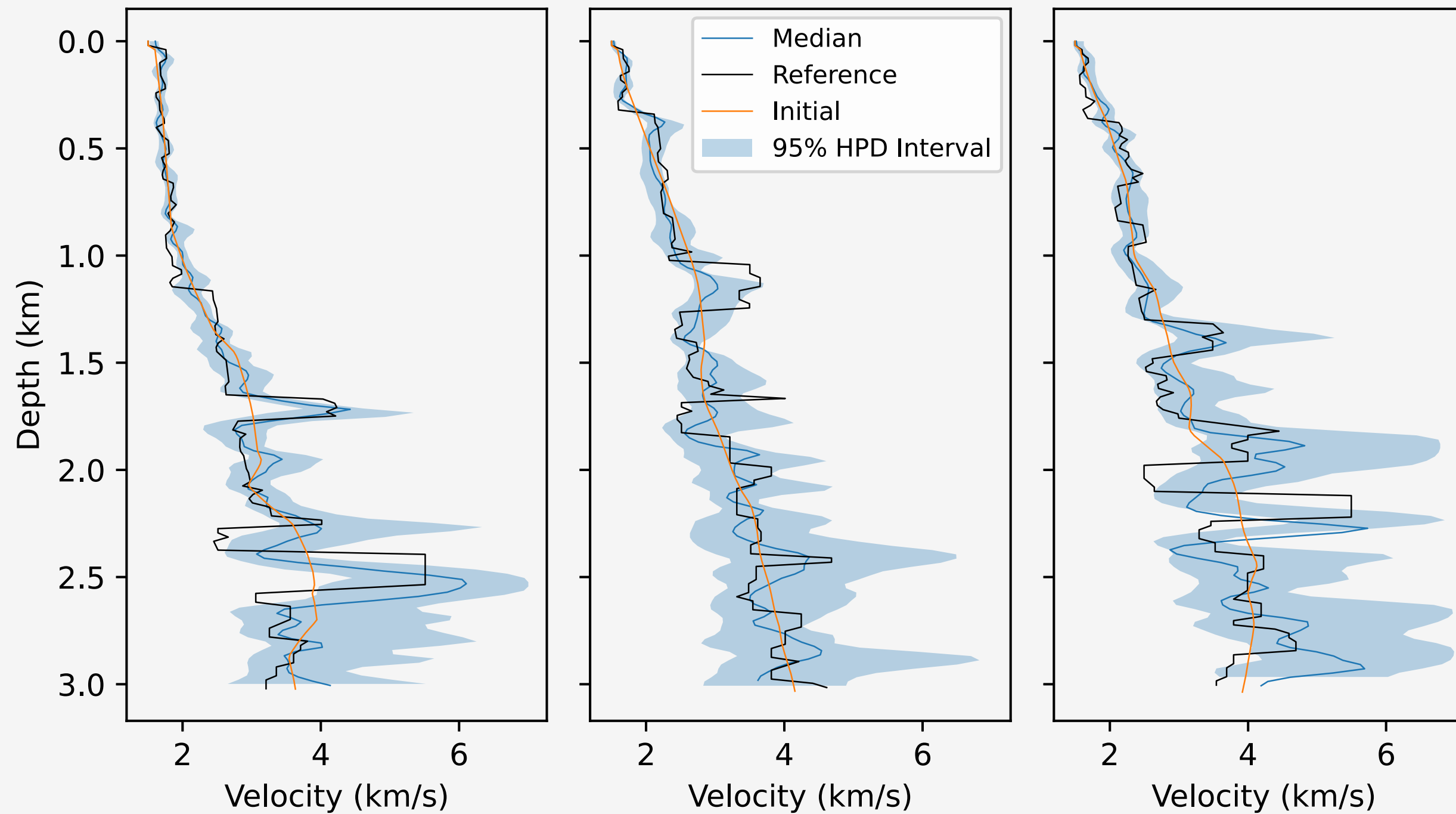


Figure: Vertical slices of the median (blue), reference (black), and initial (orange) velocity fields at locations shown on the previous slide. The blue shaded region marks the 95% HPD intervals

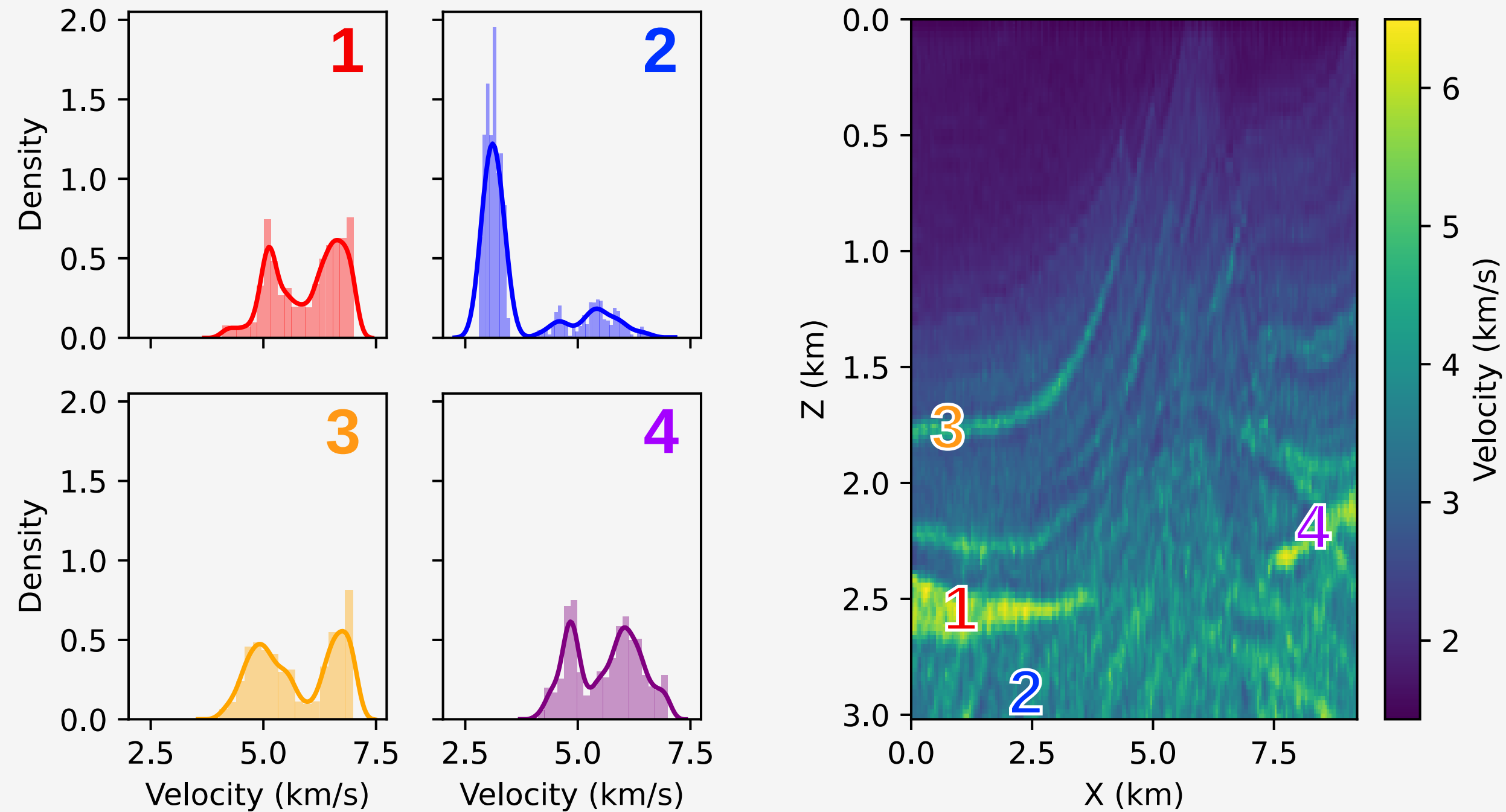


Figure: Four representative posterior distributions (left) at the marked locations shown on the right figure.









- ✚ Two-stage MCMC and HMC is an effective way to quickly reject unacceptable samples and to reduce runtime of the **expensive MCMC and HMC procedures**.
- ✚ **Operator upscaling** is a highly accurate surrogate that closely replicates the fine-grid receiver data.
- ✚ A **neural net** is an extremely inexpensive surrogate that can do a good job of approximating the exponent of the **likelihood function** and the **likelihood gradient**.
- ✚ **Neural-Net Enhanced HMC** reduces the run-time of the HMC algorithm by **over 80%** for our experiment.
- ✚ **The No-U-Turn sampler** optimizes the leapfrog trajectory length for HMC.

ACKNOWLEDGEMENTS



THANK YOU TO...

- ✿ my advisers, Sue Minkoff and Felipe Pereira.
- ✿ David Lumley and Hejun Zhu for use of the seismology group's HPC cluster.
- ✿ Chris Simmons and the Ganymede cluster team for HPC support.
- ✿ the sponsors of the UT Dallas 3D+4D seismic consortium for financial support and industry feedback throughout my research.
- ✿ the Enriched Doctoral Training (EDT) Program, DMS grant #1514808.
- ✿ Pioneer Natural Resources, in particular Rob Meek and Matt McChesney (now at Guidon Energy) for providing well log data and industry feedback.
- ✿ Tim Ullrich for supporting me and always listening to me talk about my research.

- 
 Christen, J. A., and C. Fox, 2005, [Markov chain Monte Carlo Using an Approximation](#): Journal of Computational and Graphical Statistics: A Joint Publication of American Statistical Association, Institute of Mathematical Statistics, Interface Foundation of North America, 14, 795–810.
- 
 Efendiev, Y., T. Hou, and W. Luo, 2006, [Preconditioning Markov Chain Monte Carlo Simulations Using Coarse-Scale Models](#): SIAM Journal on Scientific Computing: A Publication of the Society for Industrial and Applied Mathematics, 28, 776–803.
- 
 Hoffman, M. D., and A. Gelman, 2014, [The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo](#) : Journal of Machine Learning Research: JMLR, 15, 1593–1623.
- 
 Korostyshevskaya, O., and S. E. Minkoff, 2006, [A Matrix Analysis of Operator-Based Upscaling for the Wave Equation](#): SIAM Journal on Numerical Analysis, 44, 586–612.
- 
 Neal, R. M., 2011, [MCMC using Hamiltonian dynamics](#), in S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng, eds., Handbook of Markov Chain Monte Carlo. Handbooks of Modern Statistical Methods Chapman & Hall / CRC, 113–162.
- 
 Stuart, G. K., W. Yang, S. Minkoff, and F. Pereira, 2016, [A two-stage Markov chain Monte Carlo method for velocity estimation and uncertainty quantification](#), in SEG Technical Program Expanded Abstracts 2016, 3682–3687.
- 
 Stuart, G. K., S. E. Minkoff, and F. Pereira, 2019a, [Enhanced neural network sampling for two-stage Markov chain Monte Carlo seismic inversion](#): SEG Technical Program Expanded Abstracts 2019, 5, 1665–1669.
- 
 Stuart, G. K., S. E. Minkoff, and F. Pereira, 2019b, [A two-stage Markov chain Monte Carlo method for seismic inversion and uncertainty quantification](#): Geophysics, 84, R1015–R1032.
- 
 Vdovina, T., S. E. Minkoff, and O. Korostyshevskaya, 2005, [Operator Upscaling for the Acoustic Wave Equation: Multiscale Modeling & Simulation](#): A SIAM Interdisciplinary Journal, 4, 1305–1338.
- 
 Vdovina, T., and S. Minkoff, 2008, [An a priori error analysis of operator upscaling for the acoustic wave equation](#): International Journal of Numerical Analysis and Modeling, 5, 543–569.