Transdimensional Seismic Tomography, High Performance Computing solutions and applications

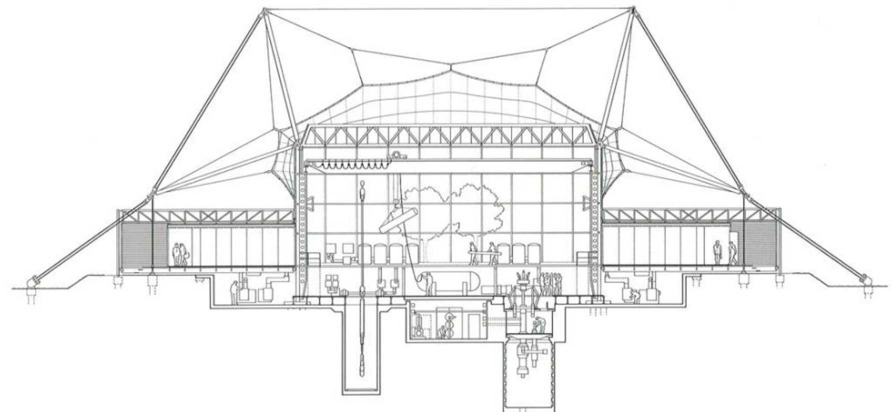
Helmut Wahanik

Ivan Vasconcelos

Erica Galetti

Andrew Curtis

Geophysics Modelling and Inversion Seminar



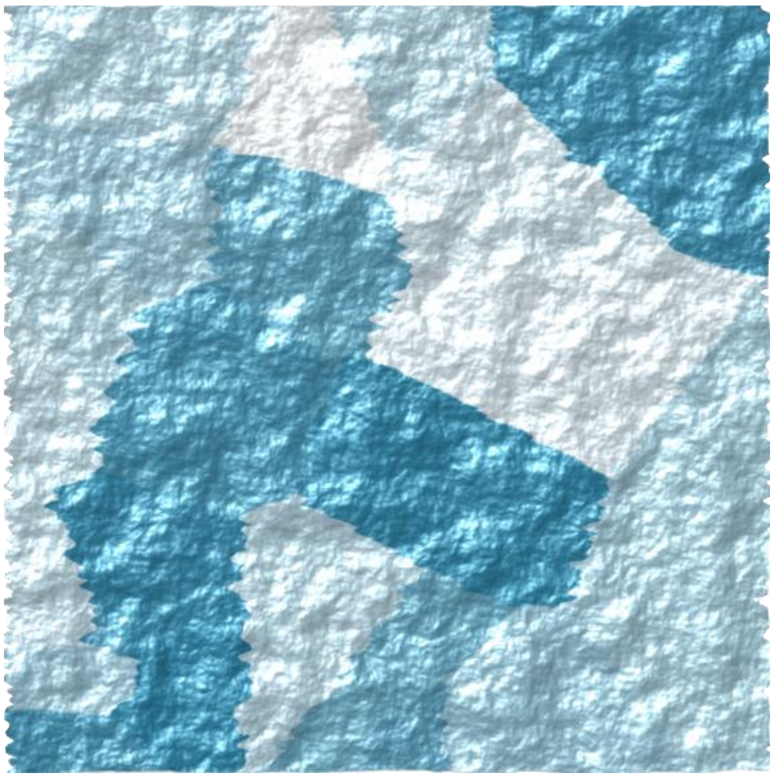
Schlumberger Gould Research

Schlumberger Geophysics Research

University of Edinburgh

Seismic Tomography

- Invert for a physical property of the medium…



Rock's seismic velocity…

Seismic Tomography

- Invert for a physical property of the medium…

…approximated by a Voronoi tesselation

Tomography ⇨ Transdimensional

- Invert for a physical property of the medium…

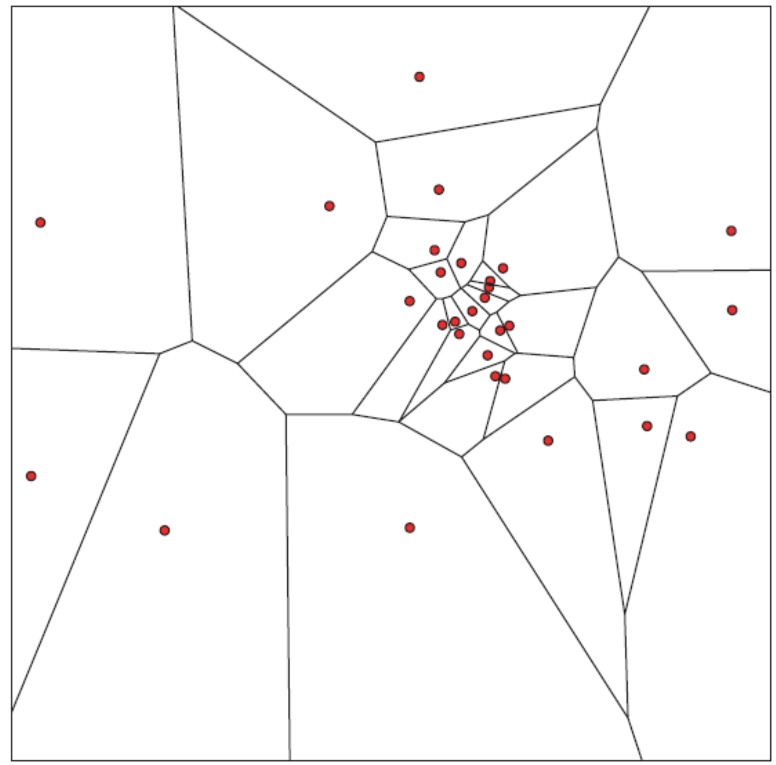
Voronoi grid of seismic velocity

Transdimensional Tomography

- Invert for a physical property of the medium… and for the dimension of the parameter space.

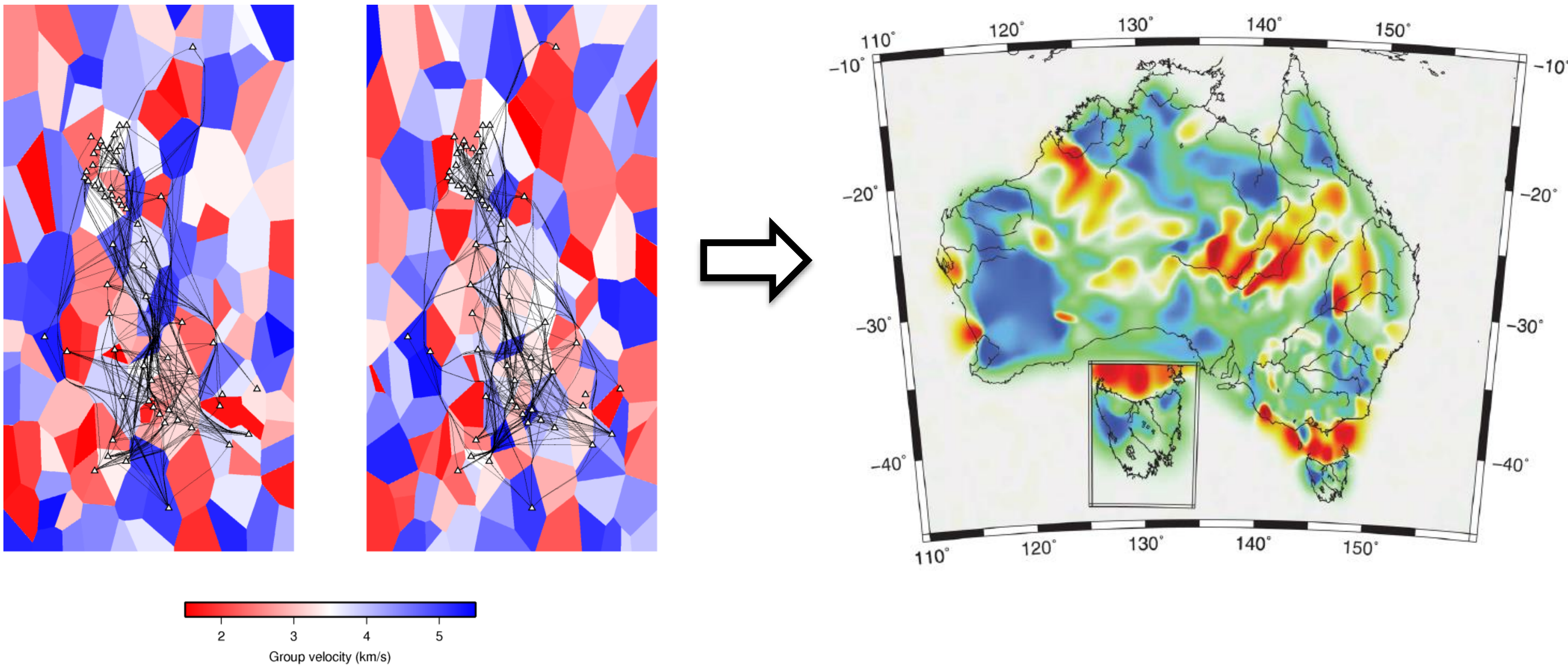
  – Extra parameter: Number of Voronoi cells for a 2-D physical model

Voronoi grid of seismic velocity



Transdimensional Tomography

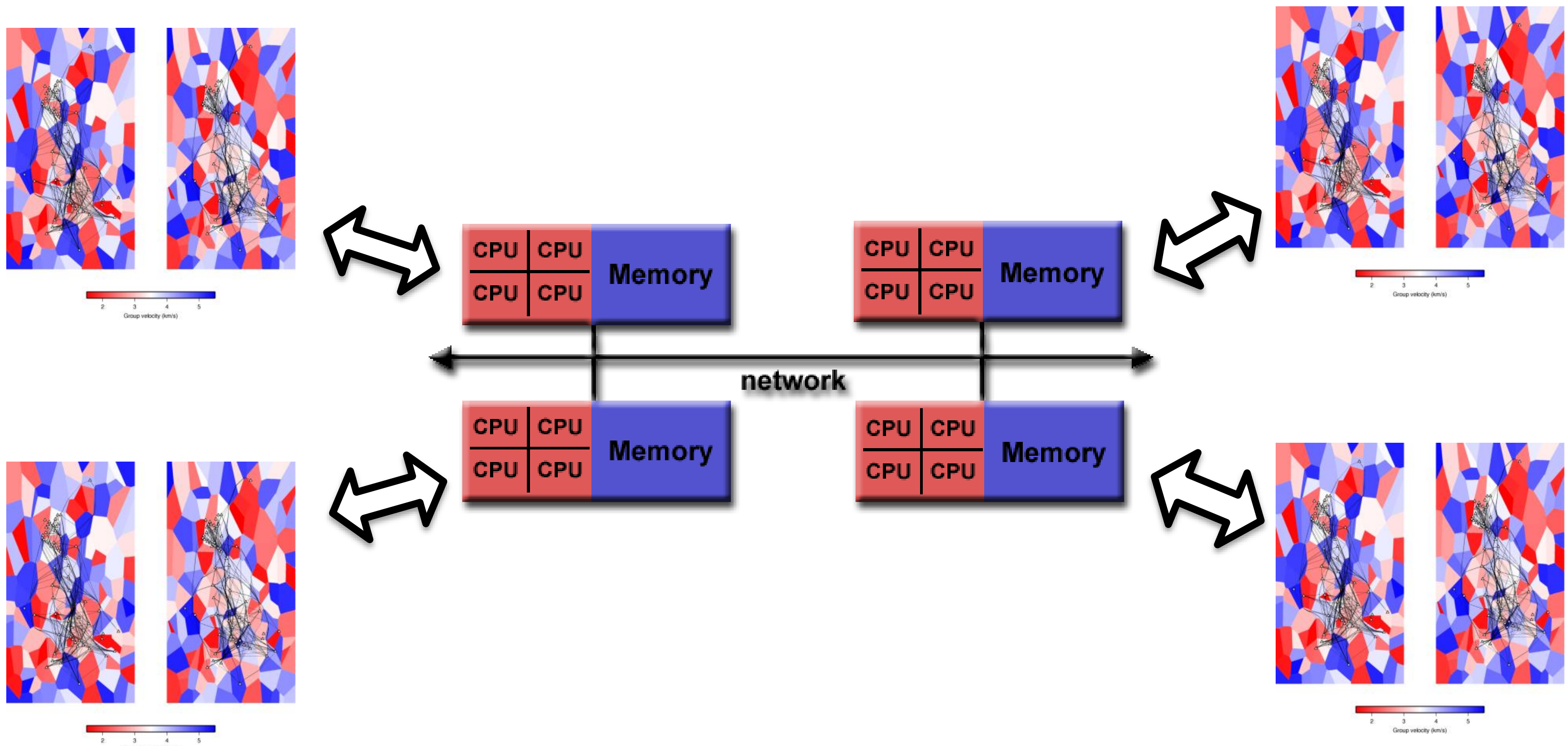
- Solutions as “ensembles” of accepted samples along MCMC chains.



Transdimensional Tomography

- Fast Marching Method is computationally expensive!

- Here we show how we used HPC solutions for increasing the performance of travel-time tomography.



Literature and history

- Green, P. J. (1995). *Reversible jump Markov chain Monte Carlo*.

  4403 Citations!

  – Transdimensional inverse theory.

- Sambridge, Bodin, Rawlinson, Gallagher et al.(2006-present)

  – Transdimensional Seismic Tomography, Geochronology, Hierarchical Bayes and Unknown data noise.

- Galleti and Curtis (2014)

  – MPI Non-linear Transdimensional Tomography: RJTOMO code.

Outline

- Transdimensional inversion in the Geosciences
- Preliminaries in Transdimensional Inversion
- Love-wave Tomography of the British Isles
- High Performance Computing Solutions
- Discussion

Outline

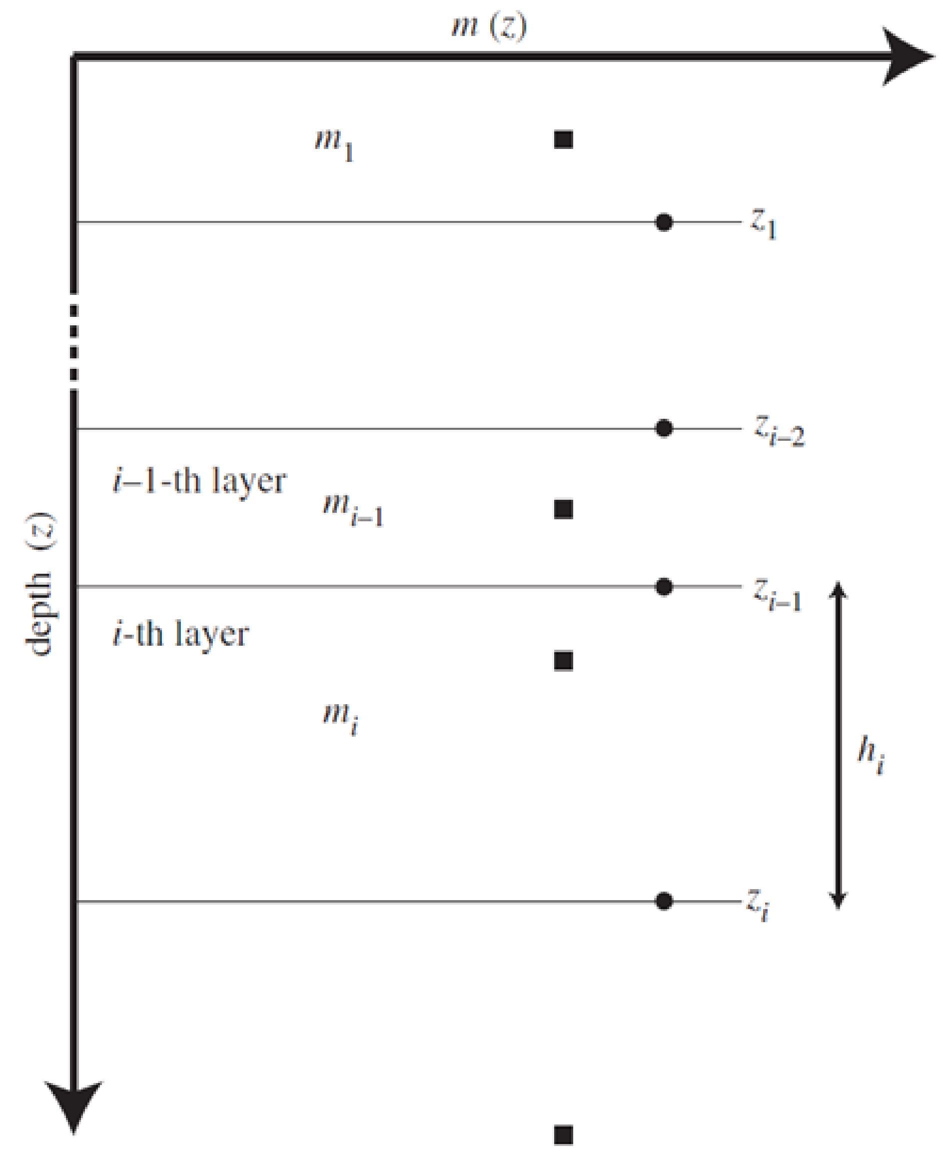
- Transdimensional inversion in the Geosciences
- Preliminaries in Transdimensional Inversion
- Love-wave Tomography of the British Isles
- High Performance Computing Solutions
- Discussion

Transdimensional inversion in the Geosciences

- Application: Rock's property identification in stratified media

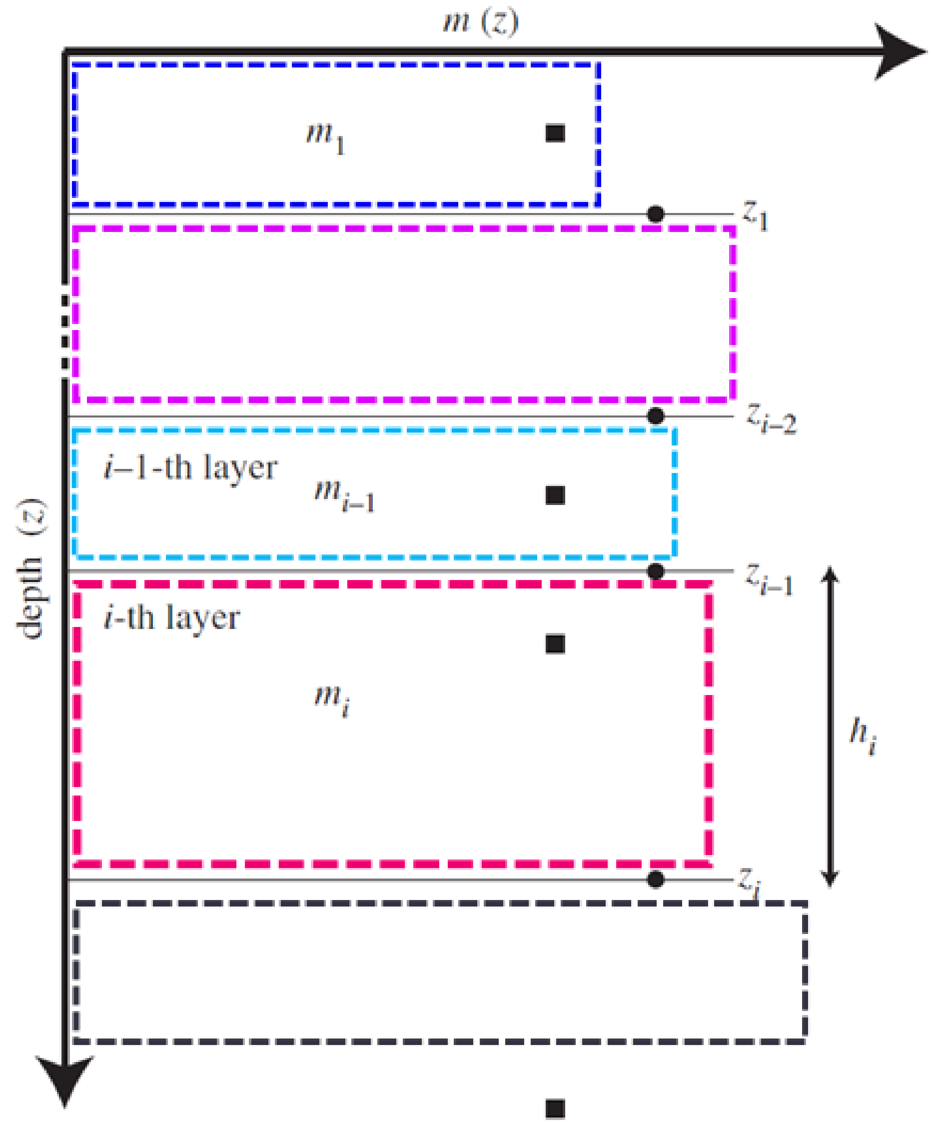
Transdimensional inversion in the Geosciences

- Parameter space dimension is unknown.
- Number of layers varies along inversion.



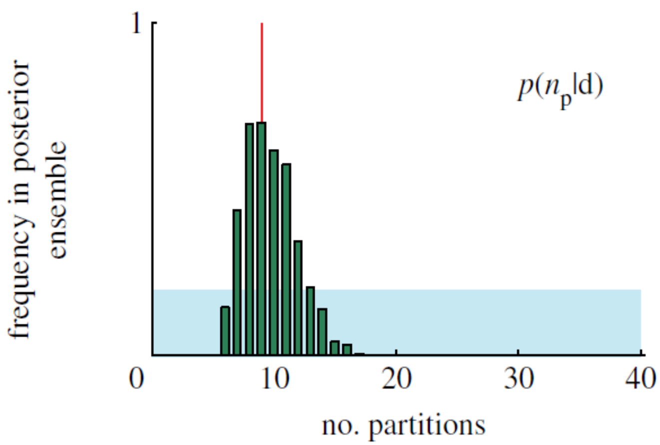
Transdimensional inversion in the Geosciences

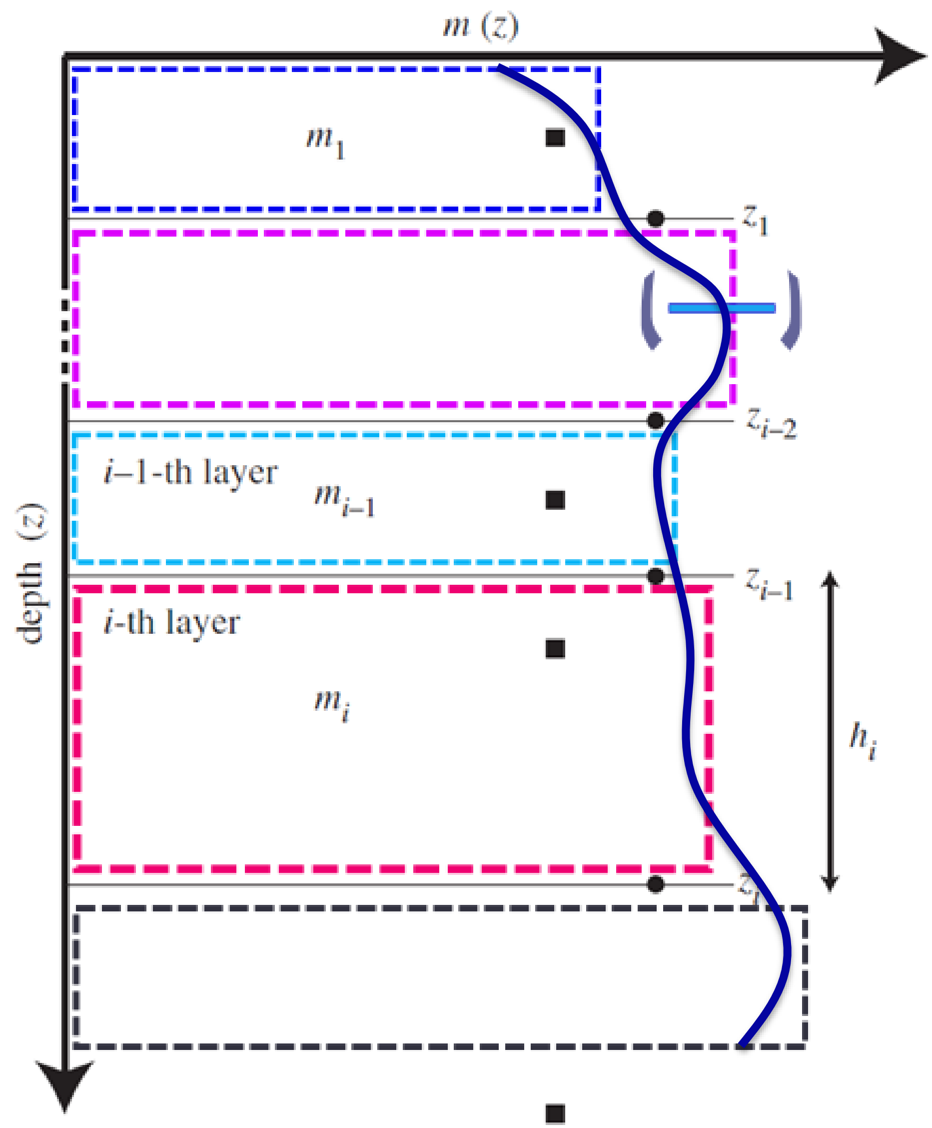
- Size of grid as well as the rock property in each layer is unknown.



Transdimensional inversion in the Geosciences

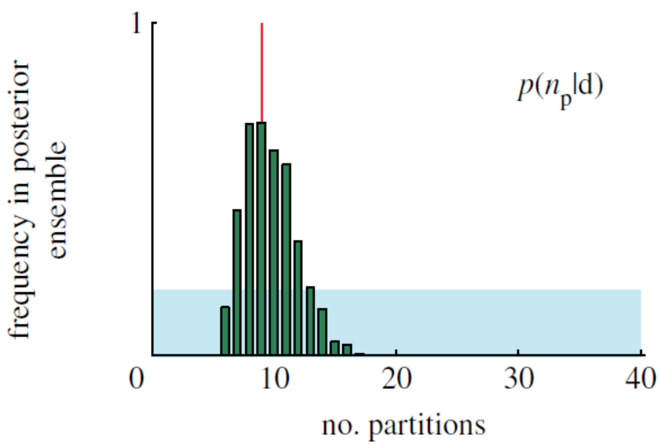
- Solution is built from the average of the accepted solutions.
- We obtain as well a posterior in the number of partitions.

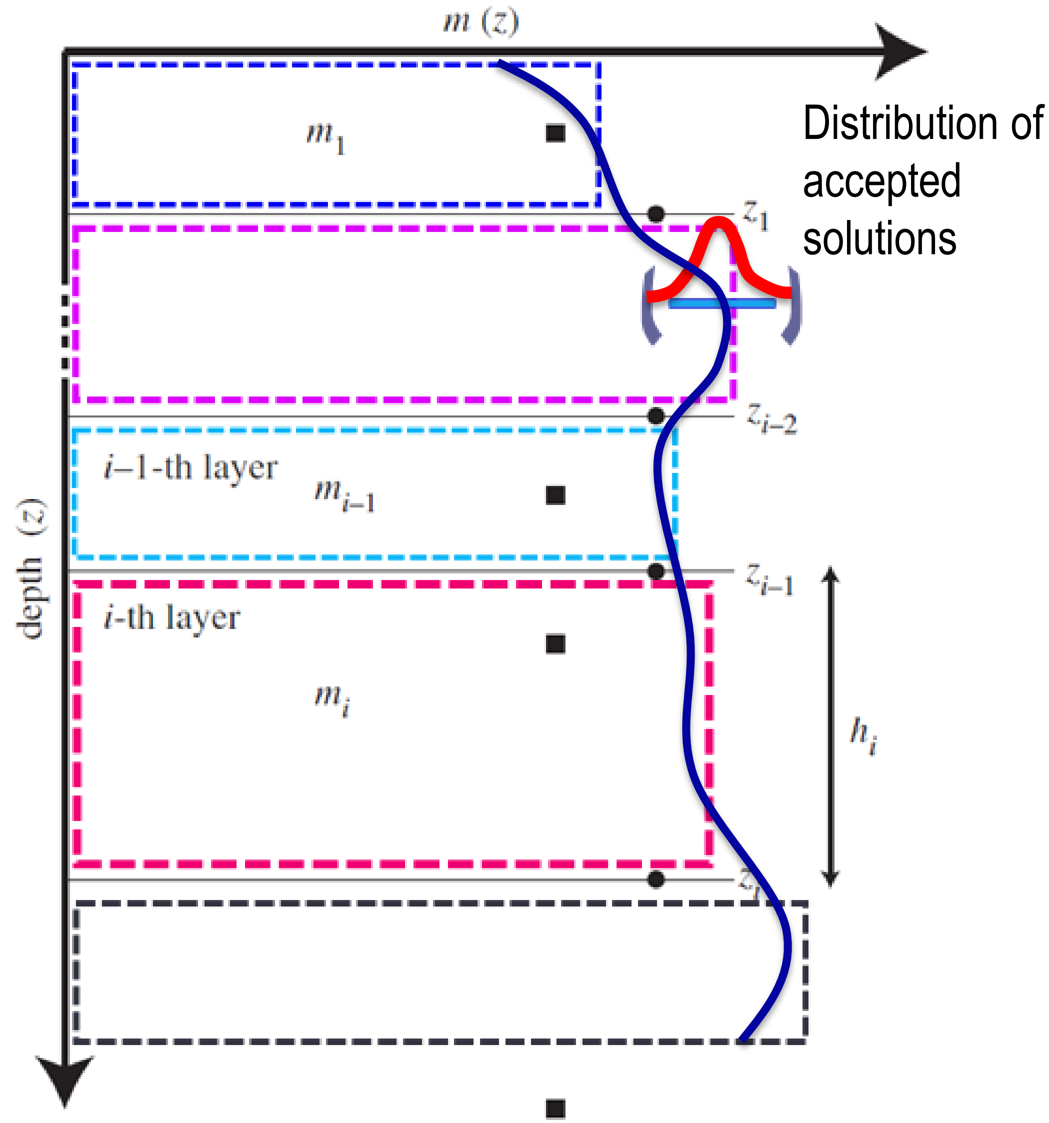




Transdimensional inversion in the Geosciences

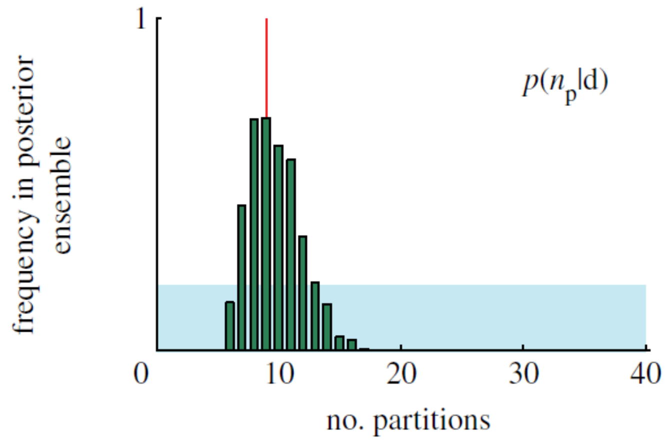
- Solution is built from the average of the accepted solutions.
- We obtain as well a posterior in the number of partitions…

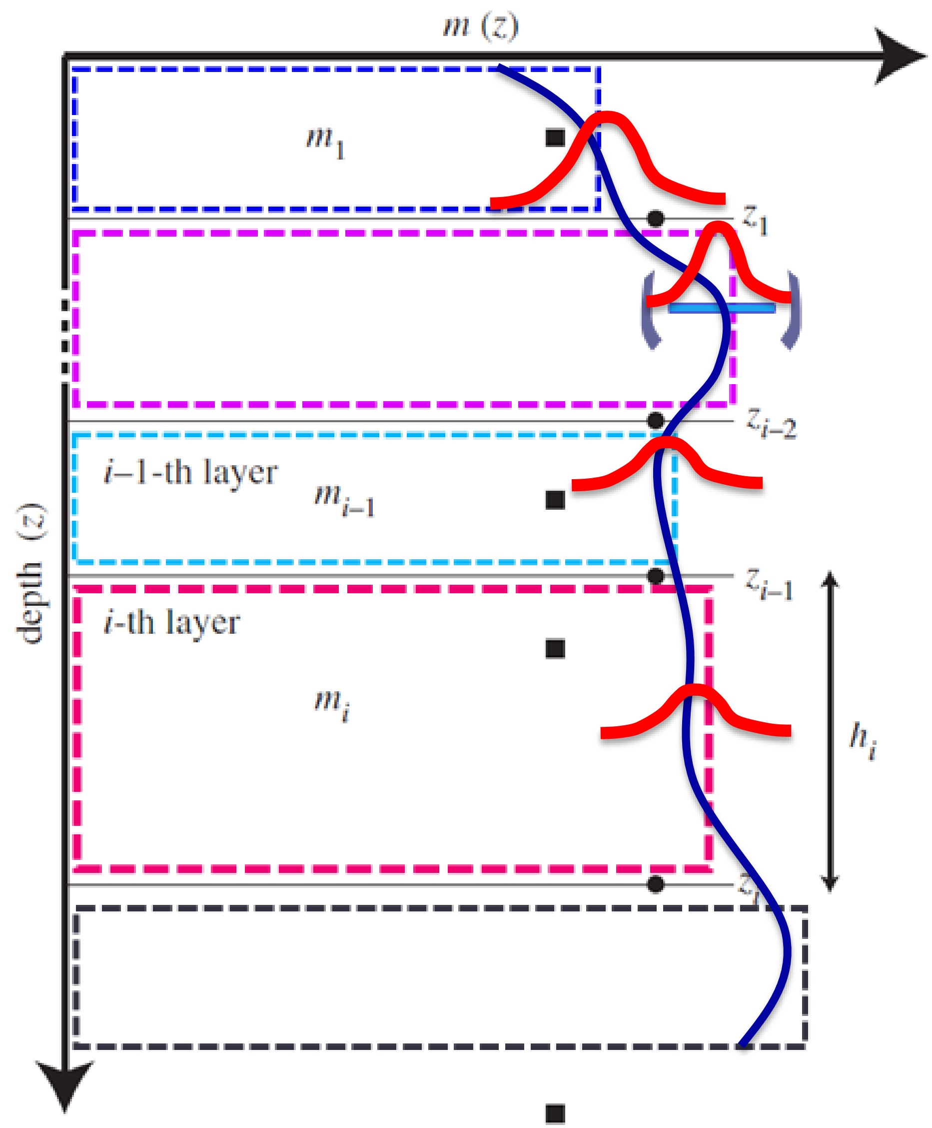




Transdimensional inversion in the Geosciences

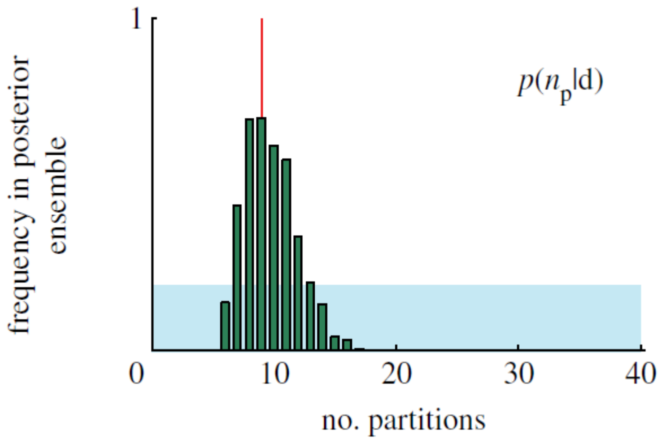
- Solution is built from the average of the accepted solutions.
- We obtain as well a posterior in the number of partitions…
- …and posteriors for all depths…

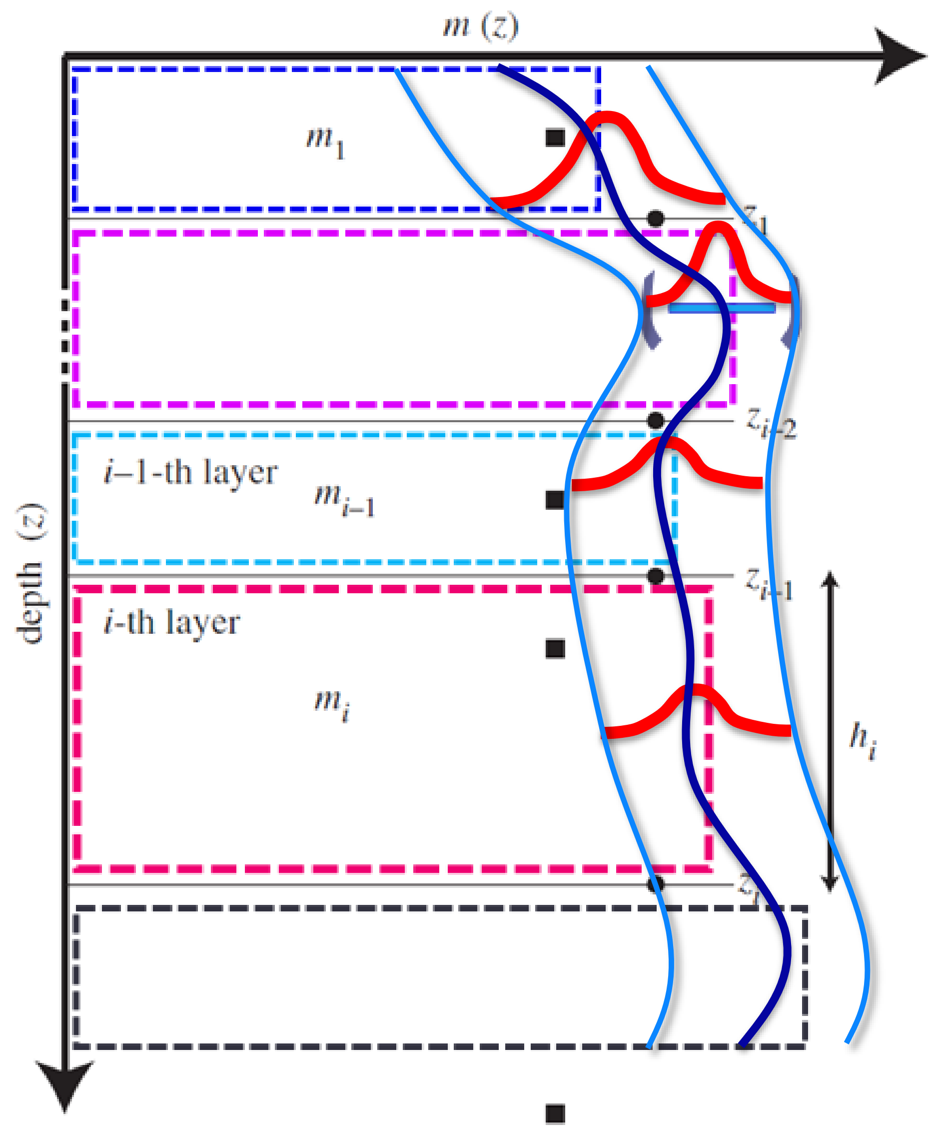




Transdimensional inversion in the Geosciences

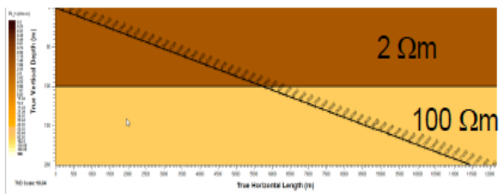
- Solution is built from the average of the accepted solutions.
- We obtain as well a posterior in the number of partitions…
- …and posteriors for all depths…
- i.e., the overall statistics.

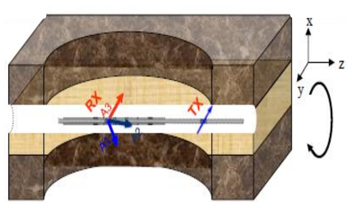


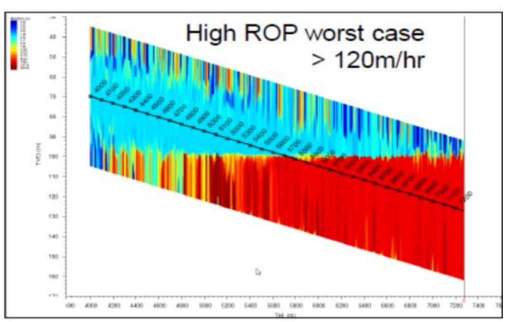


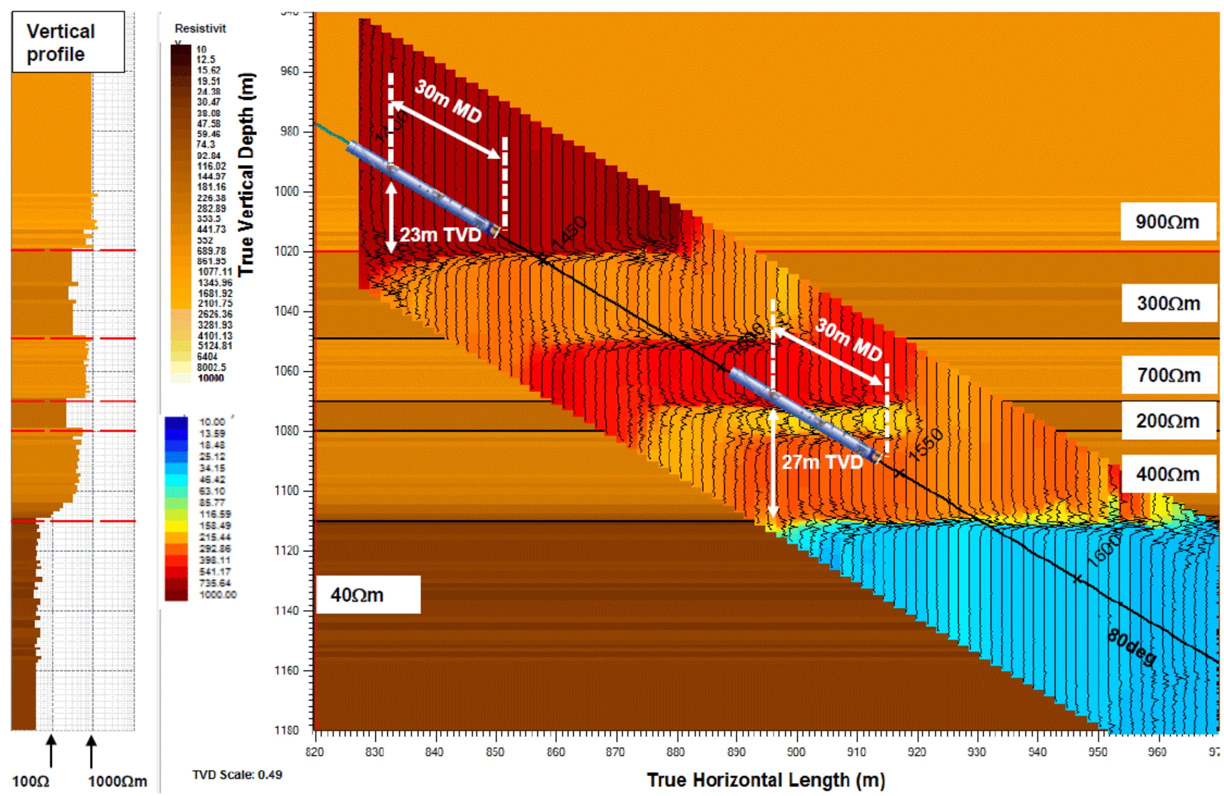
Well Placement: Direction Resistivity example

- Inversion done in 1D parameter space, at each depth.
- Fixed noise, but calibrated for acceptance ratio above 35%.









Outline

- Transdimensional inversion in the Geosciences
- Preliminaries in Transdimensional Inversion
- Love-wave Tomography of the British Isles
- High Performance Computing Solutions
- Discussion

Transdimensional Math Background

“Conventional” Bayes:

$$p(\mathbf{m}|\mathbf{d}) = \lambda p(\mathbf{d}|\mathbf{m})p(\mathbf{m})$$

$$p(\mathbf{m}|\mathbf{d}, k) \propto p(\mathbf{d}|\mathbf{m}, k)p(\mathbf{m}|k)$$

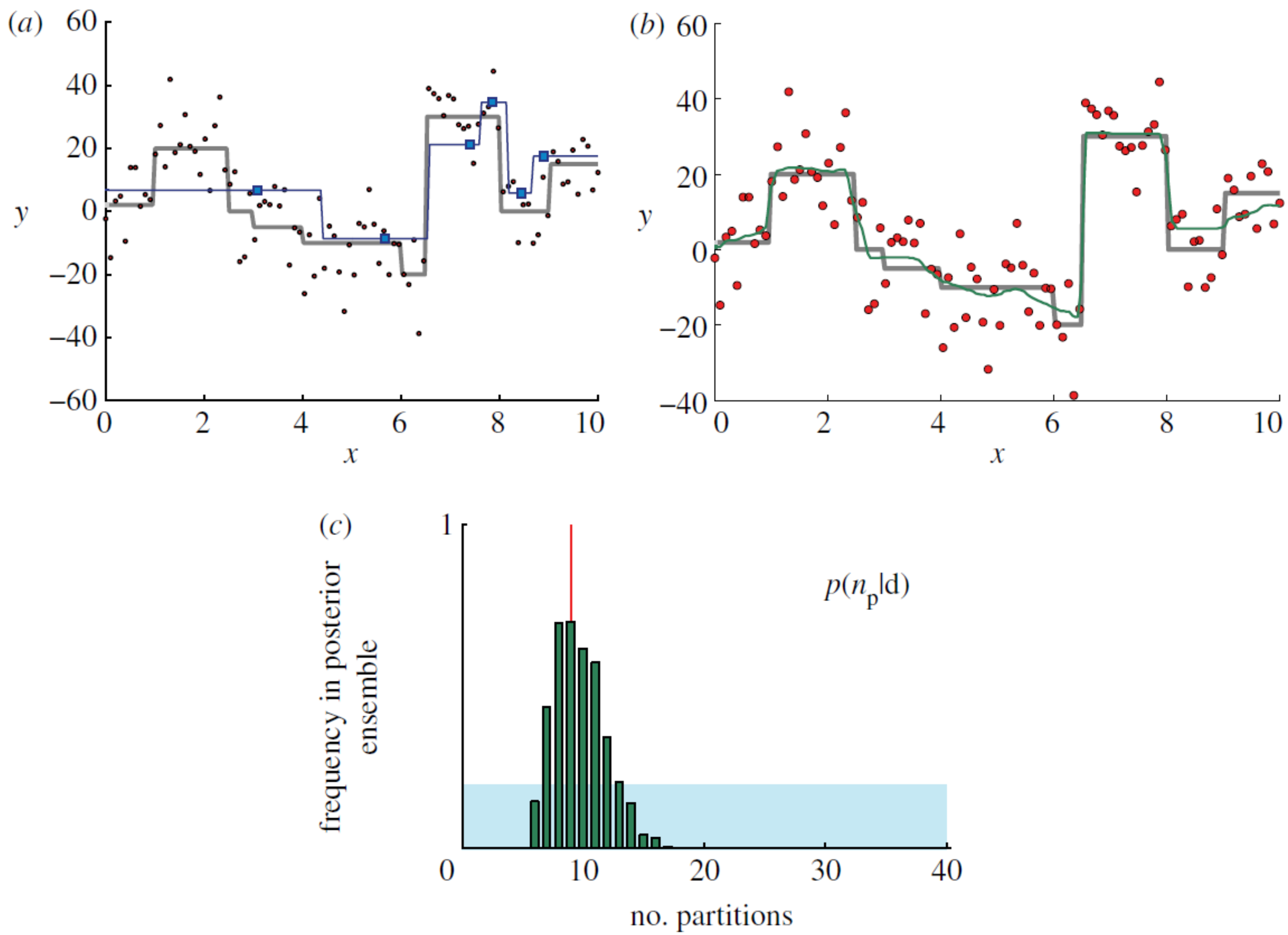
Transdimensional (Hierarchical Bayes):

$$p(\mathbf{m}, k|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m}, k)p(\mathbf{m}|k)p(k)$$

$$p(\mathbf{m}, k|\mathbf{d}) = p(\mathbf{m}|\mathbf{d}, k)p(k|\mathbf{d})$$

$$p(k|\mathbf{d}) = \int p(\mathbf{m}, k|\mathbf{d})\, \mathrm{d}\mathbf{m}$$

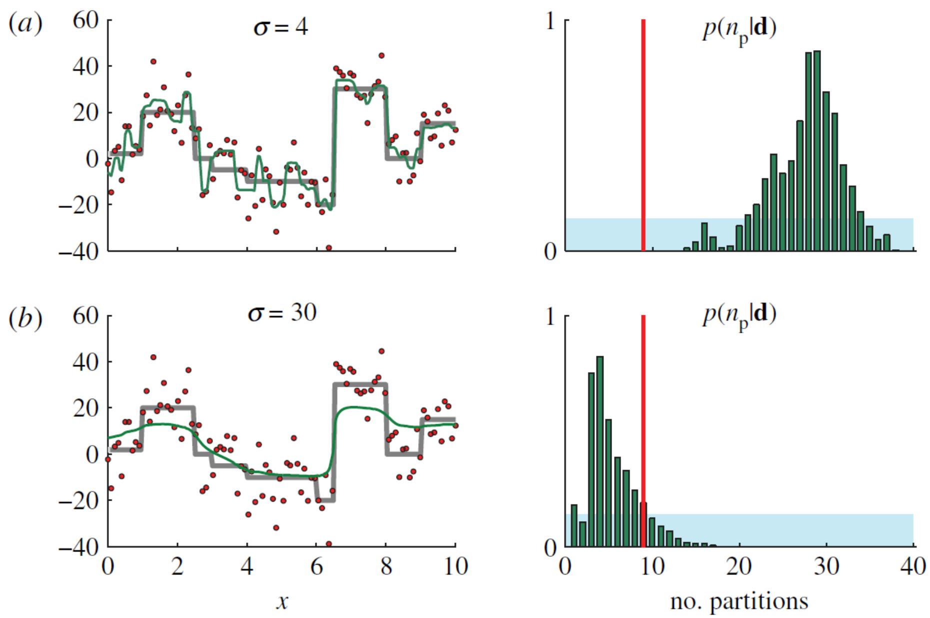
Discretization as a variable in Monte-Carlo

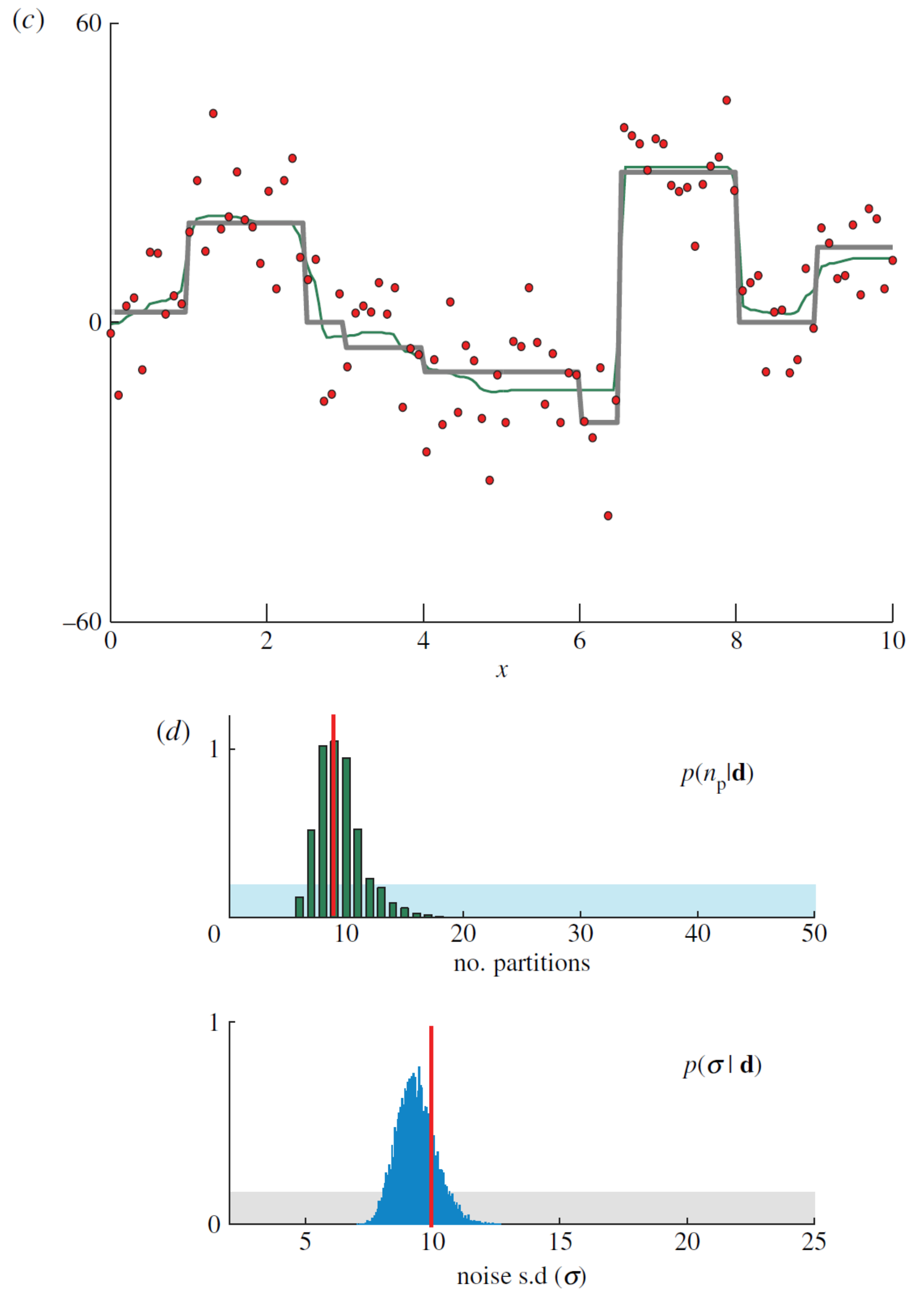


What if the data noise is another unknown

- How is the noise solved for in the solution of the inverse problem?
- Fixed vs variable noise.
- Mathematical principles of Hierarchical Bayes.
- Other filters, and advantages of “not knowing” noise.

Unknown data noise





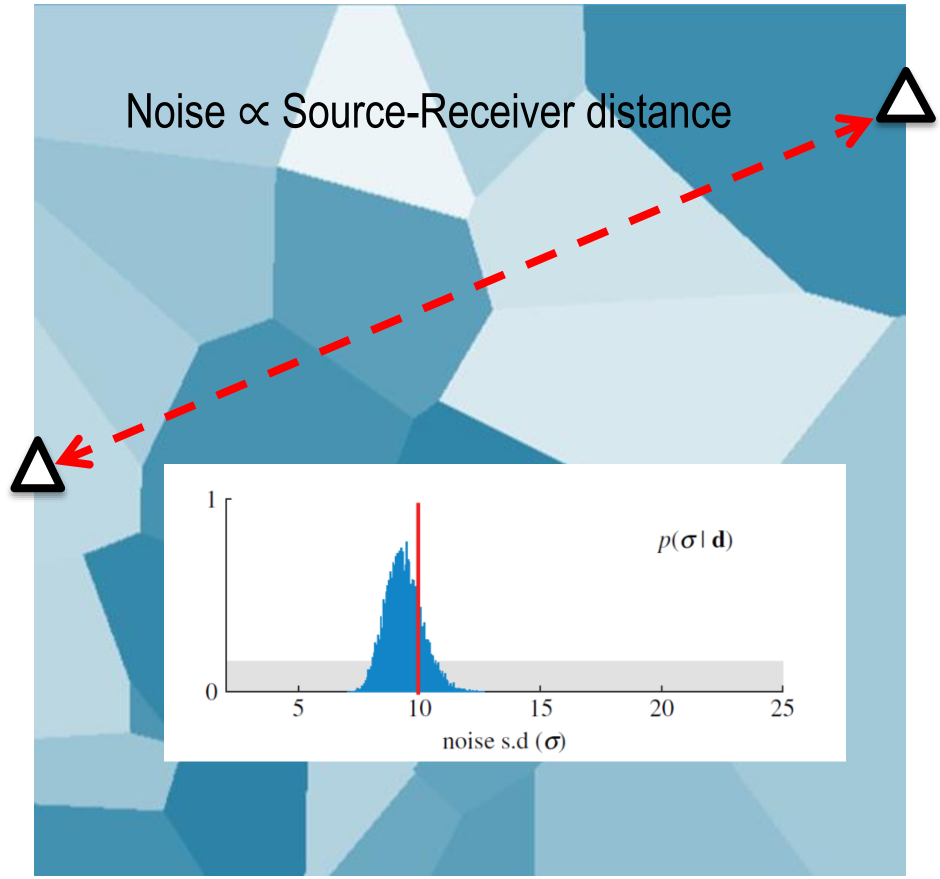
Transdimensional Seismic Tomography

Problem Setup: Data

- Setup hypotheses.
  - Number of Voronoi Cells
  - Coordinates of seismometers
  - Travel-times.
  - Other considerations

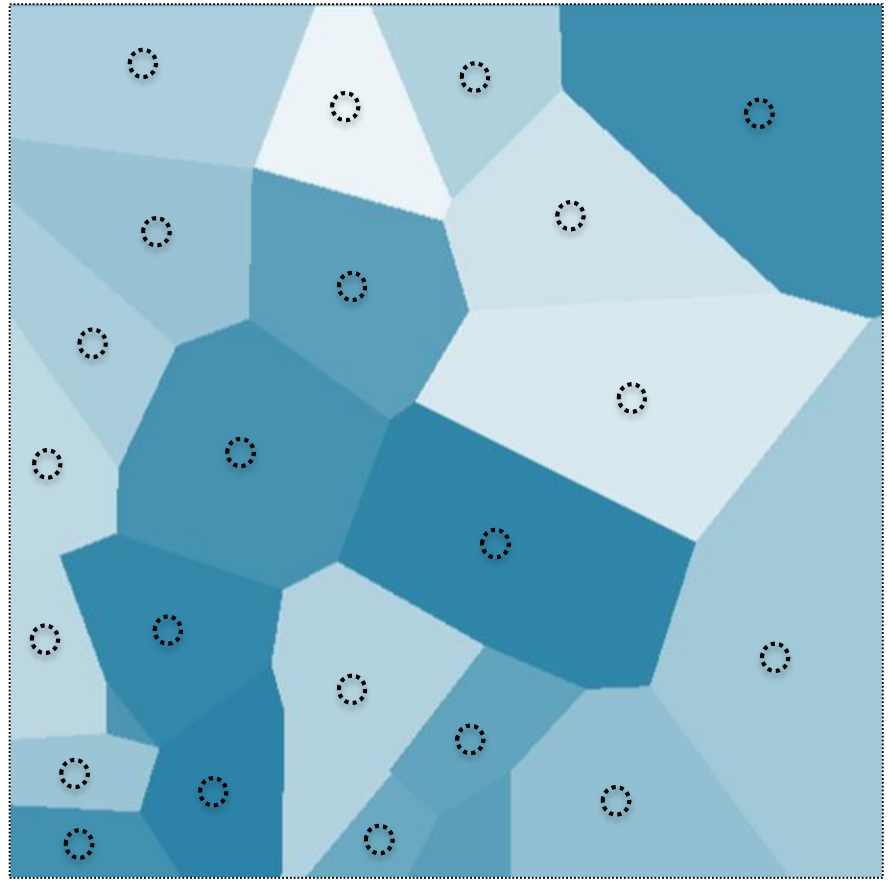
Problem Setup: How is the noise treated and included.

- Hierarchical Bayes in the Transdimensional Tomography case.
- Levels of noise:
  - Perturbations
  - Noise prior
  - Tool noise, measurement noise, line non-linear, associated to source-receiver pairs, and source-receiver distances.
- How is the noise posterior calculated?



Problem Setup: the Markov Chain

- Objective: invert for Voronoi grid slowness, number and location of grid nodes, and data noise.
- New samples of Markov Chain created with new velocity values, by creating new nodes, or by deleting nodes.
- Misfit is calculated with ray paths traveltimes.
- (TODO: new nodes or nodes deletion in figure, change velocity values).

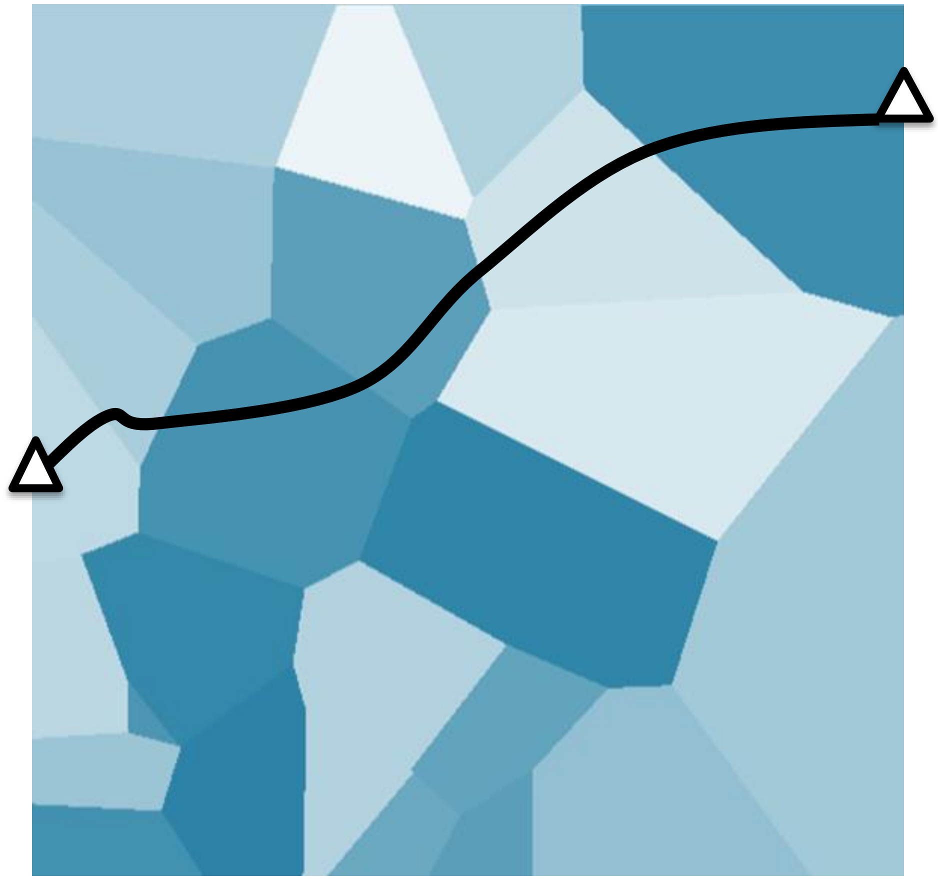


First arriving travel-times through FMM

- FMM is used for calculating first-arrival travel-times through a 2-D continuous velocity medium in spherical shell coordinates.

- We interpolate the Voronoi grid velocity values with B-splines in order to use FMM.

- Inter-seismometer ray path geometries are updated from fine grid of FMM trave times, and the corresponding ray path travel-times are calculated.

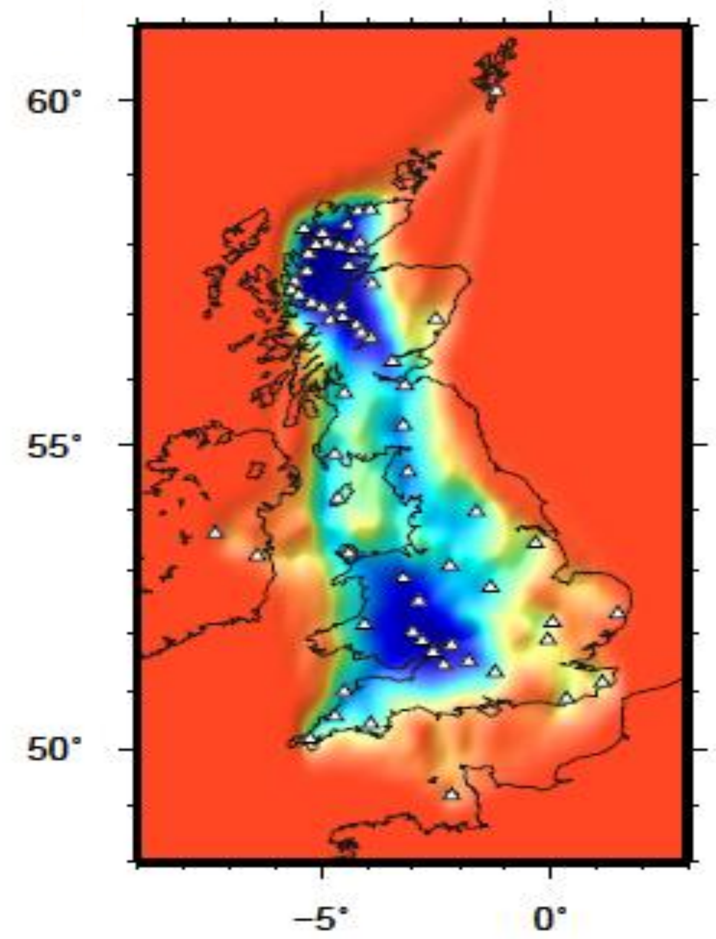


Outline

- Transdimensional inversion in the Geosciences
- Preliminaries in Transdimensional Inversion
- Love-wave Tomography of the British Isles
- High Performance Computing Solutions
- Discussion

Tomography of the British Isles

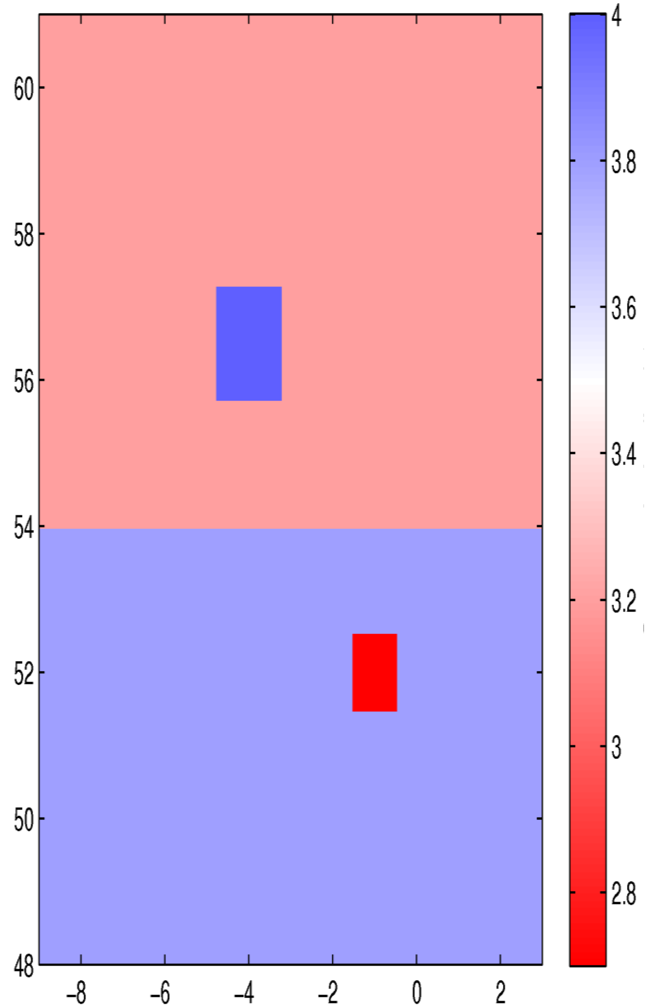
- Seismometer disposition.
- Retrieval of Love-waves traveling between pairs of seismometers calculated from ambient noise using Greens function approach (explain how).
- Authors: Galletti, Curtis, Angelo Meles. University of Edinburgh.
- (TODO: Change figure).

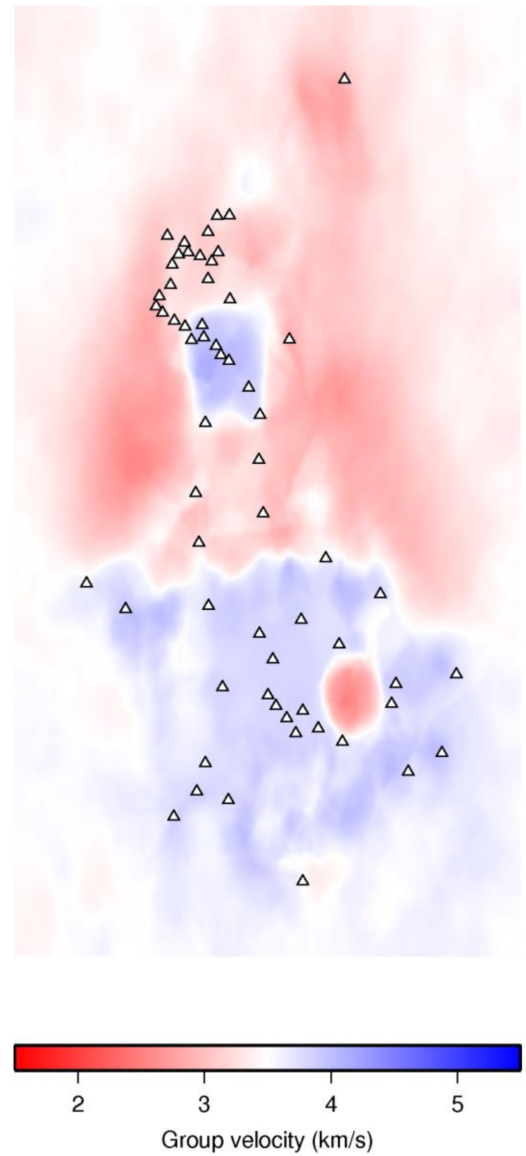


The Love-wave Synthetic Tomography

- Synthetic case design

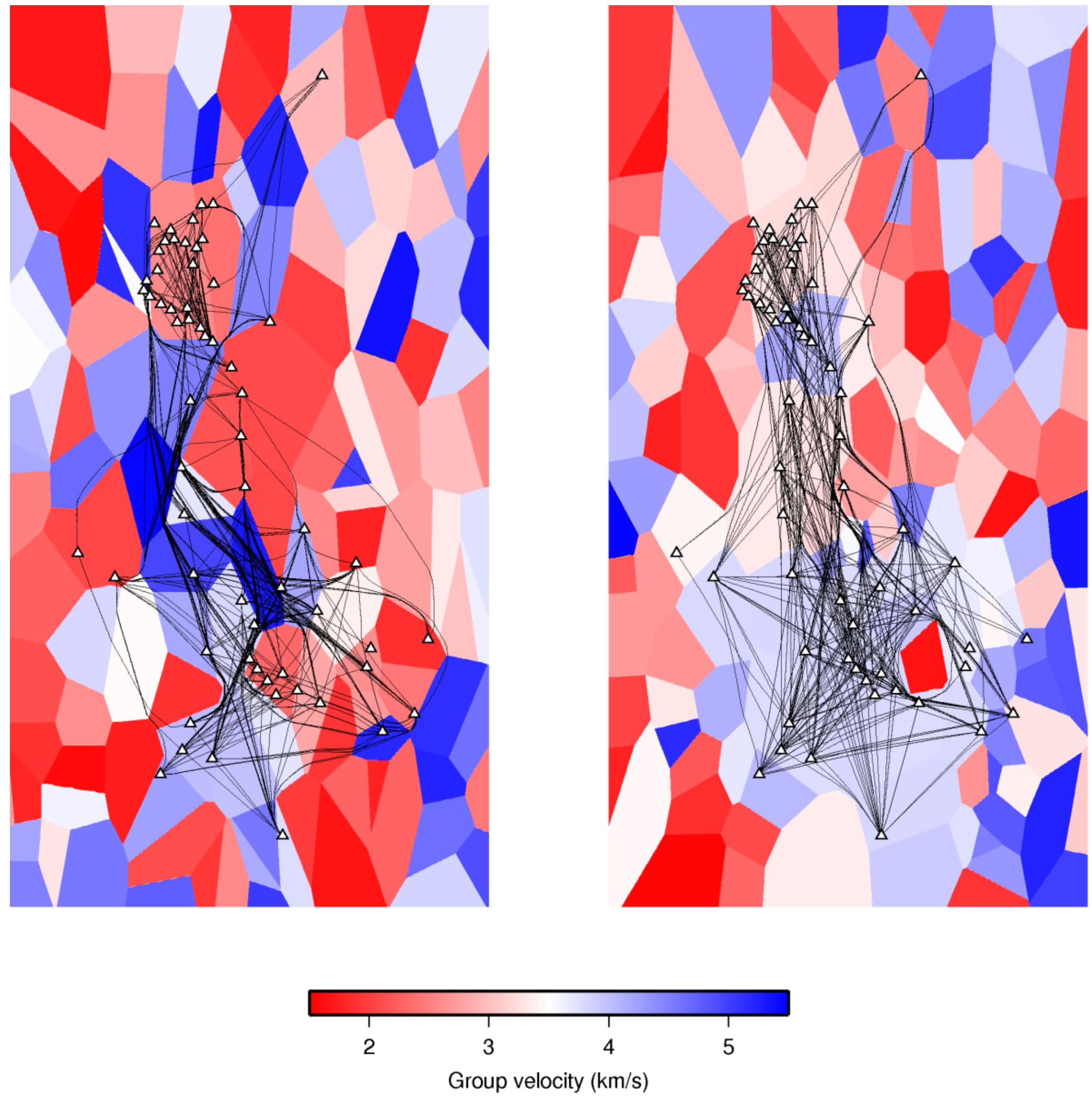
- We ran synthetic seismic velocity case keeping geographical dispositions of the seismometers.





Voronoi grid Geometry and Raypath Geometry update

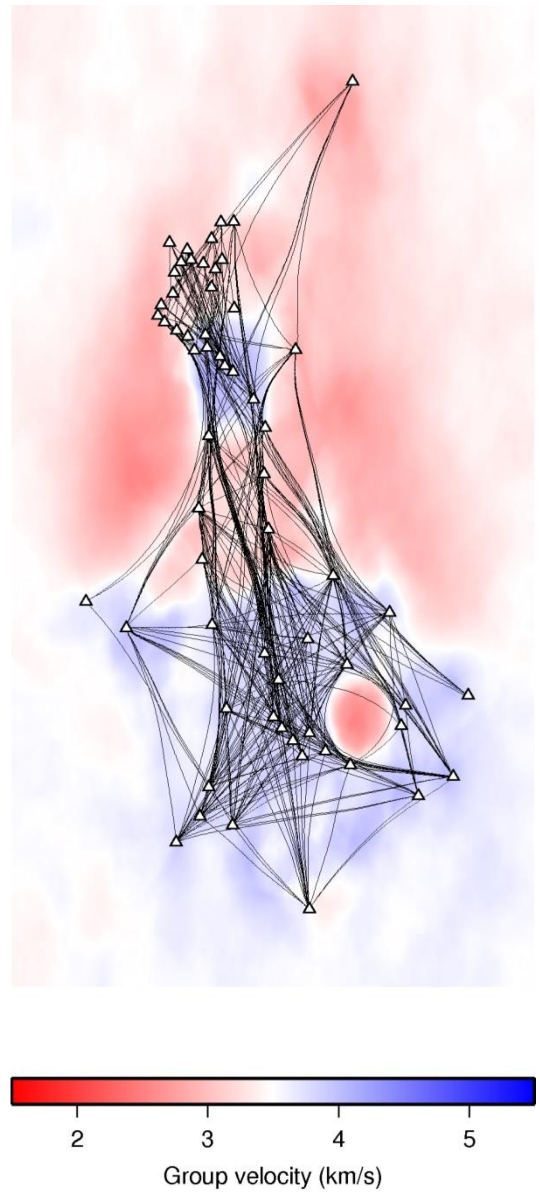
- Raypaths are updated at each step of the Markov chain, and Ray paths travel-times calculated.
- Voronoi grids of low complexity: [100, 300] cells, or high complexity [300-600] can be chosen.
- Figure: First and Last samples of Markov Chain Monte Carlo.

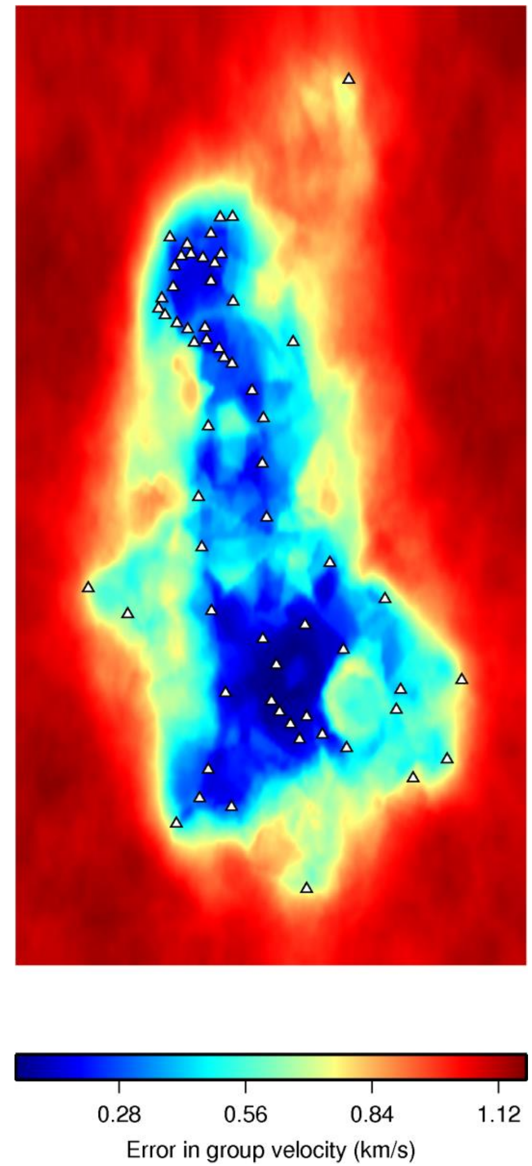


Results

- Average along the accepted samples across all chains is a statistical solution of the inverse problem.

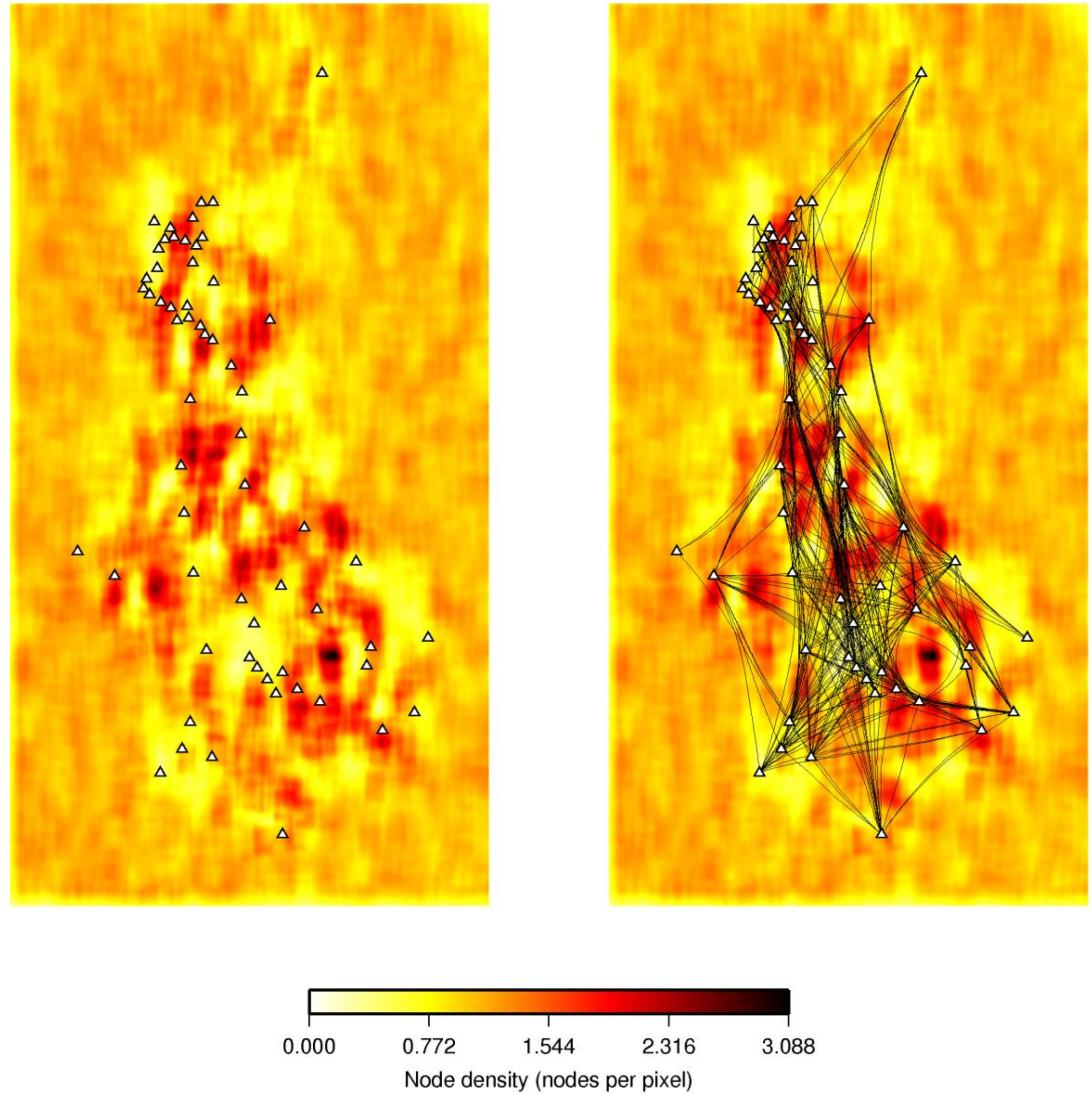
- Uncertainty can be evaluated along the accepted samples, across all chains.





Resolution of Voronoi nodes

- The transdimensional approach "resolves" the node density according to the seismometer sampling!

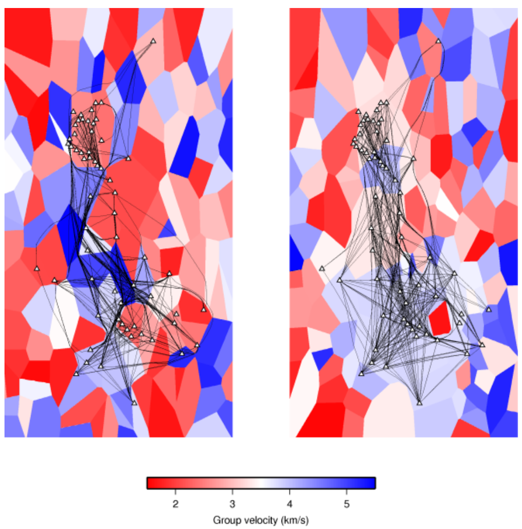


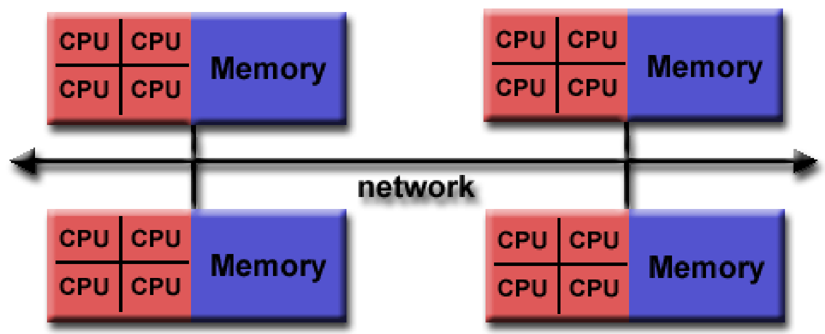
Outline

- Transdimensional inversion in the Geosciences
- Preliminaries in Transdimensional Inversion
- Love-wave Tomography of the British Isles
- High Performance Computing Solutions
- Discussion

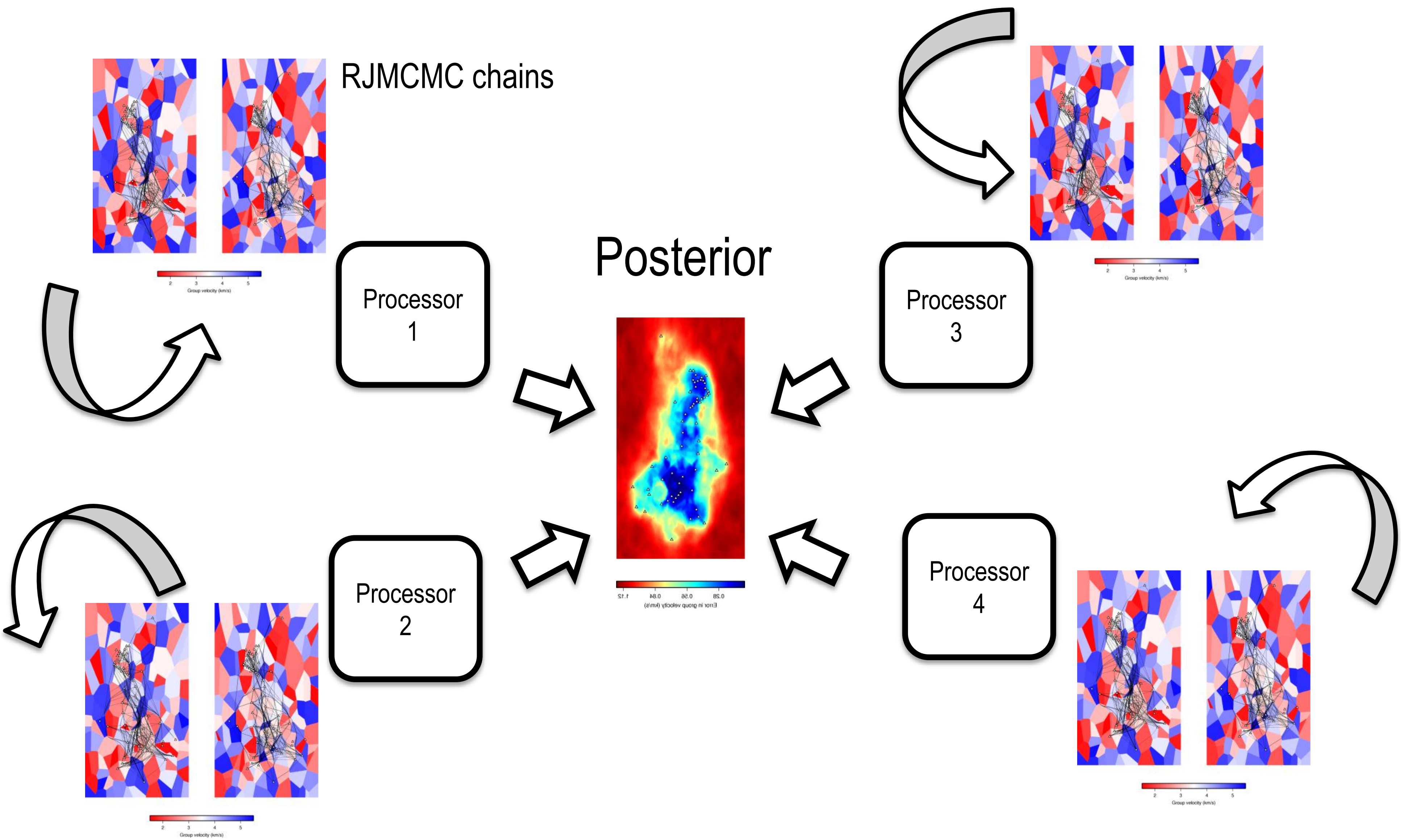
MPI architecture 1

- MPI design 1: Galletti, Curtis. We ran synthetic case design.
- Several Markov Chains (up to 16) run in parallel MPI process.
- No interaction between chains except for Message-blocking MPI commands, e.g. error messages.
- FMM is run (serially) for ALL seismometer-pairs at each MCMC samples, for EACH chains.



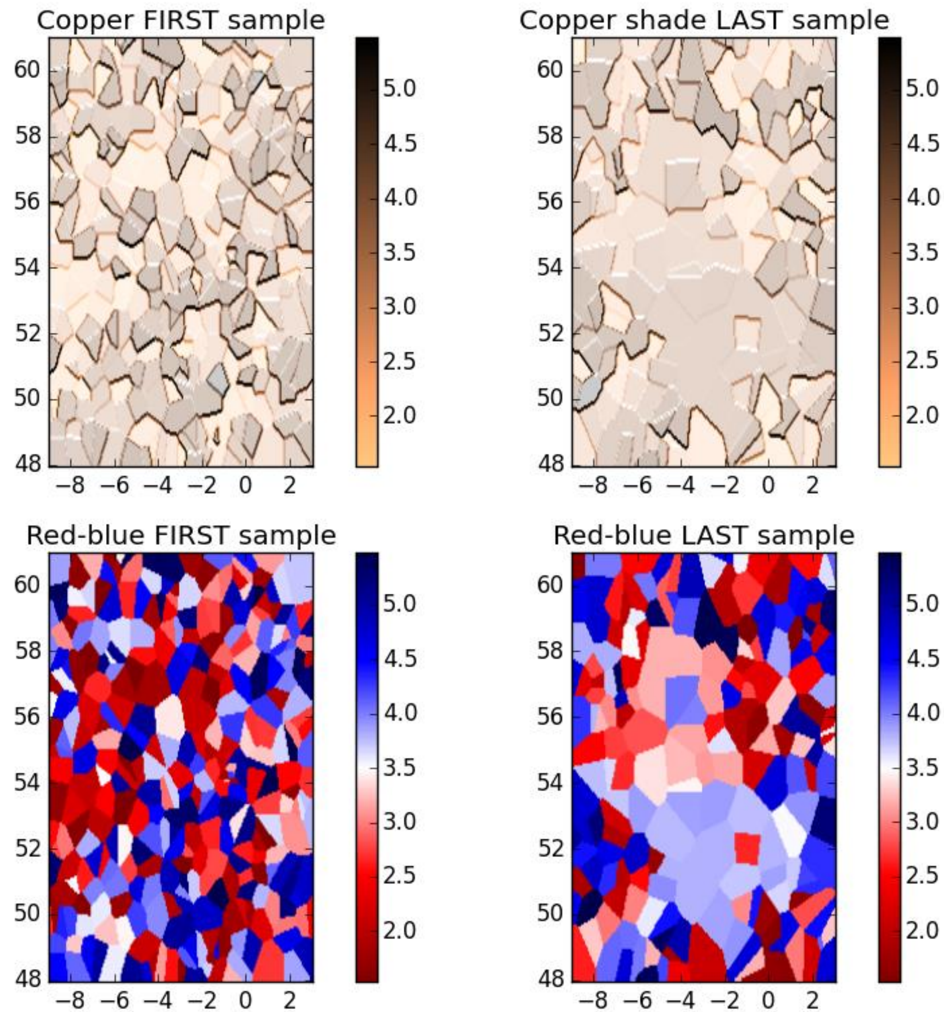


MPI architecture 1



MPI architecture 1: Expensive!

- Markov chain: 100,000 samples.
- 12 chains, acceptance, approx. 44%
- Time: approx. 24 hours
- Computationally expensive!



Computing time

- Serial code = 24 hours (Galdor) per 100,000 samples.

MPI architecture 1: Profiling the code

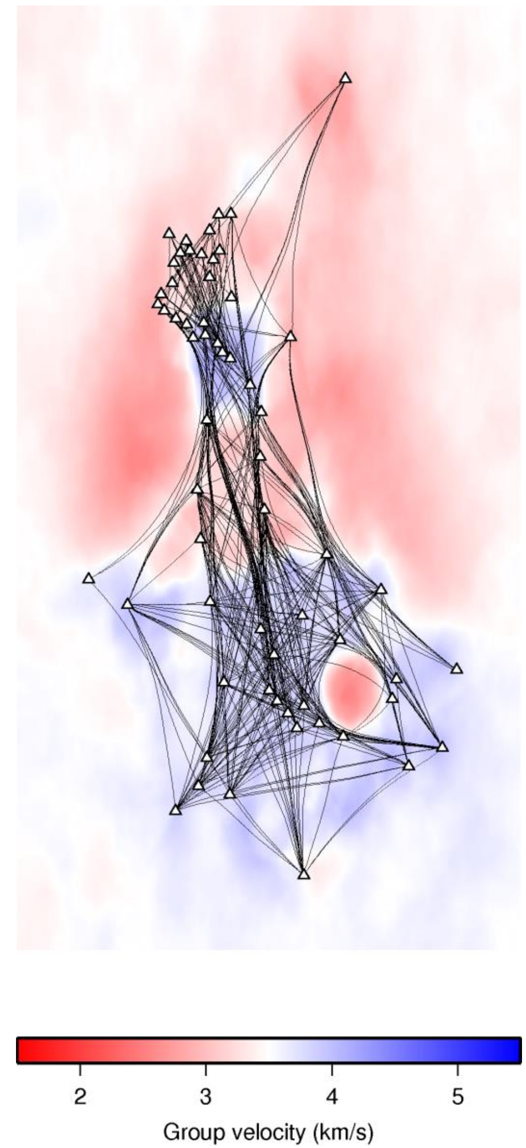
- **Code profiling results: TODO table of results.**

- Ray modifier (i.e. FMM + raypath tracers + raypath geometry calculation) spends more than 90% of code time.

- Why is the raypath modifier the perfect candidate for MPI parallelization? Is this the only candidate?

Increasing performance with HPC

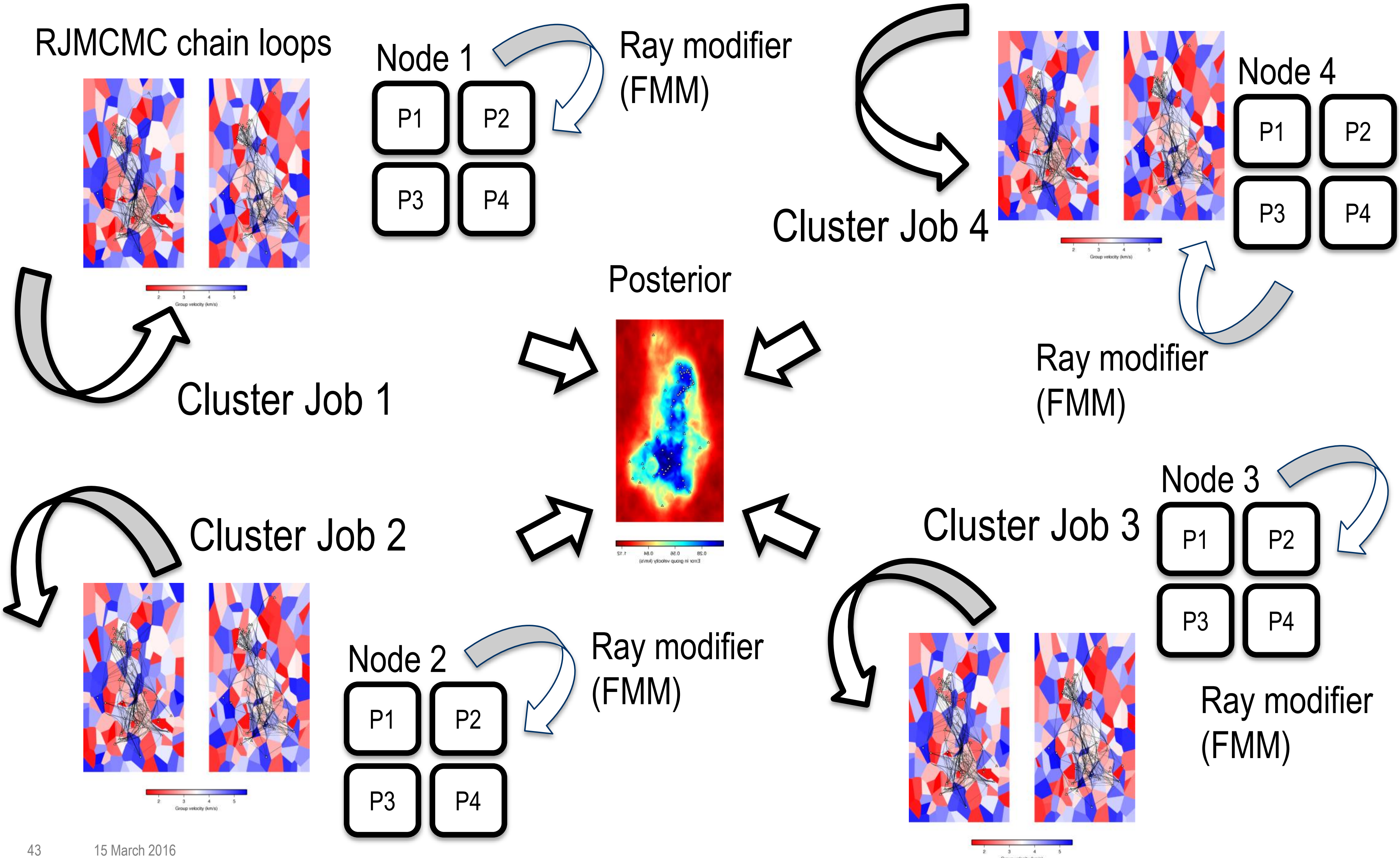
- Ray modifier is computationally expensive!
- FFM can be parallelized: Examples from the literature for HPC applications to seismic tomography?
- Proposal: HPC solutions through hybrid embarrassingly parallel Markov Chains and MPI at the ray-modifier level!



New MPI design

- Hybrid Embarrasingly Parallel Process and internal FMM and Ray modifier MPI loop.
- Candidates: places where an update in the ray-paths is used in the inverse problem solution.
- FFM can be divided for sub-groups of seismometer pairs.
- Size of group varies according to number of nodes available in the Cluster.
- Size of sub-array= Number of pairs MOD(number of nodes).
- Our case: number of pairs = 61. Number of nodes = 2 to 16 (up to 32 but not good for SGR-HPC group.
- Number of processors per node: approx 8.

MPI architecture 2: Faramir and Galdor

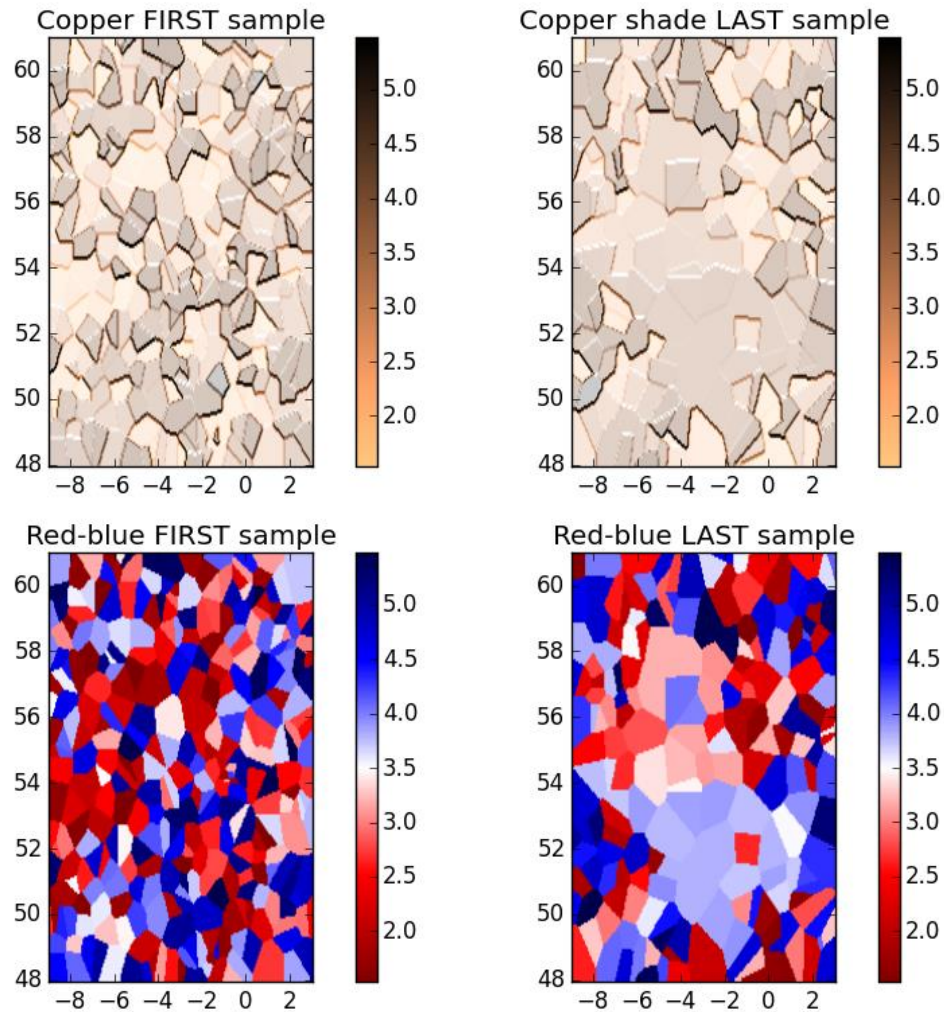


Outline

- Transdimensional inversion in the Geosciences
- Preliminaries in Transdimensional Inversion
- Love-wave Tomography of the British Isles
- High Performance Computing Solutions
- Discussion

Speed-up comparison.

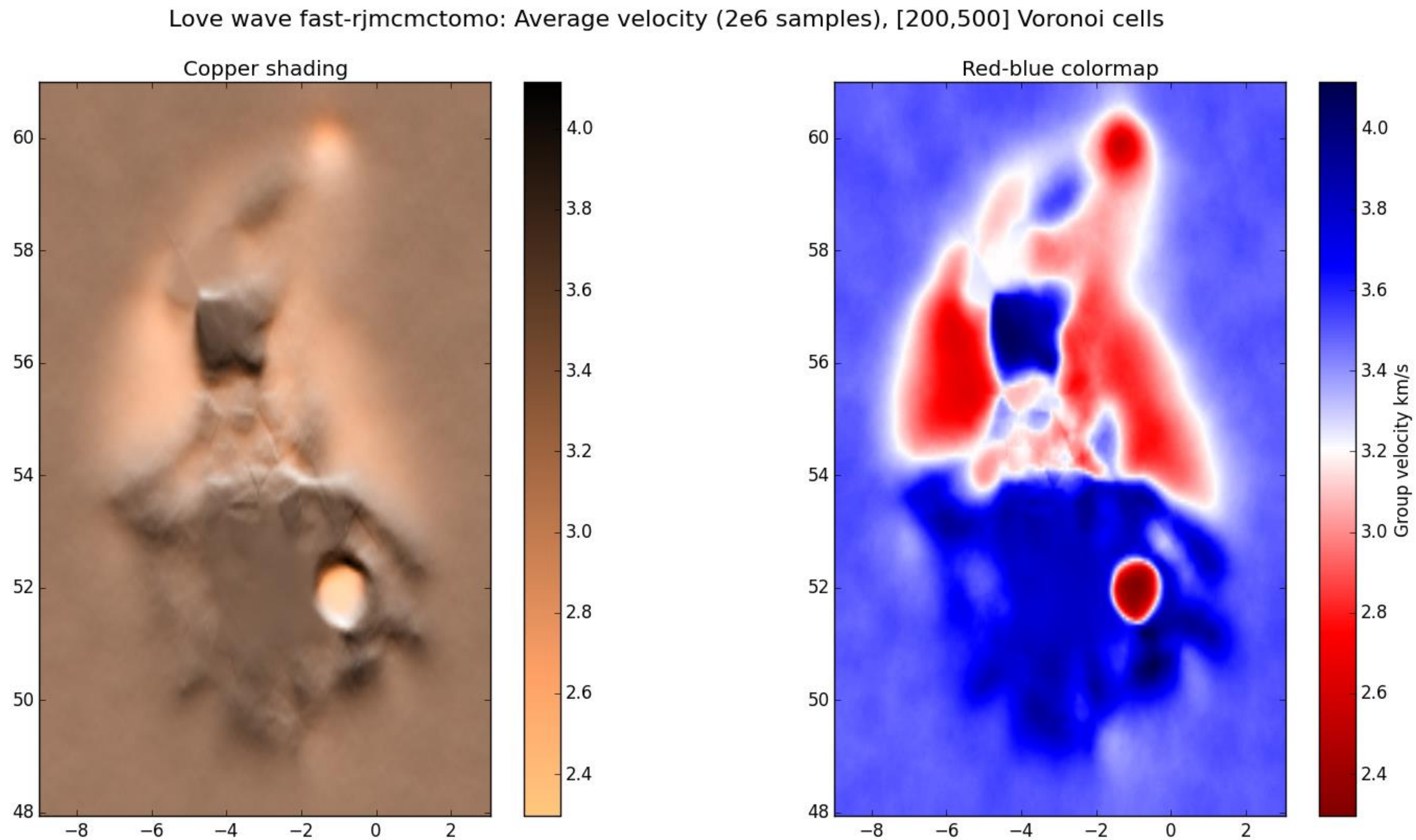
- Markov chain: 2 million samples.
- 16 chains, acceptance ~ 38%
- Fast computation!
- TODO: Table or graph with speed-up.



Computing time

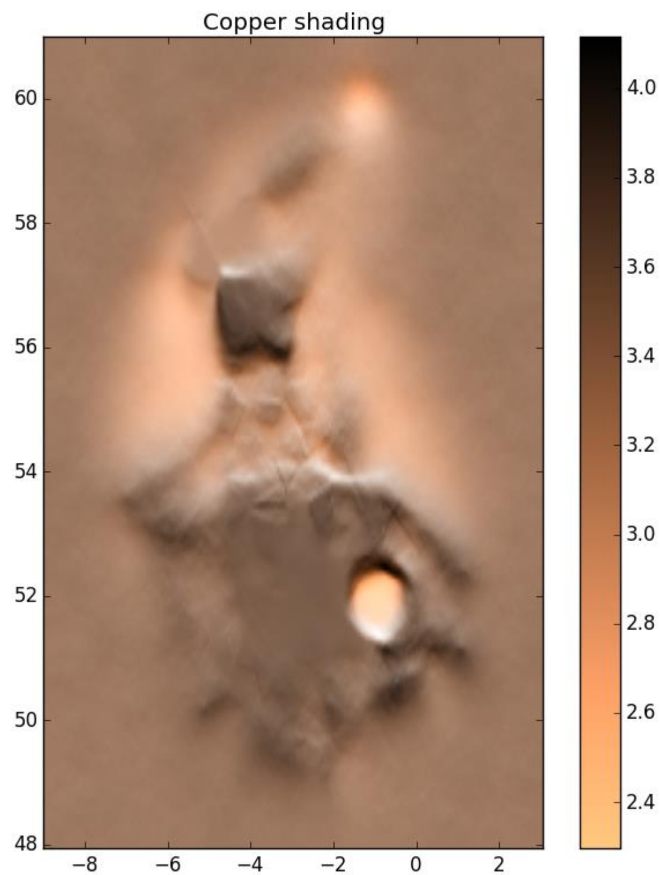
- **(before)** Semi-serial code = 480 hours for 2 million samples.
- **(after)** MPI Raytracer 24 processors, 2 million samples = 33 hours (Galdor).

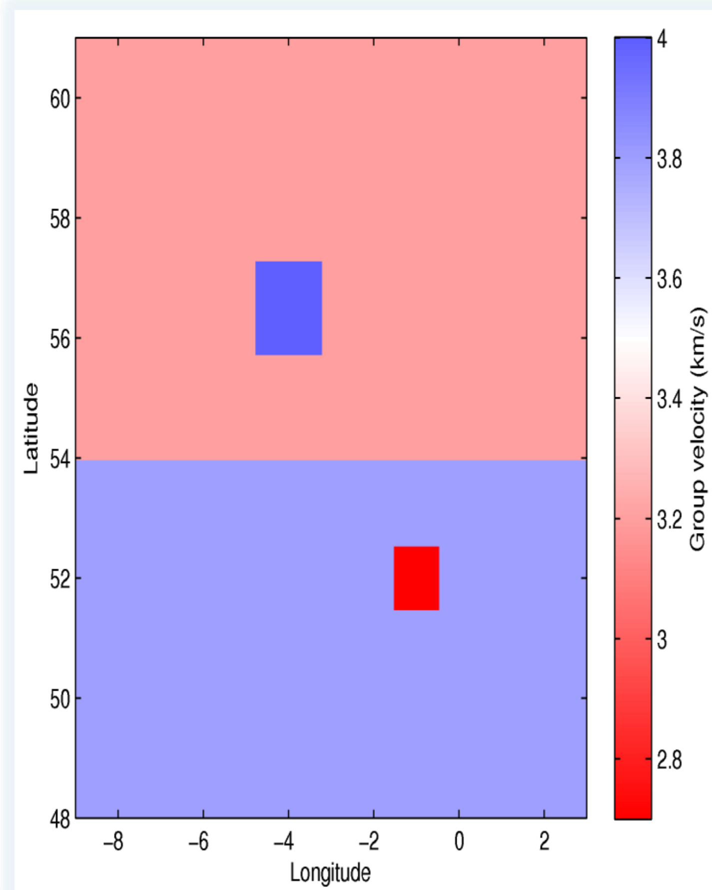
Results for Synthetic Case Average seismic velocity



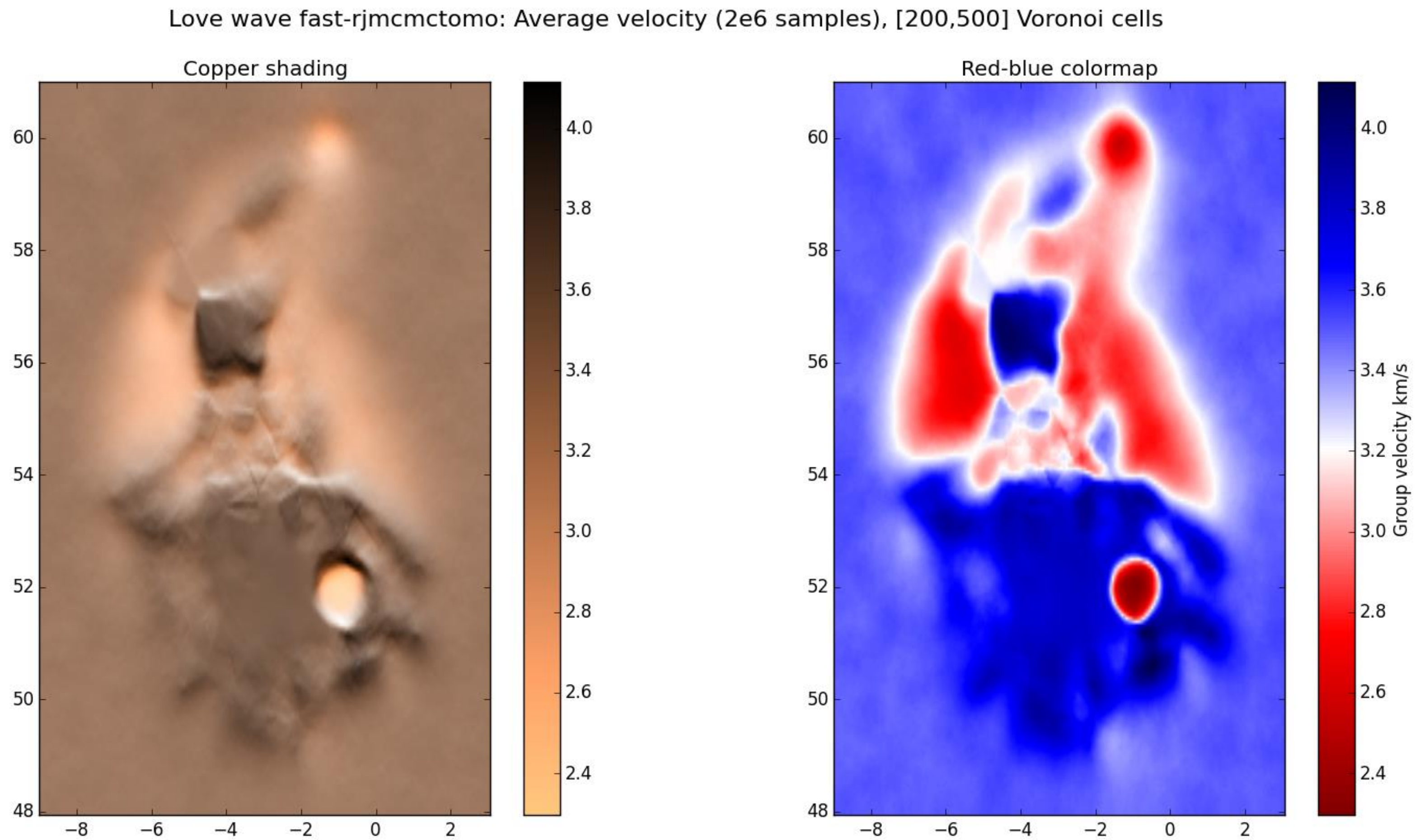
Average seismic velocity

Love wave fast-rjmcmctomo: Average velocity (2e6 samples), [200,500] Voronoi cells



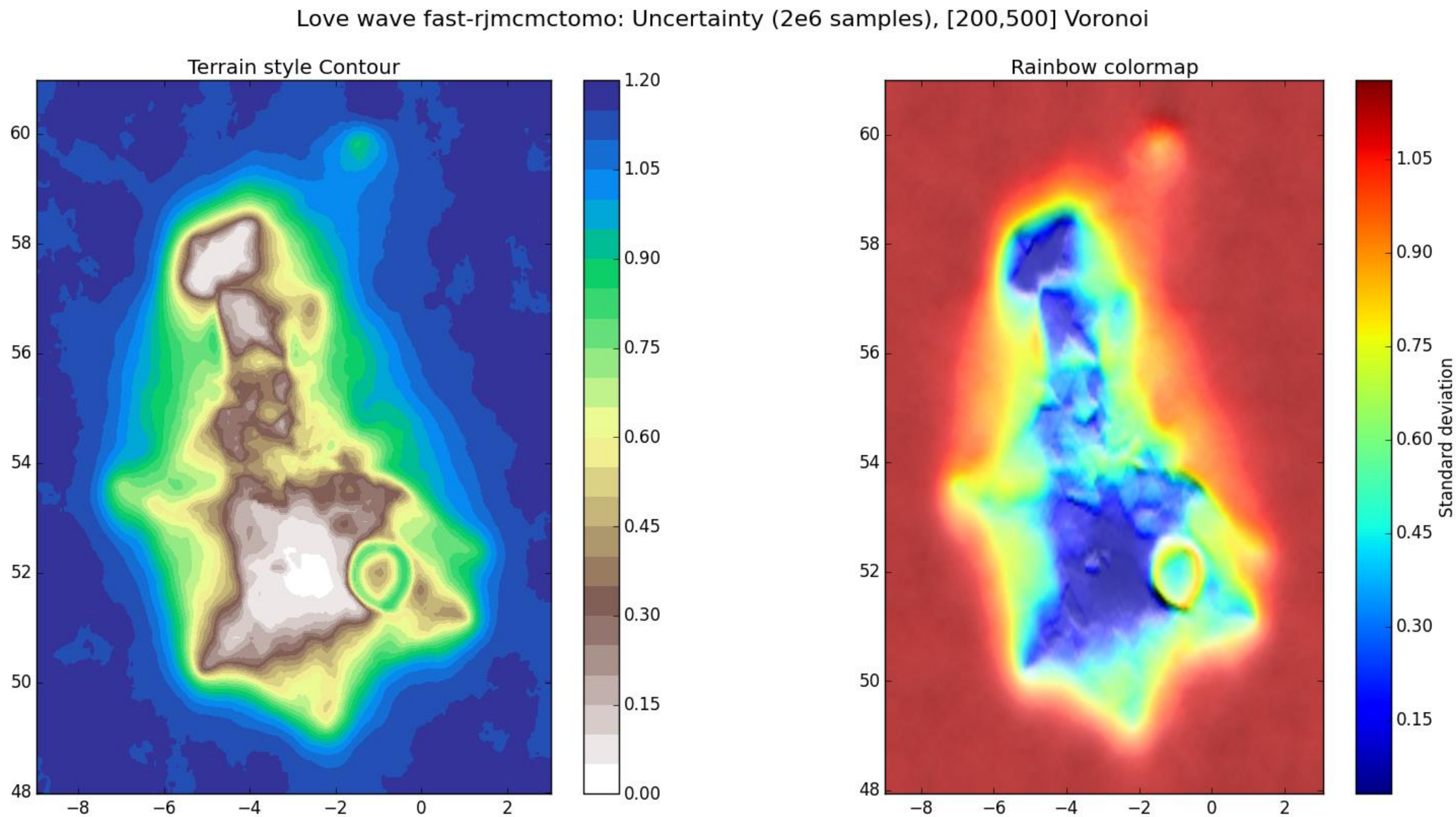


Average seismic velocity



Inversion is good for highly sampled regions!

Uncertainty quantification



Uncertainty is low for highly sampled regions!