Science & Technology Facilities Council

UCL

Proximal Markov chain Monte Carlo: Towards building a sparse Earth model

Auggie Marignier[1,2], Jason McEwen[1], Ana M. G. Ferreira[2,3], Thomas Kitching[1]

[1]Mullard Space Science Laboratory, University College London

[2]Department of Earth Sciences, University College London

[3]Instituto Superior Técnico, Universidade de Lisboa

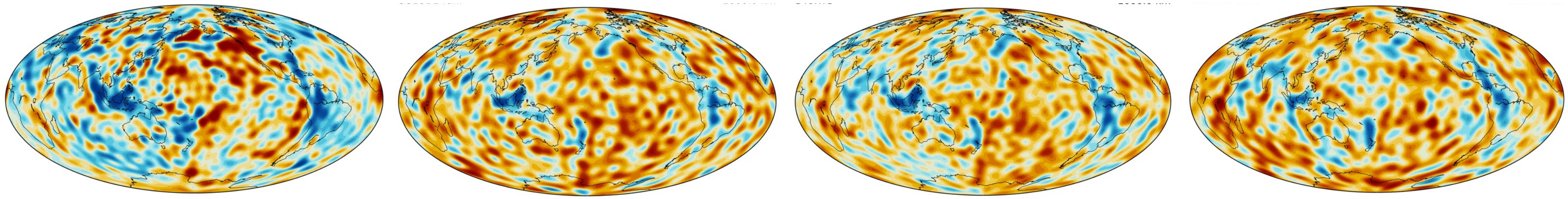
Problem statement

We want uncertainties on our tomographic models

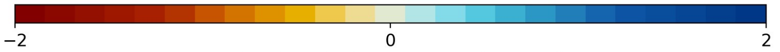
Probabilistic sampling on global scales is generally too difficult

Current efforts don't allow for certain types of desirable prior information, such as sparsity

SEISGLOB2 | SGLOBE-rani | S40RTS | SEMUCB-WM1



Durand et al., 2017 | Chang et al., 2015 | Ritsema et al., 2011 | French & Romanowicz, 2014

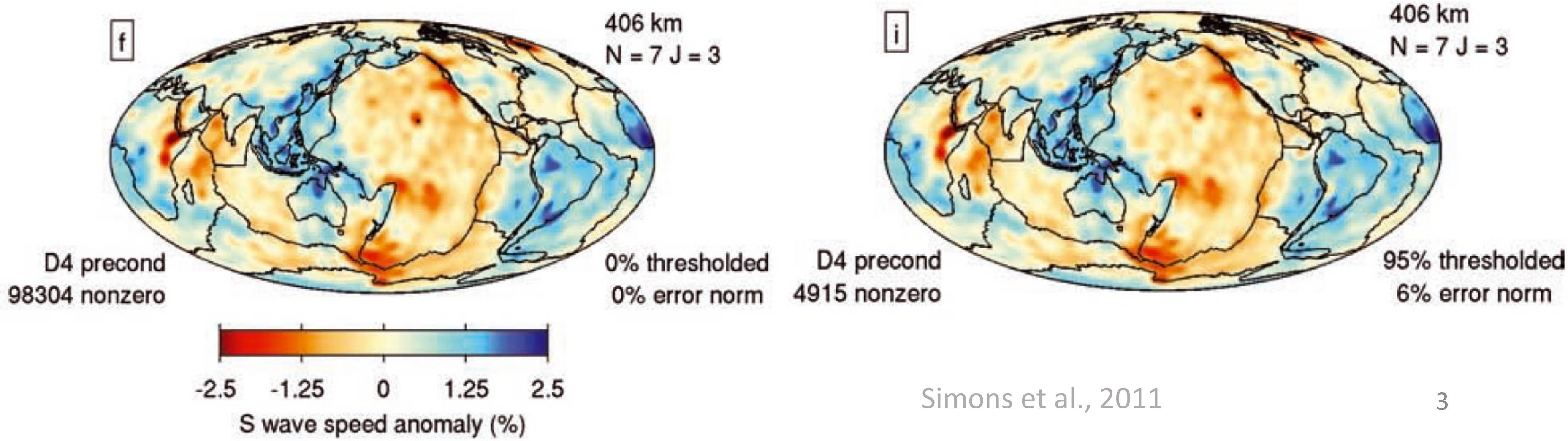


Hosseini et al., 2018

Why sparsity

A signal is “sparse” in a certain basis if many of its expansion coefficients are 0

Compressed sensing has shown that signals that are sparse in a certain basis can be accurately recovered from incomplete or poorly distributed data, which is the case in global tomography



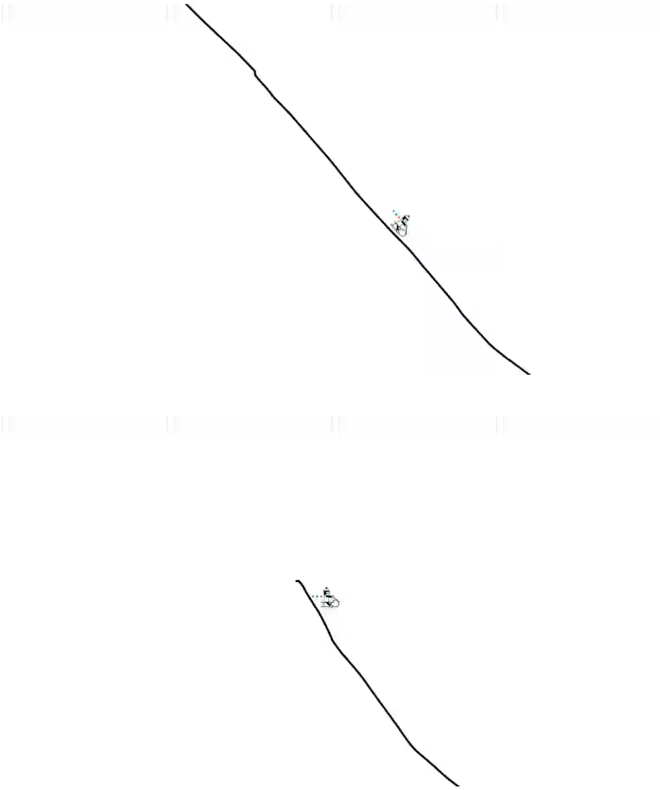
Simons et al., 2011

Proximal algorithms

Convex optimisation techniques using proximity mappings rather than gradients

Particularly well suited to high-dimensional problems like gradient-based approaches

Can be applied to non-smooth problems

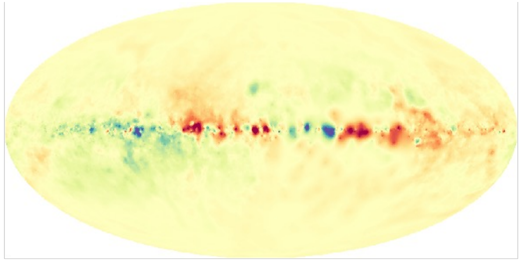


What we do

Demonstrate a recent proximal MCMC algorithm on the common problem of building global Rayleigh wave phase velocity maps

We promote sparsity in a spherical wavelet basis

This is the first use of these proximal MCMC methods on spherical problems



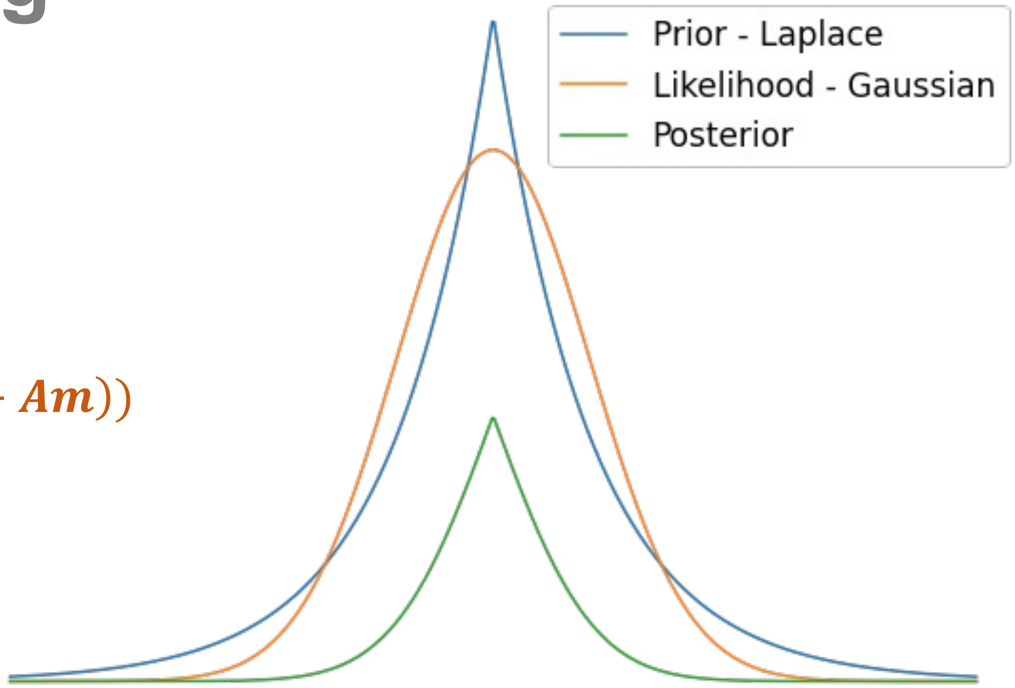
Oppermann et al., 2015

Bayesian Sampling

POSTERIOR – what we want
$p(\boldsymbol{m}|\boldsymbol{d}) \propto p(\boldsymbol{d}|\boldsymbol{m})p(\boldsymbol{m})$

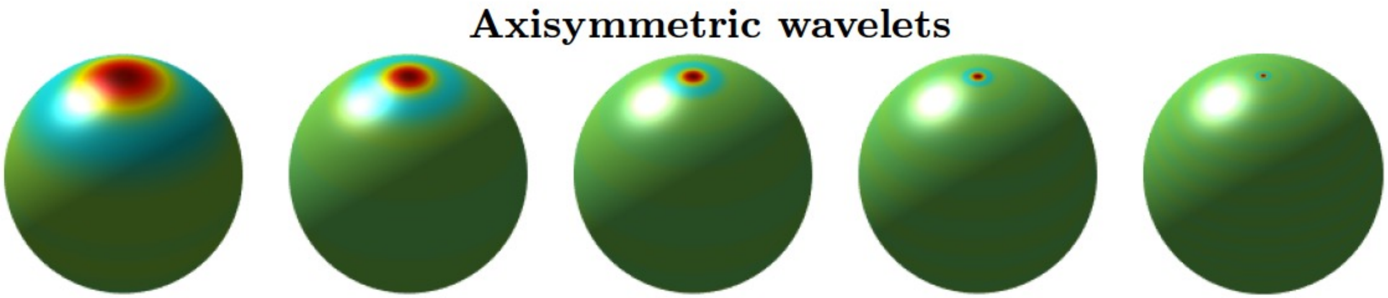
LIKELIHOOD – what we have
$p(\boldsymbol{d}|\boldsymbol{m}) \propto \exp(-(\boldsymbol{d} - \boldsymbol{Am})^T \boldsymbol{C}^{-1}(\boldsymbol{d} - \boldsymbol{Am}))$

PRIOR – what we think
$p(\boldsymbol{m}) \propto \exp(-\mu|\boldsymbol{m}|)$



Wavelet parameterisation

The parameters of our model are the wavelet coefficients in pixel space



Cai et al., 2020, Leistedt et al,. 2013

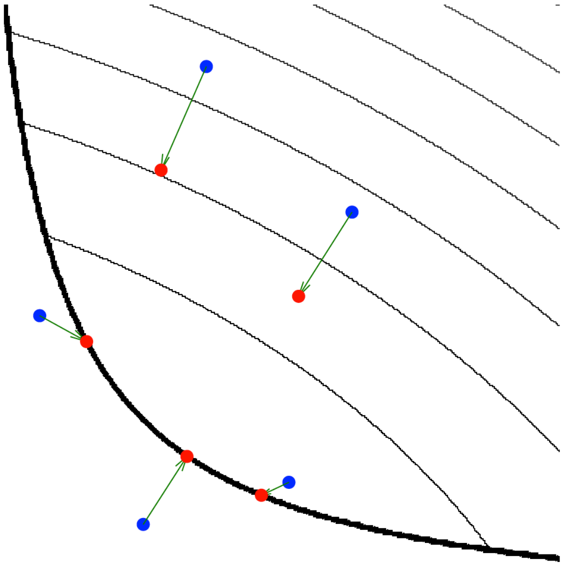
Over 10,000 parameters

Proximity mappings

A gradient step in a smoothed version of a function

This smoothed version is called the λ-MY envelope

Has very useful properties that are similar to the gradient



Parikh & Boyd, 2013

Proximal MCMC

Unadjusted Langevin Algorithm (ULA)

current chain sample

randomness

$$\boldsymbol{m}^{(n+1)} = \boldsymbol{m}^{(n)} + \frac{\delta}{2} \nabla \log \pi\left(\boldsymbol{m}^{(n)}\right) + \sqrt{\delta}\,\boldsymbol{w}^{(n)}$$

next chain sample

gradient of the posterior

Proximal MCMC

Our posterior is of the form $\pi(\boldsymbol{m}) \propto \exp(-g(\boldsymbol{m}) - f(\boldsymbol{m}))$ where f is non-differentiable

Replace f with its smooth λ-MY envelope

current chain sample

$$\boldsymbol{m}^{(n+1)} = \left(1 - \frac{\delta}{\lambda}\right)\boldsymbol{m}^{(n)}$$

randomness

next chain sample

$$-\delta\,\nabla g + \frac{\delta}{\lambda}\operatorname{prox}_f^{\lambda}\left(\boldsymbol{m}^{(n)}\right) + \sqrt{\delta}\,\boldsymbol{w}^{(n)}$$

gradient of the likelihood

prox of the prior

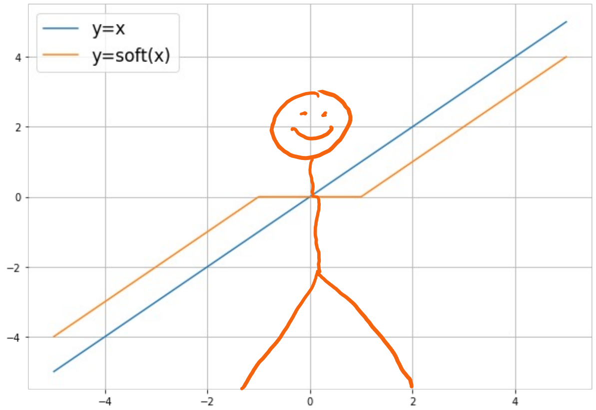
Pereyra, 2016

Moreau-Yosida Unadjusted Langevin Algorithm (MYULA)

Proximal MCMC

So how do we calculate the prox of our prior?

$$f(\boldsymbol{m}) = \mu\|\boldsymbol{m}\|_1 \Longrightarrow \text{prox}_f^{\lambda}(\boldsymbol{m}) = \text{soft}_{\mu\lambda}(\boldsymbol{m})$$



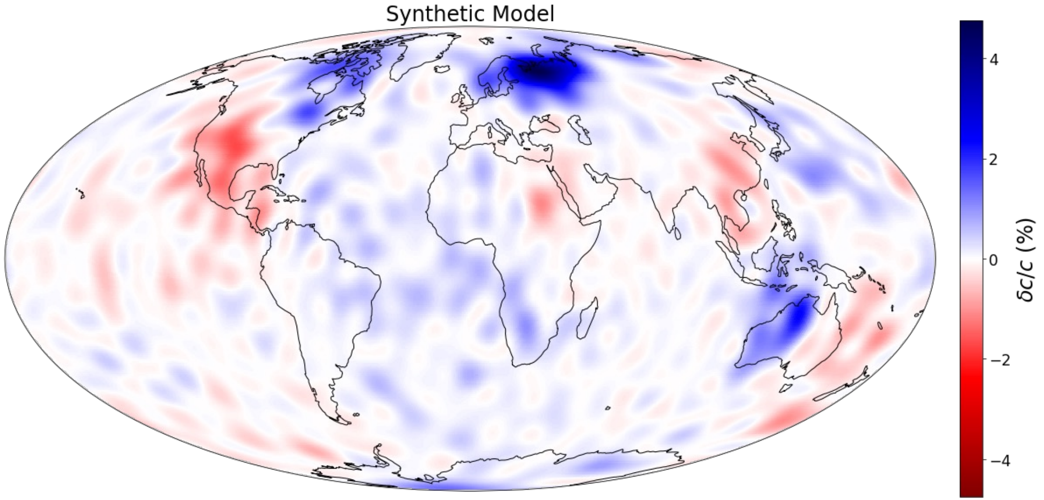
And the gradient of the likelihood?

$$g(\boldsymbol{m}) = \frac{1}{2\sigma^2}\|\boldsymbol{d} - \boldsymbol{A}\boldsymbol{m}\|_2^2 \Longrightarrow \nabla g = \boldsymbol{A}^{\dagger}(\boldsymbol{A}\boldsymbol{m} - \boldsymbol{d})/\sigma^2$$

Combettes & Pesquet, 2011

Synthetic Phase velocity experiment

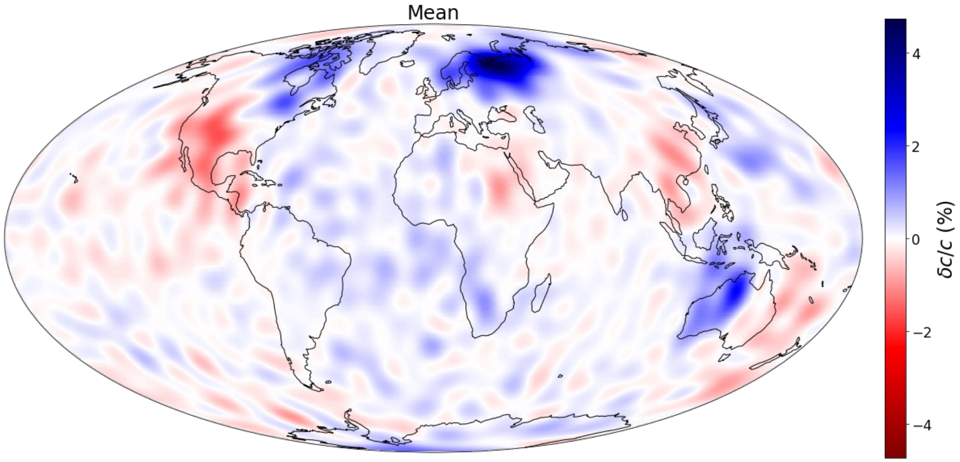
Synthetic model based on a squared phase velocity map (no Gaussian structures)

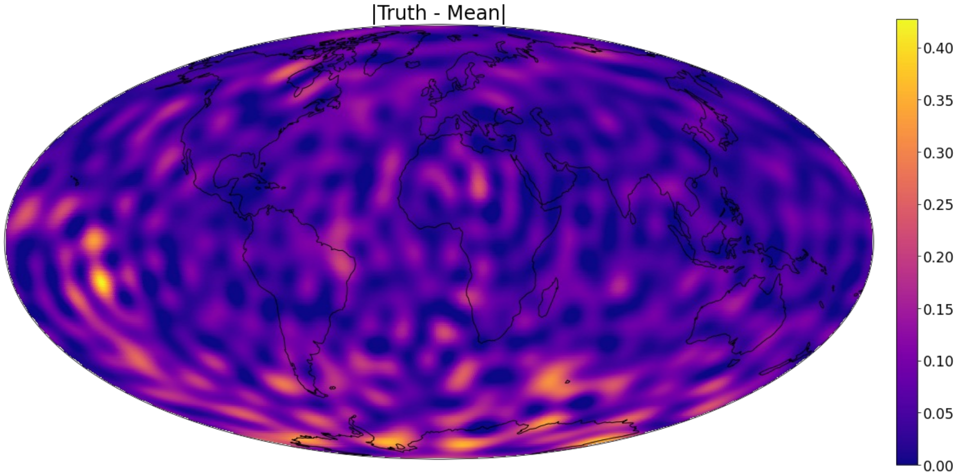


Marignier et al., *in prep*

Forward modelling (GCA) formulated as a matrix multiplication for efficient forward and adjoint modelling

Synthetic Phase velocity experiment

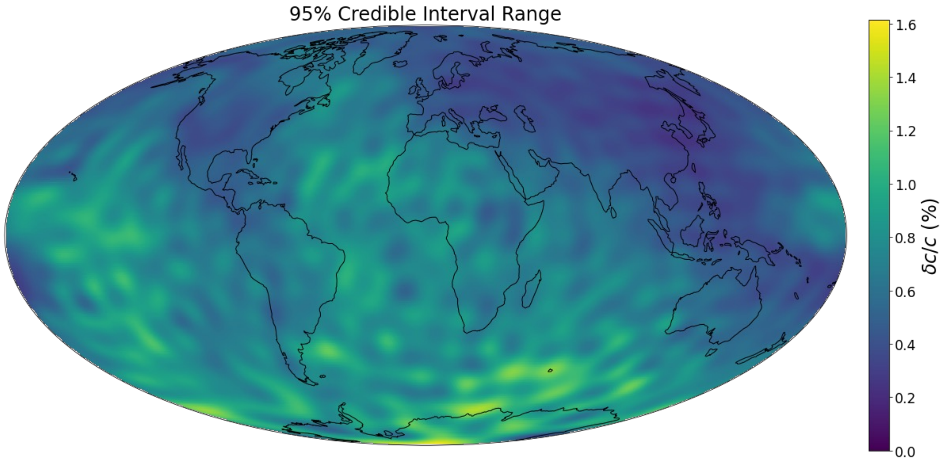


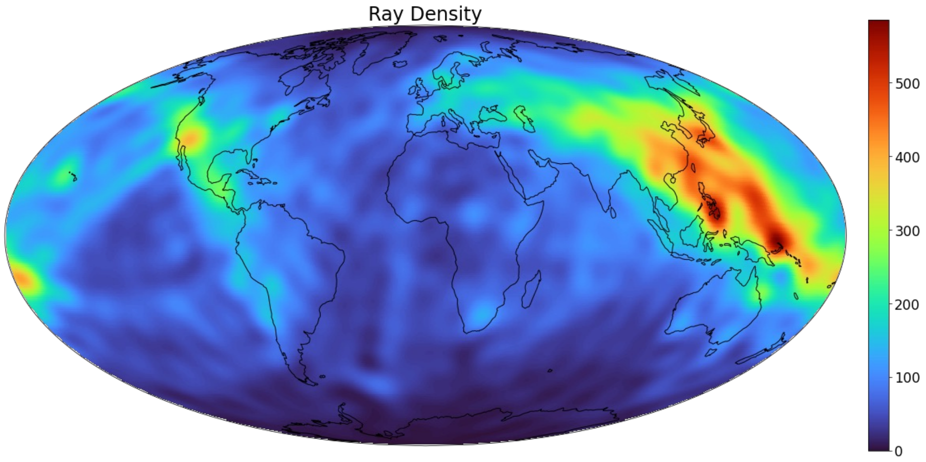


SNR = 13.45 dB

Marignier et al., *in prep*

Synthetic Phase velocity experiment





Marignier et al., *in prep*

Future prospects

Looking at using pxmcmc to invert our phase velocity maps to invert for Vs in the upper mantle with a similar wavelet parameterisation that has radial support

This would result in a 3D model with full uncertainty quantification

Hope to get sharper images from the compressed sensing approach, upon which we can perform some hypothesis testing of features of interest thanks to the uncertainties