Math in Transdimensional Bayesian Inversions: a Rigorous Formulation

Fabrizio Magrini[1], Giovanni Diaferia[2], and Luca De Siena[1]

[1]Institute of Geosciences, Johannes Gutenberg University, Mainz, Germany
[2]INGV

August 24, 2023

1 Classical, Isotropic Transdimensional Inversion

The proposed algorithm aims to find

$$\underbrace{p(\mathbf{m} \mid \mathbf{d}_{obs})}_{\text{posterior}} \propto \underbrace{p(\mathbf{d}_{obs} \mid \mathbf{m})}_{\text{likelihood}} \times \underbrace{p(\mathbf{m})}_{\text{prior}}, \tag{1}$$

by means of a reversible-jump Markov chain Monte Carlo sampling.

1.1 Free parameters

For simplicity, assume the data variance is known, i.e. the $n \times n$ matrix $\mathbf{C}_e$ can be computed before the inversion. Then, following *Bodin et al.* [2012], we only have three free parameters in the inversion; these are

1. number of layers k;
2. isotropic velocity $\mathbf{v}$, corresponding to V_S, of the k Voronoi cells;
3. position (depth) $\mathbf{c}$ of the k Voronoi cells.

1.2 Likelihood: $p(\mathbf{d}_{obs} \mid \mathbf{m})$

The general formula for the likelihood reads

$$p(\mathbf{d}_{obs} \mid \mathbf{m}) = \frac{1}{\sqrt{(2\pi)^n \, |\mathbf{C}_e|}} \exp\left\{ \frac{-\mathbf{\Phi}(\mathbf{m})}{2} \right\}, \tag{2}$$

where $\mathbf{\Phi}(\mathbf{m}) = (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})^T \cdot \mathbf{C}_e^{-1} \cdot (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})$ denotes the misfit between the observed ($\mathbf{d}_{obs}$) and modeled ($\mathbf{g}(\mathbf{m})$) data, and T transposition. As we will see in the following, since $\mathbf{C}_e$ is fixed, the multiplicative factor on the left of the exponential will cancel out.

1.3 Prior probability: $p(\mathbf{m})$

The prior probability

$$p(\mathbf{m}) = p(\mathbf{c}, \mathbf{v} \mid k)\ p(k), \tag{3}$$

is a function of the number of layers k. Assuming $\mathbf{c}$ and $\mathbf{v}$ are independent from each other, we can write

$$p(\mathbf{v} \mid k) = \prod_{i=1}^{k} \frac{1}{\Delta v} = \frac{1}{(\Delta v)^k}, \tag{4}$$

where $\Delta v = v_{max} - v_{min}$ denotes the span of the velocity range, and

$$p(\mathbf{c} \mid k) = \frac{k!\,(N-k)!}{N!}, \tag{5}$$

where N denotes the number of possible positions that can be occupied by the Voronoi cells. (Expression (5) implies that the Voronoi cells can only be placed in a discrete number of positions. We will see in the following, however, that N will cancel out in the definition of the acceptance probability, thus allowing for a uniform sampling of the depth range considered in the parameterization.)

Similar to expression (4), the prior probability of having k layers

$$p(k) = \frac{1}{\Delta k}, \tag{6}$$

where $\Delta k = k_{max} - k_{min}$. Substituting (4), (5), (6) into equation (3), we get

$$p(\mathbf{m}) = \frac{k!\,(N-k)!}{N!\ \Delta k\ \ (\Delta v)^k}. \tag{7}$$

1.4 Acceptance probability: $\alpha(\mathbf{m}' \mid \mathbf{m})$

In the inverse problem described thus far, a given model $\mathbf{m}'$ is preferred over the current model $\mathbf{m}$ if and only if its acceptance probability $\alpha(\mathbf{m}' \mid \mathbf{m}) \geq r$, where $0 \leq r < 1$ denotes a random number. It has been shown [*Green*, 2003] that the chain of sampled models will converge to the transdimensional posterior distribution $p(\mathbf{m} \mid \mathbf{d}_{obs})$ if

$$\alpha(\mathbf{m}' \mid \mathbf{m}) = \min\left[1,\ \underbrace{\frac{p\,(\mathbf{m}')}{p\,(\mathbf{m})}}_{\text{Prior ratio}}\ \underbrace{\frac{p\,(\mathbf{d}_{obs} \mid \mathbf{m}')}{p\,(\mathbf{d}_{obs} \mid \mathbf{m})}}_{\text{Likelihood ratio}}\ \underbrace{\frac{q\,(\mathbf{m} \mid \mathbf{m}')}{q\,(\mathbf{m}' \mid \mathbf{m})}}_{\text{Proposal ratio}}\right]. \tag{8}$$

It is the purpose of the following sections to robustly define each ratio in the right-hand side (RHS) of equation (8), so as to compute $\alpha(\mathbf{m}' \mid \mathbf{m})$ accurately at each iteration of the transdimensional inversion.

In general, a new model is proposed based on a random perturbation of the current model. Since we are considering only three free parameters (i.e. k, $\mathbf{c}$, and $\mathbf{v}$), the proposed model can be obtained based on four different kinds of perturbations:

1. **Change in velocity**. The velocity v_i corresponding to a Voronoi cell is randomly changed to a new value v_i' according to a Gaussian probability distribution centered

onto v_i, i.e. the proposal probability distribution

$$q_1(v_i' \mid v_i) = \frac{1}{\theta_1\sqrt{2\pi}} \exp\left\{ -\frac{(v_i' - v_i)^2}{2\theta_1^2} \right\}, \tag{9}$$

where the standard deviation of the Gaussian θ_1 should be chosen a priori and

$$v_i' = v_i + u, \tag{10}$$

with u being a random deviate from the normal distribution $\mathcal{N}(0, \theta_1)$.

2. **Change in position (depth)**. The position c_i of a Voronoi cell is randomly changed to a new value c_i' according to a Gaussian probability distribution centered onto c_i, i.e.

$$q_2(c_i' \mid c_i) = \frac{1}{\theta_2\sqrt{2\pi}} \exp\left\{ -\frac{(c_i' - c_i)^2}{2\theta_2^2} \right\}, \tag{11}$$

3. **Birth**. A new Voronoi cell is created by choosing randomly its position from the $N - k$ options available. (As anticipated, this discretizations does not need to be done in practice: a new position is drawn by sampling uniformly the depth range considered.) Similar to expression (9) a velocity v_{k+1}' is assigned to the new Voronoi cell by sampling the Gaussian distribution

$$q_3(v_{k+1}' \mid v_i) = \frac{1}{\theta_3\sqrt{2\pi}} \exp\left\{ -\frac{\left(v_{k+1}' - v_i\right)^2}{2\theta_3^2} \right\}, \tag{12}$$

where, once again, θ_3 should be chosen a priori.

4. **Death**. Remove one layer by drawing a random integer between 0 and k.

In general, the above perturbations can regrouped in three main categories, reported below, based on the dimension of $\mathbf{m}'$ with respect to $\mathbf{m}$. The only term in the RHS of equation (8) that does not depend on the belonging category is the ratio between the likelihood of the proposed model and that of the current model, and reads

$$\begin{aligned} \frac{p\left(\mathbf{d}_{obs} \mid \mathbf{m}'\right)}{p\left(\mathbf{d}_{obs} \mid \mathbf{m}\right)} &= \frac{\exp\left\{\frac{-\boldsymbol{\Phi}(\mathbf{m}')}{2}\right\}}{\sqrt{(2\pi)^n \left|\mathbf{C}_e\right|}} \frac{\sqrt{(2\pi)^n \left|\mathbf{C}_e\right|}}{\exp\left\{\frac{-\boldsymbol{\Phi}(\mathbf{m})}{2}\right\}} \\ &= \exp\left\{ -\frac{\boldsymbol{\Phi}(\mathbf{m}') - \boldsymbol{\Phi}(\mathbf{m})}{2} \right\}. \end{aligned} \tag{13}$$

1.4.1 CASE 1 (change in velocity or position): $\mathbf{m}', \mathbf{m} \in \mathbb{R}^k$

When $\mathbf{m}'$ has the same dimension of $\mathbf{m}$, the calculation of the ratios at the RHS of equation (8) is straightforward. Since Δv is a constant in expression (7) and k has not changed, $p\left(\mathbf{m}\right) = p\left(\mathbf{m}'\right)$ and $\frac{p(\mathbf{m}')}{p(\mathbf{m})} = 1$. Moreover, the proposal perturbations that do not involve a change of dimension have symmetrical probability distributions [*Bodin et al.*, 2012], i.e.

$$q_1(v_i' \mid v_i) = q_1(v_i \mid v_i')$$
$$q_2(c_i' \mid c_i) = q_2(c_i \mid c_i'), \tag{14}$$

implying that $\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})} = 1$.

The acceptance probability therefore reduces to

$$\alpha(\mathbf{m}' \mid \mathbf{m}) = \min\left[1, \exp\left\{-\frac{\mathbf{\Phi}(\mathbf{m}') - \mathbf{\Phi}(\mathbf{m})}{2}\right\}\right]. \tag{15}$$

1.4.2 CASE 2 (birth): $\mathbf{m}' \in \mathbb{R}^{k+1}$, $\mathbf{m} \in \mathbb{R}^k$

Based on equation (7), i.e. $p(\mathbf{m}) = p(\mathbf{c}, \mathbf{v} \mid k)\ p(k)$, the ratio

$$\begin{aligned}\frac{p\left(\mathbf{m}'\right)}{p\left(\mathbf{m}\right)} &= \frac{p\left(\mathbf{c}' \mid k+1\right)\ p\left(\mathbf{v}' \mid k+1\right)\ p\left(k+1\right)}{p\left(\mathbf{c} \mid k\right)\ p\left(\mathbf{v} \mid k\right)\ p\left(k\right)} \\ &= \frac{p\left(\mathbf{c}' \mid k+1\right)\ p\left(\mathbf{v}' \mid k+1\right)}{p\left(\mathbf{c} \mid k\right)\ p\left(\mathbf{v} \mid k\right)} \\ &= \frac{(k+1)!\,(N-(k+1))!}{N!\ \ (\Delta v)^{(k+1)}}\ \frac{N!\ \ (\Delta v)^k}{k!\,(N-k)!} \\ &= \frac{(k+1)}{(N-k)\Delta v},\end{aligned} \tag{16}$$

where we noticed that $p\left(k\right) = p\left(k+1\right) = {}^{1}/_{\Delta k}$ and ${}^{(a+1)!}/_{a!} = a+1\ \forall a \in \mathbb{N}$.

Following a similar reasoning, *Bodin et al.* [2012] derived the analytical expression for the proposal ratio

$$\begin{aligned}\frac{q\left(\mathbf{m} \mid \mathbf{m}'\right)}{q\left(\mathbf{m}' \mid \mathbf{m}\right)} &= \frac{q\left(\mathbf{c} \mid \mathbf{m}'\right)\ q\left(\mathbf{v} \mid \mathbf{m}'\right)}{q\left(\mathbf{c}' \mid \mathbf{m}\right)\ q\left(\mathbf{v}' \mid \mathbf{m}\right)} \\ &= \frac{N-k}{(k+1)\, q_3\left(v_{k+1}' \mid v_i\right)} \\ &= \frac{N-k}{(k+1)}\ \theta_3\sqrt{2\pi}\ \exp\left\{\frac{\left(v_{k+1}' - v_i\right)^2}{2\theta_3^2}\right\},\end{aligned} \tag{17}$$

which, together with (16), can be substituted into equation (8) to obtain, after some simplification, the acceptance probability

$$\alpha(\mathbf{m}' \mid \mathbf{m}) = \min\left[1, \frac{\theta_3\sqrt{2\pi}}{\Delta v}\exp\left\{\frac{\left(v_{k+1}' - v_i\right)^2}{2\theta_3^2} - \frac{\mathbf{\Phi}(\mathbf{m}') - \mathbf{\Phi}(\mathbf{m})}{2}\right\}\right]. \tag{18}$$

Note that the above expression does not depend on N, i.e., we are at liberty to generate the nuclei using a continuous distribution over the depth range considered, which is tantamount to $N \rightarrow \inf$.

1.4.3 CASE 3 (death): $\mathbf{m}' \in \mathbb{R}^{k-1}$, $\mathbf{m} \in \mathbb{R}^k$

For this last category of random perturbation, we can simply take the reciprocals of the RHS in equations (16) and (17) and substitute them in (8). After some algebra,

$$\alpha(\mathbf{m}' \mid \mathbf{m}) = \min\left[1, \frac{\Delta v}{\theta_3\sqrt{2\pi}}\exp\left\{-\frac{\left(v'_{k-1} - v_i\right)^2}{2\theta_3^2} - \frac{\mathbf{\Phi}(\mathbf{m}') - \mathbf{\Phi}(\mathbf{m})}{2}\right\}\right]. \tag{19}$$

2 On the effects of additional free parameters (uniformly distributed)

The previous section provides an all-around framework for writing a reversible-jump Markov chain Monte Carlo algorithm, but what happens when additional free parameters are added to the problem? To investigate this matter, let us define a general parameter, say $\mathbf{x}$, which is a function of position $\mathbf{c}$ in the model $\mathbf{m}$. ($\mathbf{x}$ may represent, for example, the ratio $^{V_P}/_{V_S}$ or the density of a given layer in our Earth model.) In general, $\mathbf{x}$ will be allowed to vary within a range defined by $\Delta x = x_{max} - x_{min}$.

2.1 Prior probability: $p(\mathbf{m})$

Under this working assumption, we shall redefine the prior probability $p\left(\mathbf{m}\right)$ so as to take into account the probability of sampling $\mathbf{x}$ uniformly within its belonging range. This is done by modifying equation (3), i.e.

$$p(\mathbf{m}) = p(\mathbf{c}, \mathbf{v}, \mathbf{x} \mid k)\ p(k). \tag{20}$$

Similar to the previous section, we can consider $\mathbf{c}$, $\mathbf{v}$, and $\mathbf{x}$ to be independent from each other, resulting in

$$p(\mathbf{m}) = p(\mathbf{c} \mid k)\ p(\mathbf{v} \mid k)\ p(\mathbf{x} \mid k)\ p(k). \tag{21}$$

Analogous to equation (4), the conditional probability $p(\mathbf{x} \mid k) = {}^{1}/_{(\Delta x)^k}$, and

$$p(\mathbf{m}) = \frac{k!\,(N-k)!}{N!\ \Delta k\ \left(\Delta v\right)^k\ \left(\Delta x\right)^k}, \tag{22}$$

i.e., the prior probability should be multiplied by a factor ${}^{1}/_{(\Delta x)^k}$. In the more general case where n parameters $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ that are only function of position $\mathbf{c}$ are inverted for, the above reads

$$p(\mathbf{m}) = \frac{k!\,(N-k)!}{N!\ \Delta k\ \left(\Delta x_1\right)^k\ \left(\Delta x_2\right)^k \ldots \left(\Delta x_n\right)^k}, \tag{23}$$

2.2 Acceptance probability: $\alpha(\mathbf{m}' \mid \mathbf{m})$

We should now evaluate the acceptance probability inherent to a proposed model $\mathbf{m}'$. As in the previous section, it is convenient to subdivide the possible random perturbations in three different categories.

2.2.1 CASE 1 (change in velocity, position, or x): $\mathbf{m}', \mathbf{m} \in \mathbb{R}^k$

As in Section 3.1, the only term in the RHS of equation (8) different than one is the ratio between the likelihood of the proposed model and of the current model. The acceptance probability, in this case, therefore coincides with that in equation (15).

2.2.2 CASE 2 (birth): $\mathbf{m}' \in \mathbb{R}^{k+1}$, $\mathbf{m} \in \mathbb{R}^k$

In this case, while the presence of an additional parameter $\mathbf{x}$ does not contribute to modifying equation (17), it should be taken into account in (16) and (19). Similar to equations (9) and (11), the proposal probability for the value x_i associated with the newly introduced layer reads

$$q_4(x_i' \mid x_i) = \frac{1}{\theta_4\sqrt{2\pi}} \exp\left\{ -\frac{(x_i' - x_i)^2}{2\theta_4^2} \right\}, \tag{24}$$

and

$$\begin{aligned} \frac{q(\mathbf{m} \mid \mathbf{m}')}{q(\mathbf{m}' \mid \mathbf{m})} &= \frac{q(\mathbf{c} \mid \mathbf{m}')\ q(\mathbf{v} \mid \mathbf{m}')\ q(\mathbf{x} \mid \mathbf{m}')}{q(\mathbf{c}' \mid \mathbf{m})\ q(\mathbf{v}' \mid \mathbf{m})\ q(\mathbf{x}' \mid \mathbf{m})} \\ &= \frac{N-k}{(k+1)\, q_3\left(v_{k+1}' \mid v_i\right) q_4\left(x_{k+1}' \mid x_i\right)} \\ &= \frac{N-k}{(k+1)}\ \theta_3\sqrt{2\pi}\ \exp\left\{ \frac{\left(v_{k+1}' - v_i\right)^2}{2\theta_3^2} \right\}\ \theta_4\sqrt{2\pi}\ \exp\left\{ \frac{\left(x_{k+1}' - x_i\right)^2}{2\theta_4^2} \right\} \\ &= \frac{N-k}{(k+1)}\ 2\pi\theta_3\theta_4 \exp\left\{ \frac{\left(v_{k+1}' - v_i\right)^2}{2\theta_3^2} + \frac{\left(x_{k+1}' - x_i\right)^2}{2\theta_4^2} \right\}. \end{aligned} \tag{25}$$

Generalizing to the case of n independent free parameters, the above can be rewritten

$$\begin{aligned} \frac{q(\mathbf{m} \mid \mathbf{m}')}{q(\mathbf{m}' \mid \mathbf{m})} &= \frac{N-k}{(k+1)}\ (2\pi)^{\frac{n}{2}} \theta_{x_1}\theta_{x_2}\ldots\theta_{x_n} \times \\ &\times \exp\left\{ \frac{\left(x_{1_{k+1}}' - x_{1_i}\right)^2}{2\theta_{x_1}^2} + \frac{\left(x_{2_{k+1}}' - x_{2_i}\right)^2}{2\theta_{x_2}^2} + \ldots + \frac{\left(x_{n_{k+1}}' - x_{n_i}\right)^2}{2\theta_{x_n}^2} \right\}, \end{aligned} \tag{26}$$

where θ_{x_1}, θ_{x_2}, …, θ_{x_n} denote the a-priori standard deviations of the Gaussians used to perturb the n free parameters.

It follows that

$$\begin{aligned} \alpha(\mathbf{m}' \mid \mathbf{m}) = \min\Bigg[1, &\frac{(2\pi)^{\frac{n}{2}}\theta_{x_1}\ldots\theta_{x_n}}{\Delta x_1 \ldots \Delta x_n} \times \\ &\times \exp\left\{ \frac{\left(x_{1_{k+1}}' - x_{1_i}\right)^2}{2\theta_{x_1}^2} + \ldots + \frac{\left(x_{n_{k+1}}' - x_{n_i}\right)^2}{2\theta_{x_n}^2} - \frac{\mathbf{\Phi}(\mathbf{m}') - \mathbf{\Phi}(\mathbf{m})}{2} \right\}\Bigg]. \end{aligned} \tag{27}$$

2.2.3 CASE 3 (death): $\mathbf{m}' \in \mathbb{R}^{k-1}$, $\mathbf{m} \in \mathbb{R}^k$

As in the previous sections, the relevant formulas can be easily obtained by taking the reciprocal of the above equations, and

$$\begin{aligned} \alpha(\mathbf{m}' \mid \mathbf{m}) = \min\Bigg[1, \frac{\Delta x_1 \dots \Delta x_n}{(2\pi)^{\frac{n}{2}} \theta_{x_1} \dots \theta_{x_n}} \times \\ \times \exp\Bigg\{ - \frac{\left(x'_{1_{k+1}} - x_{1_i}\right)^2}{2\theta_{x_1}^2} - \dots - \frac{\left(x'_{n_{k+1}} - x_{n_i}\right)^2}{2\theta_{x_n}^2} - \frac{\mathbf{\Phi}(\mathbf{m}') - \mathbf{\Phi}(\mathbf{m})}{2} \Bigg\}\Bigg]. \end{aligned} \tag{28}$$

3 Introduction of a normally-distributed additional free parameters

Let us consider now the introduction of a free parameter, say $\boldsymbol{\xi}$, whose probability of occurrence is described by a normal distribution $\mathcal{N}(\mu_\xi, \sigma_\xi)$ which is independent from position/depth. Then

$$p(\boldsymbol{\xi} \mid k) = \prod_{i=1}^{k} \frac{1}{\sigma_\xi \sqrt{2\pi}} \exp\Bigg\{ - \frac{(\xi_i - \mu_\xi)^2}{2\sigma_\xi^2} \Bigg\}, \tag{29}$$

and

$$\begin{aligned} p(\mathbf{m}) &= p(\mathbf{c}, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n, \boldsymbol{\xi} \mid k)\ p(k) \\ &= p(\mathbf{c} \mid k)\ p(\boldsymbol{\xi} \mid k)\ p(k)\ \prod_{l=1}^{n} p(\mathbf{x}_l \mid k) \\ &= \frac{k!\,(N-k)!\ \prod_{i=1}^{k} \frac{1}{\sigma_\xi \sqrt{2\pi}} \exp\Big\{ - \frac{(\xi_i - \mu_\xi)^2}{2\sigma_\xi^2} \Big\}}{N!\ \Delta k\ \prod_{l=1}^{n} (\Delta x_l)^k} \\ &= \frac{k!\,(N-k)!\ \prod_{i=1}^{k} \exp\Big\{ - \frac{(\xi_i - \mu_\xi)^2}{2\sigma_\xi^2} \Big\}}{N!\ \Delta k\ \left(\sigma_\xi \sqrt{2\pi}\right)^k\ \prod_{l=1}^{n} (\Delta x_l)^k} \end{aligned} \tag{30}$$

where, as in the previous section, $\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_n$ are free parameters described by a uniform probability of occurrence within their belonging ranges.

3.1 CASE 1 (change in $\boldsymbol{\xi}$): $\mathbf{m}', \mathbf{m} \in \mathbb{R}^k$

Similar to the previous sections, if any of the uniformly-distributed free parameters $\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_n$ are perturbed, the prior ratio $\frac{p(\mathbf{m}')}{p(\mathbf{m})} = 1$. However, this is not the case when $\boldsymbol{\xi}$ is perturbed. If we assume that only the jth element of $\boldsymbol{\xi}$ has changed in the proposed model $p(\mathbf{m}')$,

$$\begin{aligned} \frac{p(\mathbf{m}')}{p(\mathbf{m})} = \frac{p(\boldsymbol{\xi}' \mid k)}{p(\boldsymbol{\xi} \mid k)} &= \frac{\exp\Big\{ - \frac{(\xi'_j - \mu_\xi)^2}{2\sigma_\xi^2} \Big\}}{\exp\Big\{ - \frac{(\xi_j - \mu_\xi)^2}{2\sigma_\xi^2} \Big\}} \prod_{\substack{i=1 \\ i \neq j}}^{k} \frac{\exp\Big\{ - \frac{(\xi_i - \mu_\xi)^2}{2\sigma_\xi^2} \Big\}}{\exp\Big\{ - \frac{(\xi_i - \mu_\xi)^2}{2\sigma_\xi^2} \Big\}} \\ &= \exp\Bigg\{ \frac{(\xi_j - \mu_\xi)^2 - \left(\xi'_j - \mu_\xi\right)^2}{2\sigma_\xi^2} \Bigg\}. \end{aligned} \tag{31}$$

If we now define the proposal probability

$$q_\xi(\xi_i' \mid \xi_i) = \frac{1}{\theta_\xi \sqrt{2\pi}} \exp\left\{ - \frac{(\xi_i' - \xi_i)^2}{2\theta_\xi^2} \right\}, \tag{32}$$

and notice that $q_\xi(\xi_i' \mid \xi_i) = q_\xi(\xi_i \mid \xi_i')$, the acceptance probability reads

$$\begin{aligned} \alpha(\mathbf{m}' \mid \mathbf{m}) &= \min\left[1, \frac{p\,(\mathbf{m}')}{p\,(\mathbf{m})}\ \frac{p\,(\mathbf{d}_{obs} \mid \mathbf{m}')}{p\,(\mathbf{d}_{obs} \mid \mathbf{m})}\ \frac{q\,(\mathbf{m} \mid \mathbf{m}')}{q\,(\mathbf{m}' \mid \mathbf{m})}\right] \\ &= \min\left[1, \exp\left\{ - \frac{\mathbf{\Phi}(\mathbf{m}') - \mathbf{\Phi}(\mathbf{m})}{2} + \frac{(\xi_j - \mu_\xi)^2 - \left(\xi_j' - \mu_\xi\right)^2}{2\sigma_\xi^2} \right\}\right]. \end{aligned} \tag{33}$$

Note that, in the above equations, σ_ξ defines the standard deviation of the normal distribution associated with our prior knowledge of $\boldsymbol{\xi}$, while θ_ξ, as in the previous sections, is the standard deviation of the Gaussian used to perturb the model $\mathbf{m}$ by randomly changing one element of $\boldsymbol{\xi}$. (Both σ_ξ and θ_ξ should be chosen a priori.)

3.2 CASE 2 (birth): $\mathbf{m}' \in \mathbb{R}^{k+1}$, $\mathbf{m} \in \mathbb{R}^k$

Based on equation (30),

$$\begin{aligned} \frac{p(\mathbf{m}')}{p(\mathbf{m})} &= \frac{p(\mathbf{c}' \mid k+1)\ p(\boldsymbol{\xi}' \mid k+1)\ p(k+1)\ \prod_{l=1}^n p(\mathbf{x}_l \mid k+1)}{p(\mathbf{c} \mid k)\ p(\boldsymbol{\xi} \mid k)\ p(k)\ \prod_{l=1}^n p(\mathbf{x}_l \mid k)} \\ &= \frac{(k+1)\ \left(\sigma_\xi\sqrt{2\pi}\right)^k}{(N-k)\left(\sigma_\xi\sqrt{2\pi}\right)^{k+1} \prod_{l=1}^n \Delta x_l}\ \frac{\prod_{i=1}^{k+1}\ \exp\left\{ - \frac{(\xi_i' - \mu_\xi)^2}{2\sigma_\xi^2} \right\}}{\prod_{i=1}^{k}\ \exp\left\{ - \frac{(\xi_i - \mu_\xi)^2}{2\sigma_\xi^2} \right\}} \\ &= \frac{(k+1)}{(N-k)\left(\sigma_\xi\sqrt{2\pi}\right) \prod_{l=1}^n \Delta x_l} \exp\left\{ - \frac{\left(\xi_{k+1}' - \mu_\xi\right)^2}{2\sigma_\xi^2} \right\}, \end{aligned} \tag{34}$$

where we noticed that only one element of $\boldsymbol{\xi}'$ does not belong to $\boldsymbol{\xi}$, i.e. the newly born ξ_{k+1}'.

Merging equations (26) and (32), the proposal ratio

$$\begin{aligned} \frac{q\,(\mathbf{m} \mid \mathbf{m}')}{q\,(\mathbf{m}' \mid \mathbf{m})} &= \frac{q\,(\mathbf{c} \mid \mathbf{m}')\ q\,(\mathbf{x}_1 \mid \mathbf{m}')\ q\,(\mathbf{x}_2 \mid \mathbf{m}') \ldots\ q\,(\mathbf{x}_n \mid \mathbf{m}')}{q\,(\mathbf{c}' \mid \mathbf{m})\ q\,(\mathbf{x}_1' \mid \mathbf{m})\ q\,(\mathbf{x}_2' \mid \mathbf{m})\ \ldots\ q\,(\mathbf{x}_n' \mid \mathbf{m})}\ \frac{q\,(\boldsymbol{\xi} \mid \mathbf{m}')}{q\,(\boldsymbol{\xi}' \mid \mathbf{m})} \\ &= \frac{N-k}{(k+1)\ \ q_\xi(\xi_i' \mid \xi_i)\ \prod_{l=1}^n q_l\left(x_{l_{k+1}}' \mid x_{l_i}\right)} \\ &= \frac{N-k}{(k+1)}\ \theta_\xi\sqrt{2\pi}\ \exp\left\{ \frac{\left(\xi_{k+1}' - \xi_i\right)^2}{2\theta_\xi^2} \right\}\ (2\pi)^{\frac{n}{2}} \prod_{l=1}^n \theta_{x_l}\ \exp\left\{ \frac{\left(x_{l_{k+1}}' - x_{l_i}\right)^2}{2\theta_{x_l}^2} \right\} \\ &= \frac{N-k}{(k+1)}\ \theta_\xi\ (2\pi)^{\frac{n+1}{2}}\ \exp\left\{ \frac{\left(\xi_{k+1}' - \xi_i\right)^2}{2\theta_\xi^2} + \sum_{l=1}^n \frac{\left(x_{l_{k+1}}' - x_{l_i}\right)^2}{2\theta_{x_l}^2} \right\} \prod_{l=1}^n \theta_{x_l}. \end{aligned} \tag{35}$$

The acceptance probability then reads

$$
\begin{aligned}
\alpha(\mathbf{m}' \mid \mathbf{m}) &= \min\left[1, \frac{p\left(\mathbf{m}'\right)}{p\left(\mathbf{m}\right)} \frac{p\left(\mathbf{d}_{obs} \mid \mathbf{m}'\right)}{p\left(\mathbf{d}_{obs} \mid \mathbf{m}\right)} \frac{q\left(\mathbf{m} \mid \mathbf{m}'\right)}{q\left(\mathbf{m}' \mid \mathbf{m}\right)}\right] \\
&= \left(\prod_{l=1}^{n} \frac{\theta_{x_l}}{\Delta x_l}\right) \frac{\theta_\xi (2\pi)^{n/2}}{\sigma_\xi} \times \\
&\times \exp\left\{ \frac{\left(\xi'_{k+1} - \xi_i\right)^2}{2\theta_\xi^2} - \frac{\left(\xi'_{k+1} - \mu_\xi\right)^2}{2\sigma_\xi^2} - \frac{\mathbf{\Phi}(\mathbf{m}') - \mathbf{\Phi}(\mathbf{m})}{2} + \sum_{l=1}^{n} \frac{\left(x'_{l_{k+1}} - x_{l_i}\right)^2}{2\theta_{x_l}^2} \right\}
\end{aligned} \tag{36}
$$

3.3 CASE 3 (death): $\mathbf{m}' \in \mathbb{R}^{k-1}$, $\mathbf{m} \in \mathbb{R}^k$

$$
\begin{aligned}
\alpha(\mathbf{m}' \mid \mathbf{m}) = &\left(\prod_{l=1}^{n} \frac{\Delta x_l}{\theta_{x_l}}\right) \frac{\sigma_\xi}{\theta_\xi (2\pi)^{n/2}} \times \\
&\times \exp\left\{ - \frac{\left(\xi'_{k+1} - \xi_i\right)^2}{2\theta_\xi^2} + \frac{\left(\xi'_{k+1} - \mu_\xi\right)^2}{2\sigma_\xi^2} - \frac{\mathbf{\Phi}(\mathbf{m}') - \mathbf{\Phi}(\mathbf{m})}{2} - \sum_{l=1}^{n} \frac{\left(x'_{l_{k+1}} - x_{l_i}\right)^2}{2\theta_{x_l}^2} \right\}
\end{aligned} \tag{37}
$$

4 Depth-dependent parameters, uniformly distributed at each depth

Let us consider the parameter $\mathbf{x}$, whose probability distribution is dependent on depth and uniform at each depth. In a 1-D parameterization consisting of k layers, its prior probability then reads

$$
p(\mathbf{m}) = \frac{k!\,(N-k)!}{N!\ \Delta k\ \prod_{i=1}^{k} (\Delta x)_i}, \tag{38}
$$

where $(\Delta x)_i$ denotes the range of possible values spanned by $\mathbf{x}$ in the ith layer.

4.1 CASE 1: DEPTH PERTURBATION, $\mathbf{m}', \mathbf{m} \in \mathbb{R}^k$

Similar to Sections 1 and 2, a perturbation in the value of $\mathbf{x}$ in an arbitrary layer is associated with a prior ratio $\frac{p(\mathbf{m}')}{p(\mathbf{m})} = 1$. **When the perturbation involves the position of the jth Voronoi cell**, however, the above ratio reads

$$
\begin{aligned}
\frac{p(\mathbf{m}')}{p(\mathbf{m})} &= \frac{k!\,(N-k)!}{N!\ \Delta k\ (\Delta x)_{j'}\ \prod_{i=1}^{k-1} (\Delta x)_i} \frac{N!\ \Delta k\ (\Delta x)_j\ \prod_{i=1}^{k-1} (\Delta x)_i}{k!\,(N-k)!} \\
&= \frac{(\Delta x)_j}{(\Delta x)_{j'}},
\end{aligned} \tag{39}
$$

where $(\Delta x)_j$ denotes the range spanned by $\mathbf{x}$ at the depth associated with the jth layer and j' the layer corresponding to the perturbed Voronoi cell. The acceptance ratio then reads

$$
\alpha(\mathbf{m}' \mid \mathbf{m}) = \min\left[1, \frac{(\Delta x)_j}{(\Delta x)_{j'}} \exp\left\{ - \frac{\mathbf{\Phi}(\mathbf{m}') - \mathbf{\Phi}(\mathbf{m})}{2} \right\}\right]. \tag{40}
$$

4.2 CASE 2 (birth): $\mathbf{m}' \in \mathbb{R}^{k+1}$, $\mathbf{m} \in \mathbb{R}^k$

$$\alpha(\mathbf{m}' \mid \mathbf{m}) = \min\left[1, \frac{\theta_3\sqrt{2\pi}}{(\Delta v)_j}\exp\left\{\frac{\left(v'_{k+1} - v_i\right)^2}{2\theta_3^2} - \frac{\boldsymbol{\Phi}(\mathbf{m}') - \boldsymbol{\Phi}(\mathbf{m})}{2}\right\}\right], \tag{41}$$

where j is the newly added layer.

5 Hierarchical Bayesian Inversion

In the previous sections, we assumed that the data noise covariance $\mathbf{C}_e$ is known. When this is not the case, it is possible to modify the above transdimensional scheme so as to treat the noise as a free parameter of the inversion. As shown by *Bodin et al.* [2012], this is done by introducing an hyperparameter $\mathbf{h} = [\sigma, r]$, where σ denotes (period-independent) standard deviation and the correlation between two adjacent samples in the dispersion curve. Then, the determinant

$$\mid \mathbf{C}_e \mid = \sigma^{2n}(1 - r^2)^{n-1} \tag{42}$$

can be perturbed along the Markov chain and used to calculate the likelihood ratio

$$\begin{aligned} \frac{p\left(\mathbf{d}_{obs} \mid \mathbf{m}'\right)}{p\left(\mathbf{d}_{obs} \mid \mathbf{m}\right)} &= \frac{\exp\left\{\frac{-\boldsymbol{\Phi}(\mathbf{m}')}{2}\right\}}{\sqrt{(2\pi)^n\,|\mathbf{C}'_e|}}\,\frac{\sqrt{(2\pi)^n\,|\mathbf{C}_e|}}{\exp\left\{\frac{-\boldsymbol{\Phi}(\mathbf{m})}{2}\right\}} \\ &= \sqrt{\frac{\mid \mathbf{C}_e \mid}{\mid \mathbf{C}'_e \mid}}\exp\left\{-\frac{\boldsymbol{\Phi}(\mathbf{m}') - \boldsymbol{\Phi}(\mathbf{m})}{2}\right\}. \end{aligned} \tag{43}$$

References

Bodin, T., M. Sambridge, H. Tkalčić, P. Arroucau, K. Gallagher, and N. Rawlinson, Trans-dimensional inversion of receiver functions and surface wave dispersion, *Journal of Geophysical Research: Solid Earth*, *117*(B2), 2012.

Green, P. J., Trans-dimensional markov chain monte carlo, *Oxford Statistical Science Series*, pp. 179–198, 2003.