

Math in Transdimensional Bayesian Inversions: a Rigorous Formulation

Fabrizio Magrini¹, Giovanni Diaferia², and Luca De Siena¹

¹Institute of Geosciences, Johannes Gutenberg University, Mainz, Germany

²INGV

August 29, 2023

1 Classical, Isotropic Transdimensional Inversion

The proposed algorithm aims to find

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

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

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

where $\Phi(\mathbf{m}) = (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})^T \cdot \mathbf{C}_e^{-1} \cdot (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})$ 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

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

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

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

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

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

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

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

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

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

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

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

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

47 onto v_i , i.e. the proposal probability distribution

$$q_1(v'_i | v_i) = \frac{1}{\theta_1 \sqrt{2\pi}} \exp\left\{-\frac{(v'_i - v_i)^2}{2\theta_1^2}\right\}, \quad (9)$$

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

$$v'_i = v_i + u, \quad (10)$$

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

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

$$q_2(c'_i | c_i) = \frac{1}{\theta_2 \sqrt{2\pi}} \exp\left\{-\frac{(c'_i - c_i)^2}{2\theta_2^2}\right\}, \quad (11)$$

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

$$q_3(v'_{k+1} | v_i) = \frac{1}{\theta_3 \sqrt{2\pi}} \exp\left\{-\frac{(v'_{k+1} - v_i)^2}{2\theta_3^2}\right\}, \quad (12)$$

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

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

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

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

63 1.4.1 CASE 1 (change in velocity or position): $\mathbf{m}', \mathbf{m} \in \mathbb{R}^k$

64 When \mathbf{m}' has the same dimension of \mathbf{m} , the calculation of the ratios at the RHS of equation
65 (8) is straightforward. Since Δv is a constant in expression (7) and k has not changed,
66 $p(\mathbf{m}) = p(\mathbf{m}')$ 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.

$$\begin{aligned} q_1(v'_i | v_i) &= q_1(v_i | v'_i) \\ q_2(c'_i | c_i) &= q_2(c_i | c'_i), \end{aligned} \quad (14)$$

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

The acceptance probability therefore reduces to

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

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. $p(\mathbf{m}) = p(\mathbf{c}, \mathbf{v} | k) p(k)$, the ratio

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

where we noticed that $p(k) = p(k+1) = 1/\Delta k$ 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

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

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

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

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 $N \rightarrow \inf$.

80 1.4.3 CASE 3 (death): $\mathbf{m}' \in \mathbb{R}^{k-1}$, $\mathbf{m} \in \mathbb{R}^k$

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

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

83 2 On the effects of additional free parameters (uniformly dis- 84 tributed)

85 The previous section provides an all-around framework for writing a reversible-jump Markov
86 chain Monte Carlo algorithm, but what happens when additional free parameters are added
87 to the problem? To investigate this matter, let us define a general parameter, say \mathbf{x} , which
88 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
89 the density of a given layer in our Earth model.) In general, \mathbf{x} will be allowed to vary within
90 a range defined by $\Delta x = x_{max} - x_{min}$.

91 2.1 Prior probability: $p(\mathbf{m})$

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

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

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

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

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

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

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

$$p(\mathbf{m}) = \frac{k! (N - k)!}{N! \Delta k (\Delta x_1)^k (\Delta x_2)^k \dots (\Delta x_n)^k}, \quad (23)$$

101 2.2 Acceptance probability: $\alpha(\mathbf{m}' | \mathbf{m})$

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

105 **2.2.1 CASE 1 (change in velocity, position, or \mathbf{x}): $\mathbf{m}', \mathbf{m} \in \mathbb{R}^k$**

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

109 **2.2.2 CASE 2 (birth): $\mathbf{m}' \in \mathbb{R}^{k+1}$, $\mathbf{m} \in \mathbb{R}^k$**

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

$$q_4(x'_i | x_i) = \frac{1}{\theta_4 \sqrt{2\pi}} \exp \left\{ -\frac{(x'_i - x_i)^2}{2\theta_4^2} \right\}, \quad (24)$$

114 and

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

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

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

116 where $\theta_{x_1}, \theta_{x_2}, \dots, \theta_{x_n}$ denote the a-priori standard deviations of the Gaussians used to
 117 perturb the n free parameters.

118 It follows that

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

119 **2.2.3 CASE 3 (death): $\mathbf{m}' \in \mathbb{R}^{k-1}$, $\mathbf{m} \in \mathbb{R}^k$**

120 As in the previous sections, the relevant formulas can be easily obtained by taking the recip-
121 rocal of the above equations, and

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

122 **3 Introduction of a normally-distributed additional free pa-**
123 **rameters**

124 Let us consider now the introduction of a free parameter, say ξ , whose probability of oc-
125 currence is described by a normal distribution $\mathcal{N}(\mu_\xi, \sigma_\xi)$ which is independent from posi-
126 tion/depth. Then

$$p(\xi | k) = \prod_{i=1}^k \frac{1}{\sigma_\xi \sqrt{2\pi}} \exp \left\{ -\frac{(\xi_i - \mu_\xi)^2}{2\sigma_\xi^2} \right\}, \quad (29)$$

127 and

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

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

130 **3.1 CASE 1 (change in ξ): $\mathbf{m}', \mathbf{m} \in \mathbb{R}^k$**

131 Similar to the previous sections, if any of the uniformly-distributed free parameters $\mathbf{x}_1, \mathbf{x}_2,$
132 \dots, \mathbf{x}_n are perturbed, the prior ratio $\frac{p(\mathbf{m}')}{p(\mathbf{m})} = 1$. However, this is not the case when ξ is
133 perturbed. If we assume that only the j th element of ξ has changed in the proposed model
134 $p(\mathbf{m}')$,

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

135 If we now define the proposal probability

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

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

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

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

141 3.2 CASE 2 (birth): $\mathbf{m}' \in \mathbb{R}^{k+1}$, $\mathbf{m} \in \mathbb{R}^k$

142 Based on equation (30),

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

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

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

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

145 The acceptance probability then reads

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

146 **3.3 CASE 3 (death):** $\mathbf{m}' \in \mathbb{R}^{k-1}$, $\mathbf{m} \in \mathbb{R}^k$

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

147 4 Depth-dependent parameters, uniformly distributed at each 148 depth

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

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

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

153 4.1 CASE 1: DEPTH PERTURBATION, $\mathbf{m}', \mathbf{m} \in \mathbb{R}^k$

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

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

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

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

159 4.2 CASE 2 (birth): $\mathbf{m}' \in \mathbb{R}^{k+1}$, $\mathbf{m} \in \mathbb{R}^k$

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

160 where j is the newly added layer.

161 5 Hierarchical Bayesian Inversion

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

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

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

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

169 References

- 170 Bodin, T., M. Sambridge, H. Tkalčić, P. Arroucau, K. Gallagher, and N. Rawlinson, Trans-
 171 dimensional inversion of receiver functions and surface wave dispersion, *Journal of Geo-*
 172 *physical Research: Solid Earth*, 117(B2), 2012.
- 173 Green, P. J., Trans-dimensional markov chain monte carlo, *Oxford Statistical Science Series*,
 174 pp. 179–198, 2003.