Geophysical Journal International

Geophys. J. Int. (2020) **223,** 1837–1863
Advance Access publication 2020 September 1
GJI General Geophysical Methods

doi: 10.1093/gji/ggaa415

A reduced order approach for probabilistic inversions of 3-D magnetotelluric data I: general formulation

M.C. Manassero,[1] J.C. Afonso,[1,2] F. Zyserman,[3] S. Zlotnik[4] and I. Fomin[1]

[1] *Australian Research Council Centre of Excellence for Core to Crust Fluid Systems/GEMOC, Department of Earth and Environmental Sciences, Macquarie University, 2109, Sydney, Australia. E-mail: constanzamanassero@gmail.com*
[2] *Centre for Earth Evolution and Dynamics, Department of Geosciences, University of Oslo, 0315, Oslo, Norway*
[3] *CONICET, Facultad de Ciencias Astronómicas y Geofísicas, Universidad de La Plata, 1900, La Plata, Argentina*
[4] *Laboratori de Càlcul Numèric, Escola Tècnica Superior d'Enginyers de Camins, Canals i Ports, Universitat Politècnica de Catalunya, 08034, Barcelona, Spain*

Accepted 2020 August 29. Received 2020 August 18; in original form 2020 April 30

SUMMARY

Simulation-based probabilistic inversions of 3-D magnetotelluric (MT) data are arguably the best option to deal with the nonlinearity and non-uniqueness of the MT problem. However, the computational cost associated with the modelling of 3-D MT data has so far precluded the community from adopting and/or pursuing full probabilistic inversions of large MT data sets. In this contribution, we present a novel and general inversion framework, driven by Markov Chain Monte Carlo (MCMC) algorithms, which combines (i) an efficient parallel-in-parallel structure to solve the 3-D forward problem, (ii) a reduced order technique to create fast and accurate surrogate models of the forward problem and (iii) adaptive strategies for both the MCMC algorithm and the surrogate model. In particular, and contrary to traditional implementations, the adaptation of the surrogate is integrated into the MCMC inversion. This circumvents the need of costly offline stages to build the surrogate and further increases the overall efficiency of the method. We demonstrate the feasibility and performance of our approach to invert for large-scale conductivity structures with two numerical examples using different parametrizations and dimensionalities. In both cases, we report staggering gains in computational efficiency compared to traditional MCMC implementations. Our method finally removes the main bottleneck of probabilistic inversions of 3-D MT data and opens up new opportunities for both stand-alone MT inversions and multi-observable joint inversions for the physical state of the Earth's interior.

Key words: Composition and structure of the mantle; Magnetotellurics; Inverse theory; Numerical approximations and analysis; Numerical modelling.

1 INTRODUCTION

The magnetotelluric (MT) method (Tikhonov 1950; Cagniard 1953; Vozoff 1990) is a passive electromagnetic (EM) technique that aims to determine the electrical conductivity distribution of the subsurface at depths from tens of metres to hundreds of kilometres (Vozoff 1990). Given the strong sensitivity of the electrical conductivity of rocks to temperature, hydrogen and fluid/melt content inside the Earth (Karato 1990; Roberts & Tyburczy 1999; Ten Grotenhuis *et al.* 2005; Wang *et al.* 2006; Dai & Karato 2009; Yoshino *et al.* 2009, 2012; Yoshino 2010; Karato 2011; Evans 2012; Karato & Wang 2013; Selway 2014; Khan 2016), MT has become one of the most useful and cost-efficient techniques to identify deep fluid pathways and associated mineralization events (e.g. Ogawa *et al.* 2001; Wei *et al.* 2001), characterize pluming systems in active volcanic areas (e.g. Heise *et al.* 2008; Comeau *et al.* 2016), detect partial melting in the deep mantle (e.g. Evans *et al.* 2005; Selway *et al.* 2019) and study the thermochemical structure of the lithosphere in general (e.g. Fullea *et al.* 2011; Meqbel *et al.* 2014; Jones *et al.* 2017). These and other features of MT gave impetus for recent data acquisition programs over large grids such as the MT component of the USArray in US (www.usarray.org/researchers/obs/magnetotelluric), the AusLAMP program in Australia (www.ga.gov.au/about/projects/resources/auslamp) and the MT component of Sinoprobe in China (www.sinoprobe.org).

© The Author(s) 2020. Published by Oxford University Press on behalf of The Royal Astronomical Society. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

Similarly, the availability of high-quality magnetic field measurements with uniform global coverage done by satellite missions (e.g. see Sabaka *et al.* 2004; Olsen *et al.* 2006; Kuvshinov 2008; Olsen *et al.* 2013) makes it now possible to study lithospheric and whole-mantle conductivity anomalies at global scale (e.g. Kuvshinov & Olsen 2006; Kuvshinov *et al.* 2006; Kelbert *et al.* 2009; Kuvshinov & Semenov 2012; Semenov & Kuvshinov 2012).

The formal combination of MT measurements with gravity and seismic data via joint inversions (e.g. Moorkamp *et al.* 2007; Jegen *et al.* 2009; Moorkamp *et al.* 2010; Vozar *et al.* 2014; Bennington *et al.* 2015; Jones *et al.* 2017) also holds great potential, as these data sets have complementary sensitivities to the physical state of the lithosphere. For instance, specific types of seismic data can be used in combination with gravity observations to provide the background (or regional) temperature structure and compositional layering (Khan *et al.* 2008; Afonso *et al.* 2013a,b; Jegen *et al.* 2009; Afonso *et al.* 2016a). By accounting for the effect of this background in MT data, conductivity anomalies associated with factors other than temperature and bulk composition (e.g. fluid content, presence of melt and anomalous mineral assemblages) could be identified more reliably.

Due to the potential of MT methods, significant efforts have been made in the past two decades to develop robust and efficient forward and inversion algorithms (e.g. deGroot Hedlin & Constable 1990; Mackie & Madden 1993; Newman & Alumbaugh 2000; Zhdanov *et al.* 2000; Newman & Alumbaugh 2000; Sasaki 2001; Farquharson *et al.* 2002; Siripunvaraporn *et al.* 2005; Han *et al.* 2008; Avdeev & Avdeeva 2009; Siripunvaraporn & Egbert 2009; Egbert & Kelbert 2012; Key 2016). In particular, deterministic methods have been successfully applied in the inversion of high-dimensional MT problems and currently represent the most widely used techniques for inverting 3-D MT data (Mackie *et al.* 1996; Zhdanov *et al.* 2010; Tietze & Ritter 2013; Meqbel *et al.* 2014; Yang *et al.* 2015; Robertson *et al.* 2020). Some popular 3-D deterministic inversion algorithms include those of Mackie & Madden (1993), Newman & Alumbaugh (2000), Newman & Boggs (2004), Kelbert *et al.* (2008, 2014), Siripunvaraporn *et al.* (2005), Siripunvaraporn & Egbert (2009) and Egbert & Kelbert (2012), based on finite-difference forward solvers; the algorithms developed by Zhdanov & Hursan (2000), Zhdanov *et al.* (2000) and Avdeev & Avdeeva (2009), based on an integral equation formulation of the forward solver; and that of Haber *et al.* (2004), based on a finite-element (FE) formulation. A review of the different methods available for inversion of 3-D MT data can be found in Avdeev (2005) and Siripunvaraporn (2012).

The main advantage of deterministic inversions is that they typically require only a few tens or hundreds of forward evaluations to achieve convergence. This is particularly important to make large-scale 3-D MT inversions practical, as the forward problem is notoriously computationally expensive. However, deterministic solutions of nonlinear problems often depend on the choice of the initial model (to be perturbed during the inversion) and on arbitrary regularization parameters to smooth-out the solution (e.g. Robertson *et al.* 2020). It is well known that if our bias is inappropriate for the problem at hand, the regularized solution can be far from the true solution (Shen *et al.* 2012; Aster *et al.* 2018). In addition, the linearization of the problem and optimization algorithms used in determinisitc inversions are tailored for the search of a single optimal or best-fitting model. This makes it difficult, if not impossible, to perform thorough statistical analysis of model uncertainties and (perhaps more importantly) non-uniqueness.

Probabilistic inversion methods (Gregory 2005; Tarantola 2005; Mosegaard & Hansen 2016) are better suited to handle these limitations. They do not focus on single best solutions but rather on providing a full probability distribution over the whole model parameter space. This distribution, known as the *posterior distribution*, contains all the information about the unknown parameters given data and modelling assumptions and thus, it represents the most general solution to the inverse problem. The main benefits of probabilistic solutions to the MT inverse problem (including full consideration of its nonlinear nature) have been recently highlighted by Conway *et al.* (2018), Brodie & Jiang (2018), Jones *et al.* (2017), Chen *et al.* (2012) and Rosas-Carbajal *et al.* (2013) in the context of 1-D and 2-D inversions driven by Markov Chain Monte Carlo (MCMC) algorithms. However, for the case of high-dimensional 3-D MT inversions, the large number of costly forward computations required ($\sim 10^5$–10^7) has so far rendered it impractical. For instance, Rosas-Carbajal *et al.* (2015) performed a time-lapse 3-D MCMC inversion using a Legendre moment decomposition of a saline plume to reduce the dimensionality of the inverse problem to only fourteen parameters. Despite this drastic reduction of the parameter space, convergence was achieved after 60 d of computation.

It is clear that if fully probabilistic inversions of high-dimensional 3-D MT problems are to become a reality, significantly more efficient (yet accurate) forward solvers need to be developed. This is precisely the purpose of this work, which presents a new reduced order strategy to solve the probabilistic 3-D MT inverse problem. The new approach is based on a judicious combination of a reduced order technique (Reduced Basis Method) to construct fast and accurate surrogate models, an efficient parallel-in-parallel structure of the 3-D forward problem and adaptive strategies for both the surrogate and the MCMC algorithm. We will illustrate the feasibility and general performance of our approach with two numerical examples of complex, lithospheric-scale conductivity structures. In particular, we will demonstrate that probabilistic inversions of high-dimensional 3-D MT problems are now a practical option, even with modest computational resources.

The remainder of the paper is structured as follows: Section 2 reviews the general formulation of the probabilistic inverse problem. Section 3 introduces the forward problem in terms of primary and secondary fields, and the associated variational formulation. Sections 4 and 5 describe the Reduced Basis (RB) Method and our combined RB+MCMC strategy, respectively. Section 6 presents numerical examples that illustrate the benefits and limitations of the method in the context of stand-alone 3-D MT inversions for large-scale lithospheric structures. Section 7 deals with practical aspects of our approach, further improvements and remaining challenges. Lastly, the main outcomes and results of this work are summarized in Section 8.

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

2 PROBABILISTIC INVERSE PROBLEMS

2.1 Bayesian inversion

The Bayesian approach (sometimes referred to as the *statistical* or *probabilistic* approach) to the inverse problem recognizes that the data ($\mathbf{d}$) and the model parameters ($\mathbf{m}$) that characterize the system will never be known with absolute certainty and it considers them to be random variables represented by an associated probability distribution (e.g. Tarantola *et al.* 1982; Gilks *et al.* 1995; Sen & Stoffa 1996; Gregory 2005; Mosegaard & Hansen 2016). The most general solution to the inverse problem is therefore a *posterior* probability density function (PDF) over the parameter space which is formally given by the Bayes' rule

$$P(\mathbf{m}|\mathbf{d}) = \frac{P(\mathbf{d}|\mathbf{m})P(\mathbf{m})}{P(\mathbf{d})}, \tag{1}$$

where $P(\mathbf{m}|\mathbf{d})$ is the posterior conditional PDF of $\mathbf{m}$ given $\mathbf{d}$. $P(\mathbf{d}|\mathbf{m})$ is the conditional PDF of $\mathbf{d}$ given $\mathbf{m}$, $P(\mathbf{m})$ is the prior PDF of the model parameters $\mathbf{m}$ and $P(\mathbf{d})$ is the prior PDF of the data. When a particular observation is made, $P(\mathbf{d}|\mathbf{m})$ is considered a function of $\mathbf{m}$ and referred to as the *likelihood*, $\mathcal{L}(\mathbf{m})$. Considering this and the fact that $P(\mathbf{d})$ is independent of $\mathbf{m}$, eq. (1) can be written as

$$P(\mathbf{m}|\mathbf{d}) \propto \mathcal{L}(\mathbf{m})P(\mathbf{m}). \tag{2}$$

In the most general case of high-dimensional and nonlinear problems with complex priors, the only practical solution for eq. (2) is based on the construction of a Markov Chain that has $P(\mathbf{m}|\mathbf{d})$ as their equilibrium or stationary distribution (Gilks *et al.* 1995; Gregory 2005; Tarantola 2005). MCMC algorithms are designed to produce unbiased approximations of the true posterior by repeatedly drawing models $\mathbf{m}_t$ and evaluating their posterior probability $P(\mathbf{m}_t|\mathbf{d})$.

In this work, we use the Metropolis–Hastings (MH) algorithm (Metropolis *et al.* 1953; Hastings 1970) combined with the Adaptive Metropolis (AM) approach of Haario *et al.* (2001). At the core of these algorithms is a *proposal distribution*, $q(\cdot\,|\,\cdot)$, used to generate new moves or trials. In principle, the proposal distribution can be chosen rather freely from a large family of distributions (cf. Gilks *et al.* 1995; Gregory 2005). In practice, however, the actual form of $q(\cdot\,|\,\cdot)$ exerts a strong influence on the efficiency of the MCMC algorithm. The AM algorithm alleviates the problem of choosing an optimal proposal before the MCMC simulation starts by updating the proposal at regular intervals according to the statistics of the chain. The reader is referred to Haario *et al.* (2001) for further details on the AM algorithm.

At each step of the MCMC simulation, a new move in the parameter space ($\mathbf{m}_t$) is proposed from $q(\cdot\,|\,\cdot)$ and its posterior probability is compared to that of the current state of the chain ($\mathbf{m}_{t-1}$). If the new proposal has a higher probability than the current state, it is accepted as part of the chain and the current position is updated to the new state. If the new proposal has a lower probability than the current state, it is accepted with probability $\alpha(\mathbf{m}_{t-1}, \mathbf{m}_t)$, defined as

$$\alpha(\mathbf{m}_{t-1}, \mathbf{m}_t) = \min\left\{1, \frac{\mathcal{L}(\mathbf{m}_t)P(\mathbf{m}_t)\, q(\mathbf{m}_{t-1}|\mathbf{m}_t)}{\mathcal{L}(\mathbf{m}_{t-1})P(\mathbf{m}_{t-1})\, q(\mathbf{m}_t|\mathbf{m}_{t-1})}\right\}, \tag{3}$$

where we have assumed that the new proposal is conditional on the current state of the chain (Gilks *et al.* 1995; Gregory 2005). Details about the proposal distribution and the AM algorithm relevant to this work can be found in Appendix B.

2.2 The likelihood function

The likelihood function is a measure of the distance between the observed data and the solution of the forward model, that is, it is determined by the statistical distribution of the data errors. Under the common assumption of additive independent Gaussian errors, the likelihood function is given by:

$$\mathcal{L}(\mathbf{m}|\mathbf{d}) = \frac{1}{(2\pi)^{N/2}\prod_{i=1}^{N} s_i} \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\left(\frac{g_i(\mathbf{m}) - d_i}{s_i}\right)^2\right], \tag{4}$$

where N is the number of independent observations, s_i denotes the standard deviation of the ith data error and $\mathbf{g}(\mathbf{m})$ is the data predicted by the *forward problem* for the model $\mathbf{m}$.

3 THE MAGNETOTELLURIC FORWARD PROBLEM

3.1 Magnetotelluric equations

The basis of the MT method is the EM induction problem, viz. EM fields are induced within the Earth by EM fields that propagate perpendicular to the surface of the Earth as a plane monochromatic wave of frequency ω. The induced EM fields satisfy the following time-harmonic Maxwell's partial differential equations:

$$\nabla \times \mathbf{H} = \sigma \mathbf{E}, \tag{5a}$$

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

$$\nabla \times \mathbf{E} = -i\omega\mu_0\mathbf{H}, \tag{5b}$$

where $\mathbf{E}$ is the electric field ($\mathrm{V\,m^{-1}}$), $\mathbf{H}$ is the magnetic field ($\mathrm{A\,m^{-1}}$), μ_0 is the magnetic permeability of free space ($\mathrm{Vs\,A^{-1}\,m^{-1}}$) and σ is the electrical conductivity of the medium ($\mathrm{S\,m^{-1}}$).

3.2 Variational formulation

This paper follows the formulation of Zyserman & Santos (2000) where the electrical conductivity distribution in a domain $\Omega \in R^3$ is considered as the superposition of a *background* or primary conductivity of a layered Earth (σ_p) and an additional conductivity associated with the presence of 3-D anomalies (σ_s). This formulation allows to rewrite eqs (5) in terms of *primary* fields ($\mathbf{H}_p$ and $\mathbf{E}_p$) induced in the layered Earth and *secondary* fields ($\mathbf{H}_s$ and $\mathbf{E}_s$) which are generated by the presence of the conductivity anomalies. Using the absorbent boundary conditions defined by Sheen (1997), our MT problem in 3-D is formulated as follows:

Find $\mathbf{E}_s$ and $\mathbf{H}_s$ in a domain $\Omega \in R^3$ such that,

$$\sigma\mathbf{E}_s - \nabla \times \mathbf{H}_s = -\mathbf{F} = -\sigma_s\mathbf{E}_p \ \text{ in } \Omega, \tag{6a}$$

$$i\omega\mu_0\mathbf{H}_s + \nabla \times \mathbf{E}_s = 0 \ \text{ in } \Omega, \tag{6b}$$

$$(1-i)P_\tau a\mathbf{E}_s + \nu \times \mathbf{H}_s = 0 \ \text{ on } \partial\Omega \equiv \Gamma, \tag{6c}$$

where $\Gamma \equiv \partial\Omega$ is the boundary of the domain Ω and a is defined as $a = (\sigma/2\omega\mu_0)^{1/2}$. In the first term of eq. (6c), $P_\tau\boldsymbol{\varphi} = \boldsymbol{\varphi} - \nu(\nu \cdot \boldsymbol{\varphi})$ refers to the projection of the trace of any vector $\boldsymbol{\varphi}$ on Γ where ν is the unit outer normal to Γ.

Omitting the subscripts for the secondary fields, the variational formulation of the mixed problem solved in eqs (6) reads as follows (Douglas *et al.* 2000):

Find $(\mathbf{E}, \mathbf{H})$ in the suitable spaces $\mathcal{V} \times \mathcal{W}$ such that,

$$(\sigma\mathbf{E}, \boldsymbol{\varphi}) - (\mathbf{H}, \nabla \times \boldsymbol{\varphi}) + (1-i)\langle P_\tau a\mathbf{E}, P_\tau\boldsymbol{\varphi}\rangle_\Gamma = (\mathbf{F}, \boldsymbol{\varphi}) \quad \boldsymbol{\varphi} \in \mathcal{V}, \tag{7a}$$

$$i\omega\mu_0(\mathbf{H}, \boldsymbol{\psi}) + (\nabla \times \mathbf{E}, \boldsymbol{\psi}) = 0 \quad \boldsymbol{\psi} \in \mathcal{W}. \tag{7b}$$

3.3 Numerical forward solution

Numerical solutions to the variational problem in eqs (7) are obtained using a non-conforming Galerkin FE method (Douglas *et al.* 1999, 2000; Zyserman & Santos 2000). This requires finding $(\mathbf{E}^h, \mathbf{H}^h) \in \mathcal{V}^h \times \mathcal{W}^h$ such that,

$$(\sigma\mathbf{E}^h, \boldsymbol{\varphi}) - \sum_j^{N_e}(\mathbf{H}^h, \nabla \times \boldsymbol{\varphi})_{\Omega_j} + (1-i)\langle\langle P_\tau a\mathbf{E}^h, P_\tau\boldsymbol{\varphi}\rangle\rangle_\Gamma = \ell(\boldsymbol{\varphi}) \quad \forall\boldsymbol{\varphi} \in \mathcal{V}^h, \tag{8a}$$

$$i\omega\mu_0(\mathbf{H}^h, \boldsymbol{\psi}) + \sum_j^{N_e}(\nabla \times \mathbf{E}^h, \boldsymbol{\psi})_{\Omega_j} = 0 \qquad \forall\boldsymbol{\psi} \in \mathcal{W}^h, \tag{8b}$$

where N_e is the number of non-overlapping parallelepipeds Ω_j such that $\Omega = \cup_j\Omega_j, j = 1,\ldots, N_e$; $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ are the set of FE shape functions for the electric and magnetic field, respectively, and N_{FE} and M_{FE} are the degrees of freedom for each field. The FE subspaces $\mathcal{V}_h \subset \mathcal{V}$ and $\mathcal{W}_h \subset \mathcal{W}$ are defined as

$$\mathcal{V}_h = \mathrm{span}\{\varphi^1, ..., \varphi^{N_{FE}}\}, \tag{9a}$$

$$\mathcal{W}_h = \mathrm{span}\{\psi^1, ..., \psi^{M_{FE}}\}. \tag{9b}$$

The probabilistic behaviour of the 3-D MT inverse problem is introduced when we assume that the electrical conductivity is a random field of the form $\sigma(\mathbf{x}, \theta)$ where $\mathbf{x} \in \Omega$ is the vector position and $\theta \in \Theta$ refers to the randomness, being Θ the set of all possible outcomes of θ. For a specific sample of $\sigma(\mathbf{x}, \theta)$, the approximate electric and magnetic field on each Ω_j can be expressed in terms of the shape functions as follows:

$$\mathbf{E}^h(\mathbf{x}, \theta) = \sum_{\alpha=1}^{N_{FE}} \varepsilon^\alpha(\theta)\boldsymbol{\varphi}^\alpha(\mathbf{x}), \tag{10}$$

$$\mathbf{B}^h(\mathbf{x}, \theta) = \sum_{\eta=1}^{M_{FE}} h^\eta(\theta)\boldsymbol{\psi}^\eta(\mathbf{x}), \tag{11}$$

where $\varepsilon^\eta(\theta)$ and $h^\alpha(\theta)$ are the unknown coefficients to be determined for a that sample $\sigma(\mathbf{x}, \theta)$.

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

Introducing eqs (10) and (11) into eqs (8) and selecting one shape function at the time $\boldsymbol{\varphi}^n$, $n = 1, \ldots, N_{FE}$ and $\boldsymbol{\psi}^m$, $m = 1, \ldots, M_{FE}$ we obtain the following algebraic system of equations for the sample $\sigma(\mathbf{x}, \theta)$:

$$\Big(\sigma(\mathbf{x},\theta)\sum_{\alpha=1}^{N_{FE}}\varepsilon^{\alpha}(\theta)\boldsymbol{\varphi}^{\alpha}(\mathbf{x}),\boldsymbol{\varphi}^{n}(\mathbf{x})\Big)-\sum_{j}^{N_e}\Big(\sum_{\eta=1}^{M_{FE}}h^{\eta}(\theta)\boldsymbol{\psi}^{\eta}(\mathbf{x}),\nabla\times\boldsymbol{\varphi}^{n}(\mathbf{x})\Big)_{\Omega_j}+(1-i)\langle\langle P_{\tau}a\sum_{\alpha=1}^{N_{FE}}\varepsilon^{\alpha}(\theta)\boldsymbol{\varphi}^{\alpha}(\mathbf{x}),P_{\tau}\boldsymbol{\varphi}^{n}(\mathbf{x})\rangle\rangle_{\Gamma}=\ell(\boldsymbol{\varphi}^{n}(\mathbf{x}))\,\forall\boldsymbol{\varphi}^{n}, \quad (12a)$$

$$i\omega\mu_0\Big(\sum_{\eta=1}^{M_{FE}}h^{\eta}(\theta)\boldsymbol{\psi}^{\eta}(\mathbf{x}),\boldsymbol{\psi}^{m}(\mathbf{x})\Big)+\sum_{j}^{N_e}\Big(\nabla\times\sum_{\alpha=1}^{N_{FE}}\varepsilon^{\alpha}(\theta)\boldsymbol{\varphi}^{\alpha}(\mathbf{x}),\boldsymbol{\psi}^{m}(\mathbf{x})\Big)_{\Omega_j}=0\,\forall\boldsymbol{\psi}^{m}. \quad (12b)$$

Eqs (12) can be further simplified as the choice of the shape functions establishes a relationship between h^{η} and ε^{α} (see Douglas *et al.* 2000; Zyserman & Santos 2000). When computing the integral over an element Ω_j, the first term of eq. (12b) is different from zero only when $\eta = m$ allowing the coefficients h^{η} of eq. (12b) to be be expressed in terms of ε^{α} and replaced in eq. (12a). This simplification results in a linear system of equations where the unknown is a vector $\mathbf{U}$ with the coefficients ε^{α} for the approximated electric field in the whole domain. This system of equations can be represented in the following matrix form:

$$K(\theta)\mathbf{U}(\theta) = \mathbf{F}(\theta), \quad (13)$$

where K is a sparse and symmetric *stiffness matrix* of size $N_{FE} \times N_{FE}$ and $\mathbf{F}$ is the force vector of size $N_{FE} \times 1$.

In our implementation of the original code of Zyserman & Santos (2000), the system of eqs (13) is solved using the MUlti-frontal Massively Parallel Solver (MUMPS) version 5.1.2 (Amestoy *et al.* 2001, 2006). This allows us to define an efficient parallel-in-parallel structure. The first level of parallelization involves assigning a block of *np* processors to each frequency (i.e. parallelization by frequencies). The second level involves parallelizing the solution of the actual forward problem (eqs 13) for each frequency using MUMPS and the *np* processors in each block. This structure reduces the computational time of each forward solution by ~ 80 per cent, regardless of the number of frequencies used. Results for standard 3-D MT benchmarks using this implementation are provided in Section 1 of the Supporting Information.

Despite the parallel strategy described above, the computational time taken by the forward problem remains the main limitation preventing its use in MCMC-driven probabilistic inversions of 3-D MT data. The following section presents a novel strategy based on reduced order modelling that overcomes this limitation.

4 REDUCED BASIS

In order to ameliorate the computational burden of probabilistic inversions and optimization problems, various methods for reducing the cost of the full (high-fidelity) forward problem have been proposed (see reviews in Frangos *et al.* 2010; Peherstorfer *et al.* 2018). Among all available techniques, the projection-based RB method (Patera *et al.* 2007; Florentin & Díez 2012; Quarteroni *et al.* 2015; Hesthaven *et al.* 2016) is particularly well suited for our purposes due to its strong connection to FE discretizations, its implementation simplicity and its outstanding performance in situations where the same physical problem needs to be solved many times for different input parameters (as is the case in MCMC-driven inversions). In the RB method, the solution of every new model (i.e. forward problem) is sought as a projection onto a space built from previous high-fidelity solutions called *bases*. The space containing these high-fidelity solutions is referred to as the *reduced basis*. Previous work has shown that RB can speedup the numerical solutions of complex problems by several orders of magnitude without compromising the accuracy of the solutions (e.g. Rozza *et al.* 2007; Chen *et al.* 2010; Lieberman *et al.* 2010; Florentin & Díez 2012; Rozza *et al.* 2013; Cui *et al.* 2015; Ortega-Gelabert *et al.* 2020). In this paper, we will show that staggering gains in computational time can also be achieved for the 3-D MT problem. The reader is referred to Quarteroni *et al.* (2011, 2015), Benner *et al.* (2017) and Hesthaven *et al.* (2016) for more details on RB methods and their applications to a range of problems.

4.1 Reduced Basis solution

The RB approach used in this paper (cf. Florentin & Díez 2012) disassociates the discretization space (Ω) and the stochastic space (Θ) in a sense that the spatial discretization used to find the high-fidelity FE solution is not allowed to vary across the MCMC simulations, that is, the stochasticity is only assigned to the conductivity distribution, not to the FE discretization. Since the coefficients h^{η} for the magnetic field can be expressed as functions of the coefficients ε^{α} for the electric field, the low-fidelity RB solution, $\mathbf{U_{RB}}$, seeks to approximate the electric field only,

$$\mathbf{U_{RB}}(\theta) \approx \mathbf{U}. \quad (14)$$

The main idea behind the RB method is to generate a subset $\mathcal{V}_{\mathcal{RB}}$ of $\mathcal{V}_h$ with dimension $N_{RB} \ll N_{FE}$ such that,

$$\mathcal{V}_{\mathcal{RB}} = \text{span}\{\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_{N_{RB}}\} \subset \mathcal{V}_h, \quad (15)$$

where the associated basis vectors $\mathbf{V}_i$ are defined as high-fidelity FE solutions of the system of eqs (13) for certain realizations of $\sigma(\mathbf{x}, \theta)$. The generation of these bases will be described in Section 5 and we assume, for the time being, that we have an available $\mathcal{V}_{\mathcal{RB}}$. While performing

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

the probabilistic inversion, the solution of eq. (13) for a new sample $\sigma(\mathbf{x}, \theta)$ is first sought as a linear combination of the basis vectors in $\mathcal{V}_{\mathcal{RB}}$:

$$\mathbf{U_{RB}}(\mathbf{x}, \theta) = \sum_{i=1}^{N_{RB}} a_i(\theta)\mathbf{V}_i = V_{RB}\mathbf{a}(\theta), \tag{16}$$

where $V_{RB} = [\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_{N_{RB}}]^{N_{FE}\times N_{RB}}$ and $\mathbf{a}(\theta) = [a_1, a_2, \ldots, a_{N_{RB}}]$ is the vector of unknown coefficients accompanying the bases $\mathbf{V}_i$.

Introducing $\mathbf{U_{RB}}(\mathbf{x}, \theta)$ into eq. (13) and multiplying both sides of the equation by V_{RB}^T, we obtain:

$$({V_{RB}}^T K(\theta) V_{RB})\mathbf{a}(\theta) = {V_{RB}}^T \mathbf{F}(\theta). \tag{17}$$

Defining $K_{RB}(\theta) := {V_{RB}}^T K(\theta) V_{RB}$ and $\mathbf{F_{RB}}(\theta) := {V_{RB}}^T \mathbf{F}(\theta)$, we can rewrite eq. (17) to obtain the following system of equations:

$$K_{RB}(\theta)\mathbf{a}(\theta) = \mathbf{F_{RB}}(\theta), \tag{18}$$

where K_{RB} is the RB matrix of size $N_{RB} \times N_{RB}$ and $\mathbf{F_{RB}}$ is the force vector of size $N_{RB} \times 1$.

The RB solution, $\mathbf{U_{RB}}(\theta)$, for a particular sample $\sigma(\mathbf{x}, \theta)$ is then found by solving the system of eqs (18) for the coefficients $\mathbf{a}(\theta)$ and substituting them into eq. (16). Note that the linear system of eqs (18) is of size $N_{RB} \ll N_{FE}$ and thus significantly less expensive than the complete FE system of eqs (13) of size N_{FE}.

4.2 Error estimation

Once the RB solution is obtained for a particular sample $\sigma(\mathbf{x}, \theta)$ and a particular basis V_{RB}, it is necessary to evaluate whether the basis was sufficiently rich to produce an accurate RB solution. Much work has been done to define reliable error estimators and to certify the accuracy of the RB solution (Chen *et al.* 2010; Quarteroni *et al.* 2011; Hesthaven *et al.* 2012, 2016). In this paper, we define the approximation error between the RB and the FE solution at every step of the MCMC simulation as:

$$\mathbf{E_{RB}} := \mathbf{U_{RB}} - \mathbf{U}. \tag{19}$$

Since $\mathbf{U}$ is the high-fidelity FE solution that we do not want to compute, we multiply $\mathbf{E_{RB}}$ by the stiffness matrix K to obtain the so-called residual of the RB solution:

$$K\mathbf{E_{RB}} = K\mathbf{U_{RB}} - K\mathbf{U}. \tag{20}$$

Replacing eq. (13) into eq. (20), dividing by the vector of forces $\mathbf{F}$ and computing the L_2 norm, we obtain the following RB relative error:

$$\mathbf{R_{RB}} := \frac{||K\mathbf{U_{RB}} - \mathbf{F}||}{||\mathbf{F}||} \tag{21}$$

which is an adequate estimator of the energy norm of the RB solution error (Quarteroni *et al.* 2015; Hesthaven *et al.* 2016). At this stage, a criterion is necessary to evaluate the accuracy of the approximation. The simplest criterion involves defining a tolerance β in a way that the RB solution becomes admissible only if $\mathbf{R_{RB}} \leq \beta$. In the case of a relative error larger than β, the RB solution is considered unacceptable and the high-fidelity FE solution is computed. This new FE solution is added to the space $\mathcal{V}_{\mathcal{RB}}$ (i.e. enrichment of $\mathcal{V}_{\mathcal{RB}}$), which is then used in subsequent MCMC steps. This process is further discussed in the following section.

5 A COMBINED RB+MCMC METHOD FOR 3-D MT INVERSION PROBLEMS

In traditional implementations of RB methods, the creation of the basis (computationally expensive as it involves many solutions of the high-fidelity FE problem) is computed in an *offline* stage at specific locations within the parameter space defined by an heuristic criterion or greedy procedure (Rozza *et al.* 2007, 2009). During the subsequent probabilistic inversion or simulation process, known as the *online* stage, these bases are used to obtain the fast low-fidelity RB solutions (Prud'Homme *et al.* 2002; Hess & Benner 2013; Quarteroni *et al.* 2015; Hesthaven *et al.* 2016; Hess 2016). This type of approach is particularly useful when the number of parameters is small and when only a small number of high-fidelity solutions are able to represent the complete parameter space.

In the context of high- and ultrahigh-dimensional ($>10^3$ parameters) probabilistic inversions, it is practically impossible to pre-explore the parameter space in an offline stage to create relevant bases that will guarantee accurate solutions within the entire parameter space, but most importantly, within the (so far unknown) high-probability regions. Put differently, in order to guide the creation of the relevant basis, we would have to know *a priori* the regions of high probability, which implies that we already have a solution to the inverse problem. While this is in principle possible by splitting the inversion into three stages—for example, an initial low-dimensional (fast) deterministic inversion, the offline creation of the relevant basis and a final MCMC inversion with the RB surrogate—in the most general case, an adaptive MCMC approach where the basis is computed and adapted during the MCMC simulation seems to be a more effective concept. This is particularly relevant to our goals because of the intrinsic behaviour of MCMC inversions, which typically start in a region of low probability (i.e. bad models) and quickly converge to a restricted region in the parameter space where the most probable models reside. Therefore, the RB solutions that may have been relevant to the initial stages of the chain will become less relevant when the chains have converged to the

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

high probability regions. For the same reason, fewer basis enrichment (high-fidelity solutions) are needed as the chain converges, because the RB space becomes rich enough to provide good solutions within the restricted parameter space. Florentin & Díez (2012), Cui *et al.* (2015), Yan & Zhou (2019), Zhang *et al.* (2019) and Ortega-Gelabert *et al.* (2020) have introduced similar ideas. Ortega-Gelabert *et al.* (2020), in particular, demonstrated the benefits of this type of RB adaptation and applied it in MCMC inversions of complex geodynamic problems. We emphasize that this approach does not preclude the option of using bases created in previous simulations. Indeed, bases generated during a previous inversion or during a parallel MCMC chain (if different chains can exchange information) can be easily re-used to improve the performance of both subsequent and simultaneous parallel inversions.

5.1 A hybrid approach for high-dimensional probabilistic inversions

Our hybrid approach is similar to that put forward by Ortega-Gelabert *et al.* (2020) in that it automatically updates/enriches the RB space $\mathcal{V}_{RB}$ during the MCMC inversion by adding high-fidelity solutions (i.e. new bases) as needed by the evolution of the chain. The main differences between the two approaches are in the way they estimate the error introduced by the RB approximation and in the implementation of the adaptive scheme for both the basis and the MCMC algorithms. For instance, our parallel structure allows us to create and adapt the basis per frequency and orientation of the EM fields. This adaptation is also controlled by the use of variable tolerance and an orthonormalization procedure (described below). In the case of the MCMC algorithm, we have implemented an AM algorithm (Section 6.3) to update the proposal and incorporated a modified ratio that accounts for Delayed Rejection (Haario *et al.* 2006; Mira *et al.* 2001) when a new high-fidelity solution is computed and the proposed sample is rejected (described below).

The basic steps of our RB+MCMC algorithm (Algorithm 1) can be summarized as follows:

(i) In the case of having previous solutions from an offline stage or from previous probabilistic inversions, we proceed to load these solutions as the initial basis V_{RB}. Otherwise, we compute the high-fidelity solution of the starting model (i.e. starting point of the Markov Chain). In MT, each forward solution requires the computation of two (typically orthogonal) components of the EM fields per frequency i, $i = 1, \ldots, n_{\text{freq}}$. Here, these orthogonal solutions per frequency are referred to as S^i and $S^i_\perp$.
(ii) At every proposed sample, $\mathbf{m}_t = \sigma_t(\mathbf{x}, \theta)$, we compute the fast RB solutions ($\mathbf{U_{RB}}^{S^i}$ and $\mathbf{U_{RB}}^{S^i_\perp}$) and their relative RB errors ($\mathbf{R_{RB}}^{S^i}$ and $\mathbf{R_{RB}}^{S^i_\perp}$) for all frequencies.
(iii) If all the relative RB errors are smaller than a prescribed tolerance β, we accept $\mathbf{U_{RB}}^{S^i}$ and $\mathbf{U_{RB}}^{S^i_\perp}$ as good approximations of the high-fidelity solution for all frequencies. The corresponding approximate likelihood, $\overline{\mathcal{L}}(\mathbf{m}_t)$, and acceptance probability, $\overline{\alpha}(\mathbf{m}_{t-1}, \mathbf{m}_t)$, are computed and the sample is either accepted or rejected according to the MH criterion.
(iv) If any $\mathbf{R_{RB}}$ is larger than β, the high-fidelity FE solution for that frequency and component of the EM field is computed at the proposed sample $\mathbf{m}_t = \sigma_t(\mathbf{x}, \theta)$. This solution is added as a new basis into the corresponding space $\mathcal{V}^{S^i}_{RB}$ or $\mathcal{V}^{S^i_\perp}_{RB}$, thus enriching the RB surrogate.
(v) Since the posterior probabilities of the proposed sample $\mathbf{m_t}$ and that of the current sample $\mathbf{m}_{t-1}$ are no longer comparable (i.e. they were computed with different solvers, FE and RB, respectively), we recompute $\mathbf{U_{RB}}^{i_{t-1}}$ (and the associated likelihood) using the newly enriched RB space. If $\mathbf{m}_t$ is rejected by the MH criterion, a new trial $\mathbf{m}^*_t$ is proposed in the vicinity of $\mathbf{m}_t$ and its likelihood is computed with the newly enriched RB space. This new trial $\mathbf{m}^*_t$ is accepted/rejected according to a modified Metropolis ratio to account for the delayed rejection (i.e. two proposals) step (see e.g. Mira *et al.* 2001; Haario *et al.* 2006).

Step five above is critical to keep the ergodicity of the algorithm. While other choices can be implemented for this purpose, they usually require additional high-fidelity computations at the current state of the chain $\mathbf{m}_{t-1}$ (e.g. Yan & Zhou 2019; Ortega-Gelabert *et al.* 2020). Also, although the delayed rejection step is not strictly necessary from an MCMC algorithmic point of view, it is rather important in practice. This is because, despite the fact that the posteriors of both models $\mathbf{m}_t$ and $\mathbf{m}_{t-1}$ have been computed with the same updated surrogate, the surrogate solution for $\mathbf{m}_t$ still outputs a 'more accurate' high-fidelity (FEM) solution (i.e. the surrogate is effectively an FE solver at $\mathbf{m}_t$, but still an approximate surrogate at $\mathbf{m}_{t-1}$). We have observed that this tends to decrease the acceptance probability of $\mathbf{m}_t$ and using a delayed rejection step with a small variance counteracts this undesired effect.

The success of our approach relies on a judicious combination of the RB method, the MCMC sampling that drives the generation of bases and the prescribed tolerance β. These elements allow us to adapt and refine the low-fidelity solutions to obtain a very efficient and accurate sampling of the posterior distribution (or equivalently, $\overline{\mathcal{L}}(\mathbf{m}_t)$ tends to the true likelihood $\mathcal{L}(\mathbf{m}_t)$ as the surrogate is adapted). At the same time, it is imperative that the number of elements in the basis is kept as small as possible without compromising the quality of their predictions (cf. Ortega-Gelabert *et al.* 2020). In order to achieve this, we have implemented the use of variable tolerances and singular value decomposition (SVD) of the basis as additional functionalities.

5.1.1 Variable tolerance

During the so-called *burn-in* stage, a finite number of simulations (and associated forward solutions) is required to move away from the initial low-probability region and towards the support of the posterior PDF. Since these samples are not considered for the statistical analysis of the posterior PDFs, a relatively large tolerance $\beta_1 > \beta$ can be used during this initial stage to encourage the algorithm to move quickly towards

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

Algorithm 1 Creation of the basis

procedure REFINEMENT($\mathbf{U}_{\mathbf{RB}_{ij}}^{k}(\mathbf{m}_{t-1})$, $\mathbf{U}_{\mathbf{RB}_{ij}}^{k+1}(\mathbf{m}_t)$)
- Propose model $\mathbf{m}_t$ from current position $\mathbf{m}_{t-1}$

for $i = 1, 2...N_{freq}$ **do**
- **for** $j = S, S_\perp$ **do**
  - k basis members for that frequency i and orientation j
  - Compute RB solution $\mathbf{U}_{\mathbf{RB}_{ij}}^{k}(\mathbf{m}_t)$ and error $\mathbf{R}_{\mathbf{RB}}$
  - **if** ($\mathbf{R}_{\mathbf{RB}} \leq \beta$) **then**
    - Use $\mathbf{U}_{\mathbf{RB}_{ij}}^{k}(\mathbf{m}_t)$ in computation of $\overline{\mathcal{L}}(\mathbf{m}_t)$ and $\overline{\alpha}(\mathbf{m}_{t-1}, \mathbf{m}_t)$
  - **else if** ($\mathbf{R}_{\mathbf{RB}} > \beta$) **then**
    - Compute FE solution for $\mathbf{m}_t \rightarrow \mathbf{U}_{ij}(\mathbf{m}_t)$
    - Add $\mathbf{U}_{ij}(\mathbf{m}_t)$ to $\mathcal{V}_{RB_{ij}}^{k} \rightarrow \mathcal{V}_{RB_{ij}}^{k+1}$
    - $flag^i = 1$
  - **end if**
- **end for**

end for

if (any $flag^i = 1$) **then**
- Return to position $\mathbf{m}_{t-1}$
- Re-compute the RB solution for frequency i and $\overline{\mathcal{L}}(\mathbf{m}_{t-1})$
- Re-evaluate $\overline{\alpha}(\mathbf{m}_{t-1}, \mathbf{m}_t)$
- Accept/reject $\mathbf{m}_t$ with Metropolis-Hasting criterion
- **if** ($\mathbf{m}_t$ is accepted) **then**
  - Continue with the RB+MCMC simulations
- **else if** ($\mathbf{m}_t$ is rejected) **then**
  - Propose new sample $\mathbf{m}_t^*$ in the vicinity of $\mathbf{m}_t$
  - Compute $\overline{\alpha}_2(\mathbf{m}_{t-1}, \mathbf{m}_t, \mathbf{m}_t^*)$
  - Accept/reject $\mathbf{m}_t^*$ using a delayed rejection-Metropolis ratio.
- **end if**

end if

end procedure

the high probability regions. After this stage, the tolerance can be reduced to guarantee that the sampling within the support of the posterior PDF satisfies the user's criteria for accuracy. Numerous tests (not shown here) and the results in Section 6.3 demonstrate that this simple tactic results in noticeable gains in efficiency during the inversion. While in principle, the actual reduction of the tolerance can be done in any number of steps of decreasing β, we have not attempted a systematic exploration of all possible combinations. The results presented in Section 6.3 illustrate the use of a one-step reduction (i.e. two tolerance values).

5.1.2 Decomposition of the basis

Our RB approach incorporates new solutions into the RB spaces that cannot be explained by the previous basis within a prescribed tolerance. This, however, does not guarantee orthogonality of the individual basis members. A simple way to remove redundancy of information and/or basis members that do not contribute significantly to the RB solution is to apply an SVD procedure (Brunton & Kutz 2019). In our algorithm, the SVD is applied to any RB basis space once its size reaches a pre-defined value N_{svd}. We keep the most relevant N_{rel} orthogonal elements of the basis up to a pre-defined energy threshold (usually between 0.95 per cent and 0.999 per cent). After the SVD procedure is performed, we redefine a new N_{svd}^* as $N_{\text{svd}}^* = N_{\text{svd}} + N_{\text{rel}}$. This redefinition of N_{svd} allows the size of the basis to progressively increase as needed.

Section 2 of the Supporting Information presents an application to a simple 1-D heat transfer problem. The computational time of this simple forward problem allows us to approximate the true posterior PDFs with high accuracy (using high-fidelity FEM solutions and standard MCMC algorithms) and thus evaluate the accuracy of our RB+MCMC method.

6 NUMERICAL EXAMPLES

In this section, we illustrate the performance of our 3-D MT RB+MCMC algorithm by inverting synthetic data corresponding to two complex large-scale lithospheric models with dimensions 1600 × 1600 × 460 km (Figs 1 and 10). In both cases, the computational domain is discretized with 40 × 40 × 20 FEs.

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

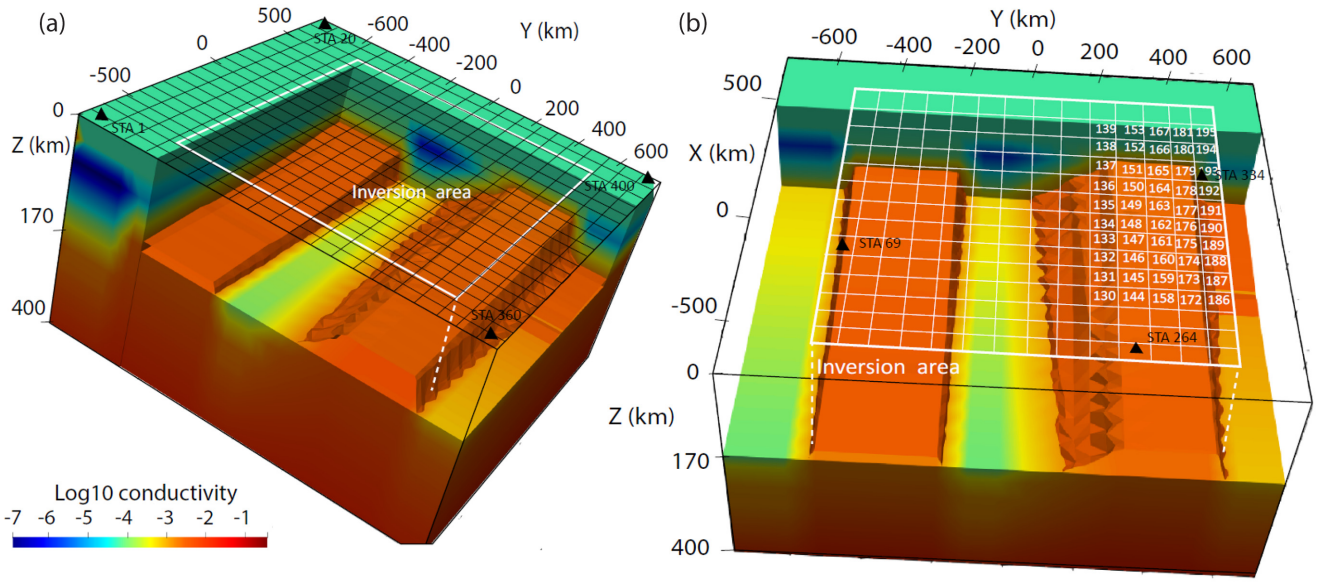

Figure 1. 3-D rendering views of the true conductivity structure where the iso-surface of -2.8 $\log_{10}$ S m^{-1} is plotted as a reference. The white rectangle indicates the region used for the inversion. Panel (a) illustrates the 20×20 station grid in black and four of the 400 stations (black triangles). The model parameters are the depths to the LAB of 196 columns. Panel (b) displays the location of these columns (white small squares) and, in particular, the 50 columns shown in Fig. 6. The reader is referred to Section 6.2.1 for details on the parametrization.

6.1 Synthetic data

The synthetic data are the off-diagonal apparent resistivities and phases computed for 12 periods between 3.2 and 10^4 s at 400 stations. The stations are located on a grid of 20 × 20 (Fig. 1a) with an interstation distance of 80 km. The data errors are assumed to be uncorrelated and normally distributed with a standard deviation of 12 per cent for the apparent resistivities and 1.5° for the phases. For the model used in Section 6.2, we avoid the so-called 'inverse crime' by generating the data with a finer FE mesh than that used in the inversion. In the second example (Section 6.3), the data are generated with the true conductivity value for each FE cell whereas the models used during the inversion are obtained via interpolation of the parameters' value. This also means that a perfect fit to data may not be achievable during the inversion.

6.2 Example 1: large-scale lithospheric structure

6.2.1 Model setup

The area selected for the inversion is subdivided into 14 × 14 columns (white squares in Fig. 1b) of size 80 × 80 × 460 km. The model parameters are the depths to the thermal lithosphere–asthenosphere boundary (LAB) of the 196 columns within the inversion area, that is, there is one model parameter per column. Here, we identify the LAB with the depth to the 1250 °C isotherm (cf. Afonso *et al.* 2016a). The goal is to retrieve the LAB depths of the model from noisy data.

The 'true' model is shown in Fig. 1. In order to obtain the conductivity structure of the true model from LAB depths, we first compute the thermal structure by solving the steady-state heat transfer problem with Dirichlet boundary conditions at the surface ($T_0 = 10$°C) and at each LAB depth ($T_{\rm LAB} = 1250$ °C). For simplicity, we assume a linear temperature gradient between the LAB and 410 km depth, where the temperature is fixed at $T_{410} = 1550$ °C. This gradient is extrapolated to the bottom of the numerical domain, located at 460 km depth. A pressure profile is also computed in each column using the following quadratic lithostatic-type approximation:

$$P(z) = 0.99 \times (4.4773 \times 10^{-3} z^2 + 3.2206 \times 10^4 z - 1.284278 \times 10^8) \quad (22)$$

where P is pressure in Pa and z is depth in m. As a further simplification, we assume a dry and homogeneous mantle composition with the following mineral modes: 0.520, 0.202 ,$9.798e^{-2}$ and 0.179 vol per cent for olivine, orthopyroxene, clinopyroxene and garnet, respectively. The electrical conductivity for each mineral phase is then obtained using eq. (A3), with parameters specified in Table A1. Finally, the bulk electrical conductivity (i.e. that of the mineral aggregate or rock) of each FE cell in the mantle is obtained using the Hashin–Shtrikman averaging scheme (Hashin & Shtrikman 1962, 1963). The conductivity in the crust (Moho at 49 km depth) is constant and equal to 20 000 Ω m.

6.2.2 Sampling strategy

The priors for the LAB depths are uniform distributions defined in a range of ±50 km, centred on the true value of each column. The proposals are Gaussian distributions centred on the current sample with a standard deviation of 10 km (see details in Appendix B). The initial

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

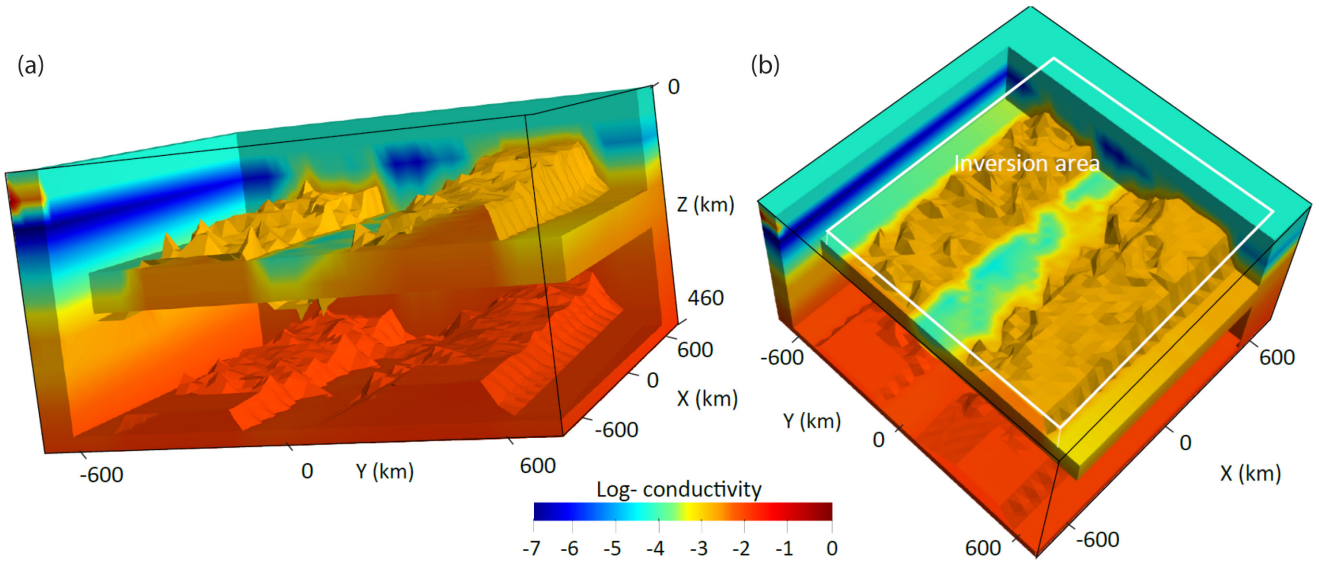

Figure 2. 3-D rendering views of the mean conductivity structure obtained after 2500 000 MCMC simulations. The iso-surfaces of −2.8 and −2 $\log_{10}$ (S m^{-1}) are plotted as a reference. The white rectangle in (b) indicates the region used for the inversion.

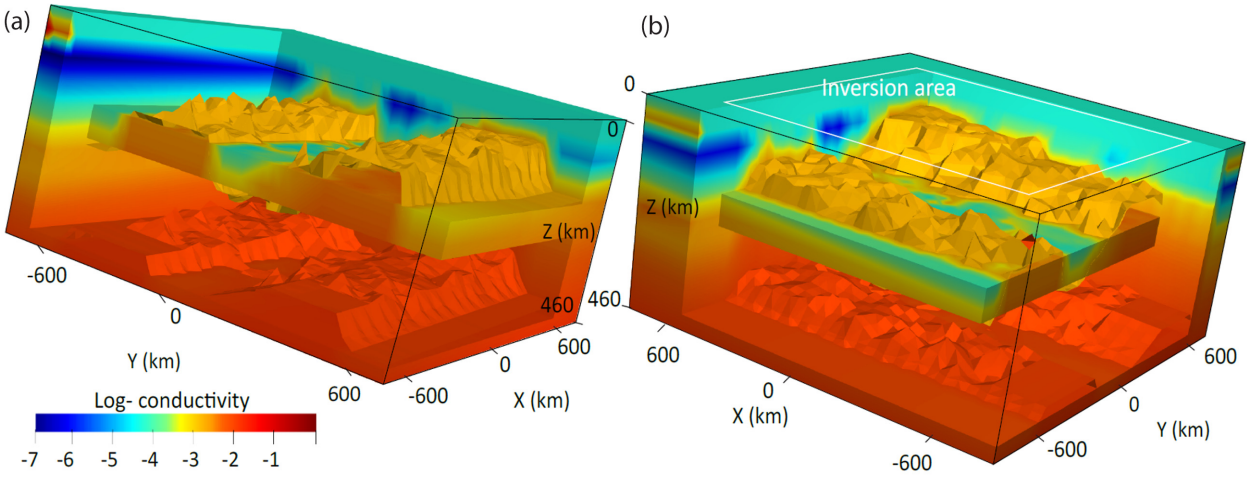

Figure 3. 3-D rendering views of the *MAP* (best-fitting) conductivity structure obtained after 2500 000 MCMC simulations. The iso-surfaces of −2.8 and −2 $\log_{10}$ S m^{-1} are plotted as a reference. The white rectangle in (b) indicates the region used for the inversion.

LAB model (i.e. starting point of the MCMC inversion) is defined by randomly choosing an LAB depth from the prior distribution of each parameter. At each step of the MCMC inversion, the algorithm randomly selects a column (metropolized independence sampler) and assigns an LAB depth from the proposal distribution. The new LAB is then used to re-compute the temperature and pressure of that column and update the conductivity model as explained in the previous section.

6.2.3 Inversion results

We ran a total of 2500 000 MCMC simulations using only two processors (Intel(R) Xeon(R) CPU E5-2680 v3 @ 2.50 GHz processors) per frequency. Even with such modest computational resources, the inversion took < 30 d (an average of 1.03 s per MCMC iteration). Since the high-fidelity solution for this model takes approximately 30 s, this represents a staggering time reduction of ∼97 per cent.

The mean conductivity model (i.e. mean values of the posterior PDFs) and the *maximum a posteriori* model (MAP) are shown in Figs 2 and 3, respectively. The electrical conductivity models corresponding to the lower and upper bound of the 95 per cent confidence interval of the posterior PDF are shown in Figs 4. No lateral smoothing has been applied in either of these plots. It can be observed that the mean conductivity model is in the vicinity of the MAP model. These results demonstrate that our algorithm succeeded in retrieving a representative solution of the inverse problem. The root-mean-square (rms) values of the mean and MAP conductivity models with respect to the true model are listed in Table 1 where we have also included the rms of the LAB depths for both models with respect to their true values.

The posterior PDFs of data for station 334 and the posterior PDFs of 50 model parameters are shown in Figs 5 and 6, respectively. Additional posterior PDFs of data for other stations and the posterior PDFs of all model parameters can be found in Section 3 of the Supporting

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

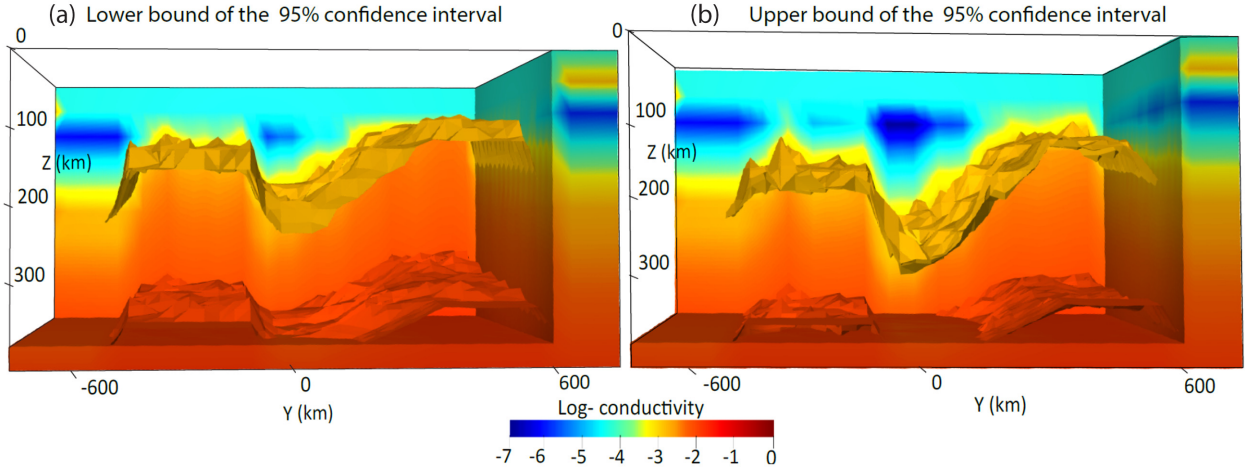

Figure 4. Electrical conductivity models corresponding to the lower (a) and upper bound (b) of the 95 per cent confidence interval of the posterior PDFs. The iso-surfaces of -2.8 and -2 $\log_{10}$ S m^{-1} are plotted as a reference.

Table 1. Rms values of the mean and MAP conductivity and LAB models with respect to the true model. The rms of the best conductivity model obtained with a ModEM deterministic inversion is also included.

	Rms conductivity ($\log_{10}$ S m^{-1})		Rms LAB depth (km)	
	Maximum a posteriori	Mean model	*Maximum a posteriori*	Mean model
RB+MCMC	0.19	0.15	21.20	17.01
ModEM	1.04			

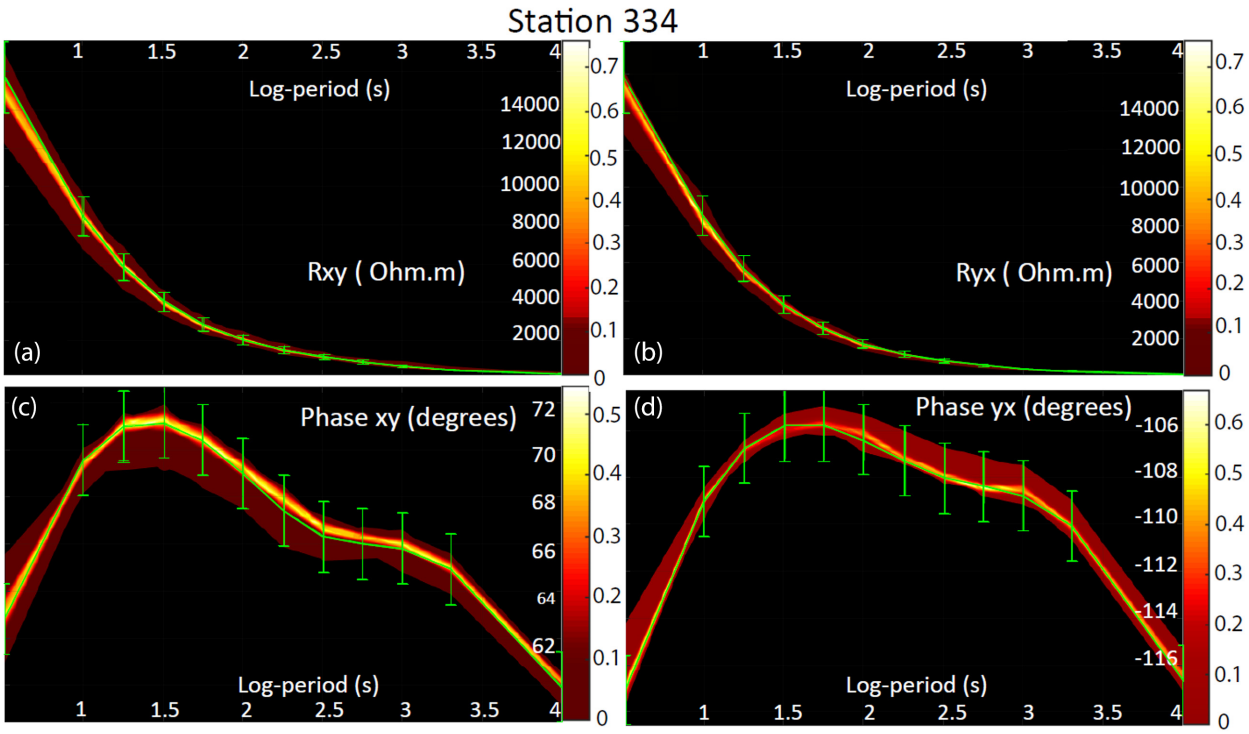

Figure 5. Posterior PDFs of data for station 334. Synthetic data and error bars are plotted in green. (a) and (b) Posterior PDFs of the off-diagonal apparent resistivity. (c) and (d) Posterior PDFs of the off-diagonal apparent phases.

Information. The results show that the great majority of observations are contained within one standard deviation of the posterior PDFs of the data. In the case of the LAB depths, most of the marginal posterior distributions have mean values that agree well with the *true* LAB depths within one standard deviation, that is, these model parameters are well resolved.

A closer examination of Fig. 6 reveals the well-known 'compensation effect' or *model equivalence* in MT and EM methods (e.g Wait 1962; Mallick & Verma 1979; Zhdanov & Keller 1994; Harinarayana 1999; Hoffmann & Dietrich 2004; Park & Ostos 2013), by which

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

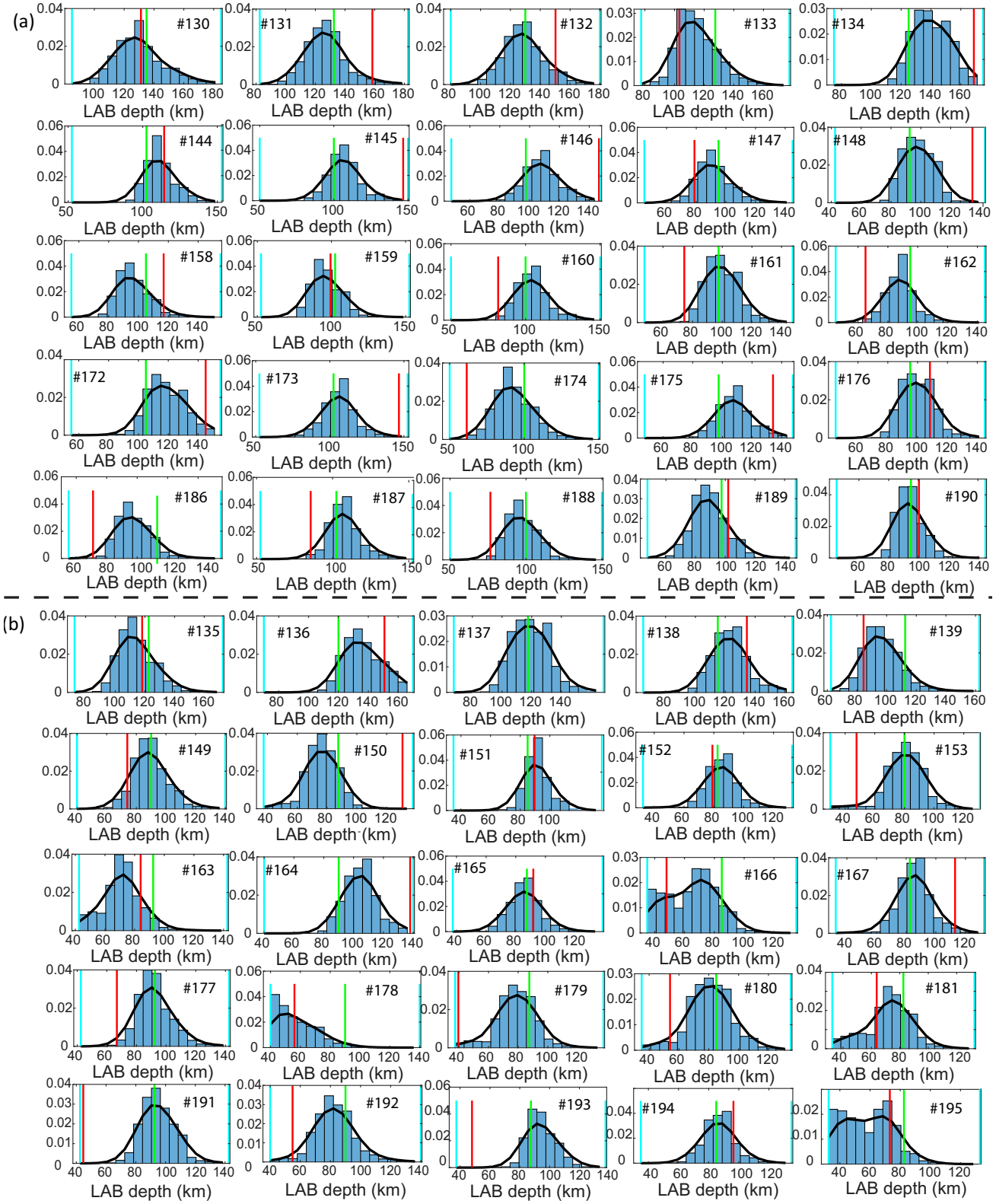

Figure 6. (a) and (b) Marginal posterior PDFs (blue bars) and best-fitting distribution (black line) of 50 model parameters obtained after 2500 000 RB+MCMC simulations. The real value, starting value and prior bounds of each parameter are shown in green, red and light blue vertical lines, respectively. The numbers within each panel correspond to the columns highlighted in Fig. 1(b).

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

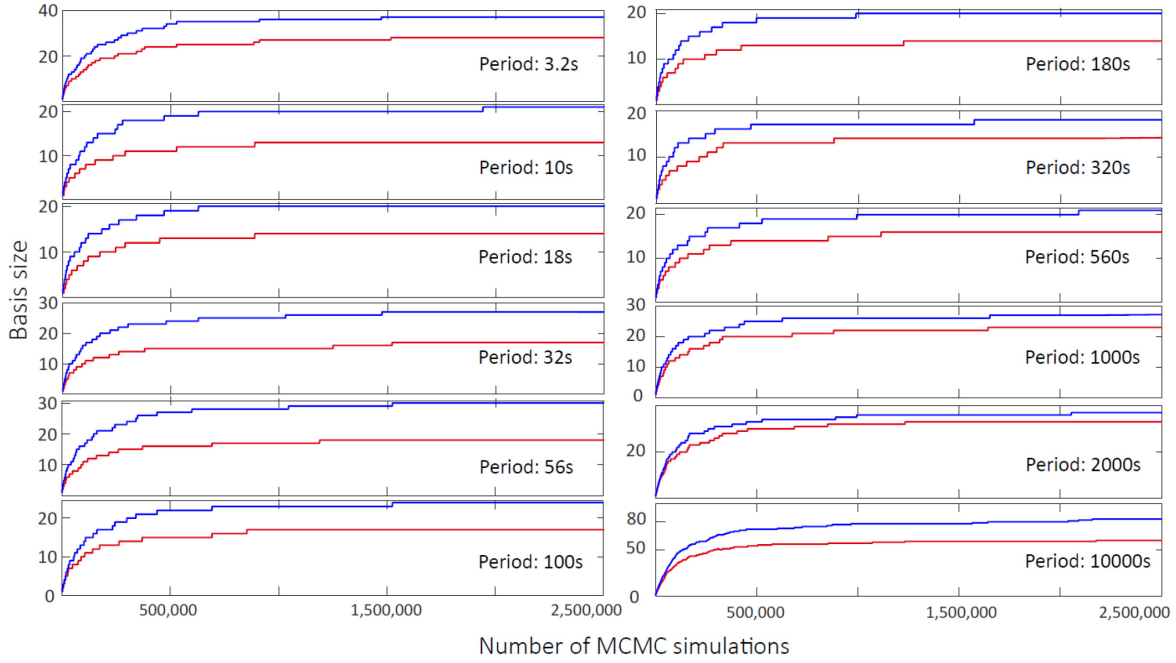

Figure 7. Basis size as a function of the MCMC simulations for different frequencies and field orientations ($S_{\perp}$ mode in blue, and S mode in red).

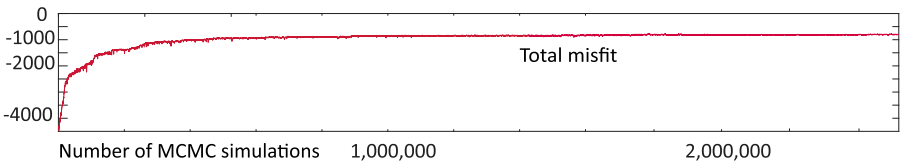

Figure 8. Data misfit for each of the 2500 000 RB+MCMC simulations.

incorrect model parameters can compensate each other and result in a combined MT signal that is close to that associated with the true solution. In the present example, the individual LAB depths within a cluster of adjacent columns can be relatively far from their true values, but their combined effect still produce a conductivity model that fits the data well. The inherent compensation effect in MT is exacerbated here by the simple parametrization and sampling strategy. For instance, we are assuming that the conductivity distribution is only controlled by the depth of the thermal LAB (Section 6.2.1), which means that decreasing the LAB depth in a column has the effect of increasing its average conductivity; the opposite is also true. Under this circumstances, a cluster of columns with LAB depths shallower and deeper than the true value for the cluster can produce a good fit to the data.

The effect of the sampling strategy can be understood as follows. Let us suppose that the current state of the MCMC chain contains a cluster of compensated columns with incorrect LAB values but with an acceptable fit to the data. Since each column is independent and we sample only one parameter (i.e. column) per MCMC iteration, if we propose a new LAB value for any of the columns in the cluster that is closer to the true value for that column, the combined MT signal will most likely produce a poorer misfit compared to the current compensated state. This is because although we are proposing a value closer to the true value for one column, it is the combined effect of the cluster that drives the misfit. In deterministic inversions, this problem is minimized by applying regularization terms. More appropriate parametrizations and sampling strategies in the context of MCMC inversions are further discussed in Section 7.

The aforementioned compensation effect can be observed, for example, in adjacent columns 172 and 186 in Fig. 6, where the mean LAB depth is ∼15 km deeper than the true value for column 172 and ∼15 km shallower than the true value for column 186. The same effect is also observed in columns 144 and 145 (with mean LAB depths deeper than their true value) and columns 158 and 159 (shallower mean LAB depths). For some parameters (178 for example), the compensation of neighbouring parameters (columns 164, 165, 177, 179, 191, 192 and 193) is such that they cannot be recovered well by the inversion given the range of data errors and the prescribed tolerance β.

The size of the basis computed per frequency and field orientation is shown in Fig. 7. A rapid increment in the basis size is observed during the *burn-in* stage (first 200 000 simulations) which also correlates with a rapid decrease in the overall misfit (Fig. 8). This increment is required to obtain representative solutions for the parameter space and, in particular, for the high-probability region. After this stage, the basis size reaches a *saturation level* or *plateau* for all frequencies, that is, the *RB* space is able to reproduce any solution within the regions of high probability. At this point, the adaptation of the RB space could be stopped without compromising subsequent solutions provided that the chain has indeed reached stationarity.

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

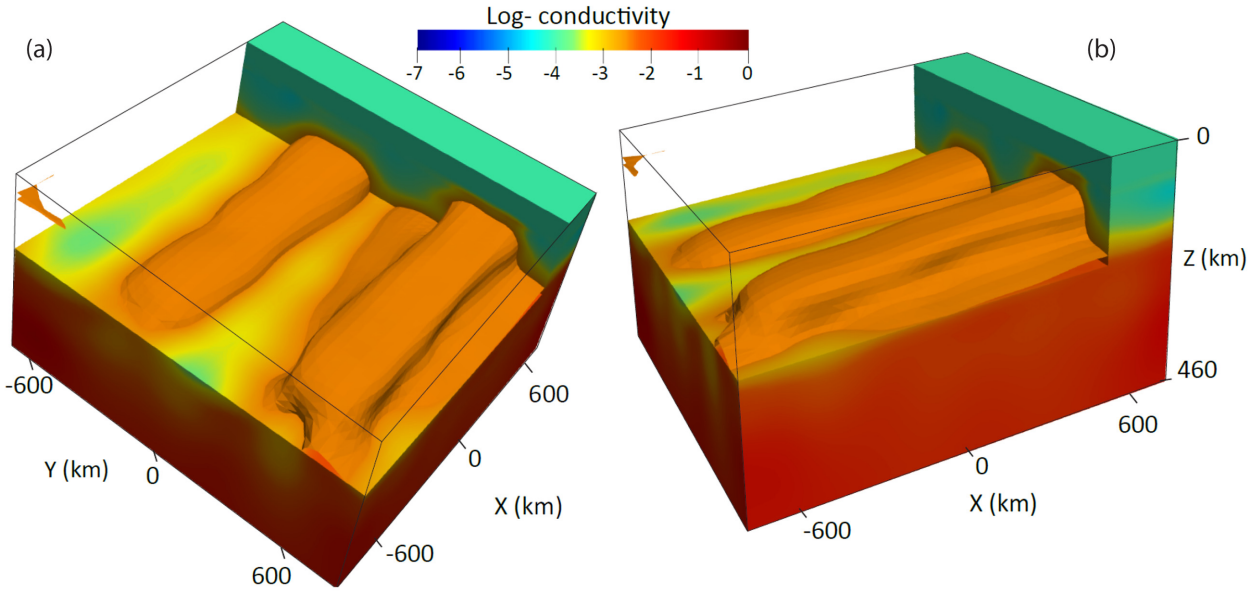

Figure 9. (a) and (b) 3-D rendering views of the best conductivity model obtained with an ModEM deterministic inversion using a horizontal and vertical smoothing factor of 0.5 and 0.2, respectively. The iso-surface of -2.8 $\log_{10}$ (S m^{-1}) is plotted as a reference.

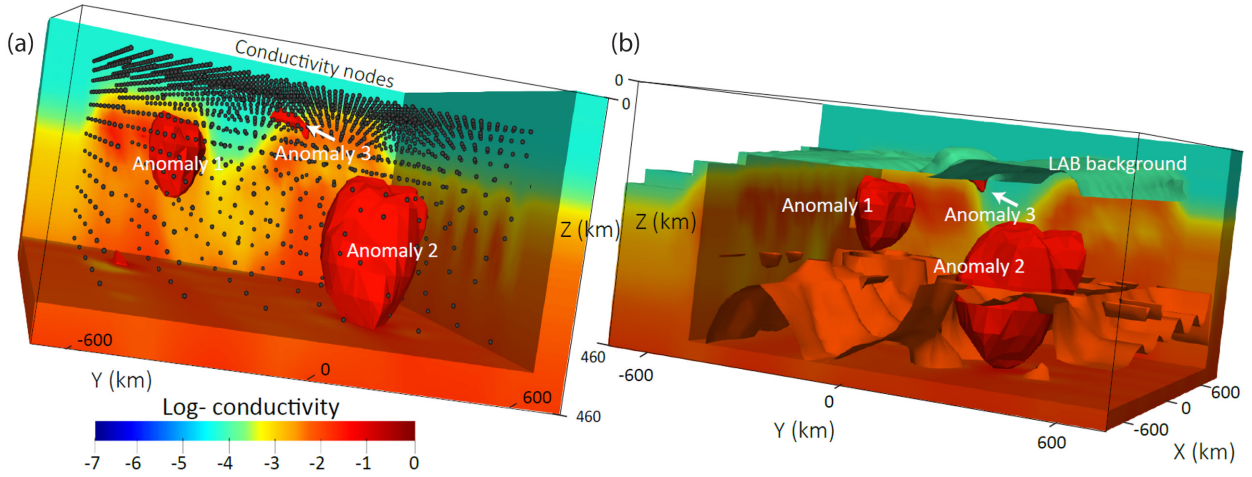

Figure 10. 3-D rendering views of the true conductivity structure after interpolation of the 1782 node parameters. Conductive anomalies are highlighted in both (a) and (b) panels. Black dots in (a) indicate the position of the node parameters within the inversion volume. Panel (a) shows the iso-surface corresponding to -1.5 $\log_{10}$ (S m^{-1}), whereas iso-surfaces of -2.15, -1.5 and -4 $\log_{10}$ (S m^{-1}) are shown in (b).

For comparison, we include the results of a ModEM (Kelbert *et al.* 2014) deterministic inversion using the same synthetic data. Fig. 9 shows the best model obtained after 42 iterations using a vertical and horizontal smoothing factor of 0.2 and 0.5, respectively. The inversion took 6 hr using 30 processors and the final data rms was 1.04. The rms value of the ModEM model with respect to the true conductivity distribution is also included in Table 1. One of the main reasons for the larger rms in the ModEM inversion is that lateral smoothing had to be imposed to stabilize the inversion (a typical requirement in determinisitc inversions). This precluded the algorithm from considering sharp discontinuities in the solution. Another contribution to the rms is the high-resistivity regions in the mantle, which could not be recovered in this example.

6.3 Example 2: large-scale lithospheric structure with conductivity anomalies

6.3.1 Model setup

The true conductivity model (Fig. 10) includes the large-scale lithospheric model of Section 6.2 as a background plus three additional and localized conductive anomalies. The conductivity structure is now parametrized by 1782 nodes (black dots in Fig. 10a) sparsely located within the inversion volume. The conductivity value of each FE cell is obtained by interpolation of the nodes' values using a quadratic Shepard

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

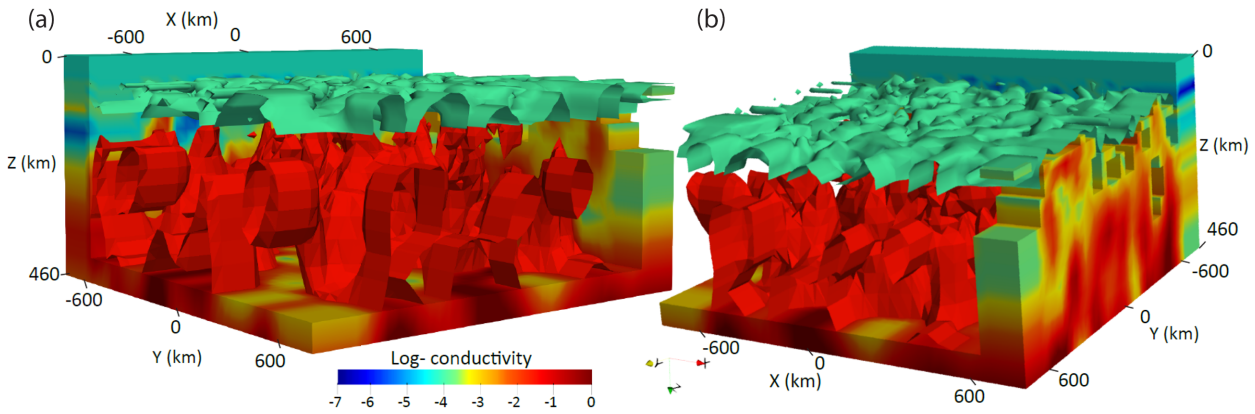

Figure 11. 3-D rendering views of the initial conductivity structure. The iso-surfaces of -4 and -1.5 $\log_{10}(\mathrm{S\,m^{-1}})$ are plotted as a reference in both (a) and (b) panels.

method for trivariate interpolation of scattered 3-D data (Renka 1988a,b). The goal here is to recover the conductivity structure from noisy MT measurements.

In order to define the location of the conductivity nodes, we first subdivide the domain into 14 horizontal layers; the vertical location of the nodes correspond to the mid-points of each layer (horizontal node layers in Fig. 10a). In this example, these layers correlate with the horizontal layers given by the FE mesh. Since bodies with dimensions smaller than the EM skin depth cannot be resolved by the data, the horizontal distance between nodes within each layer is chosen relative to the skin depth for the range of periods and apparent resistivities shown in the observed data. We discuss alternative options for dynamic parametrizations in Section 7.

6.3.2 Sampling strategy

The starting model (Fig. 11) is defined by randomly selecting a conductivity value for all nodes from their prior distributions. Since the model parameters' values can span several orders of magnitude, we define Jeffreys' prior distributions (see Appendix B) in a range of ± 2 $\log_{10}(\mathrm{S\,m^{-1}})$ centred on the true value of each node. During the *burn-in* stage, the algorithm randomly chooses one node i at a time (with all nodes having the same probability of being chosen) and assigns a conductivity value from its proposal distribution. The proposals are defined as log-normal distributions (eq. B4) centred on the node m^i_{t-1} with standard deviations of 0.5 $\log_{10}(\mathrm{S\,m^{-1}})$. For every new trial m^i_t, the full 3-D conductivity model is updated via interpolation (using a quadratic Shepard method, Section 6.3.1).

After the *burn-in* stage, we use an AM algorithm (Haario *et al.* 2001) to compute and update layerwise proposal distributions using the history and correlations of the MCMC chain (see details in Appendix B1). At each MCMC step, a *metropolised independence sampler* randomly selects a layer l and the adapted proposal (eq. B5) is used to propose a new sample $\mathbf{m}^l$ (i.e. conductivity values for all the nodes of layer l). This sampling approach results in significant gains in efficiency and makes better use of the natural correlations inferred from the data.

6.3.3 Inversion results

We ran a total of 2500 000 MCMC simulations for 11 frequencies using only two processors (Intel(R) Xeon(R) CPU E5-2680 v3 @ 2.50 GHz processors) per frequency. The tolerances used were $\beta = 0.07$ for the first 150 000 steps and $\beta = 0.05$ for the remaining of the simulation. The SVD factorization was performed every 50 new bases with an energy *cut-off* of 0.997. The first adaptation of the proposals for all node layers was done at step 200 000 and we continue to adapt them every 100 000 simulations. Again, even with modest computational resources, the inversion took $<$ 40 d with an average of 2.88 s for each of the first 300 000 simulations (burn-in and enrichment process) and 1.58 s for the rest of the simulations. This corresponds to a time reduction of $\sim$95 per cent for each forward computation after the enrichment process. It is important to note here that most of the computational time (besides the high-fidelity solutions) is now taken by the interpolation steps rather than by the surrogate itself. As discussed in Section 7, the inversion can be made significantly more efficient by changing the interpolation algorithm.

The mean conductivity model (i.e. mean values of the posterior PDFs) and those representing the lower and upper bounds of the 95 per cent confidence interval of the posterior PDFs are shown in Figs 12 and 13, respectively. It is clear that our RB+MCMC algorithm and the present parametrization result in a good approximation of the reference model, including the location and volumes of the conductive anomalies. Examples of the posterior PDFs of the data for two stations are shown in Figs 14 and 15; the marginal posterior PDFs of 50 of the 1782 model parameters (nodes) are shown in Fig. 16. Additional posterior PDFs can be found in Section 4 of the Supporting Information.

The results show that most observed data are contained within one standard deviation of the posterior PDFs. For some stations, however, a poor data fit is seen at particular periods (e.g. at shorter periods of Fig. 14). This poor fitting or bias is related to the so-called model

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

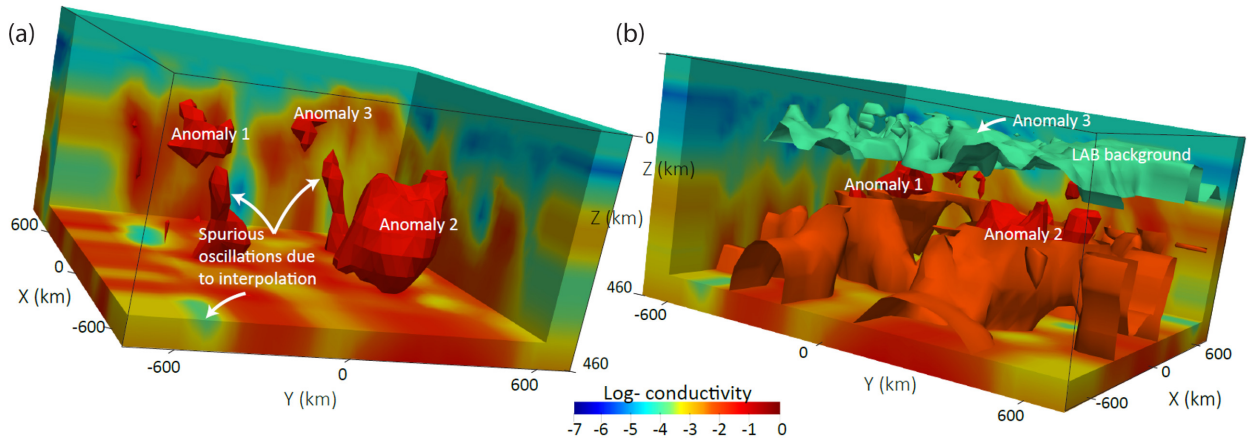

Figure 12. 3-D rendering views of the conductivity structure obtained with the mean values of the posterior PDFs. Conductive anomalies are indicated in both (a) and (b) panels. Panel (a) shows the iso-surface of -1.5 $\log_{10}(\mathrm{S\,m^{-1}})$, whereas the iso-surfaces corresponding to -2.15, -1.5 and -4 $\log_{10}(\mathrm{S\,m^{-1}})$ are shown in (b).

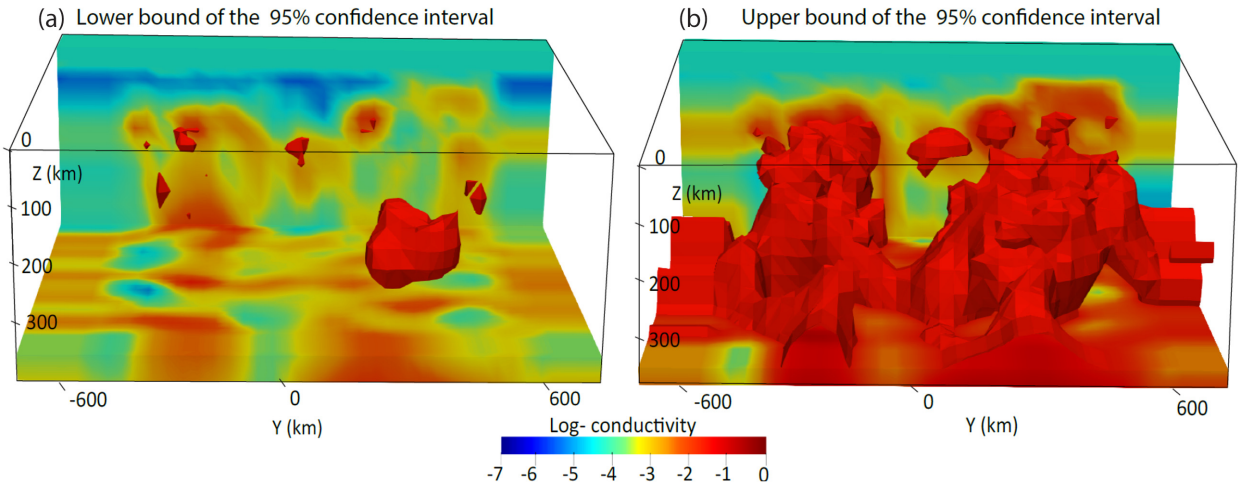

Figure 13. Electrical conductivity models corresponding to the lower (a) and upper bound (b) of the 95 per cent confidence interval of the posterior PDFs. The iso-surface of -1.5 $\log_{10}(\mathrm{S\,m^{-1}})$ is plotted as a reference.

inadequacy (e.g. Smith 2013), where the discrepancy between the conductivity models used during the inversion (derived from interpolation of nodal values) and the true model used to compute the synthetic data preclude a perfect data fit.

As in the previous example, we also observe a bias of the models with respect to the true model. This is primarily due to the compensation effect described in Section 6.2.3, which is aggravated here by the introduction of spurious oscillations produced by the Shepard interpolation (shown in Fig. 12a). In other words, since the Shepard interpolation introduces anomalous features in regions without nodes, the conductivity nodes that are located close to these spurious anomalies will most likely take incorrect (biased) conductivity values in order to produce models with better fits to the data. This bias can be further explained with a possible overparametrization of the computational domain. In overparametrized and under-regularized regions, models are expected to exhibit local variability around the true structure and still offer equally good fits to the data. We also note that the mean values of the PDFs are more likely to be biased towards resistive values. This is due to the low relative sensitivity of MT data to resistive structures (e.g. Berdichevsky & Dmitriev 2010; Park & Ostos 2013; Varentsov *et al.* 2013). This effect is well known and it has been observed in other works (e.g. Rosas-Carbajal *et al.* 2013; Brodie & Jiang 2018).

The number of basis vectors computed per frequency and field orientation is shown in Fig. 17. It can be seen that the number of bases rapidly increases during the first 120 000 steps while the inversion tries to find the regions of high probability. A second rapid increase is seen after step 150 000, which corresponds with a reduction in the tolerance of the surrogate (vertical black line in Fig. 17). The basis size reaches a saturation level for all frequencies after ∼200 000 steps. For the long periods, we see that the SVD factorization at around step 300 000 significantly decreased the size of the basis. Subsequent basis vector computation and/or SVD factorizations do not change the general pattern and therefore we only show the bases up to the simulation 500 000.

The data misfit for the 2500 000 simulations is shown in Fig. 18(a). As shown in Figs 18(b) and (c), the enrichment of the basis space correlates with jumps in the total misfit. This is related to the fact that when a high-fidelity solution is computed, the misfit of the previous sample is also re-computed with the enriched space (see Section 5.1). As this new misfit value is ‘more accurate’ and usually lower (worse)

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

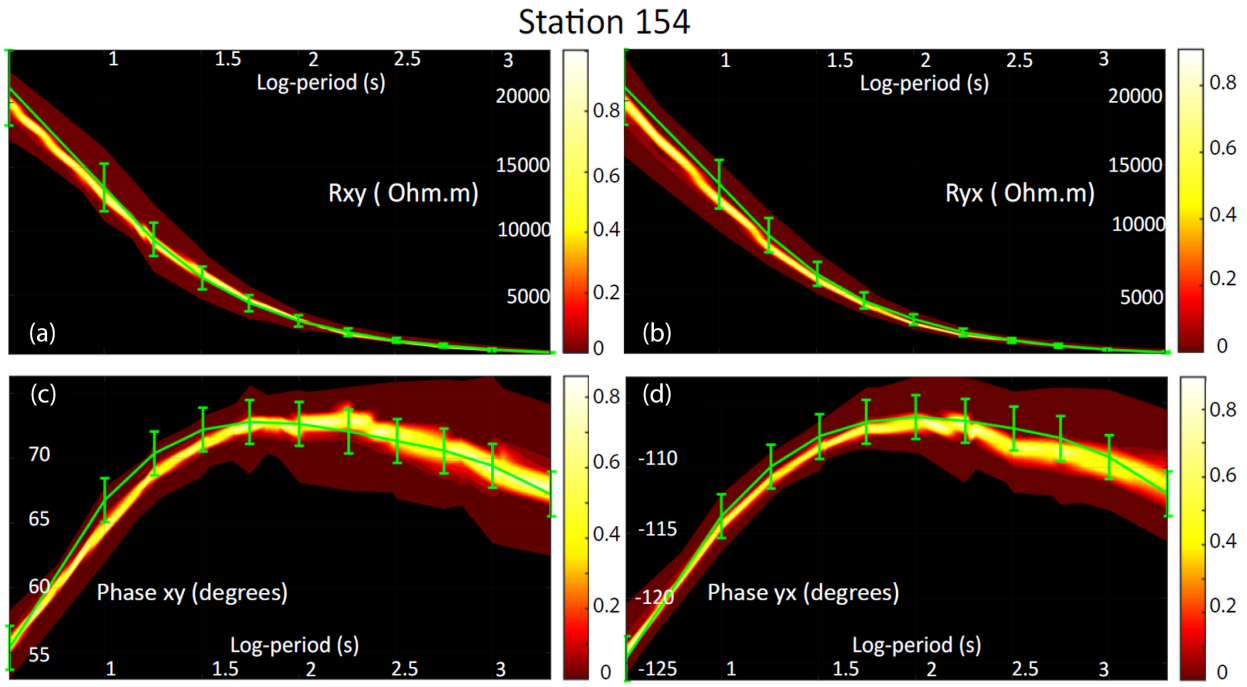

Figure 14. Posterior PDFs of data for station 154. Synthetic data and error bars are plotted in green. (a) and (b) Posterior PDFs of the off-diagonal apparent resistivity. (c) and (d) Posterior PDFs of the off-diagonal apparent phases.

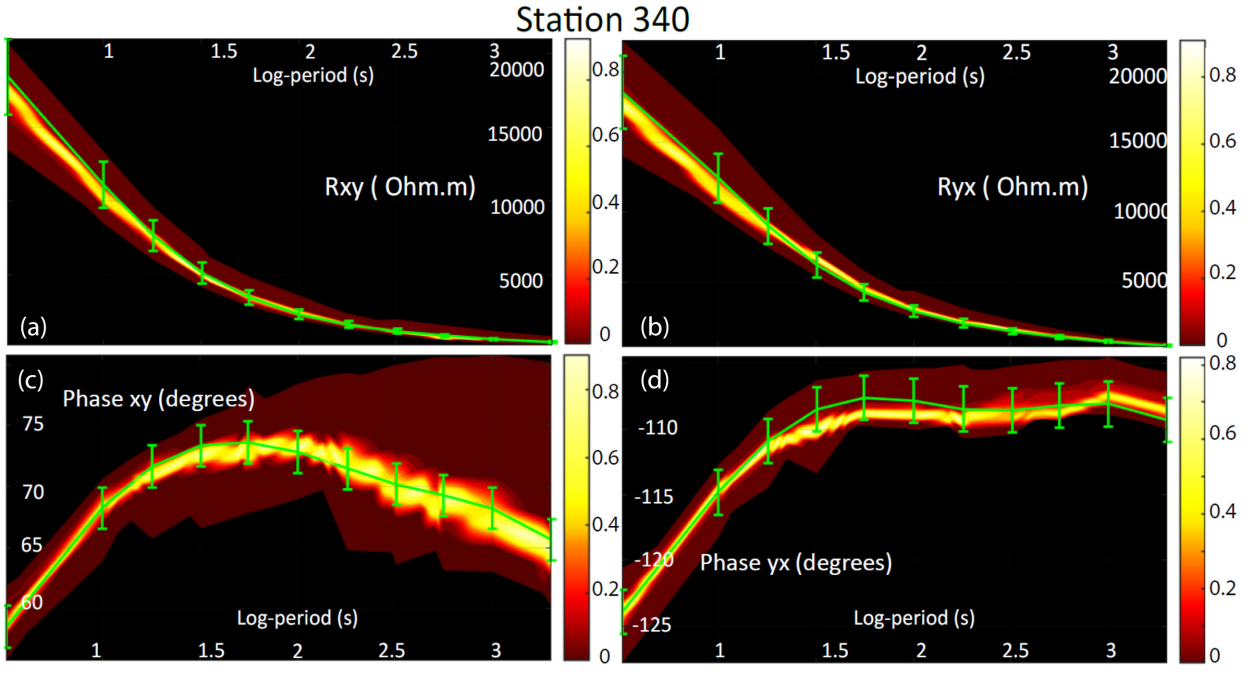

Figure 15. Posterior PDFs of data for station 340. Synthetic data and error bars are plotted in green. (a) and (b) Posterior PDFs of the off-diagonal apparent resistivity. (c) and (d) Posterior PDFs of the off-diagonal apparent phases.

than the one computed before the enrichment, jumps occur in the general decreasing tendency of the misfit (Figs 18b and c). Note that these jumps are also observed in Fig. 8 albeit smaller in amplitude compared to those in Fig. 18. These jumps are less frequent as the basis reaches a saturation level (after 200 000 simulations), which corresponds to a quasi-plateau in the misfit.

7 DISCUSSION

In this work, we have deliberately chosen to work with the simplest parametrizations and sampling strategies in order to demonstrate the viability and performance of our approach under adverse circumstances. There are, however, a large number of possible improvements that

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

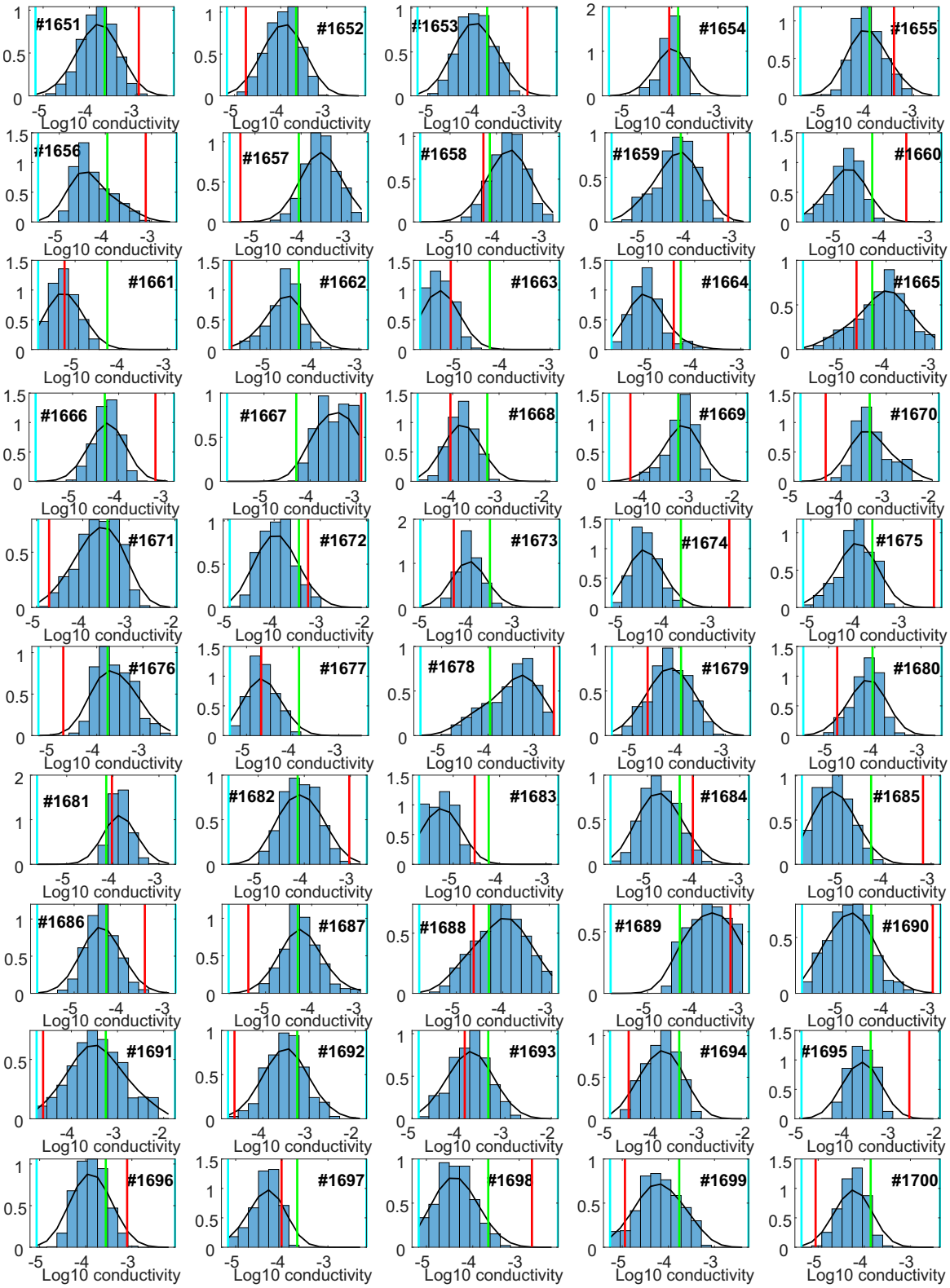

Figure 16. Marginal posterior PDFs (blue bars) and best fitting distributions (black lines) for 50 of the 1782 node parameters (corresponding to the same node layer) obtained after 2500 000 RB+MCMC simulations. The real value, starting value and bounds of each parameter are shown in green, red and light blue vertical lines, respectively. The reader is referred to the main text for details on the node-layer parametrization.

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

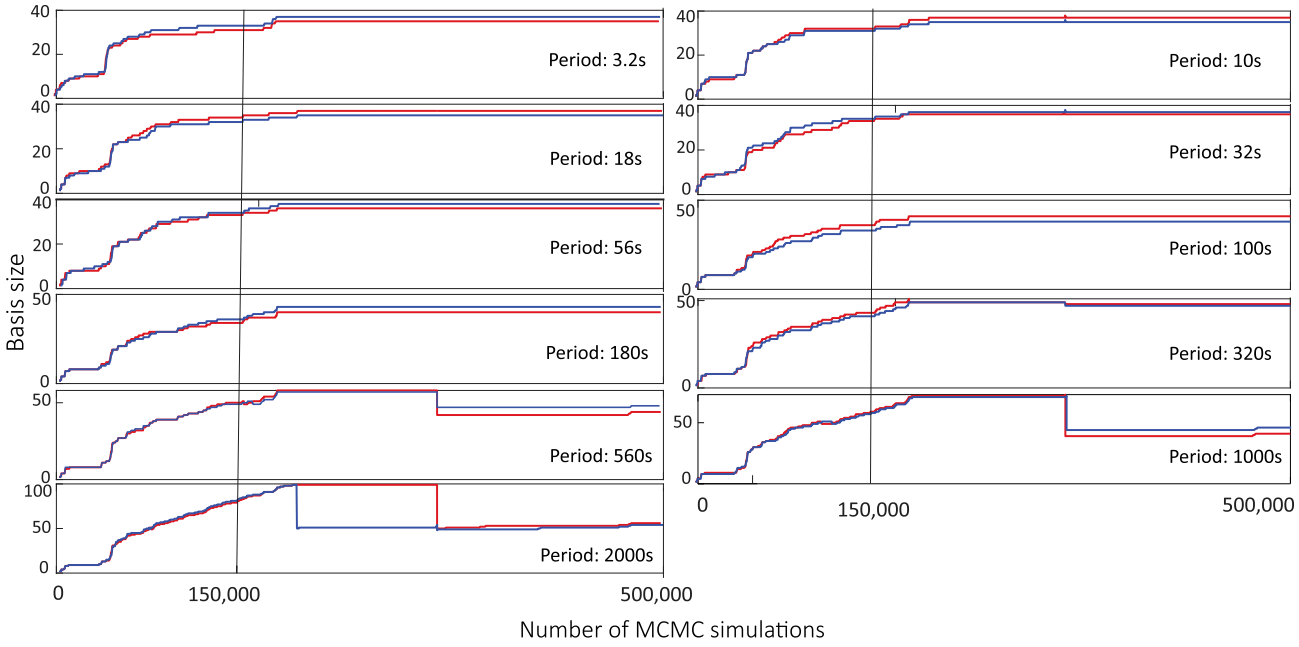

Figure 17. Basis size as a function of the MCMC simulations for different frequencies and field orientations ($S_\perp$ mode in blue, and S mode in red).

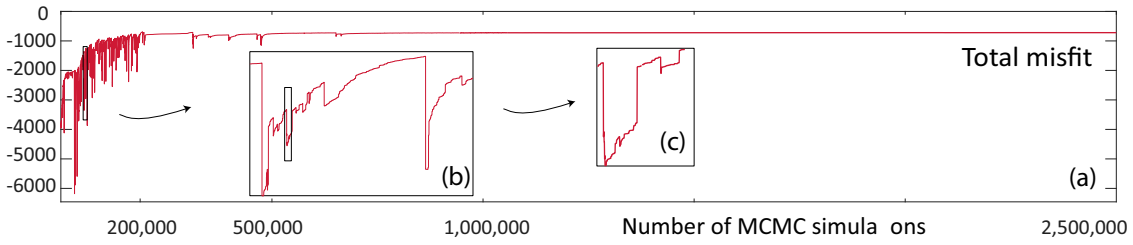

Figure 18. (a) Data misfit for each of the 2500 000 simulations. Panels (b) and (c) show the jumps in the misfit when a new basis vector is computed. The reader is referred to the main text for an explanation of this behaviour in the misfit.

can enhance the efficiency of our algorithm (e.g. the efficiency of the MCMC algorithms, the availability of computational resources, the parametrization of the model), especially when applied to real-world problems. In the following sections, we briefly discuss some of these potential improvements as well as some remaining challenges.

7.1 Parametrization of the conductivity structure

An important concept that we emphasize here is that the conductivity distribution in the lithosphere can, in principle, be treated as the superposition of two contributions: a *background* conductivity related to the long-wavelength physical state, and an *anomalous* distribution associated with the presence of smaller features (fluid pathways, melt-rich regions, hydrogen content, etc.). Assuming that the conductivity of rocks were primarily controlled by the thermal structure of the lithosphere, the inversion performed in Section 6.2 demonstrated that the long-wavelength (i.e. background) conductivity structure can be recovered satisfactorily using an LAB-based parametrization. Subsequently, we introduced a more general, node-based, parametrization in order to include smaller scale conductivity anomalies within the background (Section 6.3). While in principle this parametrization is capable of approximating any conductivity structure and allows for considerable model variance, a Shepard-type interpolation is slow (i.e. computationally inefficient) and typically introduces spurious oscillations that distort or mask the true conductivity structure (see Fig. 12a). The latter effect is exacerbated when the number and location of the nodes used in the discretization are not optimal to allow an accurate representation of the true anomalies. This is likely the case in most real-world scenarios, as it is generally impossible to know *a priori* the true variability of the conductivity in the subsurface (in practice, preliminary or nested inversions can be used for this purpose). Although it may seem reasonable at first glance to simply increase the number of nodes (and thus allow for more model variance), this can easily result in an overparametrization of the model, which can seriously compromise the convergence of the MCMC algorithm (and any inversion scheme) and increase the computational time of the interpolations.

A number of strategies can be pursued to ameliorate this problem. One possibility is to make use of well-known parameter model reduction techniques (e.g. Marzouk & Najm 2009; Lieberman *et al.* 2010). While the application of such techniques to our problem is

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

possible and likely beneficial, most of them require an exploratory offline stage to appraise the true dimensionality of the parameter space. Perhaps a more promising approach is the implementation of transdimensional algorithms (e.g. Bodin & Sambridge 2009; Brodie & Jiang 2018; Ray & Myer 2019), which have been shown to be particularly effective in identifying the minimum dimensionality of the model as required by the data itself without the need for an offline appraisal. In particular, the use of Gaussian processes, within a transdimensional context, to generate conductivity models (e.g. Ray & Myer 2019) represents a promising approach to also tackle the interpolation cost and spurious oscillations of the Shepard Method. We are currently exploring the performance of these methods with real-world data sets.

We would like to point out that, in the case of stand-alone 3-D MT inversions, one could combine both the LAB- and the node-based parametrizations to represent the long-wavelength background and the smaller scale conductivity anomalies, respectively. In this scenario, the total conductivity distribution will be the addition of the background conductivity obtained with the LAB-based parametrization and the anomalous conductivity distribution represented by nodal values. If the ultimate goal is the total conductivity field, the LAB depths and/or the conductivity nodes can simply be considered nuisance parameters, that is, they are not of immediate interest but they form part of the problem formulation. The advantage of using the combined parametrization is that rapid convergence is achieved by using LAB depths to constrain the first-order conductivity background, whereas the nodes are used to locally modify this background to fit the smaller scale features of the data.

For the case of joint inversions of MT with other data sets (e.g. Moorkamp *et al.* 2007; Khan *et al.* 2008; Jegen *et al.* 2009; Moorkamp *et al.* 2010; Afonso *et al.* 2013a,b; Vozar *et al.* 2014; Bennington *et al.* 2015; Afonso *et al.* 2016a,b; Jones *et al.* 2017) one would take advantage of the different sensitivities of the inverted observables to the conductivity structure and use complementary paramatrizations. For instance, the background conductivity distribution that results from the long-wavelength variation of temperature and bulk composition can be informed by seismic and/or potential field data using their own background parametrization, such as LAB depth, composition and pressure of individual columns of the whole domain (see e.g. Khan *et al.* 2008; Afonso *et al.* 2013a,b, 2016b; Jones *et al.* 2017). The node-based parametrization could then be used only within the MT forward problem to account for conductivity anomalies superimposed to the background and associated with factors such as fluid content, presence of melt, hydrogen content and anomalous mineral assemblages. A detailed presentation and assessment of this concept using both synthetic and real-world data sets are left for a forthcoming publication (Part II).

7.2 Efficiency of the RB+MCMC method

As mentioned above, a key factor to achieve computational efficiency is to maintain a small number of bases during the MCMC simulation without compromising the quality of their predictions. An alternative to the SVD approach described above is to identify, *on-the-fly*, the basis vectors that stop contributing to the RB solution (within a pre-defined threshold) once the MCMC begins to converge to the high probability regions of the parameter space. This information can be obtained by analysing the vector of coefficients $\mathbf{a}(\theta)$ accompanying the bases (see eq. 16). Small values for these coefficients indicate minor contributions to the RB solution of the associated basis. Once the uninformative bases (e.g. those created during the burn-in period of the inversion) are identified, they can be eliminated from the RB space. While this would be relatively easy to implement, it requires defining two additional parameters in the RB+MCMC algorithm: a minimum threshold value for the coefficients and an interval between successive assessments of the bases (doing it at every MCMC step is extremely inefficient).

Another way of further achieving a small basis size is to seek for accurate solutions only within specific regions of the numerical domain (e.g. Alvarez Aramberri 2015; Ortega-Gelabert *et al.* 2020). This stems from the fact that in many practical applications we are not interested in high accuracy at every point inside the numerical box, but rather within a restricted region. A common approach to assess the accuracy of low-fidelity solutions in particular regions of interest is by applying a standard (multi)goal-oriented criterion (Hartmann 2008; Pardo 2010; Florentin & Díez 2012; Alvarez Aramberri 2015; Alvarez-Aramberri & Pardo 2017) guided by the error E_Q in a diagnostic parameter known as the Quantity of Interest (QoI). The QoI is typically a scalar or vectorial functional of the forward solution that summarizes relevant/important information about the problem at hand. When applied to a restricted region of interest, this goal-oriented criterion results in basis of smaller size relative to those obtained by considering the entire numerical domain (cf. Ortega-Gelabert *et al.* 2020).

Before ending this section we would like to note that the assembly of the stiffness matrix K is expensive and including this step at every MCMC iteration is inefficient. When using the LAB-based parametrization, each trial of the MCMC algorithm perturbs the conductivity structure of the model only locally (i.e. within a column). In this case, only a small number of elements of the matrix K varies between proposed samples, which in turn contribute to a small and local change of the matrix K_{RB}. We have taken advantage of this and implemented an algorithm that directly updates the corresponding elements in K_{RB} (for every frequency) instead of re-computing the full K for every trial. This strategy is also possible within the node-based parametrization. However, in the present implementation, the Shepard interpolation does not provide a mapping between a node that changes value and the position of the perturbed cells in the conductivity model. This limitation can be overcome using other interpolation strategies where this mapping is available, as in the case of Gaussian processes (Ray & Myer 2019).

8 CONCLUSIONS

We have developed a novel approach for MCMC-driven probabilistic inversions of 3-D MT data. The success of the method relies on the combination of a reduced order technique (RB Method) to create fast and accurate surrogates, an efficient parallelization of the

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

forward problem and adaptive strategies for both the surrogate and the MCMC algorithm. Unlike traditional implementations, the surrogate construction and adaptation (refinement) process is integrated into the MCMC inversion, thus removing the need for costly offline stages.

The feasibility and performance of the proposed approach to invert for large-scale conductivity structures are demonstrated with two inversion examples with different parametrizations and dimensionalities. In both cases, the RB+MCMC approach successfully solves the probabilistic inverse problem using less than ∼5 per cent of the computational cost needed in a conventional MCMC approach. Furthermore, the efficiency of the method increases with the length of the MCMC simulation, as fewer basis enrichments are necessary once the chain begins to converge.

We note that the type of model parametrization and chosen MCMC scheme can significantly impact both the posterior PDFs and the convergence of the MCMC chain. In this work, we purposely adopted the simplest parametrizations and MCMC schemes. The efficiency of our strategy can therefore be further improved by making use of more advanced parametrization techniques and MCMC algorithms (e.g. transdimensional schemes, Gaussian processes, delayed rejection, parallel tempering, differential evolution, among others).

As a by-product, we have obtained an optimized version (*Max3D-G*) of the original MT FE forward code developed by Zyserman & Santos (2000). The new implementation solves the FE linear system of equations using the parallel solver MUMPS which reduces the computational time by ∼80 per cent compared to the original code.

Our new RB+MCMC approach makes 3-D MT probabilistic inversions a practical option and opens up new exciting opportunities for both stand-alone MT studies and integrated geophysical imaging of the Earth's interior via joint inversions (e.g Khan *et al.* 2008; Afonso *et al.* 2013a,b, 2016a). We explore the latter possibility in a forthcoming publication (Part II).

9 SUPPLEMENTARY MATERIAL

The online Supporting Information contains benchmarks of the optimized version (*Max3D-G*) of the original MT FE forward code developed by Zyserman & Santos (2000) and an application of our RB+MCMC approach to a heat transfer problem. We also present additional results of the numerical examples showed in Sections 6.2 and 6.3.

ACKNOWLEDGEMENTS

We thank Rob Evans, Davood Moghadas and the editor Ute Weckmann for their helpful comments on an earlier version of this manuscript. We thank Farshad Salajegheh for providing part of his Matlab codes for plotting results and Sinan Özaydin for running the ModEM inversions. Special thanks to Olga Ortega-Gelabert, Kate Selway, Stephan Thiel, Alan Jones, Anandaroop Ray, Marina Rosas-Carbajal, Juanjo Ledo, Javier Fullea and Beñat Oliveira for their comments and contributions at different stages of this work. We thank Naser Meqbel for providing the 3-D Grid tool used to create the models shown in Supporting Information. The 3-D rendering views were created using ParaView (Ahrens *et al.* 2005).

MCM thanks support from an International Macquarie Research Excellence Scholarship (iMQRES). MCM and JCA acknowledge support from ARC grant DP160103502, ARC Linkage grant LP170100233, the ARC Centre of Excellence Core to Crust Fluids Systems (http://www.ccfs.mq.edu.au) and the Centre for Earth Evolution and Dynamics, Geoscience Australia and the European Space Agency via the '3D Earth—A Dynamic Living Planet'. FZ acknowledges support from CONICET through grant PIP 112-201501-00192. SZ has been funded by the Spanish Ministry through grant DPI2017-85139-C2-2-R, by Catalan government through grant 2017-SGR-1278 and by the EU's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 777778.

REFERENCES

Afonso, J., Fullea, J., Griffin, W., Yang, Y., Jones, A., Connolly, J. & O'Reilly, S., 2013a. 3-D multiobservable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle. I: a priori petrological information and geophysical observables, *J. geophys. Res.: Solid Earth,* **118**(5), 2586–2617.

Afonso, J.C., Fullea, J., Yang, Y., Connolly, J. & Jones, A., 2013b. 3-D multi-observable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle. II: general methodology and resolution analysis, *J. geophys. Res.: Solid Earth,* **118**(4), 1650–1676.

Afonso, J.C., Moorkamp, M. & Fullea, J., 2016. Imaging the lithosphere and upper mantle: where we are at and where we are going. (Chapter) In: *Integrated imaging of the Earth,* **218,** pp 191–218. M. Moorkamp M., Lelievre P., Linde N., and Khan A. (Editors), AGU Geophysical Monograph 218, Wiley.

Afonso, J.C., Rawlinson, N., Yang, Y., Schutt, D.L., Jones, A.G., Fullea, J. & Griffin, W.L., 2016b. 3-D multiobservable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle: III. Thermochemical tomography in the Western-Central US, *J. geophys. Res.: Solid Earth,* **121**(10), 7337–7370.

Ahrens, J., Geveci, B. & Law, C., 2005. Paraview: an end-user tool for large data visualization, *The Visualization Handbook,* Vol. **2005,** 717–731, Elsevier.

Alvarez Aramberri, J., 2015. *hp-Adaptive simulation and inversion of magnetotelluric measurements,* PhD thesis, Universidad del País Vasco, Spain.

Alvarez-Aramberri, J. & Pardo, D., 2017. Dimensionally adaptive hp-finite element simulation and inversion of 2D magnetotelluric measurements, *J. Comput. Sci.,* **18,** 95–105.

Amestoy, P.R., Duff, I.S., L'Excellent, J.-Y. & Koster, J., 2001. A fully asynchronous multifrontal solver using distributed dynamic scheduling, *SIAM J. Matrix Anal. Appl.,* **23**(1), 15–41.

Amestoy, P.R., Guermouche, A., L'Excellent, J.-Y. & Pralet, S., 2006. Hybrid scheduling for the parallel solution of linear systems, *Parallel Comput.,* **32**(2), 136–156.

Aster, R.C., Borchers, B. & Thurber, C.H., 2018. *Parameter Estimation and Inverse Problems,* Elsevier.

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

Avdeev, D. & Avdeeva, A., 2009. 3D magnetotelluric inversion using a limited-memory quasi-Newton optimization, *Geophysics,* **74**(3), F45–F57.

Avdeev, D.B., 2005. Three-dimensional electromagnetic modelling and inversion from theory to application, *Surv. Geophys.,* **26**(6), 767–799.

Benner, P., Ohlberger, M., Patera, A., Rozza, G. & Urban, K., 2017. *Model Reduction of Parametrized Systems,* Springer.

Bennington, N.L., Zhang, H., Thurber, C.H. & Bedrosian, P.A., 2015. Joint inversion of seismic and magnetotelluric data in the Parkfield Region of California using the normalized cross-gradient constraint, *Pure appl. Geophys.,* **172**(5), 1033–1052.

Berdichevsky, M.N. & Dmitriev, V.I., 2010. *Models and Methods of Magnetotellurics,* Springer Science & Business Media.

Bodin, T. & Sambridge, M., 2009. Seismic tomography with the reversible jump algorithm, *Geophys. J. Int.,* **178**(3), 1411–1436.

Brodie, R. & Jiang, W., 2018. Trans-dimensional Monte Carlo inversion of short period magnetotelluric data for cover thickness estimation, *ASEG Extended Abstr.,* **2018**(1), 1–7.

Brunton, S.L. & Kutz, J.N., 2019. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control,* Cambridge University Press.

Cagniard, L., 1953. Basic theory of the magneto-telluric method of geophysical prospecting, *Geophysics,* **18**(3), 605–635.

Chen, J., Hoversten, G.M., Key, K., Nordquist, G. & Cumming, W., 2012. Stochastic inversion of magnetotelluric data using a sharp boundary parameterization and application to a geothermal site, *Geophysics,* **77**(4), E265–E279.

Chen, Y., Hesthaven, J.S., Maday, Y. & Rodríguez, J., 2010. Certified reduced basis methods and output bounds for the harmonic Maxwell's Equations, *SIAM J. Sci. Comput.,* **32**(2), 970–996.

Comeau, M.J., Unsworth, M.J. & Cordell, D., 2016. New constraints on the magma distribution and composition beneath Volcán Uturuncu and the southern Bolivian Altiplano from magnetotelluric data, *Geosphere,* **12**(5), 1391–1421.

Conway, D., Simpson, J., Didana, Y., Rugari, J. & Heinson, G., 2018. Probabilistic magnetotelluric inversion with adaptive regularisation using the No-U-Turns sampler, *Pure appl. Geophys.,* 2881–-2894, **175**.

Cui, T., Marzouk, Y.M. & Willcox, K.E., 2015. Data-driven model reduction for the Bayesian solution of inverse problems, *Int. J. Numer. Methods Eng.,* **102**(5), 966–990.

Dai, L. & Karato, S.-I., 2009. Electrical conductivity of orthopyroxene: Implications for the water content of the asthenosphere, *Proc. Japan Acad., Ser. B,* **85**(10), 466–475.

deGroot Hedlin, C. & Constable, S., 1990. Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data, *Geophysics,* **55**(12), 1613–1624.

Douglas, J., Jr, Santos, J.E., Sheen, D. & Ye, X., 1999. Nonconforming galerkin methods based on quadrilateral elements for second order elliptic problems, *ESAIM: Math. Model. Numer. Anal.,* **33**(4), 747–770.

Douglas, J., Jr, Santos, J.E. & Sheen, D., 2000. A nonconforming mixed finite element method for Maxwell's Equations, *Math. Models Methods Appl. Sci.,* **10**(04), 593–613.

Egbert, G.D. & Kelbert, A., 2012. Computational recipes for electromagnetic inverse problems, *Geophys. J. Int.,* **189**(1), 251–267.

Evans, R., 2012. Conductivity of Earth materials. (Chapter) In: Chave A. Jones A.G (Editors) *The Magnetotelluric Method, Theory and Practice,* Cambridge Univ. Press , Cambridge, 50–95.

Evans, R.L., Hirth, G., Baba, K., Forsyth, D., Chave, A. & Mackie, R., 2005. Geophysical evidence from the MELT area for compositional controls on oceanic plates, *Nature,* **437**(7056), 249–-252.

Farquharson, C.G., Oldenburg, D.W., Haber, E. & Shekhtman, R., 2002. An algorithm for the three-dimensional inversion of magnetotelluric data, in *SEG Technical Program Expanded Abstracts 2002,* pp. 649–652, Society of Exploration Geophysicists.

Florentin, E. & Díez, P., 2012. Adaptive reduced basis strategy based on goal oriented error assessment for stochastic problems, *Comput. Methods Appl. Mech. Eng.,* **225,** 116–127.

Frangos, M., Marzouk, Y., Willcox, K. & van Bloemen Waanders, B., 2010. Surrogate and reduced-order modeling: a comparison of approaches for large-scale statistical inverse problems. (Chapter) In: in Computational Methods for Large-Scale Inverse Problems and Quantification of Uncertainty, Biegler L., Biros G., Ghattas O., Heinkenschloss M., Keyes D., Mallick B., Marzouk Y., Tenorio L., van Bloemen Waanders B., and Willcox K., eds., Wiley, 2011.

Fullea, J., Muller, M. & Jones, A., 2011. Electrical conductivity of continental lithospheric mantle from integrated geophysical and petrological modeling: application to the Kaapvaal Craton and Rehoboth Terrane, southern Africa, *J. geophys. Res.: Solid Earth,* **116**(B10), doi:10.1029/2011JB008544.

Haario, H., Saksman, E. & Tamminen, J., 2001. An adaptive Metropolis algorithm, *Bernoulli,* **7**(2), 223–242,

Gilks, W.R., Richardson, S. & Spiegelhalter, D., 1995. *Markov Chain Monte Carlo in Practice,* Chapman and Hall/CRC.

Gregory, P., 2005. *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support,* Cambridge University Press.

Haario, H., Laine, M., Mira, A. & Saksman, E., 2006. DRAM: efficient adaptive MCMC, *Stat. Comput.,* **16**(4), 339–354.

Haber, E., Ascher, U.M. & Oldenburg, D.W., 2004. Inversion of 3D electromagnetic data in frequency and time domain using an inexact all-at-once approach, *Geophysics,* **69**(5), 1216–1228.

Han, N., Nam, M.J., Kim, H.J., Lee, T.J., Song, Y. & Suh, J.H., 2008. Efficient three-dimensional inversion of magnetotelluric data using approximate sensitivities, *Geophys. J. Int.,* **175**(2), 477–485.

Harinarayana, T., 1999. Combination of em and dc measurements for upper crustal studies, *Surv. Geophys.,* **20**(3–4), 257–278.

Hartmann, R., 2008. Multitarget error estimation and adaptivity in aerodynamic flow simulations, *SIAM J. Sci. Comput.,* **31**(1), 708–731.

Hashin, Z. & Shtrikman, S., 1962. A variational approach to the theory of the effective magnetic permeability of multiphase materials, *J. appl. Phys.,* **33**(10), 3125–3131.

Hashin, Z. & Shtrikman, S., 1963. A variational approach to the theory of the elastic behaviour of multiphase materials, *J. Mech. Phys. Solids,* **11**(2), 127–140.

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications, *Biometrika,* **57**(1), 97–109, doi:10.1093/biomet/57.1.97.

Heise, W., Caldwell, T., Bibby, H.M. & Bannister, S., 2008. Three-dimensional modelling of magnetotelluric data from the rotokawa geothermal field, Taupo Volcanic Zone, New Zealand, *Geophys. J. Int.,* **173**(2), 740–750.

Hess, M.W., 2016. *Reduced Basis Approximations for Electromagnetic Applications,* PhD thesis, Otto-von-Guericke Universität Magdeburg.

Hess, M.W. & Benner, P., 2013. Fast evaluation of time–harmonic Maxwell's Equations using the reduced basis method, *IEEE Trans. Microw. Theor. Tech.,* **61**(6), 2265–2274.

Hesthaven, J.S., Stamm, B. & Zhang, S., 2012. Certified reduced basis method for the electric field integral equation, *SIAM J. Sci. Comput.,* **34**(3), A1777–A1799.

Hesthaven, J.S., Rozza , G. & Stamm, B., 2016. *Certified Reduced Basis Methods for Parametrized Partial Differential Equations,* Springer.

Hoffmann, R. & Dietrich, P., 2004. An approach to determine equivalent solutions to the geoelectrical 2d inversion problem, *J. appl. Geophys.,* **56**(2), 79–91.

Jegen, M.D., Hobbs, R.W., Tarits, P. & Chave, A., 2009. Joint inversion of marine magnetotelluric and gravity data incorporating seismic constraints: preliminary results of sub-basalt imaging off the Faroe Shelf, *Earth planet. Sci. Lett.,* **282**(1–4), 47–55.

Jones, A.G., Afonso, J.C. & Fullea, J., 2017. Geochemical and geophysical constrains on the dynamic topography of the Southern African Plateau, *Geochem. Geophys. Geosyst.,* **18**(10), 3556–3575.

Karato, S.-I., 1990. The role of hydrogen in the electrical conductivity of the upper mantle, *Nature,* **347**(6290), 272.

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

Karato, S.-I., 2011. Water distribution across the mantle transition zone and its implications for global material circulation, *Earth planet. Sci. Lett.,* **301**(3–4), 413–423.

Karato, S.-I. & Wang, D., 2013. Electrical conductivity of minerals and rocks, *Phys. Chem. Deep Earth,* **5,**145–182.

Kelbert, A., Egbert, G.D. & Schultz, A., 2008. Non-linear conjugate gradient inversion for global EM induction: resolution studies, *Geophys. J. Int.,* **173**(2), 365–381.

Kelbert, A., Schultz, A. & Egbert, G., 2009. Global electromagnetic induction constraints on transition-zone water content variations, *Nature,* **460**(7258), 1003–1006.

Kelbert, A., Meqbel, N., Egbert, G.D. & Tandon, K., 2014. ModEM: a modular system for inversion of electromagnetic geophysical data, *Comput. Geosci.,* **66,** 40–53.

Key, K., 2016. MARE2DEM: a 2-D inversion code for controlled-source electromagnetic and magnetotelluric data, *Geophys. J. Int.,* **207**(1), 571–588.

Khan, A., 2016. On Earth's mantle constitution and structure from joint analysis of geophysical and laboratory-based data: An example, *Surv. Geophys.,* **37**(1), 149–189.

Khan, A., Connolly, J. & Taylor, S., 2008. Inversion of seismic and geodetic data for the major element chemistry and temperature of the Earth's mantle, *J. geophys. Res.: Solid Earth,* **11,** p B09308, doi:10.1029/2007JB005239.

Kuvshinov, A., 2008. 3-D global induction in the oceans and solid Earth: recent progress in modeling magnetic and electric fields from sources of magnetospheric, ionospheric and oceanic origin, *Surv. Geophys.,* **29**(2), 139–186.

Kuvshinov, A. & Olsen, N., 2006. A global model of mantle conductivity derived from 5 years of CHAMP, Ørsted, and SAC-C magnetic data, *Geophys. Res. Lett.,* **33**(18), doi:10.1029/2006GL027083.

Kuvshinov, A. & Semenov, A., 2012. Global 3-D imaging of mantle electrical conductivity based on inversion of observatory C-responses—I. An approach and its verification, *Geophys. J. Int.,* **189**(3), 1335–1352.

Kuvshinov, A., Sabaka, T. & Olsen, N., 2006. 3-D electromagnetic induction studies using the Swarm constellation: Mapping conductivity anomalies in the Earth's mantle, *Earth planets Space,* **58**(4), 417–427.

Lieberman, C., Willcox, K. & Ghattas, O., 2010. Parameter and state model reduction for large-scale statistical inverse problems, *SIAM J. Sci. Comput.,* **32**(5), 2523–2542.

Mackie, R.L. & Madden, T.R., 1993. Three-dimensional magnetotelluric inversion using conjugate gradients, *Geophys. J. Int.,* **115**(1), 215–229.

Mackie, R.L., Madden, T.R. & Park, S.K., 1996. A three-dimensional magnetotelluric investigation of the California Basin and Range, *J. geophys. Res.: Solid Earth,* **101**(B7), 16221–16239.

Mackie, R.L., Smith, J.T. & .Madden, T.R, 1994. Three-dimensional electromagnetic modeling using finite difference Equations: The magnetotelluric example, *Radio Science,* **29**(4), 923–935.

Mallick, K. & Verma, R., 1979. Time-domain electromagnetic sounding–computation of multi-layer response and the problem of equivalence in interpretation, *Geophys. Prospect.,* **27**(1), 137–155.

Marzouk, Y.M. & Najm, H.N., 2009. Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems, *J. Comput. Phys.,* **228**(6), 1862–1902.

Meqbel, N.M., Egbert, G.D., Wannamaker, P.E., Kelbert, A. & Schultz, A., 2014. Deep electrical resistivity structure of the northwestern us derived from 3-D inversion of USArray magnetotelluric data, *Earth planet. Sci. Lett.,* **402,** 290–304.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. & Teller, E., 1953. Equation of state calculations by fast computing machines, *J. Chem. Phys.,* **21**(6), 1087–1092.

Miensopust, M.P., Queralt, P., Jones, A.G. &3D MT modellers, 2013. Magnetotelluric 3-D inversion— a review of two successful workshops on forward and inversion code testing and comparison, *Geophysical Journal International,* **193**(3), 1216–1238.

Mira, A., 2001. On Metropolis-Hastings algorithms with delayed rejection, *Metron,* **59**(3–4), 231–241.

Moorkamp, M., Jones, A. & Eaton, D., 2007. Joint inversion of teleseismic receiver functions and magnetotelluric data using a genetic algorithm: are seismic velocities and electrical conductivities compatible?, *Geophys. Res. Lett.,* **34**(16), doi: 10.1029/2007GL030519.

Moorkamp, M., Jones, A. & Fishwick, S., 2010. Joint inversion of receiver functions, surface wave dispersion, and magnetotelluric data, *J. geophys. Res.: Solid Earth,* **115**(B4), doi:10.1029/2009JB006369.

Mosegaard, K. & Hansen, T.M., 2016. Inverse methods: problem formulation and probabilistic solutions, *Integr. Imag. Earth: Theor. Appl., Geophys. Monogr.,* **218,** 9–27.

Nam, M.J., Kim, H.J., Song, Y., Lee, T.J., Son, J.S & Suh, J.H., 2007. 3D magnetotelluric modelling including surface topography, *Geophysical Prospecting,* **55**(2):277–287.

Newman, G.A. & Alumbaugh, D.L., 2000. Three-dimensional magnetotelluric inversion using non-linear conjugate gradients, *J. geophys. Int.,* **140**(2), 410–424.

Newman, G.A. & Boggs, P.T., 2004. Solution accelerators for large-scale three-dimensional electromagnetic inverse problems, *Inverse Probl.,* **20**(6), S151–S170, doi:10.1088/0266-5611/20/6/s10.

Ogawa, Y., *et al.*, 2001. Magnetotelluric imaging of fluids in intraplate earthquake zones, NE Japan back arc, *J. geophys. Int.,* **28**(19), 3741–3744.

Olsen, N., Lühr, H., Sabaka, T.J., Mandea, M., Rother, M., Tøffner-Clausen, L. & Choi, S., 2006. CHAOS—a model of the Earth's magnetic field derived from CHAMP, Ørsted, and SAC-C magnetic satellite data, *Geophys. J. Int.,* **166**(1), 67–75.

Olsen, N., *et al.*, 2013. The Swarm satellite constellation application and research facility (SCARF) and swarm data products, *Earth Planets Space,* **65**(11), 1189–1200.

Ortega-Gelabert, O., Zlotnik, S., Afonso, J. & Diez, P., 2020. Fast Stokes flow simulations for geophysical-geodynamic inverse problems and sensitivity analyses based on reduced order modeling, *J. geophys. Res.: Solid Earth,* doi:10.1029/2019JB018314.

Pardo, D., 2010. Multigoal-oriented adaptivity for hp-finite element methods, *Proc. Comput. Sci.,* **1**(1), 1953–1961.

Park, S.K. & Ostos, L.C., 2013. Constraints from magnetotelluric measurements on magmatic processes and upper mantle structure in the vicinity of lassen volcanic center, Northern California, *Geosphere,* **9**(3), 382–393.

Patera, A.T. & Rozza, G., 2007. *Reduced basis approximation and a posteriori error estimation for parametrized partial differential equations,* MIT Cambridge, MA, USA.

Peherstorfer, B., Willcox, K. & Gunzburger, M., 2018. Survey of multifidelity methods in uncertainty propagation, inference, and optimization, *SIAM Rev.,* **60**(3), 550–591.

Pommier, A., 2014. Interpretation of magnetotelluric results using laboratory measurements, *Surv. Geophys.,* **35**(1), 41–84.

Prud'Homme, C., Rovas, D.V., Veroy, K., Machiels, L., Maday, Y., Patera, A.T. & Turinici, G., 2002. Reliable real-time solution of parametrized partial differential equations: reduced-basis output bound methods, *J. Fluids Eng.,* **124**(1), 70–80.

Quarteroni, A., Rozza, G. & Manzoni, A., 2011. Certified reduced basis approximation for parametrized partial differential equations and applications, *J. Math. Ind.,* **1**(1), 3, doi:10.1186/2190-5983-1-3.

Quarteroni, A., Manzoni, A. & Negri, F., 2015. *Reduced Basis Methods for Partial Differential Equations: An Introduction,* Vol. **92,** Springer.

Ray, A. & Myer, D., 2019. Bayesian geophysical inversion with trans-dimensional Gaussian process machine learning, *Geophys. J. Int.,* **217**(3), 1706–1726.

Renka, R.J., 1988a. Algorithm 661: QSHEP3D: quadratic Shepard method for trivariate interpolation of scattered data, *ACM Trans. Math. Softw. (TOMS),* **14**(2), 151–152.

Renka, R.J., 1988b. Multivariate interpolation of large sets of scattered data, *ACM Trans. Math. Softw. (TOMS),* **14**(2), 139–148.

Roberts, J.J. & Tyburczy, J.A., 1999. Partial-melt electrical conductivity: influence of melt composition, *J. geophys. Res.: Solid Earth,* **104**(B4), 7055–7065.

Robertson, K., Thiel, S. & Meqbel, N., 2020. Quality over quantity: on workflow and model space exploration of 3d inversion of mt data, *Earth Planets Space,* **72**(1), 1–22.

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

Rosas-Carbajal, M., Linde, N., Kalscheuer, T. & Vrugt, J.A., 2013. Two-dimensional probabilistic inversion of plane-wave electromagnetic data: methodology, model constraints and joint inversion with electrical resistivity data, *Geophys. J. Int.,* **196**(3), 1508–1524.

Rosas-Carbajal, M., Linde, N., Peacock, J., Zyserman, F.I., Kalscheuer, T. & Thiel, S., 2015. Probabilistic 3-D time-lapse inversion of magnetotelluric data: application to an enhanced geothermal system, *Geophys. Suppl. Mon. Not. R. Astron. Soc.,* **203**(3), 1946–1960.

Rozza, G., Huynh, D. B.P. & Patera, A.T., 2007. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations, *Arch. Comput. Methods Eng.,* **15**(3), 229–275.

Rozza, G., Huynh, D., Nguyen, N.C. & Patera, A.T., 2009. Real-time reliable simulation of heat transfer phenomena, in *Proceedings of HT2009, 2009 ASME Summer Heat Transfer Conference, S. Francisco, USA 2009. Paper HT 2009–88212 Paper HT 2009–88212*.

Rozza, G., Huynh, D.P. & Manzoni, A., 2013. Reduced basis approximation and a posteriori error estimation for stokes flows in parametrized geometries: roles of the inf-sup stability constants, *Numer. Math.,* **125**(1), 115–152.

Sabaka, T.J., Olsen, N. & Purucker, M.E., 2004. Extending comprehensive models of the Earth's magnetic field with Ørsted and CHAMP data, *Geophys. J. Int.,* **159**(2), 521–547.

Sasaki, Y., 2001. Full 3-D inversion of electromagnetic data on pc, *J. appl. Geophys.,* **46**(1), 45–54.

Selway, K., 2014. On the causes of electrical conductivity anomalies in tectonically stable lithosphere, *Surv. Geophys.,* **35**(1), 219–257.

Selway, K., O'Donnell, J. & Özaydin, S., 2019. Upper mantle melt distribution from petrologically constrained magnetotellurics, *Geochem. Geophys. Geosyst.,* **20**(7), 3328–3346.

Semenov, A. & Kuvshinov, A., 2012. Global 3-D imaging of mantle conductivity based on inversion of observatory C-responses–II. data analysis and results, *Geophys. J. Int.,* **191**(3), 965–992.

Sen, M.K. & Stoffa, P.L., 1996. Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion, *Geophys. Prospect.,* **44**(2), 313–350.

Sheen, D., 1997. Approximation of electromagnetic fields: Part I. Continuous problems, *SIAM J. Appl. Math.,* **57**(6), 1716–1736.

Shen, Y., Xu, P. & Li, B., 2012. Bias-corrected regularized solution to inverse ill-posed models, *J. Geod.,* **86**(8), 597–608.

Siripunvaraporn, W., 2012. Three-dimensional magnetotelluric inversion: an introductory guide for developers and users, *Surv. geophys.,* **33**(1), 5–27.

Siripunvaraporn, W. & Egbert, G., 2009. WSINV3DMT: vertical magnetic field transfer function inversion and parallel implementation, *Phys. Earth planet. Inter.,* **173**(3–4), 317–329.

Siripunvaraporn, W., Egbert, G., Lenbury, Y. & Uyeshima, M., 2005. Three-dimensional magnetotelluric inversion: data-space method, *Phys. Earth planet. Inter.,* **150**(1–3), 3–14.

Siripunvaraporn, W., Egbert, G. & Lenbury, Y., 2002. Numerical accuracy of magnetotelluric modeling: a comparison of finite difference approximations, *Earth, planets and space,* **54**(6), 721–725.

Smith, R.C., 2013. *Uncertainty Quantification: Theory, Implementation, and Applications,* Vol. **12,** SIAM.

Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation,* Vol. **89,** SIAM.

Tarantola, A. & Valette, B., 1982. Inverse problems—quest for information, *J. geophys.,* **50**(1), 159–170.

Ten Grotenhuis, S.M., Drury, M.R., Spiers, C.J. & Peach, C.J., 2005. Melt distribution in olivine rocks based on electrical conductivity measurements, *J. geophys. Res.: Solid Earth,* **110**(B12), doi:10.1029/2004JB003462.

Tietze, K. & Ritter, O., 2013. Three-dimensional magnetotelluric inversion in practice–the electrical conductivity structure of the San Andreas Fault in Central California, *Geophys. J. Int.,* **195**(1), 130–147.

Tikhonov, A., 1950. On determining electric characteristics of the deep layers of the earth's crust, *Dolk. Acad. Nauk. SSSR,* **73,** 295–297.

Varentsov, I.M., Kulikov, V., Yakovlev, A. & Yakovlev, D., 2013. Possibilities of magnetotelluric methods in geophysical exploration for ore minerals, *Izv., Phys. Solid Earth,* **49**(3), 309–328.

Vozar, J., Jones, A.G., Fullea, J., Agius, M.R., Lebedev, S., Le Pape, F. & Wei, W., 2014. Integrated geophysical-petrological modeling of lithosphere-asthenosphere boundary in central Tibet using electromagnetic and seismic data, *Geochem. Geophys. Geosyst.,* **15**(10), 3965–3988.

Vozoff, K., 1990. Magnetotellurics: principles and practice, *Proc. Indian Acad. Sci.—Earth Planet. Sci.,* **99**(4), 441–471.

Wait, J.R., 1962. Theory of magnetotelluric fields, *J. Res. NBS D,* **66**(5), 509–541.

Wang, D., Mookherjee, M., Xu, Y. & Karato, S.-I., 2006. The effect of water on the electrical conductivity of olivine, *Nature,* **443**(7114), 977–980.

Wei, W., *et al.*, 2001. Detection of widespread fluids in the Tibetan crust by magnetotelluric studies, *Science,* **292**(5517), 716–719.

Yan, L. & Zhou, T., 2019. Adaptive multi-fidelity polynomial chaos approach to bayesian inference in inverse problems, *J. Comput. Phys.,* **381,** 110–128.

Yang, B., Egbert, G.D., Kelbert, A. & Meqbel, N.M., 2015. Three-dimensional electrical resistivity of the north-central USA from Earthscope long period magnetotelluric data, *Earth planet. Sci. Lett.,* **422,** 87–93.

Yoshino, T., 2010. Laboratory electrical conductivity measurement of mantle minerals, *Surv. Geophys.,* **31**(2), 163–206.

Yoshino, T., Matsuzaki, T., Shatskiy, A. & Katsura, T., 2009. The effect of water on the electrical conductivity of olivine aggregates and its implications for the electrical structure of the upper mantle, *Earth planet. Sci. Lett.,* **288**(1–2), 291–300.

Yoshino, T., Shimojuku, A., Shan, S., Guo, X., Yamazaki, D., Ito, E., Higo, Y. & Funakoshi, K.-I., 2012. Effect of temperature, pressure and iron content on the electrical conductivity of olivine and its high-pressure polymorphs, *J. geophys. Res.: Solid Earth,* **117**(B8), doi:10.1029/2011JB008774.

Zhang, J., Zheng, Q., Chen, D., Wu, L. & Zeng, L., 2019. Surrogate-based bayesian inverse modeling of the hydrological system: an adaptive approach considering surrogate approximation error, *Water Resour. Res,* **56,** doi.org/10.1029/2019WR025721.

Zhdanov, M. & Hursan, G., 2000. 3D electromagnetic inversion based on quasi-analytical approximation, *Inverse Probl.,* **16**(5), 1297.

Zhdanov, M.S. & Keller, G.V., 1994. The geoelectrical methods in geophysical exploration, *Methods Geochem. Geophys.,* **31,** I–IX.

Zhdanov, M.S., Fang, S. & Hursán, G., 2000. Electromagnetic inversion using quasi-linear approximation, *Geophysics,* **65**(5), 1501–1513.

Zhdanov, M.S., Green, A., Gribenko, A. & Cuma, M., 2010. Large-scale three-dimensional inversion of Earthscope MT data using the integral equation method, *Izv., Phys. Solid Earth,* **46**(8), 670–678.

Zyserman, F.I. & Santos, J.E., 2000. Parallel finite element algorithm with domain decomposition for three-dimensional magnetotelluric modelling, *J. appl. Geophys.,* **44**(4), 337–351.

10 SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1 Plan view of DTM1 model. The symbol × indicates the origin of coordinates.

Figure S2 Different side views of the DTM1 model illustrating the three anomalies embedded in a 100 Ωm background.

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

Figure S3 Apparent resistivity and phase derived from the off-diagonal elements of the impedance tensor at profiles $x = 0$ and $y = 0$ for 1000 s. We use *Max3D-DDC* and other codes referred to as: *HL* (Nam *et al.* 2007), *Mackie* (Mackie *et al.* 1994), *winGLink* (Mackie *et al.* 1994), *wsinv3dmt* (Siripunvaraporn *et al.* 2002) and *mt3dinv* (Farquharson *et al.* 2002).

Figure S4 Apparent resistivity and phase derived from all elements of the impedance tensor at profile $x = 15$ km and period of 1000 s. We use *Max3D-DDC* and other codes referred to as: *HL* (Nam *et al.* 2007), *Mackie* (Mackie *et al.* 1994), *winGLink* (Mackie *et al.* 1994), *wsinv3dmt* (Siripunvaraporn *et al.* 2002) and *mt3dinv* (Farquharson *et al.* 2002).

Figure S5 Apparent resistivity and phase derived from the off-diagonal elements of the impedance tensor at profiles $x = 0$ and $y = 0$ for 1000 s. We use *Max3D-G* and other codes referred to as: *HL* (Nam *et al.* 2007), *Mackie* (Mackie *et al.* 1994), *winGLink* (Mackie *et al.* 1994), *wsinv3dmt* (Siripunvaraporn *et al.* 2002) and *mt3dinv* (Farquharson *et al.* 2002).

Figure S6 Apparent resistivity and phase derived from all elements of the impedance tensor at profile $x = 15$ km for 1000 s. We use *Max3D-G* and other codes referred to as: *HL*(Nam *et al.* 2007), *Mackie* (Mackie *et al.* 1994), *winGLink* (Mackie *et al.* 1994), *wsinv3dmt* (Siripunvaraporn *et al.* 2002) and *mt3dinv* (Farquharson *et al.* 2002).

Figure S7 List of DTM1 forward results extracted from Miensopust *et al.* (2013) and modified to include our results. The table specifies: code and its type (FD, FE or IE), user, mesh and in parentheses (if available) the lateral width of the cells at the centre of the model, required CPU time (and target tolerance TT or achieved tolerance AT) and computer used.

Figure S8 Plan view Model 2. The symbol × indicates the origin of coordinates.

Figure S9 Side view Model 2 showing three anomalies embedded in a 100 Ωm background.

Figure S10 Apparent resistivity derived from the off-diagonals components of the impedance tensor computed with (a) and (b) *Max3D-DDC* and (c) and (d) ModEM. (a)–(d) Plots for 50 cells between −125 to 125 km at $x = 0$ and for 29 frequencies (vertical axes) between 1 and 10^4 s.

Figure S11 Apparent phase derived from the off-diagonal components of the impedance tensor computed with (a) and (b) *Max3D-DDC* and (c) and (d) ModEM. (a)–(d) Plots for 50 cells between −125 to 125 km at $x = 0$ and for 29 frequencies (vertical axes) between 1 and 10^4 s.

Figure S12 Apparent resistivity derived from the off-diagonal components of the impedance tensor computed with (a) and (b) *Max3D-G* and (c) and (d) ModEM. (a)–(d) Plots for 50 cells between −125 to 125 km at $x = 0$ and for 29 frequencies (vertical axes) between 1 and 10^4 s.

Figure S13 Apparent phase derived from the off-diagonals components of the impedance tensor computed with (a) and (b) *Max3D-G* and (c) and (d) ModEM. (a)–(d) Plots for 50 cells between −125 to 125 km at $x = 0$ and for 29 frequencies (vertical axes) between 1 and 10^4 s.

Figure S14 Difference (*ModEM* - *Max3D-G*) for (a) and (b) apparent resistivity and (c) and (d) phase derived from the off-diagonals components of the impedance tensor.

Figure S15 Results for a two-parameter MCMC inversion. The panels (a)–(f) show the joint and marginal (black solid lines) PDFs. The initial and true models are indicated by red and black stars, respectively. Panel (a) shows the PDFs obtained with full FE (high-fidelity) solutions. Panels (b)–(f) show the PDFs obtained with our RB+MCMC algorithm for RB tolerances of $\beta = 1e^{-6}$, $1e^{-5}$, $6e^{-5}$, $1e^{-4}$ and $1e^{-3}$, respectively. The total number of bases computed in each case was 40, 29, 3, 2 and 1. In all cases, the total number of simulations was 4000 000.

Figure S16 Posterior PDFs of data for station 69. Synthetic data and error bars are plotted in green. (a) and (b) Posterior PDFs of the off-diagonal apparent resistivity. (c) and (d) Posterior PDFs of the off-diagonal apparent phases.

Figure S17 Posterior PDFs of data for station 264. Synthetic data and error bars are plotted in green. (a) and (b) Posterior PDFs of the off-diagonal apparent resistivity. (c) and (d) Posterior PDFs of the off-diagonal apparent phases.

Figure S18 Marginal posterior PDFs (blue bars) and best fitting distributions (black lines) of 49 of the 196 model parameters obtained after 2500 000 RB+MCMC simulations. The real value, starting value and prior bounds of each parameter are shown in green, red and light blue vertical lines, respectively. The position of each PDF in the figure corresponds with the location of the 49 columns at the south-west corner of the model (Fig. 1b in main text).

Figure S19 Marginal posterior PDFs (blue bars) and best-fitting distributions (black line) of 49 of the 196 model parameters obtained after 2500 000 RB+MCMC simulations. The real value, starting value and prior bounds of each parameter are shown in green, red and light blue vertical lines, respectively. The position of each PDF in the figure corresponds with the location of the 49 columns at the north-west corner of the model (Fig. 1b in main text).

Figure S20 Marginal posterior PDFs (blue bars) and best-fitting distributions (black lines) of 49 of the 196 model parameters obtained after 2500 000 RB+MCMC simulations. The real value, starting value and prior bounds of each parameter are shown in green, red and light blue vertical lines, respectively. The position of each PDF in the figure corresponds with the location of the 49 columns at the south-east corner of the model (Fig. 1b in main text).

Figure S21 Marginal posterior PDFs (blue bars) and best-fitting distributions (black lines) of 49 of the 196 model parameters obtained after 2500 000 RB+MCMC simulations. The real value, starting value and prior bounds of each parameter are shown in green, red and light blue vertical lines, respectively. The position of each PDF in the figure corresponds with the location of the 49 columns at the north-east corner of the model (Fig. 1b in main text).

Figure S22 Posterior PDFs of data for station 334. Synthetic data and error bars are plotted in green. (a) and (b) Posterior PDFs of the off-diagonal apparent resistivity. (c) and (d) Posterior PDFs of the off-diagonal apparent phases.

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

Figure S23 Posterior PDFs of data for station 279. Synthetic data and error bars are plotted in green. (a) and (b) Posterior PDFs of the off-diagonal apparent resistivity. (c) and (d) Posterior PDFs of the off-diagonal apparent phases.
Figure S24 Posterior PDFs of data for station 102. Synthetic data and error bars are plotted in green. (a) and (b) Posterior PDFs of the off-diagonal apparent resistivity. (c) and (d) Posterior PDFs of the off-diagonal apparent phases.
Figure S25 Marginal posterior PDFs (blue bars) and best-fitting distributions (black lines) of 50 of the 1782 node parameters obtained after 2500 000 RB+MCMC simulations. The real value, starting value and prior bounds of each parameter are shown in green, red and light blue vertical lines, respectively.
Figure S26 Marginal posterior PDFs (blue bars) and best-fitting distributions (black lines) of 50 of the 1782 node-parameters obtained after 2500 000 RB+MCMC simulations. The real value, starting value and prior bounds of each parameter are shown in green, red and light blue vertical lines, respectively.
Table S1 Control parameters used for the ModEM inversion.
Table S2 Constant values assumed during the 1-D heat transfer inversion.
Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022

APPENDIX A: MAPPING THERMOCHEMICAL PARAMETERS TO ELECTRICAL CONDUCTIVITY

The electrical conductivity of mantle rocks is primarily controlled by their thermophysical state (temperature, pressure, composition, fluid, iron and melt content). The temperature dependence of electrical conductivity can be described with an Arrhenius-type Equation:

$$\sigma = \sigma_0 \exp\left(\frac{-\Delta H}{k_B T}\right), \tag{A1}$$

Table A1. Parameters used to compute the conductivity of the mantle.

Phase	σ_0	σ_{0i}	a	b	c	d	e	f	ΔV	ΔH_i	X_{Fe}
Olivine	2.70	4.73	1.64	0.246	− 4.85	3.26			0.68	2.31	0.10
Opx	3.0		1.90	− 2.77	2.61	− 1.09					0.107
Cpx	3.25		2.07	− 2.77	2.61	− 1.09					$5.84e^{-2}$
Garnet		4.96	2.60	− 15.33	80.40	− 194.6	202.6	−75.0			0.168

where σ_0 is the so-called pre-exponential factor, T $[K]$ is absolute temperature and k_B [eV K^{-1}] the Boltzmann's constant. ΔH $[eV]$ is the activation enthalpy, which can be defined as a function of pressure P $[GPa]$ in the following way:

$$\Delta H = \Delta E + P\Delta V, \tag{A2}$$

where ΔE and ΔV are the activation energy and activation volume, respectively.

The main bulk conduction mechanisms in mantle minerals are ionic conduction, small polaron (hopping) conduction and proton conduction. Each mechanism follows an Arrhenius-type equation with particular activation energies depending on their charge mobility. These three conduction mechanisms can be integrated in a model for the electrical conductivity of mantle rocks as a function of pressure, temperature, water content, and composition (via Fe content) for each mineral phase (see also Yoshino *et al.* 2009; Fullea *et al.* 2011):

$$\sigma = \sigma_0 \exp\left(\frac{-\Delta H(X_{Fe}, P)}{k_B T}\right) + \sigma_{0i} \exp\left(\frac{-\Delta H_i}{k_B T}\right) + \sigma_p, \tag{A3a}$$

$$\sigma_p = f(C_w) \exp\left(\frac{-\Delta H_{wet}(C_w)}{k_B T}\right), \tag{A3b}$$

$$-\Delta H(X_{Fe}, P) = a + bX_{Fe} + cX_{Fe}^2 + dX_{Fe}^3 + eX_{Fe}^4 + fX_{Fe}^5 + P\Delta V, \tag{A3c}$$

where σ_0, σ_{0i} [S m^{-1}] and $f(C_w)$ are the so-called small polaron, ionic and proton pre-exponential factors, respectively; ΔV [cm^3 mol^{-1}] is activation volume; ΔH, ΔH_i [eV] and ΔH_{wet} are activation enthalpies; X_{Fe} is the bulk Fe content in wt%; T [K] is absolute temperature; P [GPa] is pressure and k_B [eV K^{-1}] the Boltzmann's constant.

The first term of the right-hand side of eq. (A3a) describes the contribution from small polaron conduction. As mentioned above, the activation enthalpy for this process depends on the iron content and pressure. This dependence is represented by a polynomial on X_{Fe} (eq. A3c) plus a term that depends on pressure (the coefficients a, b, c, d, e, f are determined experimentally). The second term of eq. (A3a) represents ionic conduction at high temperature and the third term (σ_p) represents the proton conduction due to the presence of 'water' (hydrogen diffusion). $f(C_w)$ and ΔH_{wet} are functions of the water content C_w [wt%] and they are estimated from laboratory experiments. The reader is referred to Fullea *et al.* (2011) and Pommier (2014) for a summary of results from different laboratories.

APPENDIX B: PROPOSAL AND PRIOR DISTRIBUTIONS

In the particular case of uncorrelated parameters, the proposal PDF for a sample $\mathbf{m}_t = [m_t^1, \ldots, m_t^N]$ is

$$q(\mathbf{m}_t|\mathbf{m}_{t-1}) = \prod_{i=1}^{N} q(m_t^i|m_{t-1}^i), \tag{B1}$$

where N is the number of parameters and $q(m^i|m_{t-1}^i)$ is the probability of the proposed move m_t^i for the ith parameter conditional on its current state m_{t-1}^i . In the particular case of a symmetric proposal, such as Gaussian or uniform PDFs, it satisfies $q(m_t^i|m_{t-1}^i) = q(m_{t-1}^i|m_t^i)$ for all m_t^i and m_{t-1}^i.

In the same way, assuming uncorrelated parameters as prior information, the prior probability of a sample $\mathbf{m}_t$ is

$$P(\mathbf{m}_t) = \prod_{i=1}^{N} P(m_t^i), \tag{B2}$$

where $P(m_t^i)$ corresponds to the prior for the ith parameter. In the case of uniform prior PDFs we have that $P(m_t^i) = 1$, and in the case of a Gaussian PDF:

$$P(m_t^i) = \frac{1}{s\sqrt{2\pi}} \exp\left(-\frac{(m_t^i - \mu)^2}{2s^2}\right), \tag{B3}$$

where μ and s are the mean and standard deviation, respectively.

In situations where the values of the model parameters span several orders of magnitude, it is useful to define prior and proposal distributions in logarithmic scale such as the Jeffreys' PDF (uniform in log-scale) with associated probability $y(m_t^i) = 1/m_t^i$ (Gregory 2005). Another common choice is to use lognormal distributions (Gaussian in log-scale):

$$y(m_t^i) = \frac{1}{\sqrt{2\pi} m_t^i s} \exp\left(-\frac{\ln(m_t^i) - \mu^2}{2s^2}\right), \tag{B4}$$

where μ and s are the mean and standard deviation in log-scale.

B1 Implementation of the Adaptive Metropolis algorithm

In order to enhance the efficiency of the MCMC algorithm and ease the problem of choosing an optimal proposal before the start of the MCMC simulation, we have implemented an AM algorithm (Haario *et al.* 2001) where a Gaussian proposal distribution is updated according to the statistics of the chain. Specifically, the covariance and mean of a multivariate Gaussian is re-computed at regular intervals.

In Section 6.3, we have chosen to define a multivariate Gaussian proposal per layer l. Each layerwise proposal is updated independently according to the history and correlation of the nodes within that layer. Since the nodal conductivity values can span several orders of magnitude, the Gaussian proposal is defined in log-scale and we evaluate its probability $q(\cdot\,|\,\cdot)$ in linear scale, that is, a multivariate lognormal PDF centred at the current state $\mathbf{m}_{t-1}^l$ with covariance Σ:

$$q(\mathbf{m}_t^l|\mathbf{m}_{t-1}^l) = \frac{1}{(2\pi)^{\frac{nlay}{2}} (\det\Sigma)^{\frac{1}{2}} \prod_{j=1}^{nlay} m_t^j} \exp\left(-\frac{1}{2}(\ln(\mathbf{m}_t^l) - \ln(\mathbf{m}_{t-1}^l))^t \Sigma^{-1} (\ln(\mathbf{m}_t^{lay}) - \ln(\mathbf{m}_{t-1}^l))\right), \tag{B5}$$

where $nlay$ is the number of nodes per layer and $\mathbf{m}^l$ is the proposed vector of nodes for layer l, that is, conductivity values for all the nodes in layer l.

Downloaded from https://academic.oup.com/gji/article/223/3/1837/5900140 by guest on 02 March 2022