Structure-coupled 3-D imaging of magnetotelluric and wide-angle seismic reflection/refraction data with interfaces

Miao Peng[1], Tan Handong[2], and Max Moorkamp[3]

[1]School of Geophysics and Information Technology, China University of Geosciences, Beijing, China
[2]State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Beijing, China
[3]Ludwig-Maximilians-Universitaet-Muenchen

November 23, 2022

Abstract

Magnetotelluric (MT) and wide-angle seismic reflection/refraction surveys play a fundamental role in understanding the crustal rheology and lithospheric structure of the Earth. In recent years, the integration of the two methods in order to improve the robustness of the inversion has started to gain attention. We present a new approach for joint 3-D inversion of MT and wide-angle seismic reflection/refraction data to accurately determine crustal structures and Moho depth. Based on H-$\varkappa$ stacking of teleseismic receiver functions (RFs), we estimate an initial reference Moho. This is used as input for the subsequent MT/seismic joint inversion, where the Moho interface is updated and crustal structures are added to the model. During the joint inversion process, structural similarity is facilitated through the cross-gradient constraint. Synthetic model tests show an improvement of the inversion results over separate inversions. In particular, the tests based on two geologically realistic models demonstrate that the crustal structure and even the trade-off between velocity and Moho interface can be sufficiently resolved by combined MT and seismic datasets when using the estimates from analysis of RFs. These results show that the new method can provide useful constraints on crustal structures including their geophysical properties and discontinuities.

Structure-coupled 3-D imaging of magnetotelluric and wide-angle seismic reflection/refraction data with interfaces

Miao Peng[1†], Handong Tan[1,2†], and Max Moorkamp[3*]

[1]School of Geophysics and Information Technology, China University of Geosciences, Beijing, China

[2] State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Beijing, China

[3]Department of Geology, University of Leicester, Leicester, United Kingdom

[†]Corresponding author: Miao Peng (pengm@cugb.edu.cn), Handong Tan (thd@cugb.edu.cn)

Key Points:

- A new coupled algorithm is developed for 3-D imaging of magnetotelluric and wide-angle seismic reflection/refraction data
- Two strategies are presented to mitigate the problem of accurately determining crustal structures and Moho interface simultaneously
- The inclusion of teleseismic receiver functions is proved to be effective in the new algorithm

[*] Now at University of Munich, Germany

Abstract

Magnetotelluric (MT) and wide-angle seismic reflection/refraction surveys play a fundamental role in understanding the crustal rheology and lithospheric structure of the Earth. In recent years, the integration of the two methods in order to improve the robustness of the inversion has started to gain attention. We present a new approach for joint 3-D inversion of MT and wide-angle seismic reflection/refraction data to accurately determine crustal structures and Moho depth. Based on H-κ stacking of teleseismic receiver functions (RFs), we estimate an initial reference Moho. This is used as input for the subsequent MT/seismic joint inversion, where the Moho interface is updated and crustal structures are added to the model. During the joint inversion process, structural similarity is facilitated through the cross-gradient constraint. Synthetic model tests show an improvement of the inversion results over separate inversions. In particular, the tests based on two geologically realistic models demonstrate that the crustal structure and even the trade-off between velocity and Moho interface can be sufficiently resolved by combined MT and seismic datasets when using the estimates from analysis of RFs. These results show that the new method can provide useful constraints on crustal structures including their geophysical properties and discontinuities.

1 Introduction

The magnetotelluric (MT) method provides crucial information on the conductivity of the subsurface of the Earth determined by the presence of interconnected fluids and partial melt and is widely used to image the lithosphere-asthenosphere system [e.g., Jones, 1999; Unsworth, 2010; Lin et al. 2017]. In contrast, wide-angle seismic surveys image gradual changes in velocity and distinct geological features such as interfaces between layers and unconformities in the crust using controlled or artificial sources [Rawlinson et al., 2001]. Based on such measurements, both layer velocity and interface geometry can be retrieved from reflection and/or refraction travel times using tomographic approaches [Zelt et al., 1992; Rawlinson et al., 2010; Karplus, et al. 2011].

The combination of multiple geophysical data types can help reduce the non-uniqueness of inversion results [e.g., Moorkamp, 2017]. For example, seismic methods generally have reliable vertical resolution on horizontal or dipping planar layers. They can effectively compensate the diffusive nature of electromagnetic fields that leads to decreased vertical resolution in the deeper earth and generally smooth models [Chave and Jones, 2012]. Conversely, blind areas of seismic refraction may exist in some cases but can be improved by MT surveys [Stanley et al., 1990; Bennington et al., 2015]. Another benefit of combining MT and wide-angle seismic data is that they sense structures on similar spatial scales, and therefore can be effectively compared or coupled.

In recent times, a number of studies have been performed utilizing joint inversion of MT and seismic traveltime data [e.g., Heincke et al., 2006; Gallardo, 2007; 2012; Hu et al., 2009; Moorkamp et al., 2011; Peng et al., 2013]. In terms of coupling approach, these studies can be classified into two main categories: direct petrophysical parameter coupling [e.g., Colombo and Stefano, 2007; Heincke et al., 2017] and structural coupling [e.g., Haber and Oldenburg, 1997; Gallardo and Meju, 2003]. The former approach focuses on explicit functional relationships between electrical resistivity and seismic velocity. These can be estimated from empirical laws based on rock physics such as Archie and Wyllie equations [Archie, 1942; Wyllie et al., 1956], derived from borehole data and porosity estimations [e.g., Jegen et al. 2009; Berryman et al., 2002], or calculated from composition and temperature of the mantle [Afonso et al. 2013]. However, these methods largely depend on a-priori petrophysical relationships, and if the relationships are estimated incorrectly, it is likely to lead to artifacts in the joint inversion results [Moorkamp et al. 2011].

Structural coupling approaches were first proposed to measure structural similarity using model curvature to enhance common boundaries [Zhang and Morgan; 1997; Haber and Oldenburg, 1997]. Gallardo and Meju [2003] introduced the cross-gradient method to consider direction-dependent constraints on different models. This direction-based approach is widely applicable and has been adapted to different geophysical data types [Gallardo and Meju, 2011; Linde et al., 2006; Fregoso and Gallardo, 2009; Doetsch et al., 2010; Moorkamp et al., 2016], since it provides a generalized methodology for evaluating structural similarity between diverse multidimensional models.

In wide-angle seismic surveys, crustal structure is commonly represented by a series of sub-horizontal layers separated by continuous interfaces. In comparison with traditional refraction interpretation, using both refracted and reflected phases to image seismic structure offers potentially better resolution [Rawlinson et al., 2001; Rawlinson and Urvoy, 2006]. However, although a few recent studies have taken these reflection travel times into consideration by jointly inverting multiple geophysical data sets with application to marine subsurface integrated imaging [e.g., Gallardo et al., 2012], examples of hybrid 3-D joint inversion frameworks that determine interface depths and physical properties simultaneously are harder to find. One issue with inverting reflection travel time is the difficulty to include data that involve seismic discontinuities, such as Moho reflections. Another problem is that there is a trade-off between Moho boundary location and crustal velocity which can be largely attributed to inadequate path coverage near the Moho interface [Rawlinson et al., 2010].

In this paper, we introduce a 3-D cross-gradient joint inversion algorithm that combines MT data with seismic reflection and refraction travel times. For the seismic tomography part of the algorithm, we consider a layered medium in a discontinuous velocity model intersected by an undulating Moho interface, allowing both seismic reflection and refraction phases to be tracked. Additional constraints from teleseismic RFs are also included in the new framework in order to overcome the limitation of accurately determining crustal structures and Moho interface simultaneously under geologically realistic conditions.

2 Forward Modeling Algorithms

2.1. Forward Modeling in MT Method

We use a 3-D staggered-grid finite difference method (SFD) as the forward algorithm for electromagnetic (EM) modeling [Tan et al., 2003]. Similar to previous work for computing the MT response of 3-D Earth models [Mackie et al., 1993], the algorithm is based on the integral forms of Maxwell's equations given by:

$$\oint \mathbf{H}\cdot d\mathbf{l} = \iint \mathbf{J}\cdot d\mathbf{S} = \iint \sigma\mathbf{E}\cdot d\mathbf{S} \tag{1}$$

$$\oint \mathbf{E}\cdot d\mathbf{l} = \iint i\mu\omega\mathbf{H}\cdot d\mathbf{S}, \tag{2}$$

where **J** is current density, σ is conductivity, ω is angular frequency, d**l** is the contour of integral closure, and d**S** is the area closed by the contour.

After spatial discretization on a staggered-grid with appropriate boundary conditions, integral equations (1) and (2) can be expressed as a complex system of large-scale linear equations. In this scheme [Tan et al. 2003], the linear equations are solved by using a stabilized Bi-conjugate gradient method and generate high-precision electric or magnetic field component efficiently. The complex impedance tensor (Zxx, Zxy, Zyx, and Zyy) can be computed from the electromagnetic fields along two orthogonal directions in four equations [e.g. Egbert and Kelbert, 2012].

Moreover, a parallel scheme was developed by adding Message Passing Interface (MPI) approach running on computer clusters in order to reduce computational time [Tan et al., 2006]. As the 3-D joint inversion requires a significant amount of computation, this parallel 3-D MT forward modeling scheme is particularly suited for it.

2.2. Traveltime Computation in Wide-angle Seismic Reflection/refraction Method

We use a multi-stage Fast Marching Method (FMM) [Rawlinson and Sambridge, 2005] as the forward method to compute theoretical travel times of seismic waves in a known 3-D velocity model volume, given hypocenter locations and receiving stations on the Earth surface. FMM is a grid based eikonal solver that can implicitly track the evolution of wavefronts for calculating reflection and refraction phases in layered media [Rawlinson and Sambridge, 2004; de Kool et al, 2006].

For a seismic P-wave in an isotropic elastic medium, the eikonal equation can be written as:

$$|\nabla T| = \frac{1}{\mathbf{v}}, \tag{3}$$

where ∇ represents gradient operator, T=$\boldsymbol{T}(x)$ is a time function (the traveltime) which describes surfaces of constant phase (wavefronts) when T is constant, and $\mathbf{v}$ is the seismic wave velocity as a function of the position at a spatial point. This equation describes the time of arrival of the fastest wave front at a given position.

Recently, an upwind scheme [de Kool et al, 2006] has been implemented for predicting multi-arrivals. One advantage of the practical grid-based method is that the propagation grid is defined separately from the inversion grid by specifying discrete sampling of velocity fields and implicit interface boundaries. Because it has no direct relationship with the inversion grid, we can consider both the calculation accuracy of forward modeling and the scale of the inversion grid. Also, adding interface nodes does not significantly increase the complexity of the inversion grid, which is another benefit for the joint inversion framework.

For our inversion we use the FMM algorithm with the upwind scheme implemented in the FMTOMO package developed by Rawlinson and Sambridge [2005]. Tests with a variety of velocity structures show that it is both computationally efficient and robust and thus preferable as a forward solver for reflection and refraction data for the joint inversion.

3 Inversion Algorithms

In addition to well-developed separate forward modeling algorithms, a flexible 3-D joint inversion framework needs large-scale optimization inversion schemes and a coupling approach.

In our implementation, the resistivity model and velocity model parameters are updated in an exchanging pattern that is similar to the approach by Um et al. [2015]. In addition, we also consider the determination of an initial Moho interface (**H**) and its adjustment during the velocity model updates in the algorithm.

The framework (Figure 1) is composed of two main modules (MT inversion module & seismic inversion module). Neither of them is independent of each other as they are linked by cross-gradient coupling at each iteration of a loop. Taking the MT inversion module, for example, we carry out the MT inversion first when the joint inversion starts. The parallel SFD forward solver as shown below is used to calculate theoretical electromagnetic responses from an initial resistivity model $\mathbf{m}_{\rho}^{0}$. Then, in a cross-gradient coupling module, we calculate the cross-gradient values ($\boldsymbol{\tau}$) and its Jacobian matrix (**B**) between the initial resistivity model $\mathbf{m}_{\rho}^{0}$ and the velocity model $\mathbf{m}_{s}^{0}$, respectivley. The first updated resistivity model $\mathbf{m}_{\rho}^{1}$ is generated by the model update expression of the data-space method as presented in Eq. (12). After that, we need to calculate the misfit (RMS) between the observed data and the synthetic responses and estimate if the value is lower than the desired level of misfit. We then perform similar computations for the seismic part of the inversion. The iteration loop will not end until both MT RMS and seismic RMS satisfy the requirement of error tolerances or reach the maximum number of iterations.

In the following section, we first introduce the MT and tomographic inversion schemes, and then present the model coupling approach and how the two methods are combined. We then discuss our strategy to estimate crustal thickness within the inversion and how it impacts on the tomographic results.

3.1. MT Inversion Scheme

The classical Occam's inversion seeks models fitting the data while at the same time having the "smoothest" structure [Constable et al. 1987]. The inversion uses a self-adaptive regularization scheme, which is affected little by the initial model and has a stable convergence [de Groot-Hedlin and Constable, 1990], thus it is an ideal algorithm to solve MT inverse problems. The minimization of the objective function is achieved by finding stationary points of the equation:

$$U(\mathbf{m}_{\rho},\lambda)=\left\|\mathbf{m}_{\rho}-\mathbf{m}_{\rho}^{0}\right\|_{\mathbf{C}_{\mathbf{m}}^{-1}}^{2}+\lambda^{-1}\left(\left\|\mathbf{d}_{\mathbf{obs}}-f(\mathbf{m}_{\rho})\right\|_{\mathbf{C}_{\mathbf{d}}^{-1}}^{2}-\xi\right), \quad (4)$$

where $\mathbf{m}_{\rho}$ is the vector of resistivity model parameters, $\mathbf{m}_{\rho}^{0}$ is the vector of priori model parameters, $\mathbf{C}_{\mathbf{m}}$ is the model covariance matrix, $\mathbf{d}_{\mathbf{obs}}$ is the vector of observed data, $f(\mathbf{m}_{\rho})$ is the model response, $\mathbf{C}_{\mathbf{d}}$ is the data covariance matrix, ξ is the desired level of misfit, and λ^{-1} is a Lagrange multiplier.

To find a new model we linearize the above objective function U, and obtain a series of iterative solutions:

$$\mathbf{m}_{\rho}^{k+1}(\lambda) = (\lambda \mathbf{C}_{\mathbf{m}}^{-1} + \mathbf{J}_k^T \mathbf{C}_{\mathbf{d}}^{-1} \mathbf{J}_k)^{-1} \mathbf{J}_k^T \mathbf{C}_{\mathbf{d}}^{-1} \mathbf{X}_k + \mathbf{m}_{\rho}^0, \tag{5}$$

where $\mathbf{J_k}$ is the corresponding Jacobian matrix, $\mathbf{X}_k = \mathbf{d}_{obs} - f(\mathbf{m}_{\rho}^k) + \mathbf{J}_k(\mathbf{m}_{\rho}^k - \mathbf{m}_{\rho}^0)$, and the inverse matrix $\mathbf{\Lambda}_k^{-1} = (\lambda \mathbf{C}_{\mathbf{m}}^{-1} + \mathbf{J}_k^T \mathbf{C}_{\mathbf{d}}^{-1} \mathbf{J}_k)^{-1}$ is a M×M positive semi-definite symmetric matrix in model space, where M is the number of model parameters.

Siripunvaraporn et al. [2005] reformulate Occam's inversion scheme into data-space in order to avoid the computational cost of inverting a M×M matrix. They transform the above iterative Eq. (5) mathematically into a new one that contains a N×N system of linear normal equations:

$$\mathbf{m}_{\rho}^{k+1}(\lambda) = \mathbf{C}_{\mathbf{m}} \mathbf{J}_k^T (\lambda \mathbf{C}_{\mathbf{d}} + \mathbf{J}_k \mathbf{C}_{\mathbf{m}} \mathbf{J}_k^T)^{-1} \mathbf{X}_k + \mathbf{m}_{\rho}^0. \tag{6}$$

Here N is the number of independent data, which for many geophysical surveys is orders of magnitude smaller than the number of model parameters M.

The data-space method is similar to Occam's inversion approach as the model iterations in data-space also follow the principle of minimizing a penalty function with a series of Lagrange multipliers λ. The data-space algorithm has been shown to be efficient in 3-D MT inversion in terms of reducing the matrix dimensions since the number of independent data N in realistic models is normally much less than the number of model parameters M, which results in significant savings of both memory and CPU time [Siripunvaraporn and Egbert, 2009].

Another advantage of the data-space method is that a generalized model covariance $\mathbf{C_m}$ is directly calculated, thus it avoids solving the inverse of the full matrix $\mathbf{C_m}$, which can result in large computational costs. Note that for model space inversion approaches [Constable et al., 1987; deGroot-Hedlin and Constable, 1990], it is common to utilize a sparse model roughness operator in terms of a "roughness penalty" instead of directly using the model covariance $\mathbf{C_m}$. Siripunvaraporn and Egbert [2000] discuss the differences between them in detail and how to deal with the $\mathbf{C_m}$ in the data-space algorithm.

3.2. Tomographic Inversion Scheme

We use a nonlinear subspace inversion approach to implement the tomographic inversion procedure. Similar to the MT inversion scheme, the inverse problem can be solved by specifying an objective function to adjust the model parameter values (i.e., velocity $\mathbf{m}_s$ and interface depth $\mathbf{H}$) and try to satisfy the data observations (source-receiver travel times), subject to imposed regularization constraints. The scheme is implemented by a least squares approximation to a multi-dimensional subspace of model space using the iterative subspace method [Kennett et al., 1988; Sambridge, 1990]. The model update is composed of a series of base vectors in terms of the steepest ascent vector in model space and the perturbation of model is finally written as:

$$\delta \mathbf{m}_{sh} = -\mathbf{A}[\mathbf{A}^T(\mathbf{G}^T\mathbf{C}_\mathbf{d}^{-1}\mathbf{G} + \varepsilon\mathbf{C}_\mathbf{m}^{-1})\mathbf{A}]^{-1}\mathbf{A}^T\boldsymbol{\gamma}, \quad (7)$$

where, **A** is a projection matrix [$\mathbf{a}^j$]; {$\mathbf{a}^j$} is the spanning set of basis vectors; $\boldsymbol{\gamma}$ is the gradient vector; **G** is the Jacobian matrix of partial derivatives; the variables **A**, $\boldsymbol{\gamma}$ and **G** are updated constantly in successive iteration. Here the vector $\delta\mathbf{m}_{sh}$ includes two classes of model parameters: the velocities and the depths of the interface vertices. For the derivation and simplification of the equation, see Rawlinson et al. [2001].

Once the projection matrix **A** has been computed and orthonormalized, the model update is obtained by the inversion of a relatively small matrix. The subspace inversion method is stable and efficient for large underdetermined inverse problems, and together with FMM, forms the basis of a fast and robust tomographic imaging scheme. Singular value decomposition (SVD) is used to identify and remove unnecessary basis vectors if **A** does not completely span all dimensions. Usually the value of the dimension size of **A** is small, for example, the dimension size used by Rawlinson and Urvoy [2006] is less than 20 when they implemented inversion of observed traveltime data from Tasmania, Australia. In this paper, the dimension of the subspace used in the tomographic inversion was set to 18 in that it offers a suitable compromise between the reduction of the misfit function and the computational burden.

3.3. Joint Inversion Algorithm

3.3.1. Coupling Approach

We enforce structural similarity between resistivity and P-wave velocity through the cross-gradient approach [Gallardo and Meju, 2003]. The cross-gradient function is defined as:

$$\boldsymbol{\tau}(x,y,z) = \nabla\mathbf{m}_\rho(x,y,z) \times \nabla\mathbf{m}_s(x,y,z), \quad (8)$$

where $\mathbf{m}_\rho$ and $\mathbf{m}_s$ represent the electrical resistivity model vector and the seismic P-wave velocity vector, respectively. Note that $\boldsymbol{\tau}(x,y,z)$ is a spatial vector in any grid that has three components ($\boldsymbol{\tau}^\mathbf{x}$, $\boldsymbol{\tau}^\mathbf{y}$, and $\boldsymbol{\tau}^\mathbf{z}$) along the coordinate axis in x, y and z direction, respectively. Here, the resistivity (ρ) and velocity (v) are transformed into $log\boldsymbol{\rho}$ and $1/\mathbf{v}$ when calculating cross-gradient values, to ensure that the variation in regions of high and low value of the model parameters will have a similar scale and can exert identical importance on the structural similarity.

In order to generate more precise results compared with our previous forward difference scheme [Peng et al., 2013], we use a central difference scheme here to discretize the cross-gradient in a 3-D subsurface medium (Figure 2).

The joint inversion algorithm requires the Jacobian matrix **B** associated with the cross-gradient function with regards to the model parameters. In our approach, the cross-gradient function for the current iteration is approximately linearized by a first-order Taylor expansion around their a-priori models [Gallardo and Meju, 2004]. Note that the matrix **B** is decomposed into three orthogonal components $\mathbf{B}^x$, $\mathbf{B}^y$, and $\mathbf{B}^z$. Each of the components is a sparse matrix in which every row has 10 nonzero elements in the central mesh. In addition, there are even less than 10 nonzero elements on the boundaries of the mesh depending on which region of the grid we deal with, i.e., the boundary surface, the edges or the vertices. To avoid large-scale matrix calculation, we use the compressed sparse row (CSR) format to store elements of **B** and utilize the efficient algorithms to manipulate sparse matrices implemented in the SPARSEKIT package [Saad, 1990].

3.3.2. Objective Function

In addition to the data misfit terms and the regularisation terms for separate inversion, the cross-gradient term for structural coupling [Gallardo and Meju, 2003] is added in the objective function to enforce structural similarity between resistivity and velocity. The framework of the joint inversion algorithm can be divided into two connected parts as shown in Figure 1: the MT inversion scheme and the seismic tomographic inversion scheme. With the structural coupling constraint the respective objective functions Ψ are then defined as:

$$\begin{aligned}\Psi_{\mathrm{MT}}(\mathbf{m}_\rho,\lambda)=\left\|\mathbf{m}_\rho-\mathbf{m}_\rho^0\right\|^2_{\mathbf{C}_\mathbf{m}^{-1}}+\lambda^{-1}\Big[\left\|\mathbf{d}_\rho-f(\mathbf{m}_\rho)\right\|^2_{\mathbf{C}_d^{-1}}\\ +\mu_\rho\left\|\boldsymbol{\tau}(\mathbf{m}_\rho,\tilde{\mathbf{m}}_s)-\boldsymbol{\tau}(\mathbf{m}_\rho^0,\mathbf{m}_s^0)\right\|^2_{\mathbf{C}_\tau^{-1}}-\xi\Big],\end{aligned} \tag{9}$$

$$\begin{aligned}\Psi_{\mathrm{Seis}}(\mathbf{m}_s,\mathbf{H})=\left\|\mathbf{m}_s-\mathbf{m}_s^0\right\|^2_{\mathbf{C}_\mathbf{m}^{-1}}+\varepsilon^{-1}\Big[\left\|\mathbf{d}_s-g(\mathbf{m}_s)\right\|^2_{\mathbf{C}_d^{-1}}\\ +\mu_s\left\|\boldsymbol{\tau}(\tilde{\mathbf{m}}_\rho,\mathbf{m}_s)-\boldsymbol{\tau}(\mathbf{m}_\rho^0,\mathbf{m}_s^0)\right\|^2_{\mathbf{C}_\tau^{-1}}\Big],\end{aligned} \tag{10}$$

where $\mathbf{m}_\rho$ (logarithm of resistivity) and $\mathbf{m}_s$ are the vector of current model parameters, $\tilde{\mathbf{m}}_\rho$、$\tilde{\mathbf{m}}_s$ are the vector of the model parameters updated during the latest MT and seismic tomographic inversion respectively. $f(\mathbf{m}_\rho)$、 $g(\mathbf{m}_s)$ denote the model response from the forward calculation for MT and the seismic method respectively, $\mathbf{d}_\rho$、 $\mathbf{d}_s$ are the vector of MT and seismic observed data, $\boldsymbol{\tau}(\cdot)$ is the cross-gradient operator, $\mathbf{C}_\tau$ denotes the cross-gradient covariance matrix that is an identity matrix, ε、λ are the regularization weighting factors, and μ_ρ、μ_s are the weighting factors of the cross-gradient terms.

In our joint inversion algorithm, the updated models are not used in integrated objective function. Instead the models are mutually constrained by the results of a previous iteration from the other method. As described above, we use the data-space method and the subspace method as the optimization algorithm respectively. For the MT module, we solve the inverse problem in the sense of Siripunvaraporn et al. [2005] and derive a new iterative formula to minimize the objective function of Ψ_{MT} in Eq. (9) which includes the coupling term. The expression (see Appendix A) for a series of model updates from an initial model is:

$$\mathbf{m}_\rho^{k+1}(\lambda)=\mathbf{C_m}\mathbf{J}_k^T\mathbf{\Gamma}_k^{-1}\mathbf{X}_k-\mu_\rho\mathbf{\Lambda}_k^{-1}\left[(\mathbf{B}_\rho^k)^T\mathbf{C}_\tau^{-1}\mathbf{B}_\rho^k(\mathbf{m}_\rho^k-\mathbf{m}_\rho^0)\right]+\mathbf{m}_\rho^0, \tag{11}$$

where $\mathbf{\Gamma}_k^{-1}=(\lambda\mathbf{C_d}+\mathbf{J}_k\mathbf{C_m}\mathbf{J}_k^T)^{-1}$, and $\mathbf{\Lambda}_k^{-1}=\dfrac{1}{\lambda}\left(\mathbf{I}-\mathbf{C_m}\mathbf{J}_k^T\mathbf{\Gamma}_k^{-1}\mathbf{J}_k\right)\mathbf{C_m}$.

Here, $\mathbf{B}_\rho=\mathbf{B}_\rho^x(\tilde{\mathbf{m}}_s)+\mathbf{B}_\rho^y(\tilde{\mathbf{m}}_s)+\mathbf{B}_\rho^z(\tilde{\mathbf{m}}_s)$ is the Jacobian matrix of the cross-gradient with respect to the current resistivity model $\mathbf{m}_\rho$, but only determined by the velocity model $\tilde{\mathbf{m}}_s$ that is updated during the latest seismic inversion. The definition of the vector $\mathbf{X_k}$ in Eq. (11) is identical to that given by Eq. (5). In fact, compared with the separate MT inversion in Eq. (6), the second term in Eq. (11) is an added cross-gradient constraint that can make the structural features of current resistivity model similar to the previous seismic model.

To adapt the subspace method [Kennett et al 1988, Sambridge, 1990] to our joint inversion optimisation problem for the seismic module, we rewrite the gradient vector and Hessian matrix in a new form by evaluating derivatives of objective function Ψ_{Seis} in Eq. (10). The solution is given by the following expressions (see Appendix B for more detail):

$$\delta\mathbf{m}_s=-\mathbf{A}[\mathbf{A}^T(\mathbf{G}^T\mathbf{C_d}^{-1}\mathbf{G}+\varepsilon\mathbf{C_m}^{-1}+\mu_s\mathbf{B}_s^T\mathbf{C}_\tau^{-1}\mathbf{B}_s)\mathbf{A}]^{-1}\mathbf{A}^T\boldsymbol{\gamma}, \tag{12}$$

$$\boldsymbol{\gamma}=\mathbf{G}^T\mathbf{C_d}^{-1}\left[\mathbf{d}_s-g(\mathbf{m}_s)\right]+(\varepsilon\mathbf{C_m}^{-1}+\mu_s\mathbf{B}_s^T\mathbf{C}_\tau^{-1}\mathbf{B}_s)\cdot(\mathbf{m}_s^0-\mathbf{m}_s), \tag{13}$$

where the Jacobian matrix of the cross-gradient $\mathbf{B}_s$ is also decomposed into three components and only determined by the resistivity model $\tilde{\mathbf{m}}_\rho$ that is updated during the latest MT inversion.

This form of objective function we adopt here is slightly different from the approach introduced by Gallardo and Meju [2003]. Their approach minimizes the objective function including the data misfit term and the regularisation term above but requires the cross-gradient term to be equal to zero, and thereby searches exact structural resemblance.

Although such an approach provides the benefit of fully enforcing the sought structural similarity, it poses some limitations on large-scale 3-D joint inverse problems [Meju and Gallardo, 2016]. Instead our approach only seeks partial resemblance, as the cross-gradient term can deviate from zero. In addition, updating the model in an alternating manner can, in theory, lead to less strict coupling between the methods. However, such mutually constrained inversion algorithms have been shown to be effective in previous studies [e.g., Um et al., 2014, Heincke et al., 2017] and offer high flexibility to integrate new methods.

3.3.3. Inverting for Seismic Layer Interfaces

First experiments with our coupled MT/seismic inversion show that it is important to consider the effect of variations in reflector positions in the seismic tomography. The inversion therefore also contains a module to update the position of chosen reflective interfaces.

The inclusion of interfaces can potentially increase the non-uniqueness of the solution because the amount of model parameters is larger if we invert for physical properties and depths of the interface vertices simultaneously. Also, we found that in geologically realistic situations the estimates of crustal velocity and Moho interface provided by separate tomographic inversions are likely to be inaccurate because of the trade-off between Moho boundary location and crustal velocity. This can be largely attributed to inadequate path coverage near the Moho interface. We propose two strategies in order to overcome such limitations.

The first strategy is to invert for resistivities and velocities with fixed interfaces determined by prior information. We can acquire relatively accurate information on the Moho interface and one-dimensional layered velocity model before starting a 3D inversion. This can be done by 1-D joint inversion of RFs and surface wave dispersion data [e.g., Moorkamp et al., 2010; Peng et al., 2012]. With a good constraint on the location of the Moho interface or the layering of the subsurface, an approach to jointly invert crustal structure with a fixed Moho interface can be feasible. However, the success of such a strategy largely depends on the initial model and the accuracy of the fixed Moho interface and thus is hard to judge.

Another choice is to update the location of interfaces while resistivities and velocities are jointly estimated. The advantage is that this strategy does not need rigorous requirement on the initial Moho boundary that may be not very close to the true Moho. This is therefore our preferred approach.

As RFs are sensitive to velocity discontinuities, they are considered to be an ideal tool to recover the position of the Moho interface. A grid search method [Zhu and Kanamori, 2000], H-κ stacking, is widely used to estimate crustal thickness H and Vp/Vs ratio (κ). Such an "inaccurate" Moho interface can be estimated from RFs and used as an initial reference Moho for the 3-D joint inversion.

However, if there is a 0.2 km/s deviation on average P wave velocity ($\overline{\mathrm{V}}$p) in the process of H-κ stacking, it can give rise to an average error of 1.8 km in estimating Moho depth [Zhu et al., 2006]. Thus, in some study areas with little priori information on $\overline{\mathrm{V}}$p, the method can lead to controversial results [Wölbern and Rümpker, 2017]. In this paper, we present an effective way of mitigating this problem by using an updated crustal model to calculate more accurate $\overline{\mathrm{V}}$p from the results of seismic refraction/reflection tomography. We use interface nodes (j) corresponding to different $\overline{\mathrm{V}}$p because the layer velocity V_p^i is different beneath each broadband seismic station (j). This $\overline{V}_p^j$ can be easily obtained by calculating the average velocity in the model underneath each station viz:

$$\overline{V}_p^j \approx \frac{H}{\sum_{i=1}^{N(\mathrm{j})}\left(\Delta z / V_p^i\right)} \quad (\mathrm{km/s}), \tag{14}$$

where H is the Moho depth, Δz is the interlayer spacing in the depth direction, and N(j) is the number of layer of the crust beneath the j$^{\mathrm{th}}$ seismic station. At each iteration, the Moho interface is updated by H-κ stacking with the new $\overline{\mathrm{V}}$p. Then, the updated Moho interface is used as the reference interface for the seismic joint inversion during the next iteration.

We finally note that the rates of convergence between MT and seismic inversion can vary significantly. This was also observed by Heincke et al. [2017], and may result in abnormal coupling in the process. For example, assuming that seismic inversion converges much more slowly than MT inversion, the latter would be constrained by an incomplete velocity model when the MT inversion ends. Therefore, we adjust the cross-gradient coupling flexibly by estimating which updated model should be used in the cross-gradient module. In this paper, we only have 4 iterations in MT inversion but 12 iterations in seismic inversion, so at each iteration the joint MT inversion is constrained by renewed resistivity model whose estimated iteration is multiple of 3.

4 Synthetic Examples and Discussions

We now apply the method above to synthetic examples to test the joint inversion algorithm. All computations were performed on the Alice2 cluster at the University of Leicester. Each compute node has a pair of 14-core Intel Xeon Skylake CPUs running at 2.6GHz, and 128GB of RAM. For the first two synthetic examples, the subsurface region for joint inversion is discretized with a 20×20×20 mesh yielding 8,000 equally-sized cells. Cell sizes in the interior of the mesh are 5×5×4 km.

The wide-angle synthetic dataset comprises more than 3800 crustal reflection (PmP) and refraction (Pg & Pn) travel times generated by 9 active sources and recorded by 144 receivers evenly distributed on a grid at the surface (Figure S1).

We place 36 MT sites on the surface and calculate the complex impedance tensor (Zxx, Zxy, Zyx and Zyy) for five frequencies (10, 1, 0.1, 0.01 and 0.001 Hz). Gaussian noise with 2% is added to the synthetic MT data and a standard deviation of 5 ms to the synthetic travel times. The data variance is assumed to be 2% of $|Zxy\ Zyx|^{1/2}$ for MT inversion and 5 ms for seismic inversion.

4.1. Example I: A Simple Model Inverted with Fixed Moho Interface

As a first test, we design a simple model where the resistivity and velocity of anomalies have exactly similar spatial structure features and proportional parameter values. We apply the workflow to examine how the structure-coupled 3-D joint algorithm behaves when fixing the Moho interface.

The model consists of two prisms within a given background (a 100 Ω·m half-space background for MT model and a layered background for seismic model with velocities ranging from 5.25 km/s to 8.15 km/s), and is very similar to the test model of Moorkamp et al. [2011]. Also, for the seismic model, a fixed undulated Moho boundary is defined here in order to track multiple reflection phases in the layered media. Figure 3 displays a plot of the true model used to generate the theoretical synthetic dataset.

The joint inversion starts from a simple background without two prisms: a 100 Ω·m half space and a layered velocity model (the same as the layered background for seismic model). These models were also used for the a-priori models in the joint inversion. To seek out the final optimal model, we have performed dozens of joint inversions with different cross-gradient weights. Figure 4a shows the convergence curves for the last three iterations of the joint inversion. We can see that within four iterations, the MT inversion has converged to the desired level of misfit, while the seismic inversion requires 12 iterations to achieve the target misfit. When the value of the cross-gradient weight is small (1.0~2.0), both MT and seismic data misfit change slightly due to the loose structural coupling. However, we can obtain a reasonable reduction in misfit for the last iteration when choosing somewhat larger weights (e.g., μ_ρ=2.2~3.0 for MT inversion and μ_s=4.0~4.6 for seismic inversion). When using very high coupling values the data misfit values strongly increase as the inversion only tries to create structurally similar models and ignores the data. These convergence curves help us to find a couple of corresponding cross-gradient weights for the optimal joint inversion model (shown as asterisk in Figure 4a).

We evaluate the misfit between the true model and the joint inverted model when the data misfit displays little difference. This model perturbation RMS is defined as [Moorkamp et al., 2011]:

$$\delta = \sqrt{\sum_{i=1}^{M} \left(\frac{\mathbf{m}_i^{true} - \mathbf{m}_i^{inv}}{\mathbf{m}_i^{true}}\right)^2}. \quad (15)$$

Here $\mathbf{m}_i^{true}$ and $\mathbf{m}_i^{inv}$ denote the model parameters in each cell for the true model and the inversion result respectively, and M is the total number of model parameters. If we choose μ_ρ =3.0 and μ_s =4.5 as the final cross-gradient weights we obtain δ_ρ =2.54 and δ_s =4.28, both of which are the minimum model perturbation RMS among all inverted models in Figure 4. In comparison, the values for the individual inversions are δ_ρ =2.73 and δ_s =4.82, respectively. This illustrates that the model values of the joint inversion are closer to the true values than those of the model from separate inversion.

The data misfit (χ) for the final joint inversion models are compared with separate inversion models varying with iteration (Figure 4b). The MT data misfit decreases from an initial value of 2.635 to 0.9996 at iteration 4. Correspondingly, the seismic misfit diminishes monotonically from 80.58 to 1.171, a value close to the noise level. We see a slower convergence rate in separate inversions, overtaken by joint inversion during the second half of the iterative process.

As shown in Figure 5 (a) and (c), the result of the single MT inversion differs from that of the single seismic inversion. We see that the inversion of MT data alone can recover the general shape and position of the right conductive anomaly, but the left high-resistive prism fails to be delineated in the box. The separate seismic inversion can recover the shape of the two prisms, whereas the values of the seismic velocity in the boxes cannot be fully recovered, especially at a depth of 15~20km. Compared to the separate inversions, we can see that the general shape and values of both anomalies are recovered much better in the joint inversions (Figure 5b and 5d). These jointly reconstructed resistivity and velocity models, to some extent, overcome some of the shortcomings of the separate inversions and thus improve the inversion results.

We also compare the computed values of the cross-gradients for MT and seismic models from the separate inversion with those for the joint inversion models (the third column in Figure 5). For separate inversion, the cross-gradient map shows zones of structural disparity especially near the boundary of anomalies. In comparison, some of the high values of cross-gradients in those zones decrease obviously in the joint inversion. This observation is also confirmed by the comparison of average cross-gradients value for the final models,

$$\bar{\boldsymbol{\tau}} = \frac{1}{M}\sqrt{\sum_{i=1}^{M}\left[(\boldsymbol{\tau}_i^x)^2 + (\boldsymbol{\tau}_i^y)^2 + (\boldsymbol{\tau}_i^z)^2\right]} \tag{16}$$

We obtain the average cross-gradients of $\bar{\boldsymbol{\tau}}=4.31\times10^{-4}$ for the joint inversion, which is smaller than the value ($\bar{\boldsymbol{\tau}}=5.32\times10^{-4}$) for the separate inversion. The results demonstrate that the effective cross-gradients constraint facilitate mutual structural attributes between these models.

4.2. Example II: A Complex Model with Moho Perturbations

We also test the joint inversion on a complex model in order to understand whether MT and seismics can help each other to recover more complex common structure. In this example, the true model consists of a plate-like resistive anomaly on top of a conductive and low-velocity block in the crust (Figure 6).

The a-priori models comprise a 100 Ω·m half-space resistivity model and a layered seismic background model by getting rid of the anomalies. This example is a first step towards a more realistic model, as it is designed based on the typical crustal and upper mantle structure in SE Tibet. The low resistivity anomaly is based on an observed structure which has a typical bulk resistivity of 3 Ω·m in most areas of southern Tibet [Unsworth et al., 2010] and has been interpreted as a partially molten layer. The velocity anomaly amplitude of the reference P wave velocity structure is set to 10% of the background based on results from joint inversion of RFs and Rayleigh wave dispersion [Sun et al., 2014]. Moho depth is usually not known a priori or difficult to determine accurately in some cases, and an inaccurate Moho position might have a negative impact on the inversion of the reflection travel times. Thus in this section, the second aim is to examine how the algorithm behaves if we are inverting for velocity parameters and interface depth simultaneously. Note that an electric Moho is not defined here because it is difficult to resolve in most localities [Jones, 2013]. Our joint inversion approach can be a contributing way to determine the exact electric Moho, but prior to the use it should be primarily established that there is an acceptable expectation of an electric Moho from the MT data.

Before presenting results from the test with Moho perturbations, we first conduct a test with a flat Moho at 40km depth in the starting model for the inversion. However, the results show that the crustal seismic structure is not adequately resolved by the joint inversion. Although crustal velocity at shallow depths (e.g., 0~15 km) is recovered successfully (Figure

7b and 7d, bottom), the velocity anomalies in the lower crust are poorly recovered. In addition, the true shape of the Moho interface seems to be difficult to recover (Figure 8c). Similar synthetic tests for this type of joint inversion (Figure S2) and the comparison to the results of crustal velocity and Moho depth with a fixed starting Moho boundary also confirm this conclusion.

The poor performance can be largely attributed to inadequate path coverage near the Moho interface. For separate tomographic inversion, the use of station terms or active and passive source (i.e., teleseismic events and local earthquake) datasets simultaneously can provide additional constraints on the upper-mantle velocity structure and even Moho interface [Rawlinson and Urvoy, 2006]. However, the solution not only depends on the completeness of datasets, but also suffers from the intrinsic smearing from the regularization which results in underestimating the true amplitudes of velocity structures in most tomography studies [Rawlinson et al., 2010]. Thereby, the estimates of crustal velocity and Moho interface are likely to be conservative just by separate tomographic inversion.

To overcome this shortcoming we simulate 36 broadband seismic stations evenly distributed on the surface (Figure 6a) and use teleseismic RFs from each station to estimate the Moho interface underneath. We found that the lateral offset of the Moho piercing point can be confined to the interface grid spacing of 12 km with an appropriate ray parameter (e.g., <0.04 s/km). This means each broadband station can determine the crustal thickness of the grid independently using H-κ stacking technique [Zhu et al., 2006].

In the synthetic test, a total of 72 RFs (Figure S3) are generated with two ray parameters (0.03 s/km and 0.04 s/km) to simulate the observed data. Gaussian noise with 20% is added to the synthetic RFs directly. The initial reference Moho boundary $\mathbf{H}^0$ (Figure S4) can be estimated $\overline{\mathrm{V}}\mathrm{p}$ using the H-κ stacking technique based on an estimated initial Vp value of 5.4 km/s. During the inversion, the interface position H gets updated and gets closer to the true Moho depth while simultaneously recalculating $\overline{\mathrm{V}}\mathrm{p}$ for each inversion iteration. Also, the updated Moho interface is used when theoretical travel times are computed during the seismic joint inversion scheme (Figure 1).

For the joint inversion in this example, a maximum of 18 iterations for the tomographic inversion scheme are applied to obtain solution models. The cross-gradient weights (μ_ρ=3.7 and μ_s=4.7) are selected for the optimal joint inversion model. We obtain a model misfit of δ_ρ=2.73 and δ_s=7.62 for the joint inversion with the undulating Moho interface, which is smaller than those values (2.95 and 9.88) of the separate inversion.

Figure 7 shows the joint inversion and the separate inversion results for the realistic model inverted with Moho perturbations. We can see that a significant improvement is achieved in the joint inversion (especially for the velocity parameter and interface) constrained by the RFs with an initially undulating Moho interface (the middle column), compared to the other joint inversion case with the initial flat Moho (the left column). As both inversions are run with exactly the same weights, we conclude that the additional constraints on the Moho interface from RFs provide a positive contribution to the joint inversion. Although the separate inversion constrained by the initial RFs (the right column in Figure 7b and Figure 7d) better recovers the crustal velocity structure and Moho interface than the joint inversion with a flat Moho interface, the recovery of the low-velocity anomaly in the lower crust is not successful. In contrast, for the corresponding joint inversion, the modification of $\overline{V}p$ in the joint inversion framework enables us to obtain more accurate Moho depths and thus adequately resolve crustal structures by the combined datasets. Also, the image of the final Moho geometry from joint inversion including the RF module (Figure 8d) is similar to the true Moho interface (Figure 8a). This demonstrates that the combined datasets can resolve crustal structure sufficiently and to a large degree even the trade-off between velocity and Moho interface.

We focus much of the particular discussion on the primary interface (i.e., Moho boundary), but our algorithm is applicable to a wide range of possibilities. For example, the method can be well applied in a realistic multilayered crust when we are able to identify more complicated phases, such as P_1P and P_2P for reflections associated with intra-crustal secondary interfaces (Figure S5). In practice, the upper crustal refraction (Pg) and the Moho reflection (PmP) phases are generally easier to identify than any other arrivals (e.g. P_1, P_2, Pn for refractions, P_1P, P_2P for reflections), but the identification of these later phases may be also feasible by careful picking of arrival times in record sections [Rawlinson and Urvoy, 2006; Karplus, et al. 2011].

4.3. Example III: A Realistic Model in Namche Barwa, SE Tibet

A third test is performed on a new model that is compatible with geological reality. The Namche Barwa region, located in southeastern Tibet (Figure 9), is marked by strong tectonic stress, rapid rock uplift and exhumation, extremely young cooling ages, and intense metamorphism and anatexis [Xu et al., 2012; Zeitler et al., 2014]. The formation mechanism is still debated although several plausible geodynamic models (e.g., indenter corner, crustal folding, and channel flow) have been proposed [Koons, 1995; Burg and Schmalholz, 2008; Jamieson et al., 2004]. Our previous 3D teleseismic tomographic P-wave model showed a complex Indian subduction style with slab fragmentation beneath the eastern Himalaya on a large-scale image [Peng et al., 2016]. MT data around the Namche Barwa was inverted to produce 3D resistivity images and revealed the electrical structure of the crust [Lin et al. 2017].

Although MT and teleseismic data has been measured over the region, no wide-angle reflection/refraction seismic surveys are available in the study area and we are currently not aware of other suitable datasets. To circumvent this issue, we use synthetic data for this test. We can obtain a resistivity model (Figure S6a) from the individual inversion of the MT field data [Lin et al. 2017]. The images exhibit several conductors in the middle crust which are mainly distributed along Indus-Yarlung suture zone (IYSZ) and to the northeast of the NB. These characteristics appear to be prominent and the anomalous region with a resistivity of approximately 1–5 Ω·m has been interpreted as an accumulation of partial melt [Lin et al. 2017]. To highlight the features of these anomalous bodies, the background of the MT model is modified by setting the varying background resistivities ($1.8\ \Omega\cdot m < \log(\rho) < 3.2\ \Omega\cdot m$) to a constant value ($\rho$=310 Ω·m). We then calculate MT synthetic data from the resulting resistivity model (Figure 9c).

Compared with the resistivity model, the teleseismic tomographic results [Peng et al., 2016] show a similar distribution of low-velocity anomalies in the lower crust, but they have inferior lateral resolution to the shallow structures. Therefore, the resistivity model is firstly translated into a velocity model via an empirical relationship [Moorkamp, 2017], $\log\rho = \kappa v - 1$, where κ is set to 1/2500. A reasonable seismic model (Figure 9d) is then designed by mixing the above translated model and the pre-existing teleseismic P-wave model of Namche Barwa (Figure S6b). We use a weighted summation method to combine these two models with the same weighting factor (0.5). As a result, resistivity and velocity images show some common structure characteristics, but some details (e.g., a seismic low velocity anomaly near Motuo as shown in Figure 9d) are not matched. The velocity background is adopted from a 1-D reference model that comprises two layers over a half space with the Moho at 64 km [Zhu and Helmberger, 1996]. The real data of Moho depth (Table S1) is derived from the previous analysis of teleseismic RFs in Namche Barwa [Peng et al., 2017]. The synthetic wide-angle seismic datasets were generated from the realistic seismic model with 8 shots and 50 receivers distributed in the realistic acquisition geometry (Figure 9a).

We apply our joint inversion approach to the synthetic datasets from the realistic model. In this joint inversion test, the MT inversion starts from a 310 Ω·m half-space, and all impedance tensor elements are included at 17 periods between 0.05 s and 364 s. For the seismic inversion, the initial reference Moho boundary $\mathbf{H}^0$ (Figure 11b) is derived from a 1-

degree global model of the Earth's crust (CRUST1.0) [Laske et al., 2013], and the interface **H** is updated by recalculating $\overline{V}p$ for each iteration. The joint inversion finally reaches the desired RMS (χ_ρ =1.886 and χ_s =1.236) when we select the optimal cross-gradient weights (μ_ρ =1.8 and μ_s =5.2).

A significant difficulty when performing joint inversion with real data is the selection of appropriate weights for the datasets, constraints and regularization [e.g. Moorkamp 2017]. By separating the selection procedure into three steps, we have found that we can produce robust results with the joint inversion framework presented here. In the first step, we perform constrained seismic inversions with a fixed conductivity model based on the final individual MT inversion and a series of seismic cross-gradient weights (e.g., μ_s = 2.0, 2.4, 2.8, …, 7.2). This allows us to focus on the seismic inversion and the impact of the cross-gradient on the velocity models. From these experiments we can choose a reasonable cross-gradient weight (e.g., μ_s =5.4) for the optimal seismic joint inversion model. We then repeat the procedure for the MT inversion and use the final velocity model of the constrained inversion from the first step as a constraint to determine the optimal cross-gradient weight for the MT part of the inversion. In the final step, we fixed the MT cross-gradient weight and conduct a mutually coupled MT-seismic inversion with a new series of seismic cross-gradient weights (e.g., μ_s = 5.0, 5.1, 5.2, …, 5.6) in order to fine-tune the seismic part. The resulting weights provide a good balance between structural similarity and fitting the data. Furthermore, the stepwise selection procedure reduces the mutual influence of the different weights which makes assessing the result difficult.

The solution models obtained from the separate inversions are shown in the third row of Figure 10. We observe that the separate MT inversion senses the general presence of the conductive anomalous bodies, but it fails to accurately delineate the boundaries of the anomalies. For example, the two separate low-resistivity anomalous bodies near Bomi (i.e., L3 and L4 at 25 km depth in the true model) merge into one anomaly in the separate MT inversion model as shown in Figure 10a. In contrast, the separate seismic inversion is able to delineate the boundaries of some low-velocity anomalies well (e.g., the anomaly L2), but it is almost insensitive to anomaly L4. This appears to be a blind area of seismic refraction where ray coverage is sparse. In general, the seismic and MT results exhibit different imaging sensitivity to subsurface anomalies.

In the joint inversion, we first perform a test with Moho perturbations and a workflow without RFs. Figure 10 (the left column) shows the resulting images including at 25 km depth and a cross-section. Compared to the separate inversion, the low-resistivity anomalies for the joint inversion are relatively similar in shape, but they are much more intensive in space. Also, the maximum resistivity of the joint inversion model is relatively higher as shown in the cross-section. The recovered velocity attributes of some anomalies (e.g., the anomaly L4) are relatively closer to the true value due to the contributions of the MT data. However, the steep zone in the middle crust associated with the low velocity values from profile position 140 km to position 200 km (Figure 10c) is not clearly resolved by the joint inversion. This can be probably attributed to the unsuccessful recovery of the Moho interface.

To produce a more consistent image of the crustal structure, we apply the coupled joint inversion algorithm with the RF module to the same datasets. In the middle column of Figure 10, we present the resulting MT resistivity and seismic velocity images of the joint inversion models. We can see a clear structural resemblance between these two models. This demonstrates that the two types of parameters are effectively coupled by the algorithm. The improvement in resolution of the resistivity image is clear in comparison with the separate inversion. In the cross-sections in Figure 10c, the low-resistivity anomalies are difficult to distinguish based on the individual inversion results. For example, conductive anomaly L2 is incapable of separating from anomaly L3. In the joint inversion models, by contrast, the low-resistivity anomalies are much more isolated in space and one of these anomalies (e.g., the anomaly L3) is significantly concentrated.

The seismic shallow structure in the upper crust is also better resolved in the joint inversion results. For example, the improvement of the ball-shaped low-velocity anomaly L4 is visible compared to the separate inversion. In addition, a minor low-velocity anomaly (i.e., anomaly L3) in the upper ~20 km emerges in the image as shown in the middle column of Figure 10d. Overall, we achieve a better recovery of the geometry of the low-resistivity anomalies and the distribution of the velocity values.

Because the interface perturbation in this test is smaller than in the previous example, we cannot observe a distinct improvement of the estimate of Moho interface depth from the cross-section images (Figure 10c). For the purpose of comparing the Moho interfaces, we therefore investigate maps of Moho depth variation for the example. We can see that an accurate Moho interface is achieved by the coupled joint inversion algorithm with the RF module (Figure 11d). This Moho interface retrieved by joint inversion is very similar to the true Moho interface (Figure 11a). Note that although we obtain a seemingly satisfactory result of the Moho structure by only jointly inverting the MT and seismic data (Figure 11c) with a prior Moho estimate, the recovery of crustal anomalies is impacted. This highlights the importance of the teleseismic RF datasets in the accurate recovery of crustal structure.

5 Conclusions

We have developed a new algorithm for 3-D structure-coupled joint inversion of MT and wide-angle seismic reflection/refraction data by assembling pre-existing MT and seismic refraction travel-time tomography schemes and incorporating reflectors into the framework. Our results show that an accurate estimation of the position of the Moho interface is critical for this type of joint inversion. We therefore present two strategies to solve the problem of accurately determining velocity parameters and Moho depth simultaneously. One strategy is to independently acquire an accurate Moho interface or layered velocity model, for example by 1-D joint inversion of RFs and surface wave dispersion. However, in these cases the inversion results strongly depend on the quality of this prior information. Thus our preferred approach is to adjust the average velocity and position of the Moho interface within the joint inversion and utilize teleseismic receiver functions to update these estimates.

Analysis of synthetic test results reveals that the joint inversion improves the inversion results in comparison with the separate inversion. Also, the combined datasets can resolve crustal structure sufficiently and even, to a large degree, the trade-off between velocity and Moho interface. For the realistic examples, we can conclude that the inclusion of teleseismic RF data can effectively decrease the non-uniqueness for the 3-D coupled MT/seimic joint inversion. The Moho interfaces need to be updated by H-κ stacking with estimates of average V_P from the tomography. Then the new V_P that can be used to update the crustal velocity model and the interface during the seismic inversion. This significantly contributes to a more accurate estimation of Moho depth.

The new algorithm can be used as a tool for integrated imaging of crustal structures including the distribution of resistivity and velocity parameters, the Moho interface and the secondary crustal discontinuities. With suitable experiments in the future, we envisage significant improvements in imaging such structures within the Earth.

Acknowledgments

This study is jointly supported by research grants from the National Natural Science Foundation of China (41374078, 41804130), the State Key Program of National Natural Science Foundation of China (41830429), and the Natural Science Foundation of Beijing, China (8192041). We thank associate editor M. Savage (Victoria University of Wellington) and reviewers L. A. Gallardo (Earth Science Division, CICESE) and M. Jegen (Geomar) for valuable comments that helped us greatly improve the manuscript. Professor N. Rawlinson and W. Siripunvarapon are thanked for providing, respectively, the FMTOMO and

WSINV3DMT codes. Most of the figures in the paper are plotted using the Generic Mapping Tools (GMT) of Wessel and Smith [1998]. Supplementary data for generating the 3D realistic crustal models of Namche Barwa in SE Tibet is freely available accompanying this paper as supporting information.

Appendix A

To minimize the objective function for MT joint inversion in Eq. (9), we evaluate the partial derivatives of Ψ_{MT} and yield a series of model updates from an initial model:

$$\begin{aligned}\mathbf{m}_{\rho}^{k+1}-\mathbf{m}_{\rho}^{0}=(\lambda\mathbf{C}_{\mathbf{m}}^{-1}+\mathbf{J}_{k}^{T}\mathbf{C}_{\mathbf{d}}^{-1}\mathbf{J}_{k})^{-1}\left\{\mathbf{J}_{k}^{T}\mathbf{C}_{\mathbf{d}}^{-1}\mathbf{X}_{k}-\right.\\ \left.\mu_{\rho}(\mathbf{B}_{\rho}^{k})^{T}\mathbf{C}_{\tau}^{-1}[\boldsymbol{\tau}(\mathbf{m}_{\rho}^{k},\tilde{\mathbf{m}}_{s}^{k})-\boldsymbol{\tau}(\mathbf{m}_{\rho}^{0},\mathbf{m}_{s}^{0})]\right\}\end{aligned} \tag{A1}$$

As presented by Siripunvaraporn et al. [2005], and summarized in Eq. (5) and Eq. (6), the normal equation in the model space is replaced by a system in the data-space:

$$\boldsymbol{\Lambda}_{k}^{-1}\mathbf{J}_{k}^{T}\mathbf{C}_{\mathbf{d}}^{-1}=\mathbf{C}_{\mathbf{m}}\mathbf{J}_{k}^{T}\boldsymbol{\Gamma}_{k}^{-1} \tag{A2}$$

Here, $\boldsymbol{\Lambda}_{k}=\lambda\mathbf{C}_{\mathbf{m}}^{-1}+\mathbf{J}_{k}^{T}\mathbf{C}_{\mathbf{d}}^{-1}\mathbf{J}_{k}$ is an M×M positive symmetric matrix in the model space, and $\boldsymbol{\Gamma}_{k}=\lambda\mathbf{C}_{\mathbf{d}}+\mathbf{J}_{k}\mathbf{C}_{\mathbf{m}}\mathbf{J}_{k}^{T}$ is an N×N positive symmetric matrix in the data-space.

In order to apply the data-space approach to joint inversion, we need to solve the solution of $\boldsymbol{\Lambda}_{k}^{-1}$. This M×M inverse matrix can be transformed into data-space by Woodbury matrix identity and it follows that:

$$\begin{aligned}\boldsymbol{\Lambda}_{k}^{-1}&=(\lambda\mathbf{C}_{\mathbf{m}}^{-1}+\mathbf{J}_{k}^{T}\mathbf{C}_{\mathbf{d}}^{-1}\mathbf{J}_{k})^{-1}\\ &=\frac{1}{\lambda}\mathbf{C}_{\mathbf{m}}-\frac{1}{\lambda}\mathbf{C}_{\mathbf{m}}\mathbf{J}_{k}^{T}\left[\mathbf{C}_{\mathbf{d}}+\mathbf{J}_{k}(\frac{1}{\lambda}\mathbf{C}_{\mathbf{m}})\mathbf{J}_{k}^{T}\right]^{-1}\mathbf{J}_{k}(\frac{1}{\lambda}\mathbf{C}_{\mathbf{m}})\\ &=\frac{1}{\lambda}\mathbf{C}_{\mathbf{m}}-\mathbf{C}_{\mathbf{m}}\mathbf{J}_{k}^{T}\left(\lambda\mathbf{C}_{\mathbf{d}}+\mathbf{J}_{k}\mathbf{C}_{\mathbf{m}}\mathbf{J}_{k}^{T}\right)^{-1}\mathbf{J}_{k}(\frac{1}{\lambda}\mathbf{C}_{\mathbf{m}})\\ &=\frac{1}{\lambda}\left(\mathbf{I}-\mathbf{C}_{\mathbf{m}}\mathbf{J}_{k}^{T}\boldsymbol{\Gamma}_{k}^{-1}\mathbf{J}_{k}\right)\mathbf{C}_{\mathbf{m}}\end{aligned} \tag{A3}$$

where $\mathbf{I}$ is the identity matrix, and by substituting Eq. (A2, A3) into Eq. (A1), the (k+1)-th model update can be expressed as:

$$\mathbf{m}_{\rho}^{k+1}=\mathbf{C}_{\mathbf{m}}\mathbf{J}_{k}^{T}\boldsymbol{\Gamma}_{k}^{-1}\mathbf{X}_{k}-\mu_{\rho}\boldsymbol{\Lambda}_{k}^{-1}\left[(\mathbf{B}_{\rho}^{k})^{T}\mathbf{C}_{\tau}^{-1}\mathbf{B}_{\rho}^{k}(\mathbf{m}_{\rho}^{k}-\mathbf{m}_{\rho}^{0})\right]+\mathbf{m}_{\rho}^{0} \tag{A4}$$

Appendix B

The subspace inversion method requires the partial derivatives of objective function Ψ_{Seis} at a specified point in model space and it satisfies a basic assumption that Ψ_{Seis} is adequately smooth to validate a locally quadratic approximation about some current model:

$$\Psi_{\text{Seis}}(\mathbf{m}_s+\delta\mathbf{m}_s)\approx\Psi_{\text{Seis}}(\mathbf{m}_s)+\boldsymbol{\gamma}\delta\mathbf{m}_s+\frac{1}{2}\delta\mathbf{m}_s{}^T\mathbf{H}\delta\mathbf{m}_s \quad \text{(B1)}$$

where $\delta\mathbf{m}_s$ is a perturbation to the current model and $\boldsymbol{\gamma}=\partial\Psi_{\text{Seis}}/\partial\mathbf{m}_s$ and $\mathbf{H}=\partial^2\Psi_{\text{Seis}}/\partial\mathbf{m}_s^2$ are the gradient vector and Hessian matrix respectively.

Disturbance $\delta\mathbf{m}_s$ is composed of a series of multi-dimensional base vectors:

$$\delta\mathbf{m}_s=\sum_{j=1}^{N}\mu_j a^j=\mathbf{A}\mu \quad \text{(B2)}$$

where, $\mathbf{A}$ is a projection matrix and the weighting coefficient μ_j is the length of corresponding base vector a^j that minimises the quadratic form of Ψ_{Seis}.

In order to determine the coefficient μ_j, the following Eq. (B3) in a summation form can be obtained by substituting Eq. (B2) into Eq. (B1):

$$\Psi_{\text{Seis}}(\mathbf{m}_s+\delta\mathbf{m}_s)=\phi(\mathbf{m}_s)+\sum_{j=1}^{n}\mu_j\boldsymbol{\gamma}^T a^j+\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}\mu_j\mu_k[a^k]^T\mathbf{H}[a^k] \quad \text{(B3)}$$

The partial derivative of Eq. (B3) with respect to μ is:

$$\frac{\partial\phi(\mathbf{m}_s)}{\partial\mu_q}=\boldsymbol{\gamma}^T a^q+\sum_{k=1}^{N}\mu_k[a^k]^T\mathbf{H}[a^k]=0 \quad \text{(B4)}$$

Rearranging Eq. (B4) for μ gives:

$$\mu=-[\mathbf{A}^T\mathbf{H}\mathbf{A}]^{-1}\mathbf{A}^T\boldsymbol{\gamma} \quad \text{(B5)}$$

Evaluating partial derivatives $\boldsymbol{\gamma}$ and $\mathbf{H}$ in Eq. (10) gives:

$$\boldsymbol{\gamma}=\mathbf{G}^T\mathbf{C}_{\mathbf{d}}^{-1}\left[\mathbf{d}_s-g(\mathbf{m}_s)\right]+(\varepsilon\mathbf{C}_{\mathbf{m}}^{-1}+\mu_s\mathbf{B}_s^T\mathbf{C}_\tau^{-1}\mathbf{B}_s)\cdot(\mathbf{m}_s^0-\mathbf{m}_s) \quad \text{(B6)}$$

$$\mathbf{H}=\mathbf{G}^T\mathbf{C}_{\mathbf{d}}^{-1}\mathbf{G}+\varepsilon\mathbf{C}_{\mathbf{m}}^{-1}+\mu_s\mathbf{B}_s^T\mathbf{C}_\tau^{-1}\mathbf{B}_s \quad \text{(B7)}$$

and since $\delta\mathbf{m}=\mathbf{A}\mu$ in Eq. (B2), the solution is:

$$\delta\mathbf{m}_s=-\mathbf{A}[\mathbf{A}^T(\mathbf{G}^T\mathbf{C}_{\mathbf{d}}^{-1}\mathbf{G}+\varepsilon\mathbf{C}_{\mathbf{m}}^{-1}+\mu_s\mathbf{B}_s^T\mathbf{C}_\tau^{-1}\mathbf{B}_s)\mathbf{A}]^{-1}\mathbf{A}^T\boldsymbol{\gamma} \quad \text{(B8)}$$

References

Afonso, J. C., J. Fullea, W. L. Griffin, Y. Yang, A. G. Jones, J. A. D. Connolly, et al. (2013), 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, *Journal of Geophysical Research: Solid Earth*, 118(5), 2586-2617, doi:10.1002/jgrb.50124.

Archie, G. E. (1942), The electrical resistivity log as an aid in determining some reservoir characteristics, AIME Trans. (Petroleum Development and Technology) 146, 54-62.

Bennington, N. L., H. Zhang, C. H. Thurber, and P. A. Bedrosian (2015), Joint Inversion of Seismic and Magnetotelluric Data in the Parkfield Region of California Using the Normalized Cross-Gradient Constraint, *Pure and Applied Geophysics*, 172(5), 1033–1052, doi:10.1007/s00024-014-1002-9.

Berryman, J. G., P. A. Berge, and B. P. Bonner (2002), Estimating rock porosity and fluid saturation using only seismic velocities, *Geophysics*, 67(2), 391–404, doi:10.1190/1.1468599.

Burg, J. P., and S. M. Schmalholz (2008), Viscous heating allows thrusting to overcome crustalscale buckling: numerical investigation with application to the Himalayan syntaxes, Earth and Planetary Science Letters, 274, 189–203.

Chave, A. D., and A. G. Jones (2012), The magnetotelluric method: theory and practice, Cambridge University Press, Cambridge, 1–884, doi:10.1017/CBO9781139020138.

Colombo, D., and M. D. Stefano (2007), Geophysical modeling via simultaneous joint inversion of seismic, gravity, and electromagnetic data: Application to prestack depth imaging, *Leading Edge*, 26(3), 326–331, doi:10.1190/1.2715057.

Constable, S. C., R. L. Parker, and C. G. Constable (1987), Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, 52(3), 289–300, doi:10.1190/1.1442303.

de Groot-Hedlin, C., and S. Constable (1990), Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data, *Geophysics*, 55(55), 1613–1624, doi:10.1190/1.1442813.

de Kool, M., N. Rawlinson, and M. Sambridge (2006), A practical grid-based method for tracking multiple refraction and reflection phases in three-dimensional heterogeneous media, *Geophysical Journal International*, 167(1), 253–270, doi:10.1111/j.1365-246X.2006.03078.x.

Doetsch, J., N. Linde, I. Coscia, S. A. Greenhalgh, and A. G. Green (2010), Zonation for 3D aquifer characterization based on joint inversions of multimethod crosshole geophysical data, *Geophysics*, 75(6), G53–G64, doi:10.1190/1.3496476.

Egbert, G. D., and A. Kelbert (2012), Computational recipes for electromagnetic inverse problems, *Geophysical Journal International*, 189(1), 251-267, doi:10.1111/j.1365-246X.2011.05347.x.

Fregoso, E., and L. A. Gallardo (2009), Cross-gradients joint 3D inversion with applications to gravity and magnetic data, *Geophysics*, 74(4), L31–L42, doi:10.1190/1.3119263.

Gallardo, L. A. (2007), Multiple cross-gradient joint inversion for geospectral imaging, *Geophysical Research Letters*, 34, L19301, doi:10.1029/2007gl030409.

Gallardo, L. A., and M. A. Meju (2003), Characterization of heterogeneous near-surface materials by joint 2D inversion of dc resistivity and seismic data, *Geophysical Research Letters*, 30(13), 1658–1658, doi:10.1029/2003GL017370.

Gallardo, L. A., and M. A. Meju (2004), Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints, *Journal of Geophysical Research Solid Earth*, 109, B03311, doi:10.1029/2003JB002716.

Gallardo, L. A., and M. A. Meju (2011), Structure-coupled multiphysics imaging in geophysical sciences, *Reviews of Geophysics*, 49, RG1003, doi:10.1029/2010RG000330.

Gallardo, L. A., S. L. Fontes, M. A. Meju, M. P. Buonora, and P. P. D. Lugao (2012), Robust geophysical integration through structure-coupled joint inversion and multispectral fusion of seismic reflection, magnetotelluric, magnetic, and gravity images: Example from Santos Basin, offshore Brazil, *Geophysics*, 77(5), 237-251, doi:10.1190/geo2011-0394.1.

Haber, E., and D. Oldenburg (1997), Joint inversion: a structural approach, *Inverse Problems*, 13(1), 63, doi:10.1088/0266-5611/13/1/006.

Heincke, B., M. Jegen, and R. Hobbs (2006), Joint inversion of MT, gravity and seismic data applied to sub-basalt imaging, in *SEG Technical Program Expanded Abstracts*, pp. 770–775, doi: 10.1190/1.2370374.

Heincke, B., M. Jegen, M. Moorkamp, R. W. Hobbs, and C. Jin (2017), An adaptive coupling strategy for joint inversions that use petrophysical information as constraints, *Journal of Applied Geophysics*, 136(1), 279-297, doi:10.1016/j.jappgeo.2016.10.028.

Hu, W., A. Abubakar, and T. M. Habashy (2009), Joint electromagnetic and seismic inversion using structural constraints, *Geophysics*, 74(6), R99–R109, doi:10.1190/1.3246586.

Jamieson, R. A., C. Beaumont, S. Medvedev, and M. H. Nguyen (2004), Crustal channel flows: 2. Numerical models with implications for metamorphism in the Himalayan–Tibetan orogen, Journal of Geophysical Research, 109, B06407, http://dx.doi.org/10.1029/2003JB002811.

Jegen, M. D., R. W. Hobbs, P. Tarits, and A. Chave (2009), Joint inversion of marine magnetotelluric and gravity data incorporating seismic constraints: Preliminary results of sub-basalt imaging off the Faroe Shelf, *Earth & Planetary Science Letters*, 282(1), 47–55, doi:10.1016/j.epsl.2009.02.018.

Jones, A. G. (1999), Imaging the continental upper mantle using electromagnetic methods, *Lithos*, 48(1–4), 57–80, doi:10.1016/S0024-4937(99)00022-5.

Jones, A. G. (2013), Imaging and observing the electrical Moho, Tectonophysics, 609, 423–436, doi:10.1016/j.tecto.2013.02.025

Karplus, M. S.,W. Zhao, S. L. Klemperer, Z. Wu, J. Mechie, D. Shi, et al. (2011), Injection of Tibetan crust beneath the south Qaidam Basin: Evidence from INDEPTH IV wide-angle seismic data, *Journal of Geophysical Research Solid Earth*, 116(B07301), doi:10.1029/2010JB007911.

Kennett, B. L. N., M. S. Sambridge, and P. R. Williamson (1988), Subspace methods for large inverse problems with multiple parameter classes, *Geophysical Journal International*, 94(2), 237–247, doi:10.1111/j.1365-246X.1988.tb05898.x.

Koons, P. O. (1995), Modeling the topographic evolution of collisional mountain belts, Annual Review of Earth and Planetary Science, 23, 375–408, doi:10.1146/annurev.ea.23.050195.002111.

Laske, G., G. Masters, Z. Ma, and M. Pasyanos (2013), Update on CRUST1.0 - A 1-degree global model of Earth's crust, paper presented at Egu General Assembly Conference. Geophys. Res. Abs. 15, EGU2013–2658.

Lin, C., M. Peng, H. Tan, Z. Xu, Z. H. Li, W. Kong, et al. (2017), Crustal structure beneath Namche Barwa, eastern Himalayan syntaxis: New insights from three-dimensional magnetotelluric imaging, *Journal of Geophysical Research Solid Earth*, 122(7), 5082–5100, doi:10.1002/2016JB013825.

Linde, N., A. Binley, A. Tryggvason, L. B. Pedersen, A. Revil, and Eacute (2006), Improved hydrogeophysical characterization using joint inversion of cross-hole electrical resistance and ground-penetrating radar traveltime data, *Water Resources Research*, 42, W12404, doi:10.1029/2006WR005131.

Mackie, R. L., T. R. Madden, and P. E. Wannamaker (1993), Three-dimensional magnetotelluric modeling using difference equations—Theory and comparisons to integral equation solutions, *Geophysics*, 58(2), 215–226, doi:10.1190/1.1443407.

Mackie, R.L., T.R. Madden, and P.E. Wannamaker (1993). Three-dimensional magnetotelluric modeling using difference equations: Theory and comparisons to integral equation solutions. *Geophysics*, 58(2), 215–226, doi: 10.1190/1.1443407

Meju, M. A., and L. A. Gallardo (2016), Structural coupling approaches in integrated geophysical imaging. In M. Moorkamp, P. G. Lelièvre, N. Linde, A. Khan (Eds.), Integrated Imaging of the Earth: Theory and Applications, Geophysical Monograph Series (Vol. 218, pp. 49–67), Washington, DC: American Geophysical Union.

Meju, M. A., L. A. Gallardo, and A. K. Mohamed (2003), Evidence for correlation of electrical resistivity and seismic velocity in heterogeneous near-surface materials, *Geophysical Research Letters*, 30(7), 1373, doi:10.1029/2002GL016048.

Moorkamp, M. (2017), Integrating electromagnetic data with other geophysical observations for enhanced imaging of the Earth: A tutorial and review, *Surveys in Geophysics*, 38(5), 935-962, doi:10.1007/s10712-017-9413-7.

Moorkamp, M., A. G. Jones, and S. Fishwick (2010), Joint inversion of receiver functions, surface wave dispersion, and magnetotelluric data, *Journal of Geophysical Research Solid Earth*, 115, B04318, doi:10.1029/2009JB006369.

Moorkamp, M., A. W. Roberts, M. Jegen, B. Heincke, and R. W. Hobbs (2013), Verification of velocity-resistivity relationships derived from structural joint inversion with borehole data, *Geophysical Research Letters*, 40(14), 3596-3601, doi:10.1002/grl.50696.

Moorkamp, M., B. Heincke, M. Jegen, A. W. Roberts, and R. W. Hobbs (2011), A framework for 3-D joint inversion of MT, gravity and seismic refraction data, *Geophysical Journal International*, 184(1), 477–493, doi:10.1111/j.1365-246X.2010.04856.x.

Moorkamp, M., P. G. Lelièvre, N. Linde, and A. Khan (2016), Integrated Imaging of the Earth: Theory and Applications, *Geophysical Monograph Series*, 218, Washington, DC: American Geophysical Union.

Peng, M., H. D. Tan, M. Jiang, H. Qian, and J. Y. Tan (2013), Three-dimensional joint inversion of magnetotelluric and seismic travel time data with cross-gradient constraints, *Chinese Journal of Geophysics*, 56(8), 2728–2738, doi:10.6038/cjg20130821.

Peng, M., H. D. Tan, M. Jiang, W. Wang, Q. Q. Li, and L. S. Zhang (2012), Joint inversion of receiver functions and magnetotelluric data: Application to crustal and mantle structure beneath central Namche Barwa, eastern Himalayan syntaxis, *Chinese Journal of Geophysics*, 55(7), 2281–2291, doi:10.6038/j.issn.0001-5733.2012.07.014.

Peng, M., M. Jiang, Y. L. Chen, H. D. Tan, L. I. Qing Qing, L. S. Zhang, et al. (2017), Crustal structure under the eastern Himalayan Syntaxis seismic array and its geodynamic implications derived from receiver functions, *Chinese Journal of Geophysics*, doi:10.6038/cjg20170107.

Peng, M., M. Jiang, Z. H. Li, Z. Xu, L. Zhu, W. Chan, et al. (2016), Complex Indian subduction style with slab fragmentation beneath the Eastern Himalayan Syntaxis revealed by teleseismic P-wave tomography, *Tectonophysics*, 667, 77-86, doi:10.1016/j.tecto.2015.11.012.

Rawlinson, N., and M. Sambridge (2004), Wave front evolution in strongly heterogeneous layered media using the fast marching method, *Geophysical Journal International*, 156(3), 631–647, doi:10.1111/j.1365-246X.2004.02153.x.

Rawlinson, N., and M. Sambridge (2005), The fast marching method: an effective tool for tomographic imaging and tracking multiple phases in complex layered media, *Exploration Geophysics*, 36(4), 341–350, doi:10.1071/EG05341.

Rawlinson, N., and M. Urvoy (2006), Simultaneous inversion of active and passive source datasets for 3-D seismic structure with application to Tasmania, *Geophysical Research Letters*, 33, L24313, doi:10.1029/2006GL028105.

Rawlinson, N., G. A. Houseman, and C. D. N. Collins (2001), Inversion of seismic refraction and wide-angle reflection traveltimes for 3-D layered crustal structure, *Geophys. J. Int.*, 145, 381–401, doi:10.1046/j.1365-246x.2001.01383.x

Rawlinson, N., S. Pozgay, and S. Fishwick (2010), Seismic tomography: A window into deep Earth, *Physics of the Earth & Planetary Interiors*, 178(3), 101–135, doi:10.1016/j.pepi.2009.10.002.

Saad, Y. (1990), Sparsekit: a basic tool kit for sparse matrix computations, Technical Report, pp. 1–27. Moffett Field, CA: Advanced Computer Science. Retrieved from https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19910023551.pdf

Sambridge, M. S., and B. L. N. Kennett (1990), Boundary value ray tracing in a heterogeneous medium: a simple and versatile algorithm, *Geophysical Journal International*, 101(1), 157–168, doi:10.1111/j.1365-246X.1990.tb00765.x.

Siripunvaraporn, W., and G. Egbert (2000), An efficient data-subspace inversion method for 2-D magnetotelluric data, *Geophysics*, 65, 791–803, doi:10.1190/1.1444778.

Siripunvaraporn, W., and G. Egbert (2009), WSINV3DMT: Vertical magnetic field transfer function inversion and parallel implementation, *Physics of the Earth & Planetary Interiors*, 173(3), 317-329, doi:10.1016/j.pepi.2009.01.013.

Siripunvaraporn, W., G. Egbert, Y. Lenbury, and M. Uyeshima (2005), Three-dimensional magnetotelluric inversion: data-space method, *Physics of the Earth & Planetary Interiors*, 150(1), 3–14, doi:10.1016/j.pepi.2004.08.023.

Stanley, W. D., W. D. Mooney, and G. S. Fuis (1990), Deep crustal structure of the Cascade Range and surrounding regions from seismic refraction and magnetotelluric data, *Journal of Geophysical Research Solid Earth*, 95(B12), 19419–19438, doi:10.1029/JB095iB12p19419.

Sun, X., X. Bao, M. Xu, D. W. Eaton, X. Song, L. Wang, et al. (2014), Crustal structure beneath SE Tibet from joint analysis of receiver functions and Rayleigh wave dispersion, *Geophysical Research Letters*, 41(5), 1479–1484, doi:10.1002/2014GL059269.

Tan, H., Q. Yu, J. Booker, and W. Wei (2003), Three-Dimensional Magnetotelluric Modeling using the Staggered-Grid Finite Difference Method, *Chinese Journal of Geophysics*, 46(5), 705–711, doi:10.1002/cjg2.420.

Tan, H., T. Tong, and C. Lin (2006), The parallel 3D magnetotelluric forward modeling algorithm, *Appl. Geophys.*, 3(4), 197–202, doi:10.1007/s11770-006-4001-5.

Um, E. S., M. Commer, and G. A. Newman (2014), A strategy for coupled 3D imaging of large-scale seismic and electromagnetic data sets: Application to subsalt imaging, *Geophysics*, 79(3), ID1-ID13, doi:10.1190/geo2013-0053.1.

Unsworth, M. (2010), Magnetotelluric Studies of Active Continent–Continent Collisions, *Surveys in Geophysics*, 31(2), 137–161, doi:10.1007/s10712-009-9086-y.

Wessel, P., and Smith, W. H. F. (1998), New, improved version of the generic mapping tools released, EOS Trans. AGU, 79(47), 579–579.

Wölbern, I., and G. Rümpker (2017), Limitations of H-κ stacking: ambiguous results caused by crustal layering, *Journal of Seismology*, 21(1), 221-235, doi:10.1007/s10950-016-9599-z.

Wyllie, M. R. J., A. R. Gregory, and L. W. Gardner (1956), Elastic Wave Velocities in Heterogeneous and Porous Media, *Geophysics*, 21(1), 41–70, doi:10.1190/1.1438217.

Xu, Z. Q., S. C. Ji, Z. H. Cai, L. S. Zeng, Q. R. Geng, and H. Cao (2012), Kinematics and dynamics of the Namche Barwa Syntaxis, eastern Himalaya: constraints from deformation, fabrics and geochronology, Gondwana Res., 21(1), 19–36, doi:10.1016/j.gr.2011.06.010.

Zeitler, P. K., A. S. Meltzer, L. Brown, W. S. F. Kidd, C. Lim, and E. Enkelmann (2014), Tectonics and topographic evolution of Namche Barwa and the easternmost Lhasa Block, Tibet, Special Paper of the Geological Society of America, 507, 23-58, doi:10.1130/2014.2507(02).

Zelt, C. A., and R. B. Smith (1992), Seismic traveltime inversion for 2-D crustal velocity structure, *Geophysical Journal International*, 108(1), 16–34, doi:10.1111/j.1365-246X.1992.tb00836.x.

Zhang, J., and F. D. Morgan (1997), Joint Seismic and Electrical Tomography, in *Symposium on the Application of Geophysics to Engineering and Environmental Problems*, edited, pp. 391–396, doi:10.4133/1.2922412.

Zhu, L., and D. V. Helmberger (1996), Intermediate depth earthquakes beneath the India-Tibet Collision Zone, *Geophysical Research Letters*, 23(5), 435-438, doi:10.1029/96gl00385.

Zhu, L., and H. Kanamori (2000), Moho depth variation in southern California from teleseismic receiver functions, *Journal of Geophysical Research Solid Earth*, 105(B2), 2969-2980, doi:10.1029/1999JB900322.

Zhu, L., B. J. Mitchell, N. Akyol, I. Cemen, and K. Kekovali (2006), Crustal thickness variations in the Aegean region and implications for the extension of continental crust, *Journal of Geophysical Research: Solid Earth*, 111(B1), doi:10.1029/2005jb003770.

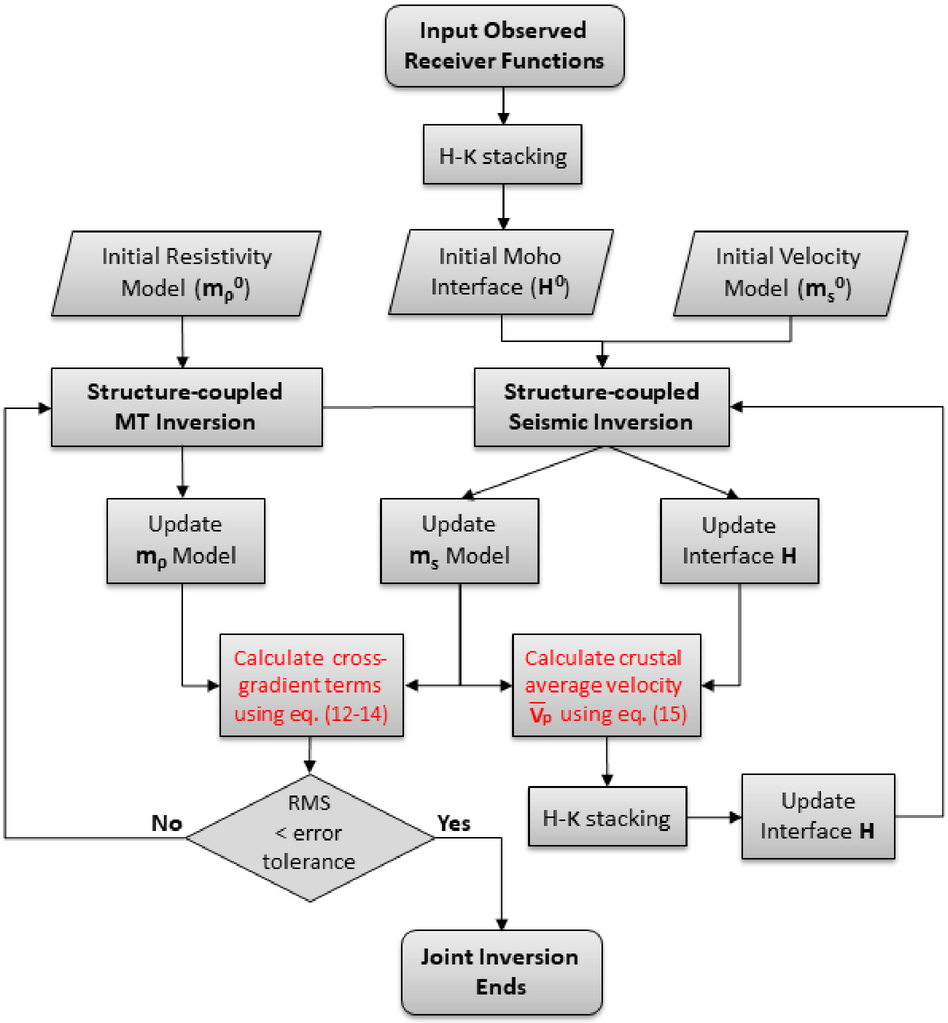

Figure 1. Flow chart for the joint inversion algorithm including the estimation of the Moho interface through H-k stacking of RFs

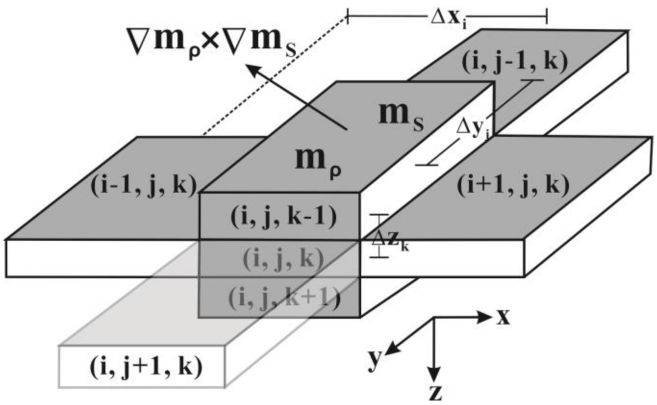

Figure 2. Diagram of 3-D grid discretization for cross-gradient.

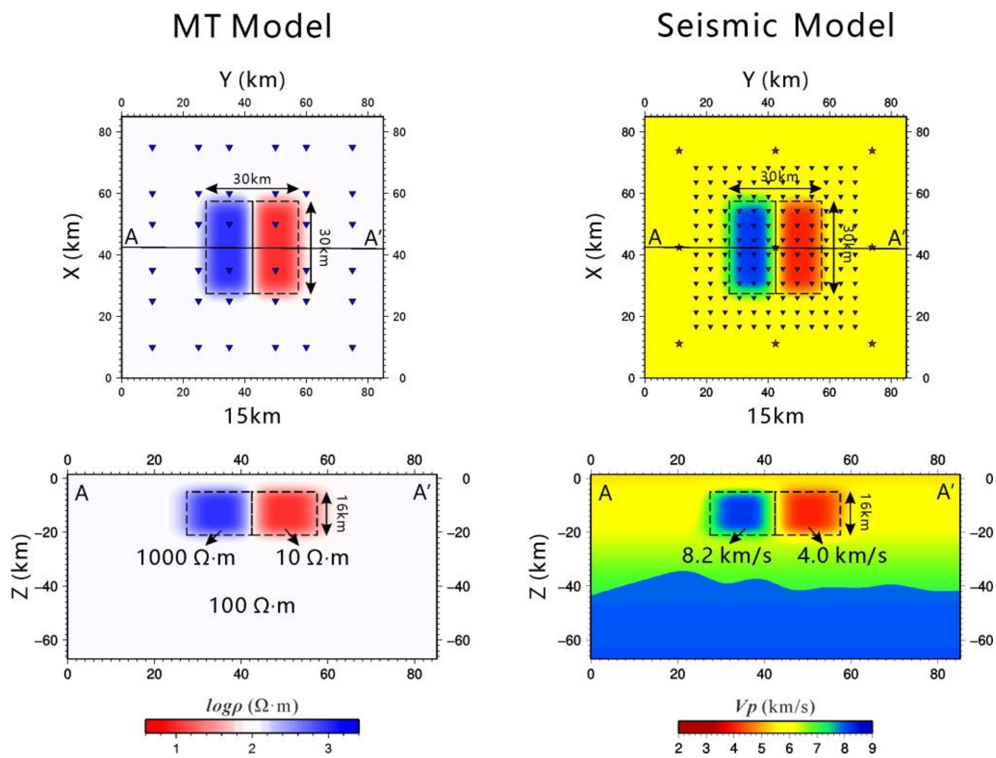

Figure 3. Sections of true model in synthetic example I, consisting of two prisms (8 km burial depth) buried in a background. The top panel is a plan section at 15 km depth and the bottom panel is a cross-section (AA') cutting across the anomalies along E-W direction. The red stars denote sources and blue triangles denote seismic receivers or MT sites.

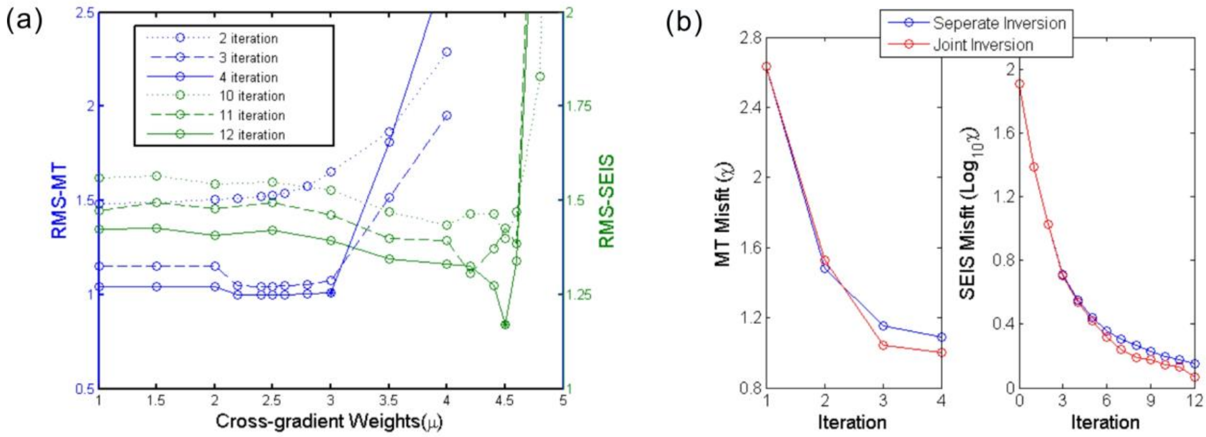

Figure 4. Convergence curves for the last three iterations of joint inversion versus cross-gradient weights (a) and for the final inverted model versus iterations (b). Both MT and seismic joint inversion exhibit faster convergence than the separate inversion. The definition of RMS can be found in the supporting information (Text S1)

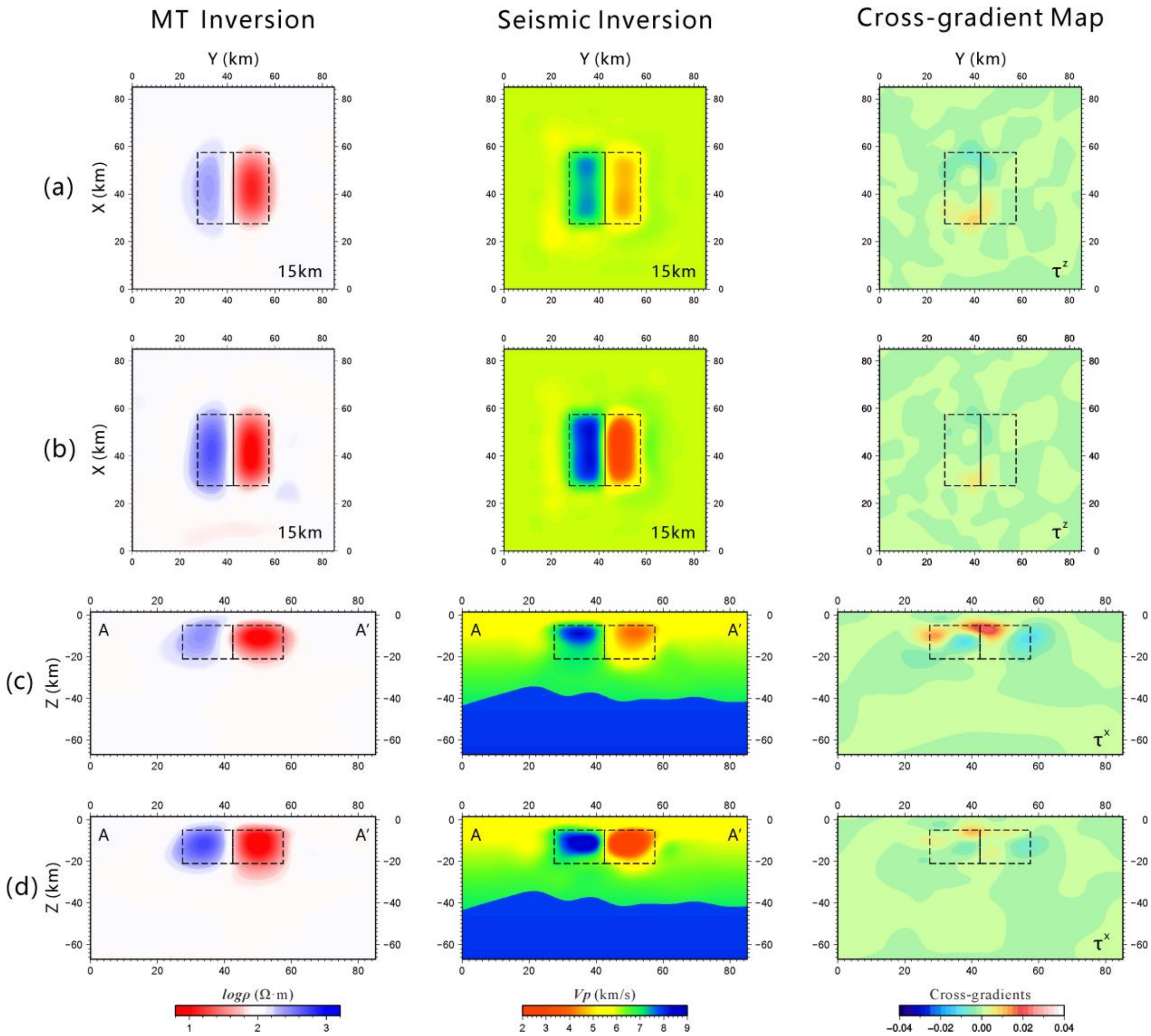

Figure 5. Imaging results from the separate inversion and joint inversion of synthetic MT and wide-angle reflection/refraction data for a compatible model. The top panels (a)-(b) is a plan section at 15 km depth and the bottom panels (c)-(d) is the cross-section (AA') shown in Figure 3. The separate inversion results are displayed in (a) and (c) and the joint inversion results are shown in (b) and (d). Cross-gradient maps shown in the third column are plotted to examine the cross-gradients constraint and structural similarity between resistivity model and velocity model.

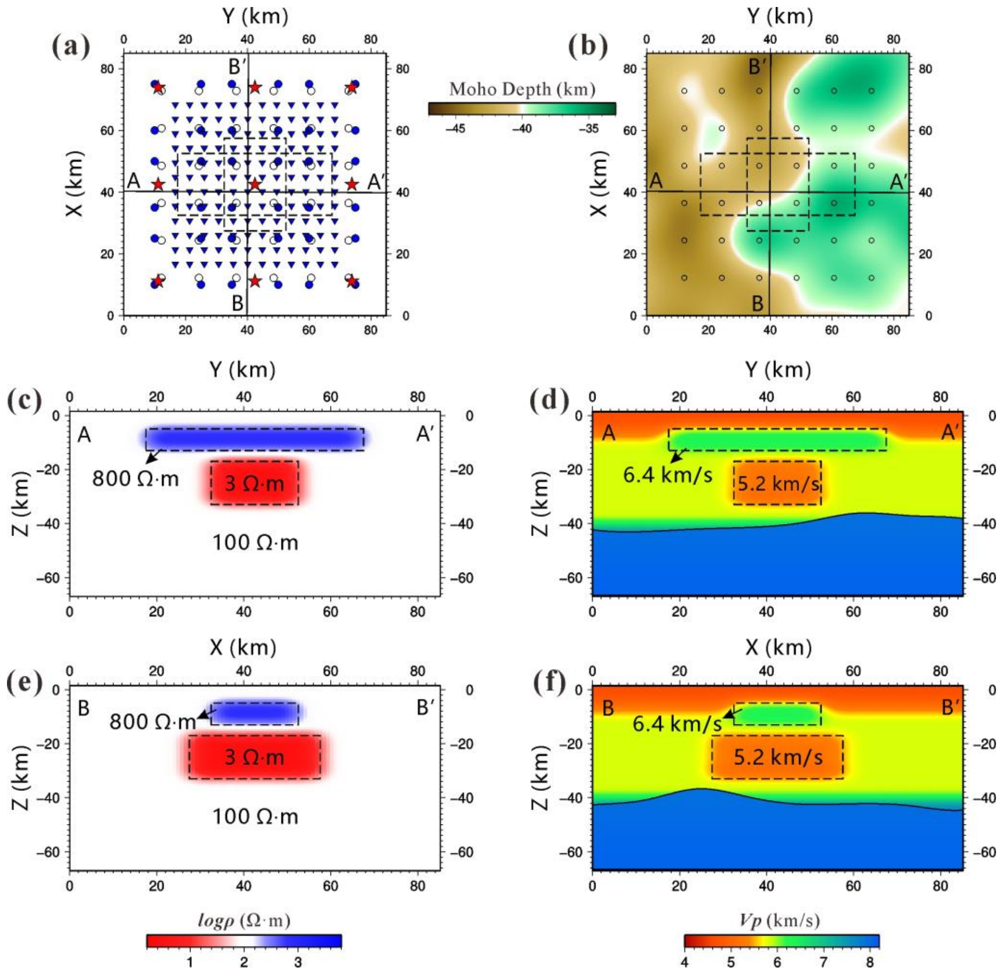

Figure 6. Sections of true model for a realistic model test in synthetic example II, consisting of a thin layer and an underlying prism buried in a background. The top panel (a) is a plan section on the surface showing MT sites (solid blue circle), broadband seismic stations (black circle), active sources receivers (blue triangles) and the sources (red stars) respectively; the top panel (b) shows a random Moho structure for the synthetic test; the middle panels (c)-(d) and the bottom panels (e)-(f) are cross-sections AA’ and BB’ along E-W direction respectively.

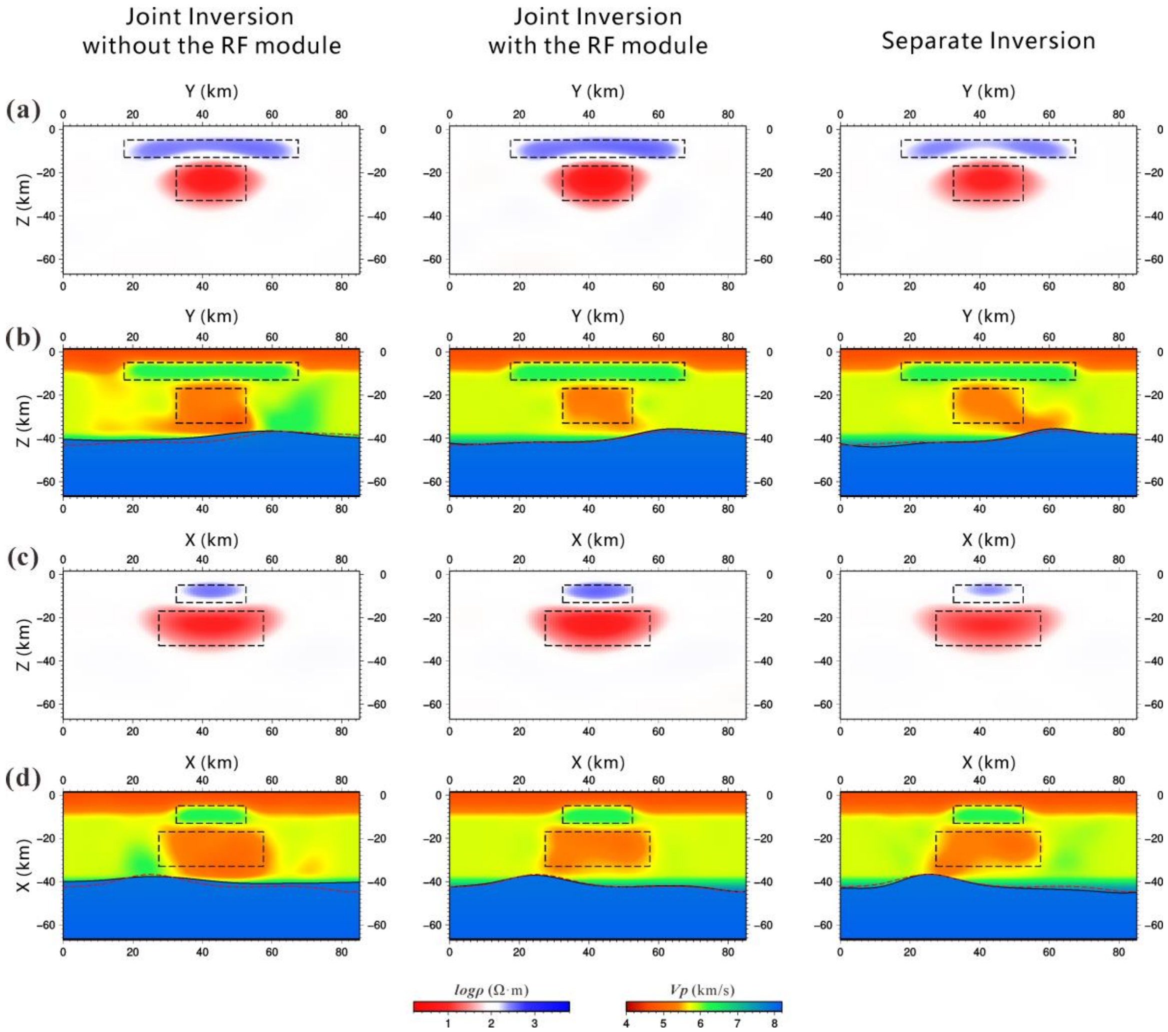

Figure 7. Imaging results from the separate inversion and joint inversion of synthetic data generated from a complex model inverted with Moho perturbations. The panels (a) and (b) are the MT and seismic results of cross-section AA' along E-W direction, respectively; the panels (c) and (d) are the MT and seismic results of cross-section BB' along S-N direction, respectively. In the seismic images, black curve denotes inverted Moho interface and red dashed line denotes true Moho interface.

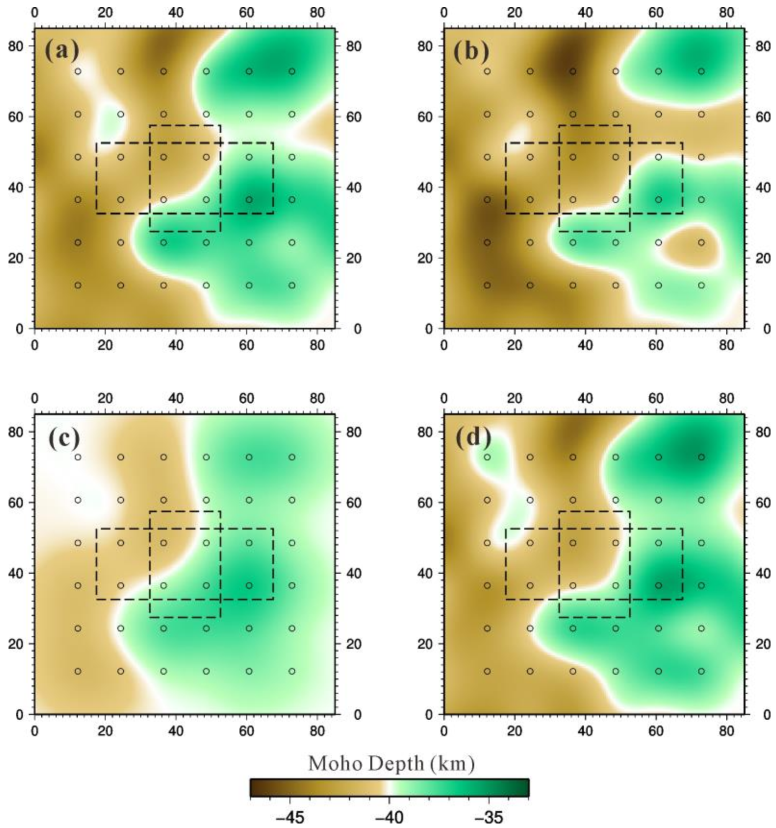

Figure 8. Moho depth variation in synthetic example II, including (a) true Moho interface; (b) initial reference Moho depth from H-κ stacking of RFs; (c) inverted Moho interface from joint inversion without the receiver function (RF) module, and (d) inverted Moho interface from joint inversion with the RF module. The black dashed curve denotes the true shape of two anomalous prims and Moho interface and black circle denotes broadband seismic stations.

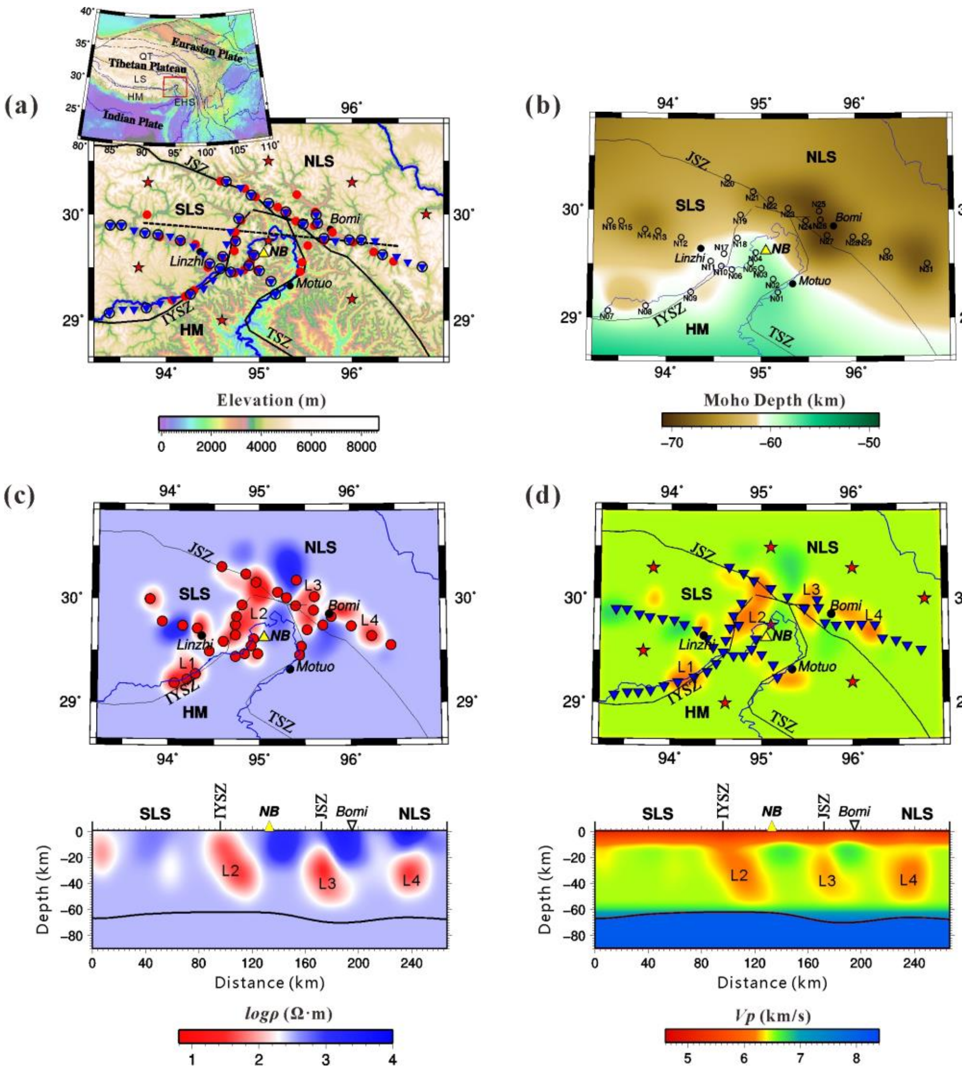

Figure 9. (a) Locations of 35 MT sites (red circle), 31 broadband seismic stations (black circle), 50 reflection/refraction seismic receivers (blue triangle), and 8 shot points (red pentagon) on the Earth surface used in this study. The top inset map displays the location of the study area in a red box. The map shows main suture zones and large faults including: the Qiangtang terrain (QT), Lhasa terrain (LS), Eastern Himalayan syntaxis (EHS), North Lhasa terrain (NLS), Namche Barwa mountain (NB), South Lhasa terrain (SLS), Himalaya terrain (HM), Tidding suture zone (TSZ), Jiali shear zone (JSZ), and Indus-Yarlung suture zone (IYSZ). (b) True Moho depth for the realistic model interpolated by the data from CRUST1.0 [Laske et al., 2013] and H-κ stacking analysis of RFs [Peng et al., 2017]. (c) Map views of true resistivity model derived from the previous 3D MT inversion results [Lin et al. 2017]. The top panel is a plan section at 25 km depth and the bottom panel is the cross-section shown in Figure 9a. L1, L2, L3 and L4 are low-resistivity anomalies. (d) Map views of true velocity model modified from the true resistivity model and the 3D teleseismic tomographic P-wave model [Peng et al., 2016]. The position of the panels is the same as Figure 9c

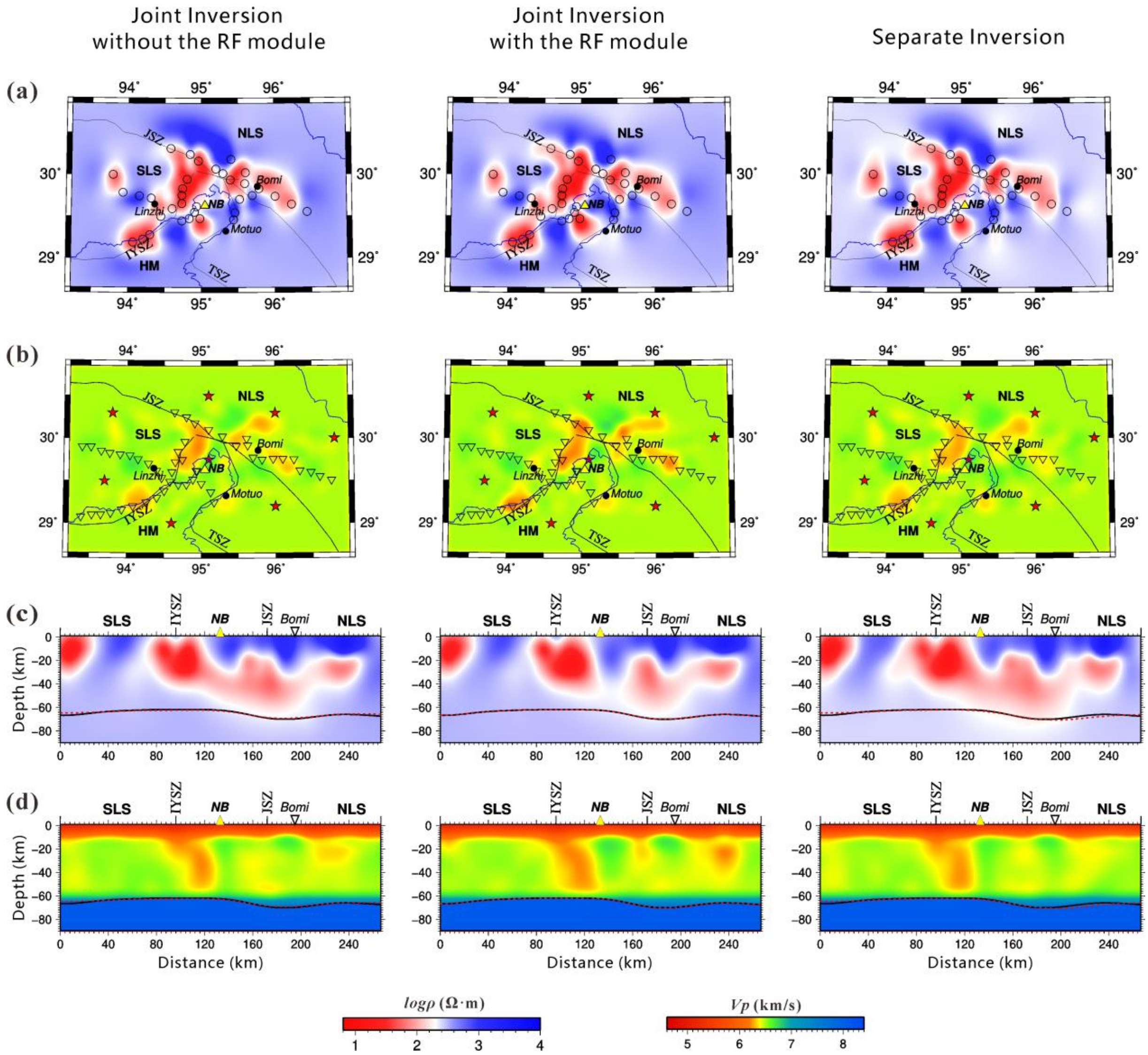

Figure 10. Imaging results from the separate inversion and joint inversion of synthetic MT data and wide-angle seismic data in Namche Barwa. The top panels (a) and (b) are plan sections of the MT and seismic results at 25 km depth, respectively. The bottom panels (c) and (d) show the MT and seismic results as a cross-section along the line shown in Figure 9a. In the cross-section images, the black curve denotes inverted Moho interface and the red dashed line denotes the true Moho interface.

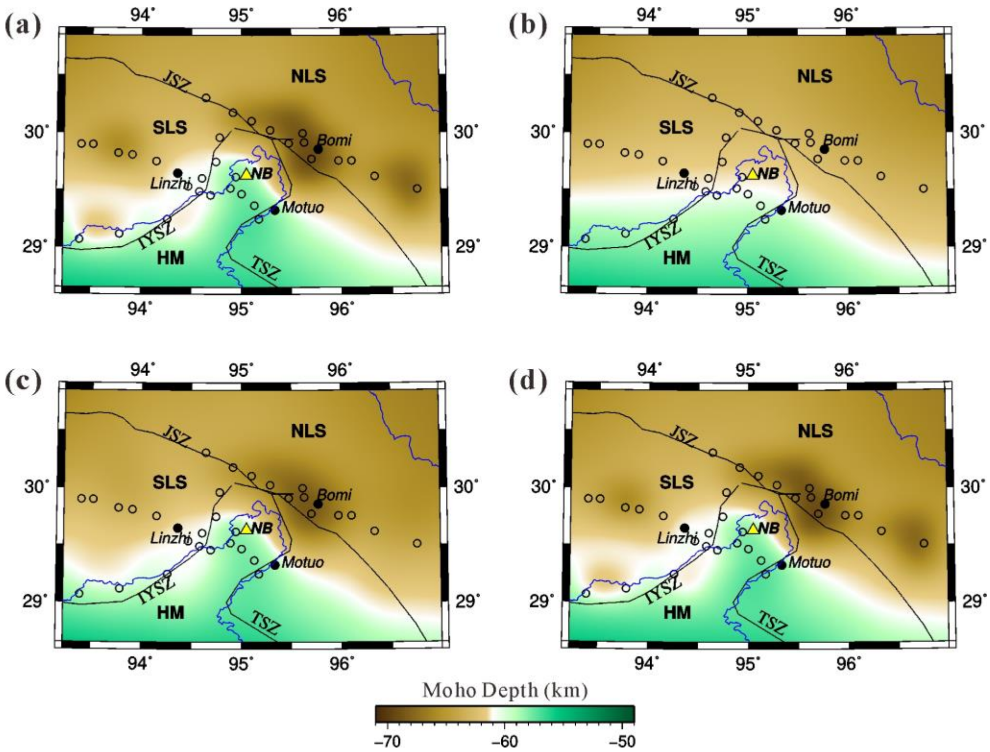

Figure 11. Maps of Moho depth variation for the realistic example III, including (a) true Moho interface; (b) initial reference Moho depth from CRUST1.0 [Laske et al., 2013]; (c) inverted Moho interface from joint inversion without the RF module and (d) inverted Moho interface from joint inversion with the RF module. The black circle denotes broadband seismic stations.