Geophysical Prospecting

EAGE EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

Geophysical Prospecting, 2020 doi: 10.1111/1365-2478.12932

Incorporating known petrophysical model in the seismic full-waveform inversion using the Gramian constraint

M. Malovichko[1,2]*, N. Khokhlov[2], N. Yavich[1,2] and M. S. Zhdanov[3,4,2]

[1]*CDISE, Skolkovo Institute of Science and Technology, Moscow, Russia,* [2]*Moscow Institute of Physics and Technology, Dolgoprudny, Russia,* [3]*University of Utah, Salt Lake City, Utah, USA, and* [4]*TechnoImaging, Salt Lake City, Utah, USA*

Received October 2018, revision accepted December 2019

ABSTRACT

In this paper, we develop a general approach to integrating petrophysical models in three-dimensional seismic full-waveform inversion based on the Gramian constraints. In the framework of this approach, we present an example of the frequency-domain P-wave velocity inversion guided by an electrical conductivity model. In order to introduce a coupling between the two models, we minimize the corresponding Gramian functional, which is included in the Tikhonov parametric functional. We demonstrate that in the case of a single-physics inversion guided by a model of different physical type, the general expressions of the Gramian functional and its gradients become simple and easy to program. We also prove that the Gramian functional has a non-negative quadratic form, so it can be easily incorporated in a standard gradient-based minimization scheme. The developed new approach of seismic inversion guided by the known petrophysical model has been validated by three-dimensional inversion of synthetic seismic data generated for a realistic three-dimensional model of the subsurface.

Key words: Full-waveform inversion, Joint inversion, Electromagnetics, Seismics, Gramian constraints.

INTRODUCTION

The methods of integrated interpretation of multiphysics geophysical data have been advanced considerably in recent years. One of the most widely used methods in geophysical community is based on imposing so-called structural constraints. The concept itself and its numerical implementation, known as the cross-gradient inversion, have been introduced in a seminal paper of Gallardo and Meju (2003a). For a recent review of this approach in a broad geophysical context, we refer the reader to the paper by Moorkamp *et al.* (2016). The cross-gradient method, as well as its numerous modifications, has some difficulties both from theoretical and practical standpoints. For example, the cross-gradient functional does not provide a unique solution; it is non-quadratic, which makes it difficult to find the Fréchet derivative of this functional without an approximation; it is a non-trivial task to extend the method to more than two types of data. Nevertheless, the most prominent examples of geophysical multiphysics inversion have been obtained by this approach. Several examples, which demonstrate recent developments, can be found in the papers by Hu, Abubakar and Habashy (2009), De Stefano *et al.* (2011), Molodtsov *et al.* (2013), Um, Commer and Newman (2014), Gessner *et al.* (2016), Rittgers *et al.* (2016), Manukyan, Maurer and Nuber (2018), Gao and Zhang (2018), Gross (2019), Lan *et al.* (2019), Zhang *et al.* (2019), among many others.

Another approach to the coupling joint inversion is based on the Gramian constraints introduced in Zhdanov, Gribenko and Wilson (2012) and Zhdanov (2015). This method exploits both the existence of petrophysical and/or structural relationships between the different physical properties of the subsurface, assuming that they reflect the same geologic formation; however, the relations between different

*E-mail: m.malovichko@skoltech.ru

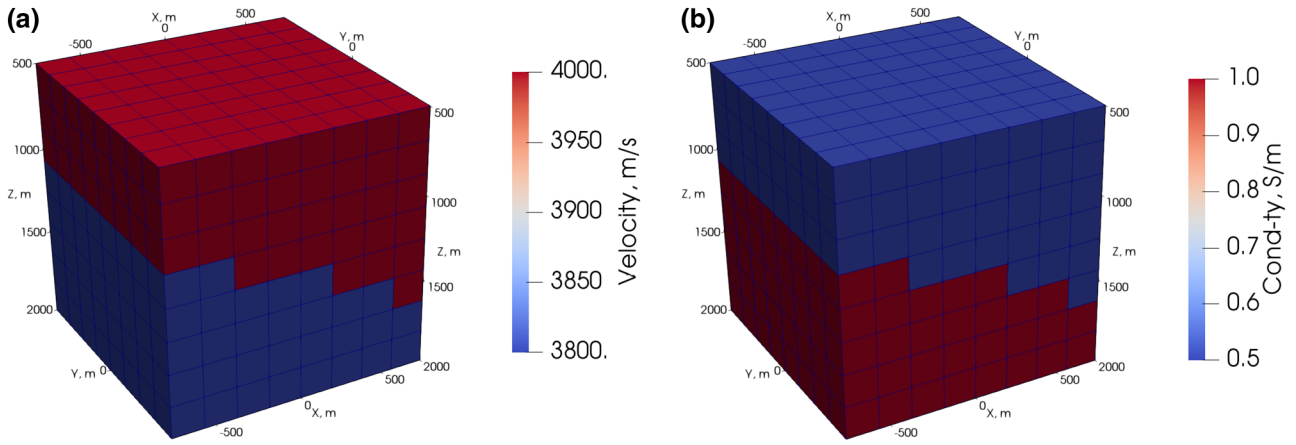

Figure 1 The true models used in the numerical experiment 1: (a) velocity model and (b) conductivity model.

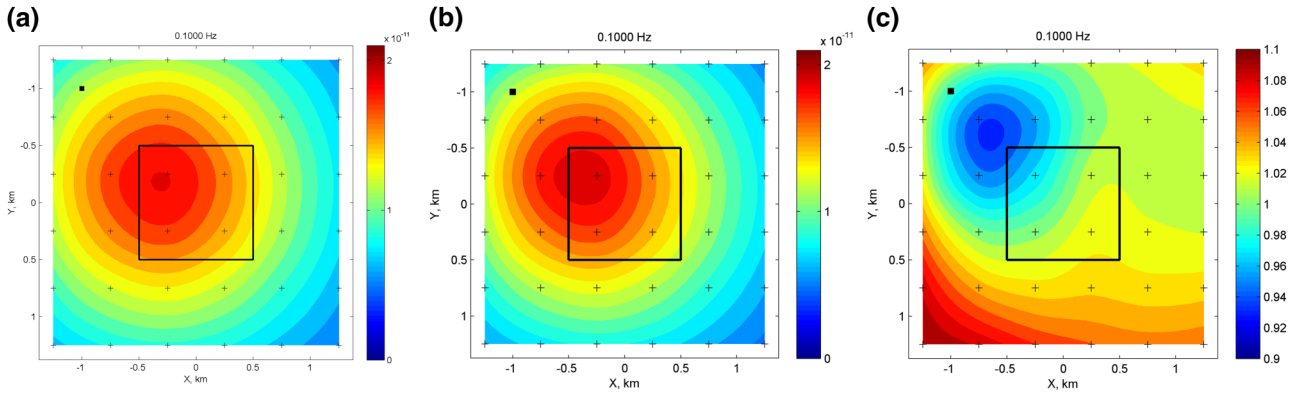

Figure 2 Data for a representative source in the first numerical experiment. The source was located at (–1, 1); receivers are marked with crosses. The black square in the middle outlines the border of the anomalous domain. (a) The real part of anomalous pressure response, $Re(p_a)$, (b) $Re(p_a)$ at the end of inversion with 90% impact of the Gramian term (see forth row in Table 1) and (c) normalized difference between input data and data at the end of inversion, that is $|\frac{p_a^{\text{end}} - p_a^{\text{ini}}}{p_a^{\text{ini}}}|$.

physical parameters are not necessary to be known and are restored during inversion. The mathematical formulation of the Gramian method is very general, which allows us to incorporate both conditions of the petrophysical and structural relationships in the Gramian stabilizing functional of a unified type. The other advantage of this approach is that the Gramian stabilizing functional is always quadratic, similar to the conventional minimum norm functional. This makes it easier to calculate the derivatives of the Gramian functional and provides the basis for a relatively simple numerical implementation. Another advantage of the Gramian constraints is that they can be easily extended to any number of the model parameters by simple construction of the Gramian matrices of size $N \times N$, where N is the number of different physical domains used in the inversion. The approach has been applied to joint inversion of potential (Zhdanov *et al.* 2012; Zhu *et al.* 2015; Čuma and Zhdanov 2017; Zhdanov and Cai 2017; Lin and Zhdanov 2018), electromagnetic (Zhu and Zhdanov 2015; Jorgensen and Zhdanov 2019; Ogunbo 2019), and seismic field data (Lin and Zhdanov 2019).

Table 1 Parameters of four inversion runs

Run Number	c_1	c_2	c_3	No. of Iterations
1	0.9	0	0.1	45
2	0.7	0	0.3	43
3	0.3	0	0.7	53
4	0.1	0	0.9	56

Note that, in this paper, for simplicity, we use the Gramian constraints to enforce the petrophysical correlations only. However, the Gramian approach can be easily extended

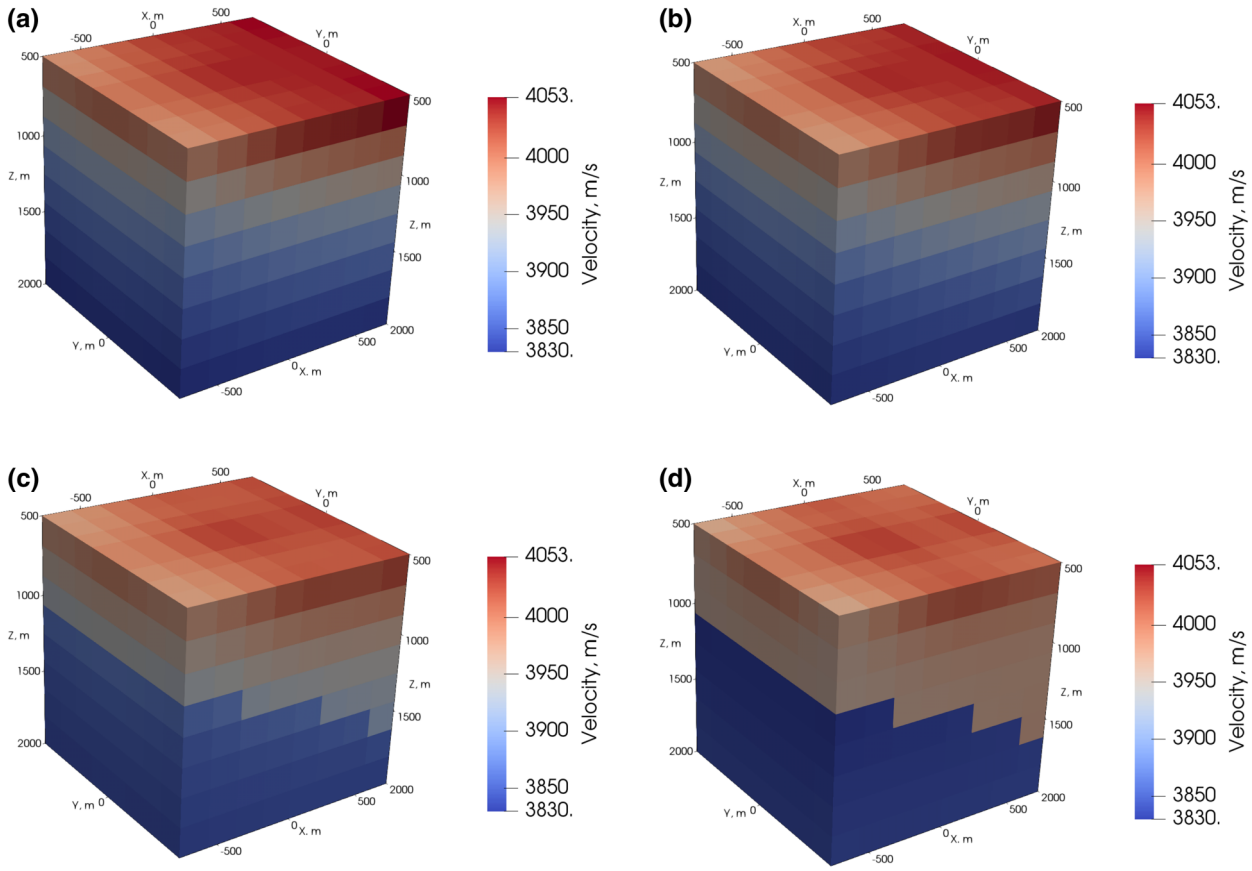

Figure 3 Results of seismic inversion stabilized with the conductivity model with different contributions of the Gramian term. Each cube has a side length of 1.5 km; it is divided into 187.5 m cubical cells. (a) Run #1, Gramian contribution was 10%; (b) run #2, Gramian contribution was 30%; (c) run #3, Gramian contribution was 70% and (d) run #4, Gramian contribution was 90%.

to enforce the structural constraints if one would use the gradients of physical properties instead of physical parameters themselves (e.g. Jorgensen and Zhdanov 2019; Lin and Zhdanov 2019).

The full joint inversion, that is, with all involved models being updated simultaneously, is a very promising but challenging problem. However, *the guided* inversion, when some parameters of petrophysical models are known beforehand and are fixed, is also of significant practical importance. It can be considered as a tool to incorporating *a priori* information in the inversion. Electromagnetic inversion tied to a known velocity model through the Gramian term has been reported in Zhdanov *et al.* (2016). However, the question of applicability of this approach to seismic data has not been studied in full extent yet (an introduction to this approach was presented by the authors at the EAGE conference (Malovichko, Yavich and Zdanov 2017b)).

In this paper, we use the concept of guided inversion to incorporate the known petrophysical model in the three-dimensional (3D) seismic frequency-domain full-waveform inversion (FWI). As an illustration of this approach only, we consider the electrical conductivity distribution as a reference petrophysical model. We also assume that the geoelectrical model is well defined by the resistivity well-logging data or by a very detailed and dense controlled-source electromagnetic survey. In practical applications, the reference petrophysical model can be constructed based on all available surface and borehole geological/geophysical data. The goal of our paper is to demonstrate how the available subsurface information can be incorporated in the full-waveform seismic inversion using the Gramian constraints. At the same time, if for any reason the available *a priori* information is not consistent with the seismic data, the Gramian term will not decrease during the inversion, thus manifesting that integration is not feasible in that case.

We have derived simple explicit expressions for the gradient of the Gramian functional for the models characterized by two physical properties. We demonstrate that the Gramian

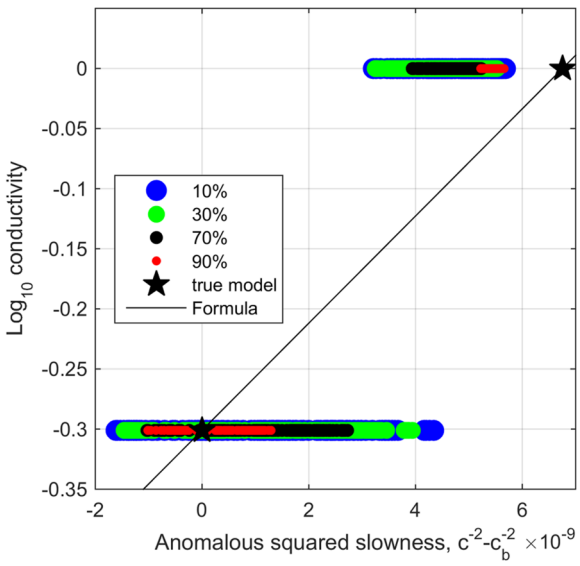

Figure 4 Cross plots of anomalous square slowness against logarithmic conductivity. The percentage in the legend stands for relative contribution of the Gramian term and corresponds to those in Fig. 3. The two stars indicate the correspondence between the true velocity and true conductivity. The black line connecting these two stars indicates a linear relationship between the true model parameters described by formula (28).

stabilizing functional has an appealing property from numerical standpoint – it has a non-negative quadratic form, so it can be easily incorporated in a standard gradient-based minimization scheme. The forward modelling of seismic data is performed by a parallel implementation of the method of integral equations. The inversion is based on the regularized conjugate gradient algorithm. The developed method and computer code were tested on 3D models containing both seismic and geoelectric inhomogeneities. The results of this study provide the building blocks for the full joint inversion and also provide the means of introducing an *a priori* information in the seismic FWI.

FORWARD MODELLING

The time-harmonic acoustic wavefield of angular frequency ω propagating in a medium of primary velocity $c = c(\mathbf{r})$ satisfies the following partial differential equation:

$$-\Delta p - (\omega^2/c^2)p = \delta\left(\mathbf{r} - \mathbf{r}'\right), \tag{1}$$

where $p = p(\mathbf{r})$ is the pressure, Δ is the Laplacian; δ is the Dirac delta-function; $\mathbf{r}$ is the position vector, primed coordinates denote the source location. Throughout this paper we use the right-handed coordinate system with Z-axis pointing downward. Equation (1) is supplemented with the radiation conditions on infinity or with boundary conditions in a bounded domain, D. In this study, we solve problem (1) by the integral equations (IE) method. This choice is motivated by the fact that the system matrix arising from discretization of IE is better conditioned comparing to the finite-difference system matrix. Also, this choice of the forward modelling operator allows us to use the same discretization both for forward and inverse problems, simplifying the numerical implementation (this point is commented in the next section).

We assume that domain D with arbitrary velocity distribution, $c(\mathbf{r})$, is submersed into a layered background model with a piecewise constant velocity, c_b, varying in the vertical direction, z, only: $c_b = c_b(z)$. The pressure response, p, can be represented as a sum of two parts: the background (incident) part, p_b, due to the background medium, and the anomalous (scattered) part, p_a, due to the anomalous velocity

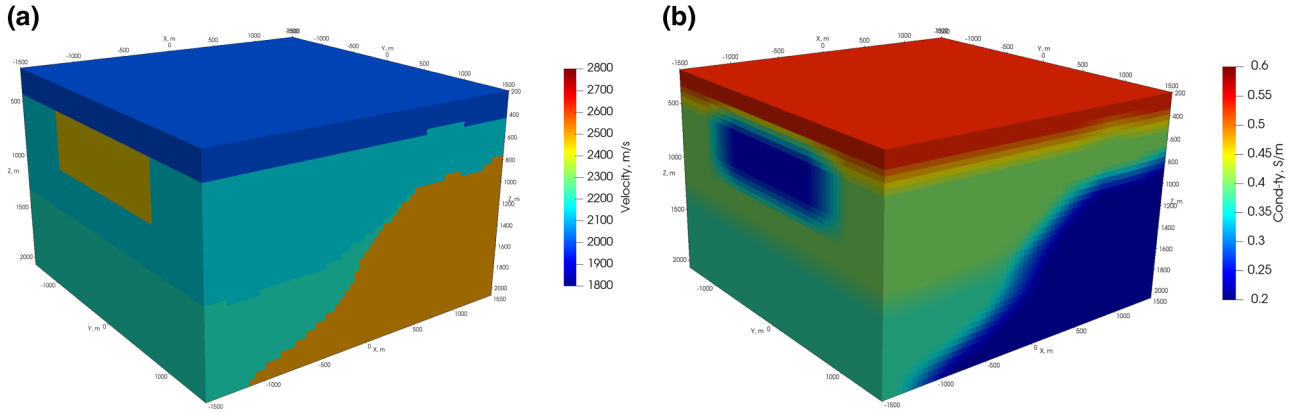

Figure 5 True 3 km × 3 km × 1.875 km models: (a) the velocity model and (b) the conductivity model obtained from the velocities after smoothing and applying the conversion formula.

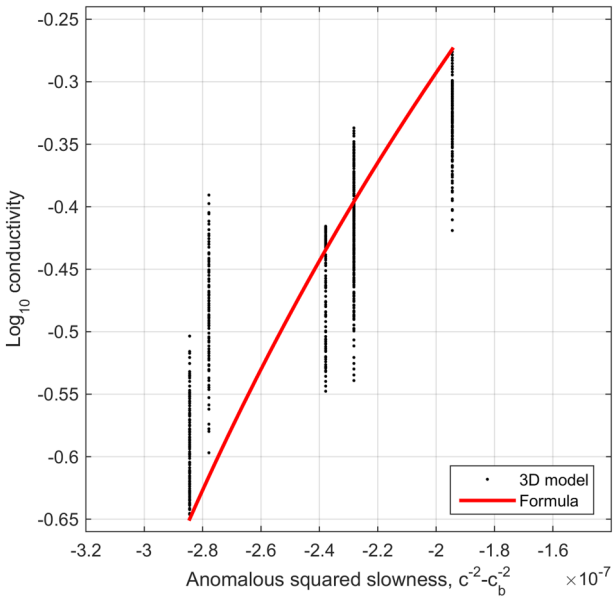

Figure 6 Cross plot of the true conductivity versus true velocity models against formula (30).

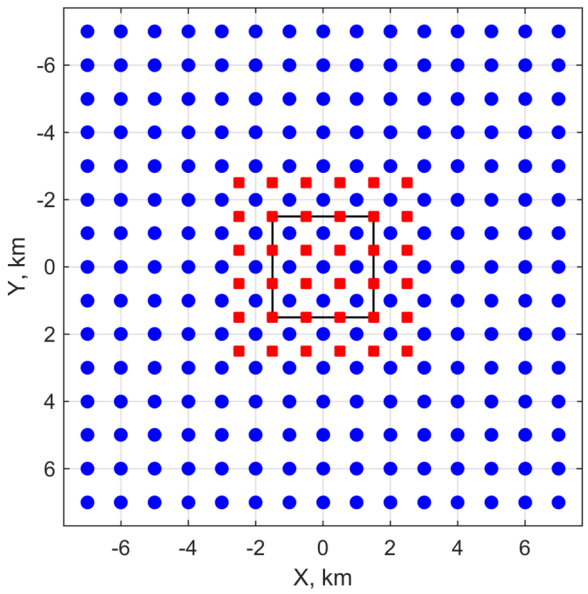

Figure 7 Acquisition grids. The red squares are sources and the blue circles are receivers. The black square in the middle outlines the boundaries of the true model.

distribution. The solution for p is given by the following expression (Zhdanov 2002):

$$p(\mathbf{r}) = p_b(\mathbf{r}) + \omega^2 \int_D g(\mathbf{r}|\mathbf{r}') \left(\frac{1}{c^2(\mathbf{r}')} - \frac{1}{c_b^2(z')} \right) p(\mathbf{r}')\, \mathrm{d}V', \quad (2)$$

where g is the background Green's function. This expression becomes an integral equation when $\mathbf{r},\ \mathbf{r}' \in D$ (termed *a domain equation*). After determining the pressure filed, p, in D, expression (2) is used again to map the pressure field to the receivers (in this case it is termed *a field equation*).

We emphasize that in the framework of the IE method the background velocity, c_b, must allow for analytical or semi-analytical solution for the Green's function, while $c_a = c - c_b$ can be a function of x, y and z. In exploration geophysics, the obvious choice is a horizontally layered Earth. In this case, the Green's function is given by a numerical Hankel transform of an analytically computed kernel (for example, Malovichko *et al.* 2018; the Appendix). In the two special cases, which are the constant velocity model, $c_b = \text{const}$, and the half-space model, $c_b(z \geq 0) = \text{const}$ and $c_b(z < 0) = 0$, the solution to the Green's function simplifies even more to a closed-form expressions. In this paper, we use the half-space background model.

The numerical method consists of iterative solution of the system of linear equations arising from (2). The details of the numerical method and its parallel implementation can be found in Malovichko *et al.* (2018).

FORMULATION OF THE INVERSE PROBLEM

Let $V \subset D$ be an inversion domain, discretized into M rectangular cells. The inverse model is described by the values of P-wave velocity $c(\mathbf{r})$ and electrical conductivity, $\sigma(\mathbf{r})$, sampled at the cell's centres. The goal is to find c based on the measured values of p imposing the usual smoothness constraint on c and also the structural coupling between c and σ. We define a model vector, $\mathbf{m} \in \mathbb{R}^M$, which contains the values of anomalous squared slowness, $1/c^2 - 1/c_b^2$. This parametrization of the velocity model is motivated by the choice of the forward modelling operator. Additionally, we define vector $\mathbf{s} \in \mathbb{R}^M$ containing the values of decimal logarithm of electrical conductivities in the cell's centres, $\log \sigma$. Let vector $\mathbf{d} \in \mathbb{C}^N$ be the values of the pressure response, measured at a finite number of source–receiver positions and temporal frequencies, where N is the total number of the data points. Formally, the correspondence between $\mathbf{d}$ and $\mathbf{m}$ is established by a forward operator, $\mathcal{A}$, as $\mathbf{d} = \mathcal{A}(\mathbf{m})$, which implies solving problem (1).

Here we comment on our choice of discretization for both the forward and inverse problems. The forward modelling grid is selected such that a discrete solution is a good approximation to the continuous solution. For example, for a second-order finite-difference scheme, a well-known rule is to have not less than 12–14 grid points per shortest wavelength. In contrast, the step size of the inversion grid is

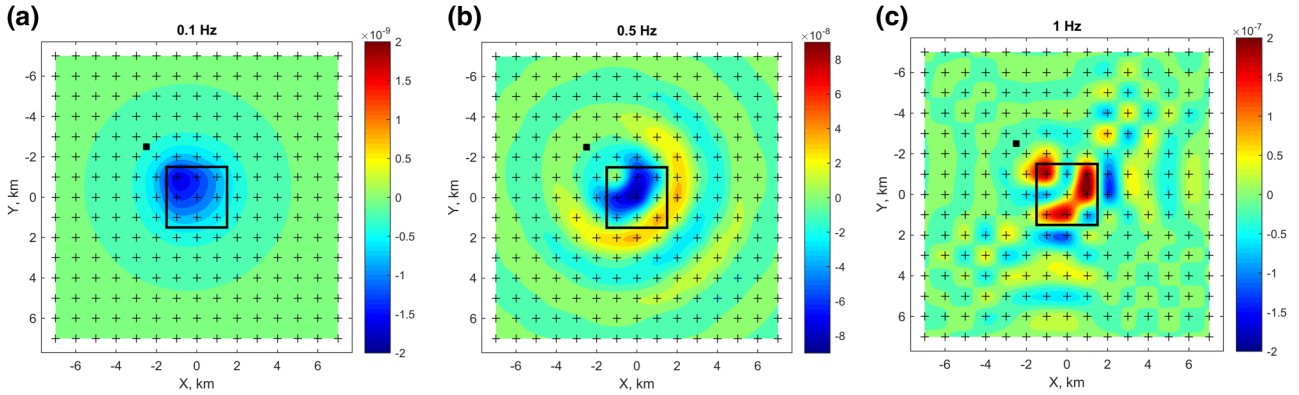

Figure 8 Example of simulated data for a single source location at three frequencies: (a) 0.1 Hz, (b) 0.5 Hz and (c) 1 Hz. The real part of anomalous pressure response, p_a, is shown in colour. Note the strong aliasing that is clearly visible in the 1 Hz data.

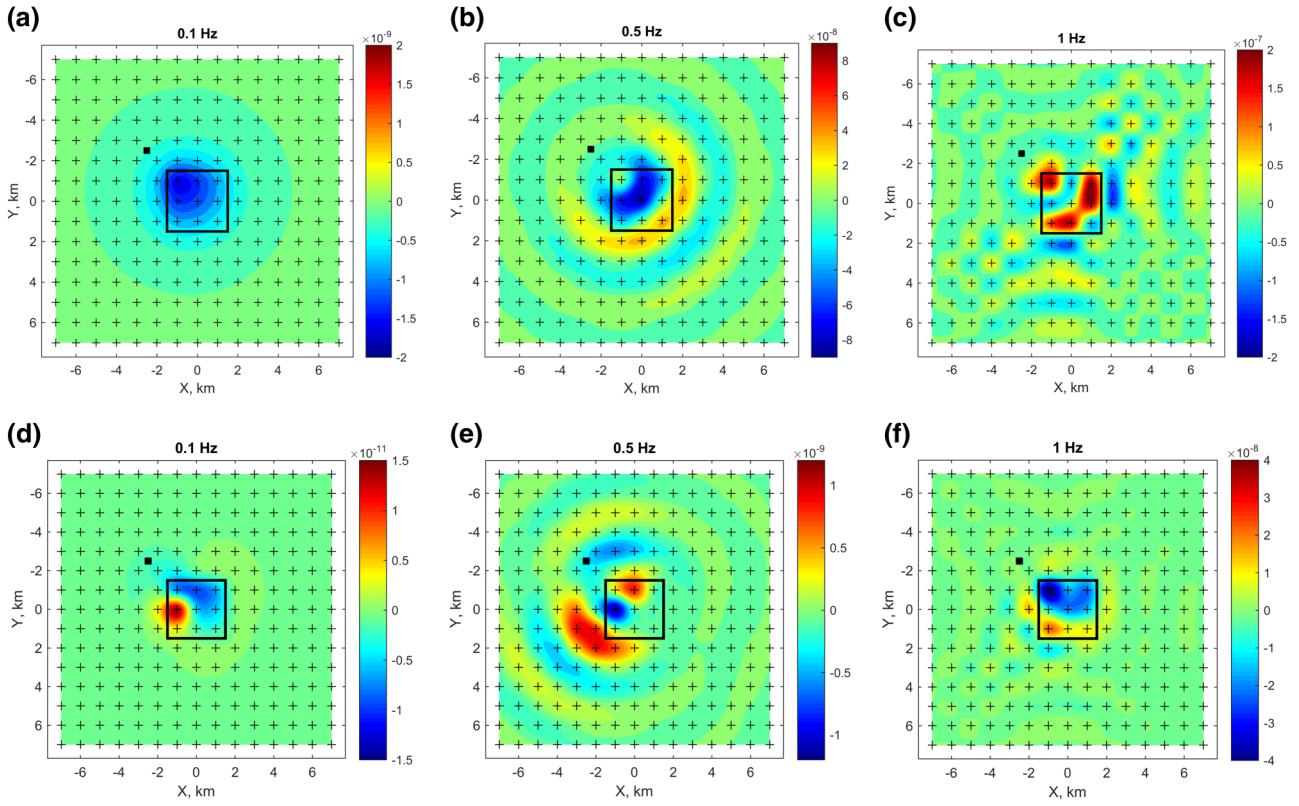

Figure 9 Final synthetic data at the end of the smooth inversion with $\alpha_{\text{rel}} = 10^{-3}$. Panels (a) and (b) show the real part the final synthetic wavefield, $Re(p_a^{\text{end}})$ at three representative frequencies for a single source position. Panels (d)–(f) show the real part of the difference between the input and final acoustic fields, $Re(p_a^{\text{end}} - p_a)$. The data correspond to the model shown in Fig. 12(d).

usually motivated by far looser requirements. For example, by a required resolution of the final velocity model or by a limit on the size of the normal system matrix in case when optimization is performed by Newton's method with direct matrix factorization. Restrictions on the forward modelling grid are usually much stricter, so it is generally accepted that the forward modelling numerical grid should be finer than the inversion one. Nevertheless, in this study we use the same discretization for both forward and inverse problems. Both grids had not less than 32 cells per shortest wavelength at the highest frequency (see Numerical Experiments section). These grids are excessively fine, which, apparently, increased the

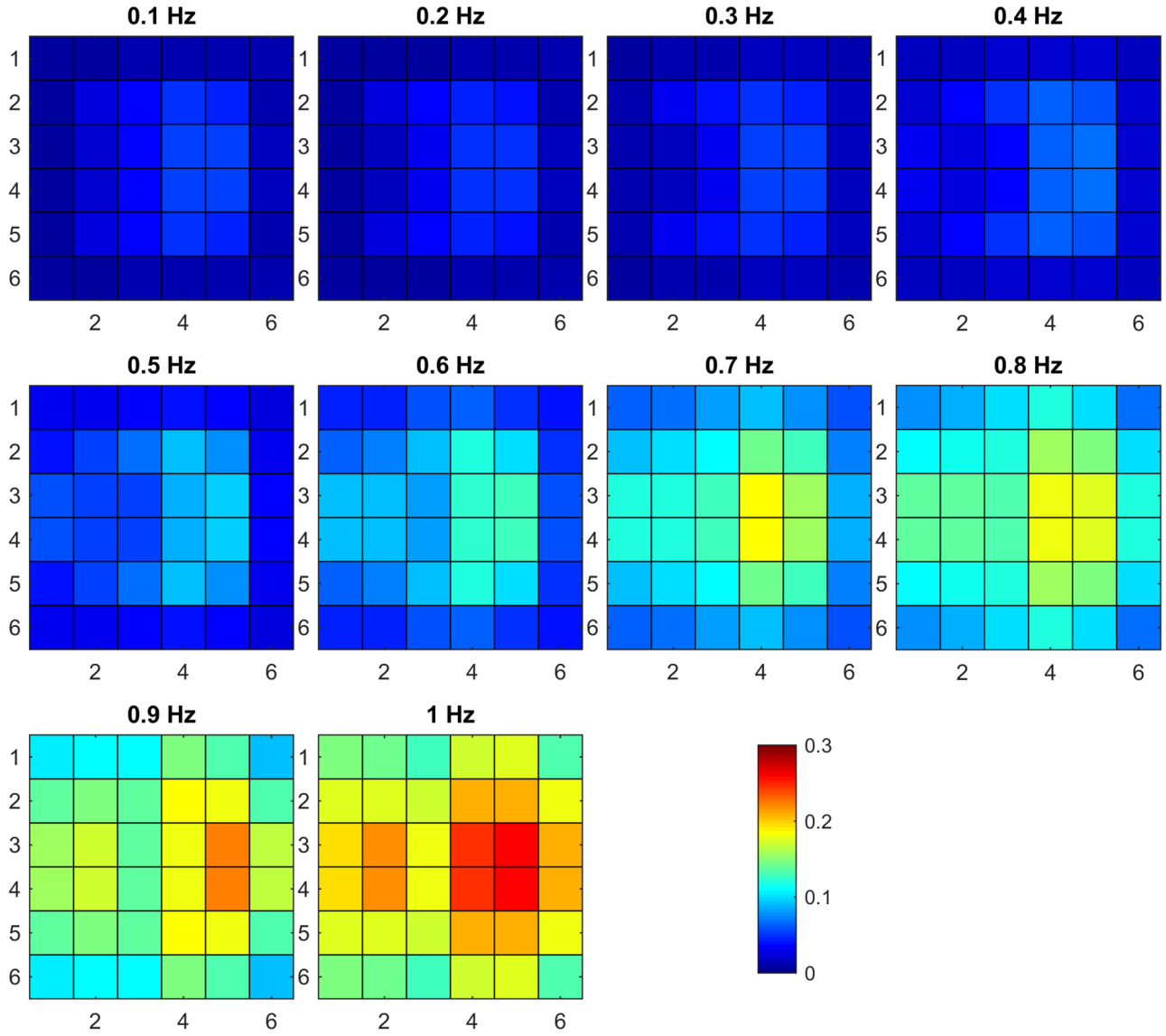

Figure 10 Normalized misfit at the end of the smooth inversion with $\alpha_{\text{rel}} = 10^{-3}$. Each panel shows quantity $\|p_a{}^{\text{end}} - p_a\|/\|p_a\|$ for all 36 sources at a single frequency.

inversion run-time. On the other hand, this way we avoided the use of interpolation operators (and their transposes) acting from a forward modelling grid to the inversion grid and from the receivers to a forward modelling grid (see, for example, Malovichko *et al.* 2019; the Appendix). It greatly simplified the programming implementation and made our presentation clearer, leaving out unnecessary technical details.

We consider the following unconstrained minimization problem for the parametric functional, P^α:

$$\begin{aligned} &\inf_{\mathbf{m}\in\mathbb{R}^M} P^\alpha(\mathbf{m}), \\ &P^\alpha = \phi(\mathbf{m}) + \alpha\psi(\mathbf{m}), \end{aligned} \tag{3}$$

with

$$\phi(\mathbf{m}) = \|\mathbf{W}(\mathbf{d} - \mathcal{A}(\mathbf{m}))\|^2, \tag{4}$$

$$\psi(\mathbf{m}) = q_1\|\mathbf{L}^{1/2}(\mathbf{m} - \mathbf{m}_0)\|^2 + q_2\|\mathbf{m} - \mathbf{m}_0\|^2 + q_3\gamma(\mathbf{m}, \mathbf{s}), \tag{5}$$

where ϕ is the misfit functional, ψ is the stabilizing functional, $\mathbf{W}$ is the diagonal matrix with inverse standard deviations of the noise on its diagonal, $\mathcal{A}$ is the forward modelling operator discussed above, α is the regularization parameter, $\mathbf{L}$ is the finite-difference approximation of the Laplacian operator with zero Dirichlet boundary conditions, $\mathbf{m}_0$ is the starting model, q_1, q_2 and q_3 are three positive numbers, which are

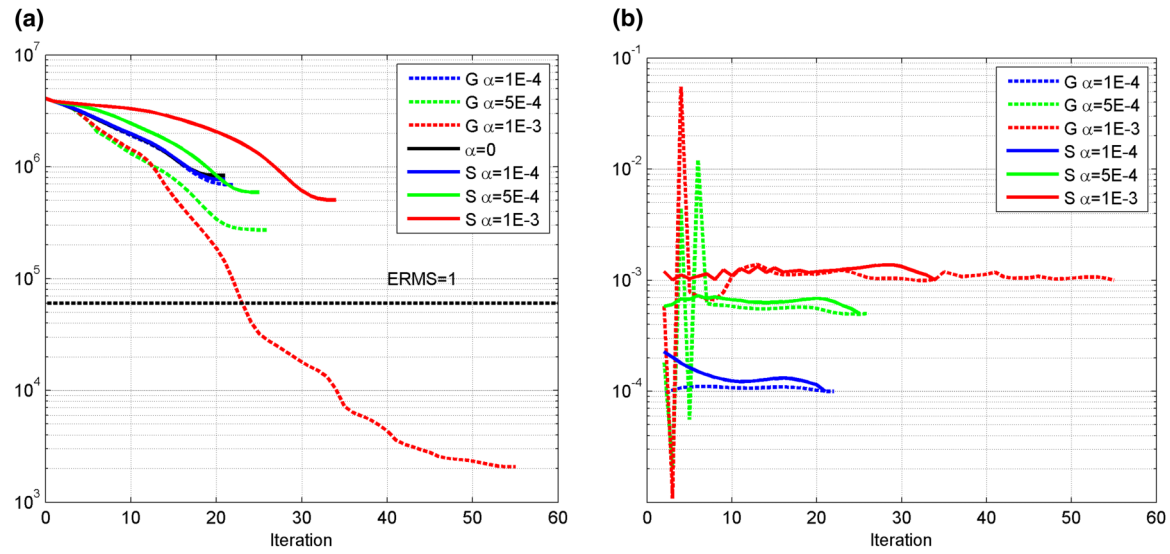

Figure 11 Inversion progress in numerical experiment 2. (a) Plot of the misfit functional, ϕ_n, versus iteration number. Letter 'S' means the smoothness constraint was enabled only ($c_1 = 1$); letter 'G' means the Gramian constraint was enabled only ($c_3 = 1$). The horizontal dashed line corresponds to ERMS = 1. (b) Plot of the ratio $\alpha_n \psi_n / \phi_n$ versus iteration number for several different inversion runs.

commented below and γ is the determinant of the Gram matrix for a set of vectors $\{\mathbf{m}, \mathbf{s}\}$ (Zhdanov 2015):

$$\gamma(\mathbf{m}, \mathbf{s}) = \det \begin{pmatrix} \|T[\mathbf{m}]\|^2 & (T[\mathbf{m}], T[\mathbf{s}]) \\ (T[\mathbf{m}], T[\mathbf{s}]) & \|T[\mathbf{s}]\|^2 \end{pmatrix}, \tag{6}$$

where $T[\cdot]$ is a given model transform, discussed later; $(\cdot, \cdot)$ is the inner product; $\|\cdot\|$ is the Euclidian norm. The main advantage of the Gramian mutual constraint (6) over some other approaches is that the stabilizing functional, ψ, is convex given transformation T is linear (see the Appendix). It should be emphasized that in (6) $\mathbf{m}$ and $\mathbf{s}$ depend non-linearly on physical parameters c and σ, respectively. Thus, minimization of (6) imposes a linear relation between $\mathbf{m}$ and $\mathbf{s}$ but a non-linear one between c and σ. Furthermore, in the numerical experiments presented below, the relation between $\mathbf{m}$ and $\mathbf{s}$ was, in fact, non-linear, so the Gramian constraint enforced its linear approximation (see Numerical Experiments section).

We apply the non-linear conjugate gradient technique to problem (3)–(5) as described in Malovichko *et al.* (2017a). The algorithm can be defined by the following sequence:

$$\mathbf{y}_n = \mathcal{A}(\mathbf{m}_n) - \mathbf{d}, \tag{7}$$

$$\mathbf{u}_n = \xi \mathbf{F}_n^* \mathbf{W}^2 \mathbf{y}_n, \tag{8}$$

$$\mathbf{g}_n = u_n + \alpha_n q_{1,n} \mathbf{l}_1 + \alpha_n q_{2,n} \mathbf{l}_2 + \alpha_n q_{3,n} \mathbf{l}_3, \tag{9}$$

$$\beta_n = \|\mathbf{g}_n\|^2 / \|\mathbf{g}_{n-1}\|^2, \tag{10}$$

$$\mathbf{p}_n = -\mathbf{g}_n + \beta_n \mathbf{p}_{n-1}, \quad \mathbf{p}_1 = -\mathbf{g}_1, \tag{11}$$

$$k_n = \frac{(\mathbf{p}_n, \mathbf{g}_n)}{\|\mathbf{F}_n \mathbf{W} \mathbf{p}_n\|^2 + \alpha_n q_{1,n} \mathbf{p}_n^* \mathbf{L} \mathbf{p}_n + \alpha_n q_{2,n} \|\mathbf{p}_n\|^2}, \tag{12}$$

$$\mathbf{m}_n = \mathbf{m}_{n-1} + k_n \mathbf{p}_n. \tag{13}$$

Here $\mathbf{u}_n$ is the gradient of squared norm of the weighted data residual, $\mathbf{y}_n$, the subscript denotes iteration number; $\mathbf{F}$ is the matrix of Frechet derivative of operator $\mathcal{A}$, $\mathbf{F} = \partial \mathcal{A} / \partial \mathbf{m}$; asterisk '*' denotes Hermitian conjugation.

Scalar factor, ξ, in (8) balances impact of data at different frequencies: $\xi = \omega^{-4}$. This choice is explained in Sirgue and Pratt (2004), so we breifly restated it here. The first ω^{-2} factor comes from differentiating the integral operator in (2) with respect to $\mathbf{m}$, which is essentially the anomalous squared slowness, $1/c^2 - 1/c_b^2$. Another ω^{-2} factor appears when the Born approximation is used to calculate the data gradient, or, in our case, from the use of the quasi-analytical (QA) approximation, which has similar effect, see (Malovichko *et al.* 2017a). In other words, the second ω^{-2} factor is needed to compensate the impact of the approximation used in calculation of the gradient $\mathbf{y}_n$.

Vector $\mathbf{g}_n$ is the gradient of the parametric functional. Real positive value k_n is the step in the search direction, $\mathbf{p}_n$. Vectors $\mathbf{l}_1$ and $\mathbf{l}_2$ are gradients of the smoothness and damping terms, respectively:

$$\mathbf{l}_1 = \mathbf{L}(\mathbf{m}_n - \mathbf{m}_0), \quad \mathbf{l}_2 = \mathbf{m}_n - \mathbf{m}_0. \tag{14}$$

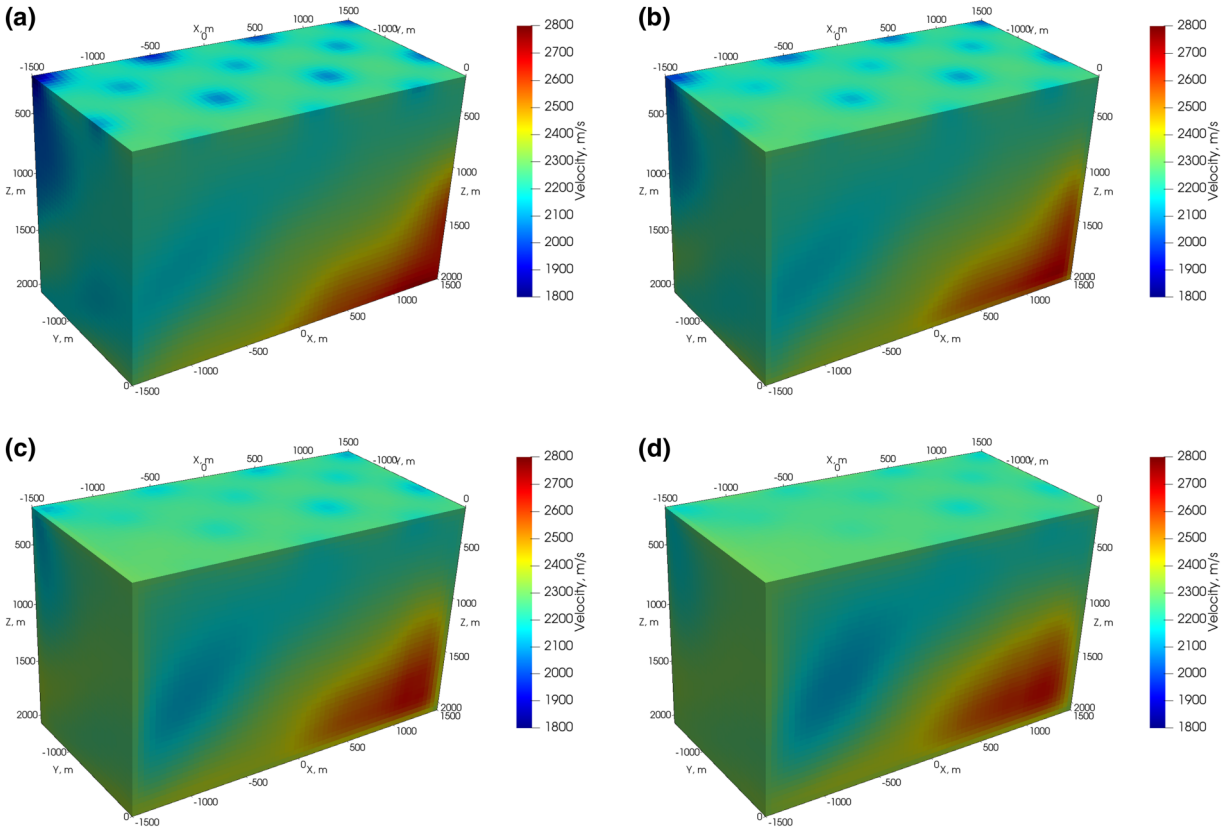

Figure 12 Final velocity models for smoothness regularization. The cells with $y > 0$ are cut out for better visualization. (a) No regularization, $\alpha_{\text{rel}} = 0$, (b) smoothness $\alpha_{\text{rel}} = 10^{-4}$, (c) smoothness $\alpha_{\text{rel}} = 5 \times 10^{-4}$ and (d) smoothness $\alpha_{\text{rel}} = 10^{-3}$.

Vector $\mathbf{l}_G$ represents gradient of Gramian, γ, which will be discussed in detail in the next section.

In order to compute product (8), we employ the QA approximation, described in Malovichko *et al.* (2017a) and Zhdanov (2015), based on approximate solution of the corresponding integral equation. The QA approximation is fast and much more accurate comparing to the widely used Born approximation. It reduces the computational time of the data gradient, thus accelerating the overall inversion procedure. The regularization parameter, α_n, is selected on each iteration in order to keep the ratio $\alpha\psi/\phi$ equal to some predefined constant α_{rel}:

$$\alpha_{n+1} = \alpha_{\text{rel}}\phi_n/\psi_n. \tag{15}$$

The three positive constants, q_i, are selected on each iteration by the following rule:

$$q_{i,n+1} = \frac{c_i/w_{i,n}}{\sum_{j=1}^{3} c_j/w_{j,n}}, \quad i = 1, 2, 3, \tag{16}$$

where c_i are three positive predefined constants such as $c_1 + c_2 + c_3 = 1$ and w_i is ith member of the regularization functional on nth iteration:

$$w_{1,n} = \|\mathbf{L}^{1/2}(\mathbf{m}_n - \mathbf{m}_0)\|^2, \quad w_{2,n} = \|\mathbf{m}_n - \mathbf{m}_0\|^2, \quad w_{3,n} = \gamma_n. \tag{17}$$

Algorithm (16) and (17) takes into account different scales of different stabilizing terms making their contribution correspond to c_1, c_2 and c_3, which can be regarded as weights of corresponding model constraints. The progress of inversion is controlled by the value of normalized misfit, $||\mathbf{W}y_n||/\mathrm{N}$, referred as the root-mean-square error (ERMS) in the text.

DIRECTIONS OF THE STEEPEST DESCENT

The steepest descent directions (gradients) of the first two terms in equation (5) can be found using a standard technique (e.g. Zhdanov 2002, 2015). The gradient of the third term is treated in this section. We assume that model transform T is a linear operator and is described by its symmetric matrix, $\mathbf{T}$. The choice of operator T is important in determining the form

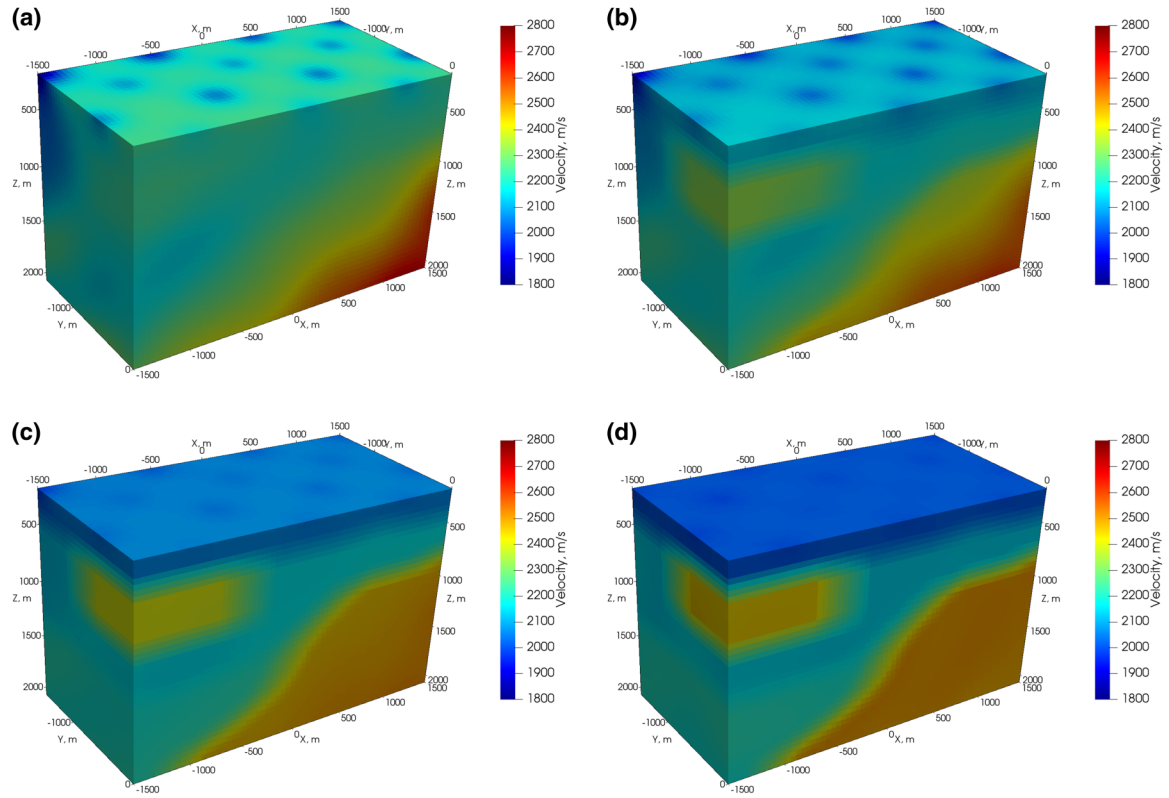

Figure 13 Final velocity models for the Gramian regularization. The cells with $y > 0$ are cut out for better visualization. (a) Gramian $\alpha_{\text{rel}} = 10^{-4}$, (b) Gramian $\alpha_{\text{rel}} = 5 \times 10^{-4}$, (c) Gramian $\alpha_{\text{rel}} = 10^{-3}$ (iteration 23) and (d) Gramian $\alpha_{\text{rel}} = 10^{-3}$ (iteration 55).

of correlation between the model parameters, **m** and **s**. In this study, we choose operator T as follows:

$$T[\mathbf{v}] = (\mathbf{v} - \hat{v})/\sigma_v, \qquad \mathbf{v} = \mathbf{m} \text{ or } \mathbf{s}, \tag{18}$$

where $\hat{v}$ and σ_v are the mean and standard deviation of **v**, respectively. Thus, we impose the correlation between the standardized quantities: anomalous squared slowness, $1/c^2 - 1/c_b^2$, and decimal logarithm of total conductivity, $\log \sigma$. In order to make this transform linear, we set $\hat{v}$ and σ_v to their values calculated on the previous iteration of the inversion. Thus, T becomes a linear transform, $T(\mathbf{v}) = \mathbf{T}\mathbf{v}$, with matrix **T** defined as follows:

$$\mathbf{T} = \sigma_v^{-1} (\mathbf{I} - \mathbf{U}/N), \tag{19}$$

where **I** is an $N \times N$ identity matrix and **U** is an $N \times N$ matrix of ones.

We note that it is possible to change model parametrization to $\tilde{\mathbf{m}} = T[\mathbf{m}]$ and $\tilde{\mathbf{s}} = T[\mathbf{s}]$, and then invert for $\tilde{\mathbf{m}}$ and $\tilde{\mathbf{s}}$. This simplifies definition of model transform to $\mathbf{T} = \mathbf{I} - \mathbf{U}/N$ as compared to (19) and allows to accurately implement arbitrary non-linear data transforms. However, this would require modifications in program modules responsible for gradient computations for each particular non-linear transform. This is undesirable because we wish to isolate the usage of model transform from the rest of the inversion code, which itself may include some non-linear transformations of the data and model parameters for numerical reasons. With this definition, we can write for Gramian:

$$\gamma = (\mathbf{m}^T\mathbf{T}^2\mathbf{m})(\mathbf{s}^T\mathbf{T}^2\mathbf{s}) - (\mathbf{m}^T\mathbf{T}^2\mathbf{s})^2. \tag{20}$$

Let us calculate the gradient of Gramian:

$$\begin{aligned} \nabla\gamma &= 2\begin{pmatrix} \mathbf{T}^2\mathbf{m}(\mathbf{s}^T\mathbf{T}^2\mathbf{s}) - \mathbf{T}^2\mathbf{s}(\mathbf{m}^T\mathbf{T}^2\mathbf{s}) \\ -\mathbf{T}^2\mathbf{m}(\mathbf{m}^T\mathbf{T}^2\mathbf{s}) + \mathbf{T}^2\mathbf{s}(\mathbf{m}^T\mathbf{T}^2\mathbf{m}) \end{pmatrix} \\ &= 2\begin{pmatrix} \phi_2\mathbf{T}^2 & -\phi_3\mathbf{T}^2 \\ -\phi_3\mathbf{T}^2 & \phi_1\mathbf{T}^2 \end{pmatrix}\begin{pmatrix} \mathbf{m} \\ \mathbf{s} \end{pmatrix}, \end{aligned} \tag{21}$$

where $\phi_1 = \mathbf{m}^T\mathbf{T}^2\mathbf{m}$, $\phi_2 = \mathbf{s}^T\mathbf{T}^2\mathbf{s}$, $\phi_3 = \mathbf{m}^T\mathbf{T}^2\mathbf{s}$. Since in this study we do not update conductivity model, **s**, we use only the gradient of Gramian with respect to **m**:

$$\mathbf{l}_G = 2\mathbf{T}^2(\phi_2\mathbf{m} - \phi_3\mathbf{s}). \tag{22}$$

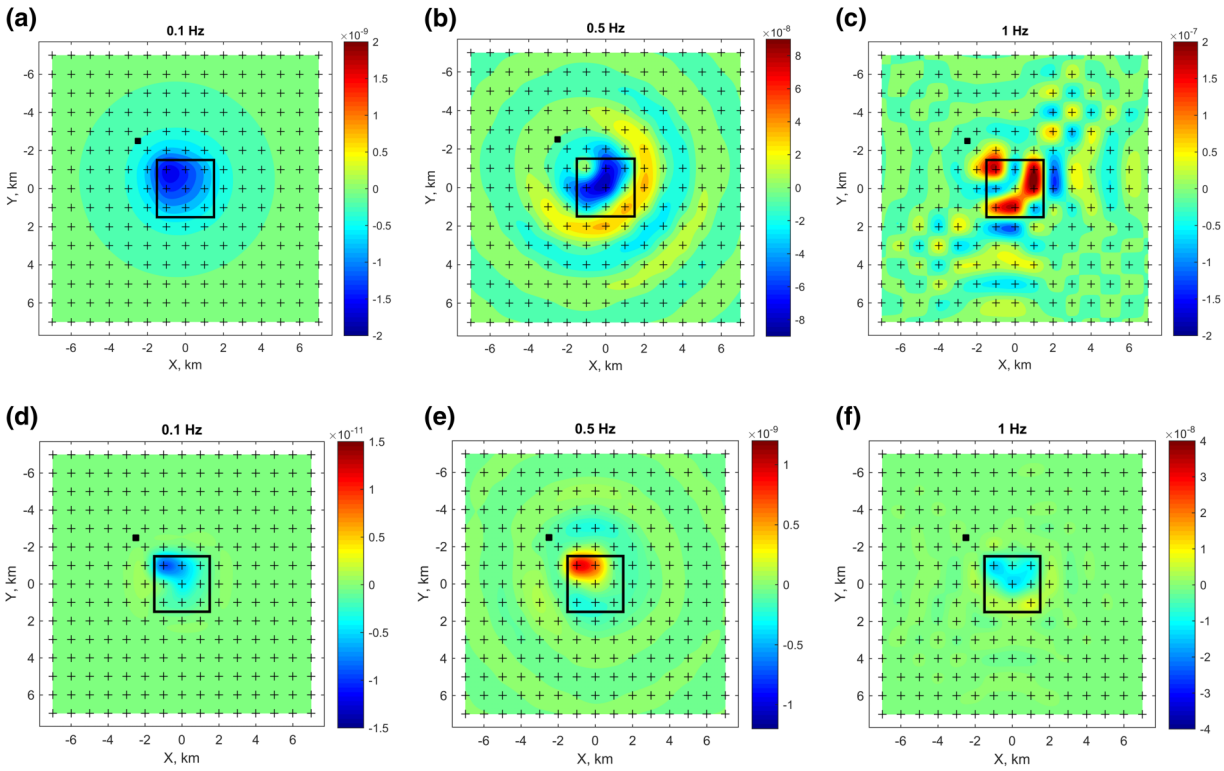

Figure 14 Final synthetic data at the end of the Gramian inversion with $\alpha_{\text{rel}} = 10^{-3}$ (iteration 23). (a) and (b) show the real part the final synthetic wavefield, $Re(p_a^{\text{end}})$ at three representative frequencies for a single source position. Panels (d)–(f) show the real part of the difference between the input and final acoustic fields, $Re(p_a^{\text{end}} - p_a)$. The data correspond to the model shown in Fig. 13(c).

We can also provide another derivation for $\nabla\gamma$, which helps to understand the structure of the resulting gradient vector. The Gramian can be rewritten as follows:

$$\gamma = (\mathbf{Qx}, \mathbf{W}_G\mathbf{Qx}), \tag{23}$$

where an auxiliary vector $\mathbf{x}$ combines both model vectors, $\mathbf{x} = \binom{\mathbf{m}}{\mathbf{s}}$; $\mathbf{Q}$ is the matrix of model transform, which can be presented in the following form:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{T} & \mathbf{O} \\ \mathbf{O} & \mathbf{T} \end{pmatrix}, \tag{24}$$

where $\mathbf{T}$ is defined above and $\mathbf{O}$ is $N \times N$ zero block; matrix $\mathbf{W}_G$ has the following representation:

$$\mathbf{W}_G = \begin{pmatrix} \phi_2\mathbf{I} & -\phi_3\mathbf{I} \\ -\phi_3\mathbf{I} & \phi_1\mathbf{I} \end{pmatrix}. \tag{25}$$

Matrix $\mathbf{W}_G$ is positive semi-definite and depends on $\mathbf{m}$ and $\mathbf{s}$. Gradient of γ can be expressed as follows:

$$\nabla\gamma = 2\mathbf{Q}^T\mathbf{W}_G\mathbf{Qx}. \tag{26}$$

Expressions (21) and (26) are identical. However, in (26) the model transform appears more isolated from the rest of computations allowing for clearer understanding of its impact. Using these results, we finally arrive at the following expression for the gradient, $\mathbf{g}$, of the parametric functional, P^α:

$$\mathbf{g} = \mathbf{F}_n^*\mathbf{W}^2(\mathbf{d} - \mathcal{A}(\mathbf{m})) + \alpha_n q_1 \mathbf{L}(\mathbf{m} - \mathbf{m}_0) + \alpha_n q_2 (\mathbf{m} - \mathbf{m}_0) + \alpha_n q_3 \mathbf{l}_G. \tag{27}$$

NUMERICAL EXPERIMENTS

We have conducted several numerical experiments to illustrate the effectiveness and robustness of the developed method. The seismic inversion was run in frequency domain, and the reference petrophysical model was presented by a conductivity section, as an example. We should note that any type of *a priori* information could be used to construct the reference model. In a case of petrophysical correlations between the seismic velocity and known petrophysical parameters, the

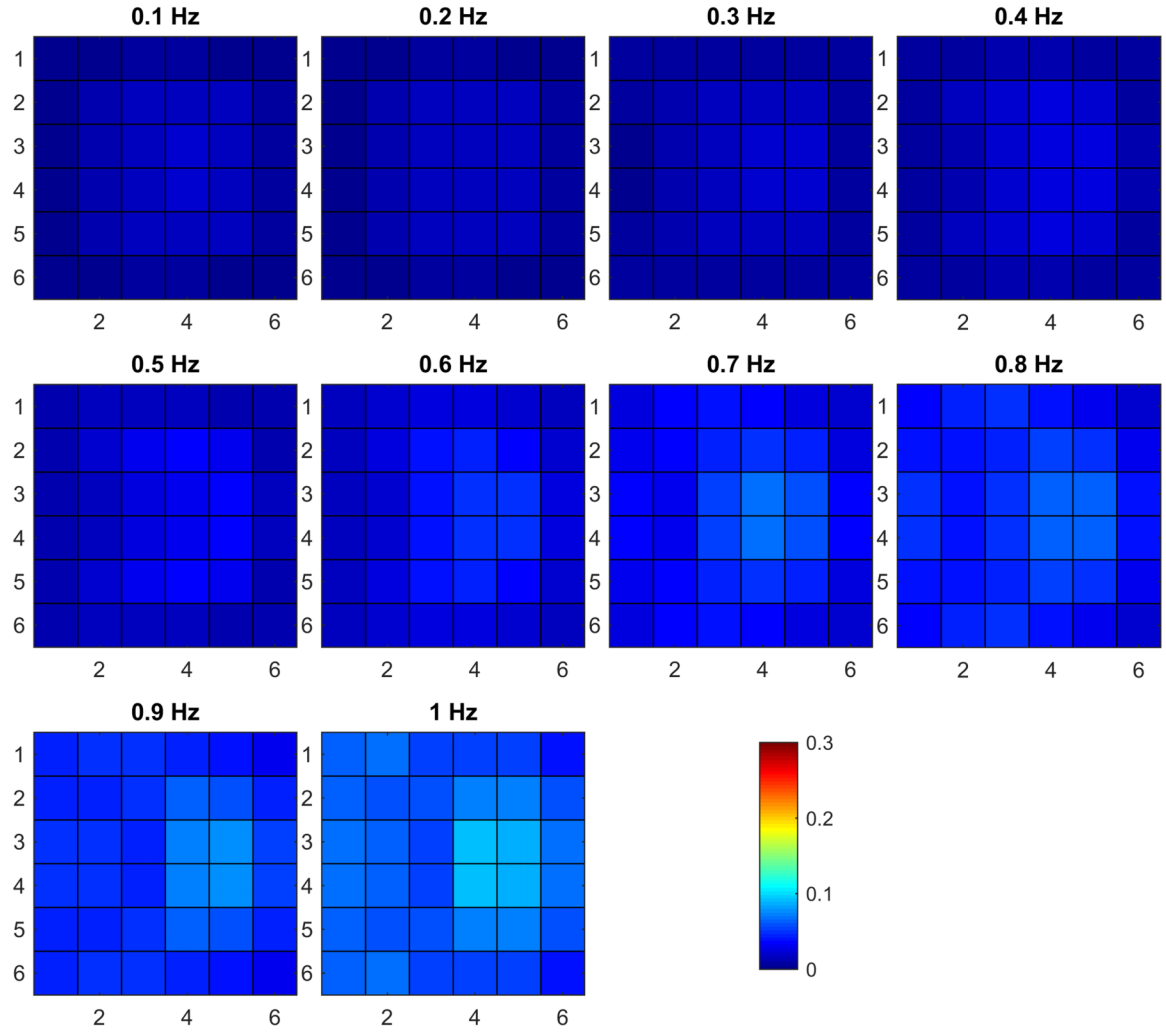

Figure 15 Normalized misfit at the end of the Gramian inversion with $\alpha_{\text{rel}} = 10^{-3}$, iteration 23. Each panel shows quantity $\|p_a^{\text{end}} - p_a\|/\|p_a\|$ for all 36 sources at a single frequency.

Gramian term would enforce these correlations, if they exist. In the same way, in a case of structural similarity between the interfaces of seismic model and known petrophysical model, the Gramian term would enforce these relationships.

The first numerical experiment

In the first numerical experiment, we studied the effect of the Gramian term on the reconstructed velocity model. We inverted a small model consisted of 512 cubical cells (Fig. 1a). The model had dimensions 1.5 km × 1.5 km × 1.5 km with its top at a depth of 500 m. The anomalous domain was immersed into a half space with $c_b = 4000$ m/s. The reference conductivity model is presented in Fig. 1(b). The models were adjusted to ensure the perfect linear correlation between S and χ according to the following formula:

$$S(\chi) = \frac{0.3010}{6.7521 \times 10^{-9}}(\chi - 6.7521 \times 10^{-9}). \tag{28}$$

This is the correlation the Gramian constraint would enforce. In the model shown in Fig. 1, every cell can take one of the two combinations of velocity and conductivity: (3800 m/s, 1 S/m) or (4000 m/s, 0.5 S/m) according to equation (28).

There were nine sources uniformly distributed in the square $-1500 \le x, y \le 1500$ m with a step of 1 km, and 36 receivers, distributed in the square $-1250 \le x, y \le 1250$ m with a step of 500 m. The source and receivers were placed at a depth of 50 m below the surface. The data were simulated at a frequency of 0.1 Hz.

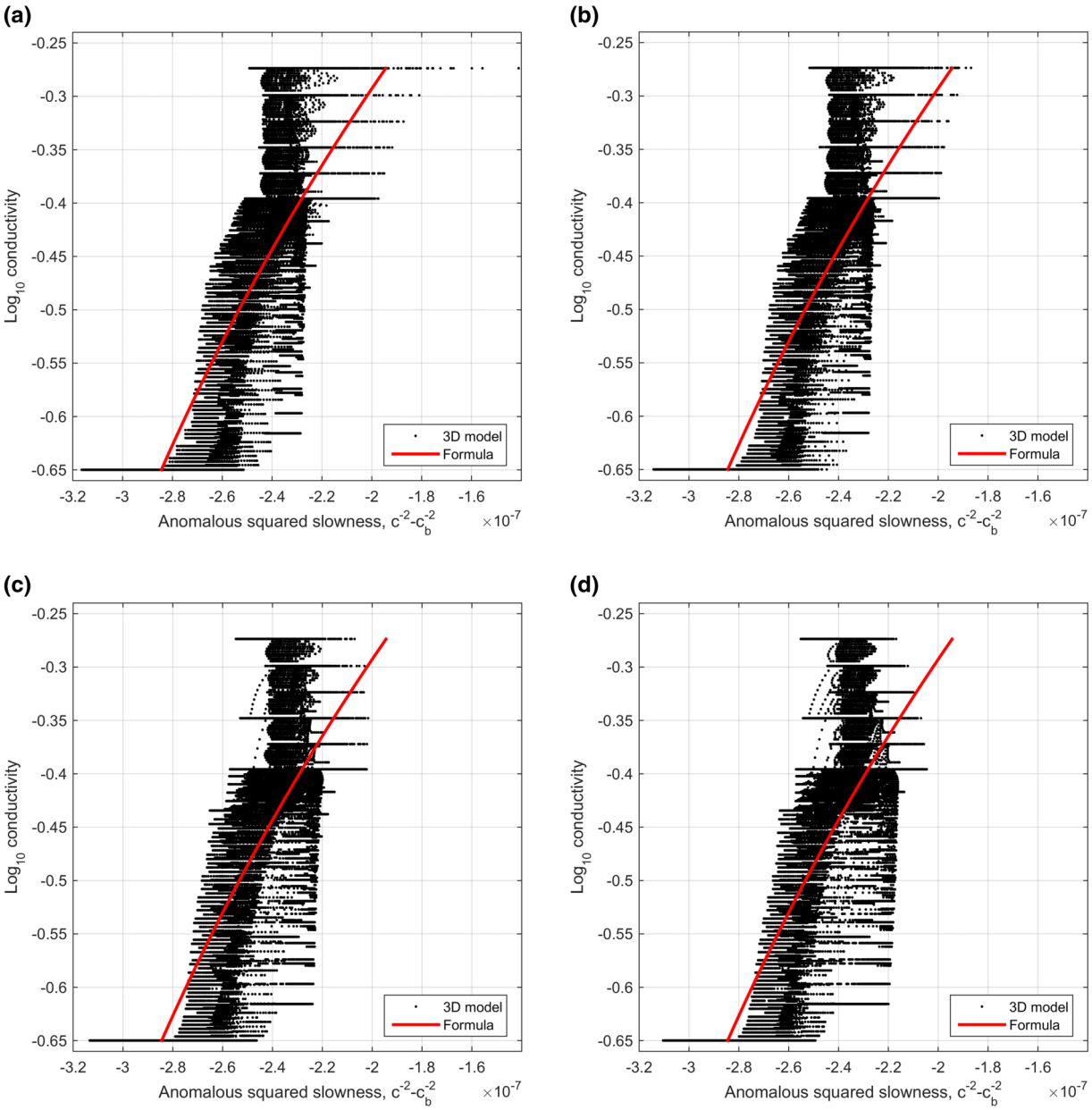

Figure 16 Cross-plots between the true conductivity model and inverse models, recovered by a separate inversion of seismic data only with different regularization parameters. (a) No regularization, $\alpha_{\text{rel}} = 0$, (b) smoothness $\alpha_{\text{rel}} = 10^{-4}$, (c) smoothness $\alpha_{\text{rel}} = 5 \times 10^{-4}$ and (d) smoothness $\alpha_{\text{rel}} = 10^{-3}$. The lines labelled as Formula correspond to relation (30).

The synthetic data for a single source are presented in Fig. 2(a).

Inversion started with the uniform velocity model of 3800 m/s. We assume the 5% noise in the input data. We did not use the smoothness term in the stabilizing functional, ψ, in this inversion, for simplicity. The inversion was terminated when RMS misfit is reached unity. We performed four inversion runs, with parameters summarized in Table 1.

The data at the end of inversion look quite similar since they fit the input data to a specified 5% tolerance. We presented the final data for inversion run #4 in Fig. 2(b,c). The final velocity distributions presented in Fig. 3 confirm that the Gramian constrains effectively guide velocity model during inversion. Figure 4 shows the cross plots of the inverted models. As the contribution of the Gramian term increases, the reconstructed velocities shift to their true values

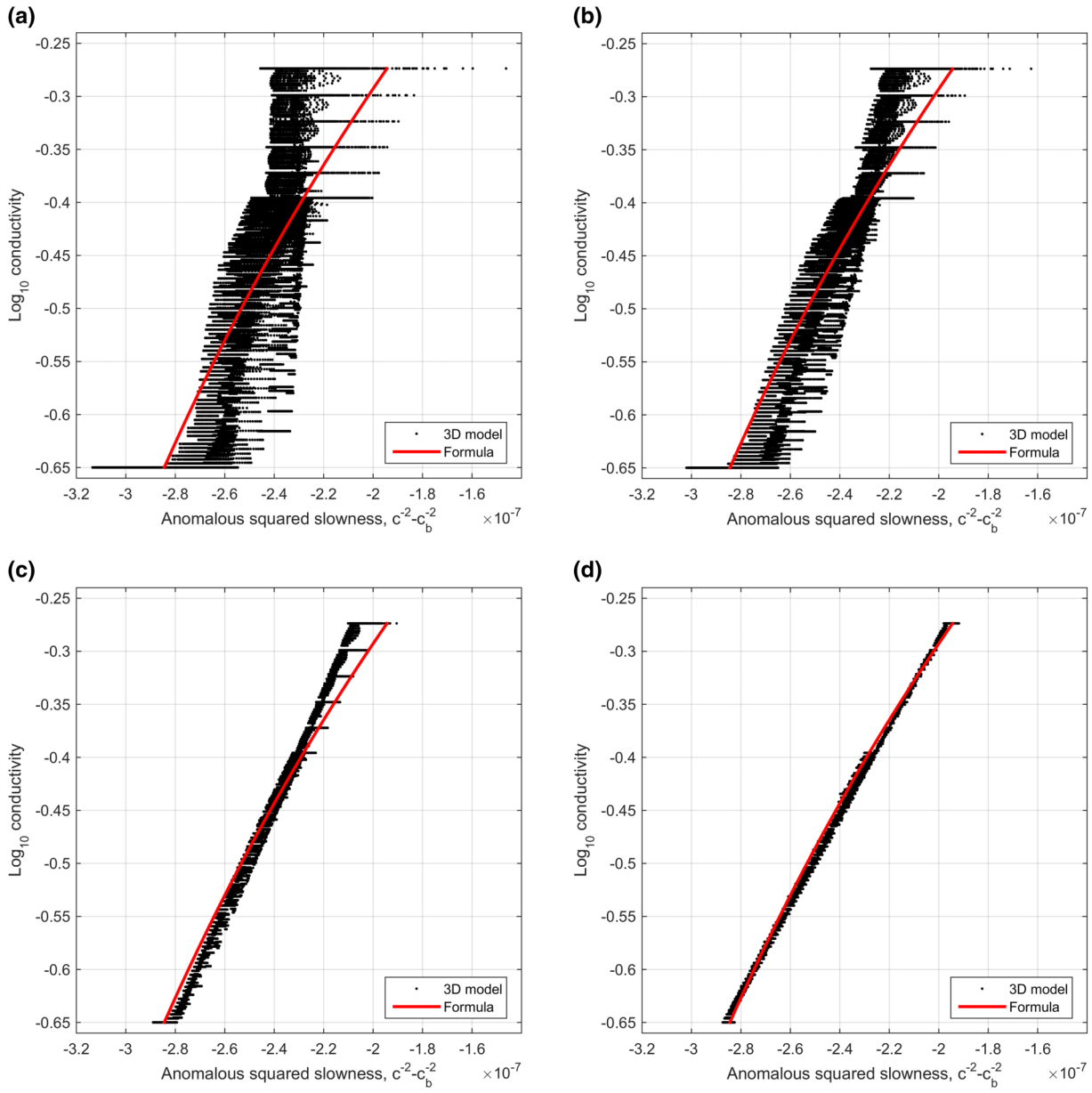

Figure 17 Cross plots between the true conductivity model and inversion models, recovered by the Gramian-based guided inversion with different regularization parameters. (a) Gramian $\alpha_{\text{rel}} = 10^{-4}$, (b) Gramian $\alpha_{\text{rel}} = 5 \times 10^{-4}$, (c) Gramian $\alpha_{\text{rel}} = 10^{-3}$ (iteration 23) and (d) Gramian $\alpha_{\text{rel}} = 10^{-3}$ (iteration 55). The lines labelled as Formula correspond to relation (30).

and the velocity spread in each compartment of the model decreases. We note that the velocities in the upper model compartment (at $\log_{10}\sigma = 0$) are biased from the true model, because the data have more impact on the shallow cells.

The second numerical experiment

In the second numerical experiment, we have simulated low-frequency seismic data and deliberately removed some transmitters to model a real-world situation where some area of the survey was inaccessable for some reason (e.g., diffcult terrain and private land) which resulted in a decrease of the spatial resolution of the data. To make our study more realistic, we have also assumed that the models are not perfectly correlated, and a smoothness term is enabled in the inversion. We inverted the seismic data for a model with 3 km × 3 km × 1.875 km anomalous domain, discretized into 64 × 64 × 40 cubical cells of 46.875 m each. The model contains a high-velocity block

at the right-bottom side, and a high-velocity prism at the left-middle side (Fig. 5a). The anomalous domain was immersed into a half space. Its top face was located at a depth of 200 m. The halfspace velocity was set to $c_b = 1500$ m/s.

The conductivity model, which was considered known during the inversion, is presented in Fig. 5(b). This model was constructed by smoothing velocities with the Gaussian filter and applying the following transform (Meju, Gallardo and Mohamed 2003):

$$\sigma = (c/c_{\text{ref}})^{-3.88}, \tag{29}$$

where $c_{\text{ref}} = 1700$ m/s.

The relation between the logarithmic conductivity, S, and squared anomalous slowness, χ, is given by the following formula:

$$S(\chi) = 3.88 \times \log_{10}\left(c_{\text{ref}}\sqrt{\chi + c_b^{-2}}\right). \tag{30}$$

This relation is not linear, but it is quite close to a linear one within the considered velocity range. The cross plot of the two true models is given in Fig. 6. The true velocity model consists of five compartments with constant velocity, whereas the conductivity model is smooth, which forms the five vertical segments on the plot.

The velocity model shown in Fig. 5(a) has been successfully reconstructed with the pure acoustic inversion in Malovichko *et al.* (2017a). In this study, we use a more sparsely sampled input data to mimic a common case, when the measured seismic data cannot resolve specific targets due to the lack of resolution. We have employed the Gramian constraint approach to obtain missing information from a known conductivity model, by imposing correlation between anomalous squired slowness and logarithmic conductivities.

There were 36 sources, distributed inside the XY square $-2750 \le x,\ y \le 2750$ m with a separation of 500 m, and 225 receivers distributed in the XY square $-7000 \le x,\ y \le 7000$ m , with a separation of 500 m (Fig. 7). The synthetic data were simulated at 10 frequencies from 0.1 Hz to 1 Hz with a step of 0.1 Hz. The spectrum of the signal in the source was formed by the Ricker wavelet with the central frequency of 0.5 Hz. During inversion, the level of data uncertainty was set to 5% of the amplitude of corresponding data point. The simulated data for a representative shot are plotted in Fig. 8. The checkerboard pattern, visible in the 1 Hz data, indicates a strong data aliasing due to a sparse receiver grid. Since the source grid is sparse too, it has a similar effect.

In the first set of inversions, the smoothness constraint only was active, that is $c_1 = 1$, $c_2 = c_3 = 0$. In all cases, inversion started from the uniform velocity model of 2300 m/s. We tried the following values of the regularization parameter: $\alpha_{\text{rel}} = 0$ (meaning no regularization), $\alpha_{\text{rel}} = 10^{-4}$, $\alpha_{\text{rel}} = 5 \times 10^{-4}$ and $\alpha_{\text{rel}} = 10^{-3}$. The data fit at the end of inversion is evaluated in Figs. 9 and 10.

Convergence plots and ratio $\alpha_n \psi_n / \phi_n$ are shown in Fig. 11 (see curves labelled by 'S'). As one could see, all inversions stagnated prematurely by being trapped in the local minima. Higher values of α_{rel} caused inversion to progress more slowly. However, the values of the final misfit were close to each other and approximately corresponded to ERMS = 3.

The final models (Fig. 12) look quite similar. All of them image the high-velocity block but show no indication of the high-velocity prism. The strong acquisition footprint, clearly visible in the inversion model with low regularization, is caused by aliasing in higher frequency data.

In the second set of inversions, the mutual Gramian constraint was enabled only, that is $c_1 = c_2 = 0$, $c_3 = 1$. We tested the following values of the regularization parameter: $\alpha_{\text{rel}} = 10^{-4}$, $\alpha_{\text{rel}} = 5 \times 10^{-4}$ and $\alpha_{\text{rel}} = 10^{-3}$. Convergence plots are presented in Fig. 11 (see curves labelled with letter 'G'). The final inversion models are presented in Fig. 13. Value $\alpha_{\text{rel}} = 10^{-4}$ caused the inversion to behave very similar to one with the smoothness regularization only. The misfit curve is close to that of the inversion with the smoothness regularization with $\alpha_{\text{rel}} = 10^{-4}$. The final model (Fig. 13a) is almost the same as well. This indicates that the inversion was primarily driven by the misfit with negligible contribution of regularization.

Value $\alpha_{\text{rel}} = 5 \times 10^{-4}$ produced a notable dispatch from this pattern. The convergence considerably accelerated (Fig. 11a, curve 'G, 5E-4'). The misfit curve on iteration 26 corresponded to ERMS = 1.83. The high-velocity prism is clearly visible in the final model (Fig. 13b). The inversion with regularization parameter $\alpha_{\text{rel}} = 10^{-3}$ performed even better. The misfit functional reached prescribed ERMS = 1 in 23 iterations (Fig. 11a, curve 'G, 1E-3'). After 55 iterations, the misfit reduced to 2×10^3. This corresponds to ERMS = 0.16, or, equivalently, to ERMS = 1 with 0.8% noise in the input data. The data fit for this inversion on iteration 23 is examined in Figs. 14 and 15.

The resolution of the final models was drastically improved comparing to the previous runs (Fig. 13c,d). The high-velocity block, as well as the high-velocity prism, was accurately imaged. The strong acquisition footprint has the same nature as in the previous test. We also note that the norm of the stabilizing functional behaves rather different in

cases of the smoothness or Gramian constraint. In the latter case, the norm of the model varies rapidly on the first iterations, as indicated by the zigzag pattern of ratio $\alpha_n \psi_n / \phi_n$ (Fig. 11b).

The cross plots, produced for the results of the separate inversion of seismic data only and of the Gramian-based guided inversion, are presented in Figs. 16 and 17. We observe that the Gramian constraint visibly improves correlation between the two parameter distributions by steering inversion towards the conductivity model. The correlation is improved with the increase of the value of the regularization parameter from 10^{-4} to 5×10^{-3} (see Fig. 17a–d)

CONCLUSIONS

We have developed a new strategy for integrating an *a priori* petrophysical model in the seismic full-waveform inversion by means of coupling between velocity and known petrophysical models via the Gramian constraints. Many inversion algorithms are based on an assumption that we know some analytical relationships between different physical properties of the rocks. However, in most cases, even if one has a linear relationship, the coefficients of this relationship are unknown unless one would conduct a petrophysical study of the rock samples from the boreholes in the surveyed areas. The advantage of the Gramian approach is that it does not require a specific knowledge of the coefficients of the analytic relationships and can produce this information as a result of the inversion. Indeed, the cross plot between the inverse model parameters presented in Fig. 17(d) provides almost exact reconstruction of analytical relationship described by formula (30). One can see from the results of this numerical experiment that the Gramian-based approach to a joint inversion makes it possible to reconstruct the form of a relationship between the different physical properties without any rock sample analysis. In other words, we can establish the relationships between different petrophysical properties from the remotely observed geophysical data. This is an important observation which may find a wide range of applications in geophysics.

We have presented an explicit expression of the gradient of the Gramian term in the case of two-parameter inversion, which can be easily incorporated in a standard optimization scheme. This approach opens a way for effective use of *a priori* data in seismic inversion. The numerical experiments suggest that our approach works reasonably well, at least on synthetic models. Future research will be focused on applying this method to more complex models and to field seismic data.

ACKNOWLEDGEMENTS

This research was supported by the Russian Science Foundation, project No. 16-11-10188. The authors acknowledge the support of the University of Utah Consortium for Electromagnetic Modeling and Inversion (CEMI) and TechnoImaging. This work has been carried out using computing resources of the federal collective usage center Complex for Simulation and Data Processing for Mega-science Facilities at NRC 'Kurchatov Institute', http://ckp.nrcki.ru/. We would like to acknowledge the Skoltech CDISE's high performance computing cluster, Zhores (Zacharov *et al.* 2019), for providing the computing resources that have contributed to the results reported herein.

ORCID

M. Malovichko https://orcid.org/0000-0002-7618-0998

REFERENCES

Čuma M. and Zhdanov M.S. 2017. Continental-scale joint inversion of Alaska and Yukon gravity and magnetic data. *First Break* **35**, 55–62.

De Stefano M., Andreasi F.G., Re S., Virgilio M. and Snyder F.F. 2011. Multi-domain, simultaneous joint inversion of geophysical data with application to subsalt imaging. *Geophysics* **76**, R69–R80.

Gallardo L.A. and Meju M.A. 2003a. Characterization of heterogeneous near-surface materials by joint 2D inversion of DC resistivity and seismic data. *Geophysical Research Letters* **30**, 1658–1661.

Gao J. and Zhang H. 2018. An efficient sequential strategy for realizing cross-gradient joint inversion: method and its application to 2-D cross borehole seismic traveltime and DC resistivity tomography. *Geophysical Journal International* **213**, 1044–1055.

Gessner K.L., Gallardo A., Wedin F. and Sener K. 2016. Crustal structure of the northern Menderes Massif, western Turkey, imaged by joint gravity and magnetic inversion. *International Journal of Earth Sciences* **105**, 2133–2148.

Gross L. 2019. Weighted cross-gradient function for joint inversion with the application to regional 3-D gravity and magnetic anomalies. *Geophysical Journal International* **217**, 2035–2046.

Hu W., Abubakar A. and Habashy T.M. 2009. Joint electromagnetic and seismic inversion using structural constraints. *Geophysics* **74**, R99–R109.

Jorgensen M. and Zhdanov M.S. 2019. Imaging Yellowstone magmatic system by the joint Gramian inversion of gravity and magnetotelluric data. *Physics of the Earth and Planetary Interiors* **292**, 12–20.

Lan T., Liu N., Han F. and Liu Q.H. 2019. Joint petrophysical and structural inversion of electromagnetic and seismic data based on volume integral equation method. *IEEE transactions on geoscience and remote sensing* **57**, 2075–2086.

Lin W. and Zhdanov M.S. 2018. Joint multinary inversion of gravity and magnetic data using Gramian constraints. *Geophysical Journal International* **215**, 1540–1557.

Lin W. and Zhdanov M.S. 2019. The Gramian method of joint inversion of the gravity gradiometry and seismic data. *Pure and Applied Geophysics* **176**, 1659–1672.

Malovichko M., Khokhlov N., Yavich N. and Zhdanov M. 2017a. Approximate solutions of acoustic 3D integral equation and their application to seismic modeling and full-waveform inversion. *Journal of Computational Physics* **346**, 318–339.

Malovichko M., Khokhlov N., Yavich N. and Zhdanov M. 2018. Acoustic 3D modeling by the method of integral equations. *Computers & Geosciences* **111**, 223–234.

Malovichko M., Tarasov A., Yavich N. and Zhdanov M. 2019. Mineral exploration with 3D controlled-source electromagnetic method: a synthetic study of Sukhoi Log gold deposit. *Geophysical Journal International* **219**, 1698–1716.

Malovichko M., Yavich N. and Zdanov M.S. 2017b. Integrating electrical conductivity in 3D seismic inversion with Gramian constraints. 79th EAGE Conference and Exhibition, 2017.

Manukyan E., Maurer H. and Nuber A. 2018. Improvements to elastic full-waveform inversion using cross-gradient constraints. *Geophysics* **83**, 1–46.

Meju M.A., Gallardo L.A. and Mohamed A.K. 2003. Evidence for correlation of electrical resistivity and seismic velocity in heterogeneous near-surface materials. *Geophysical Research Letters* **30**, 1373.

Molodtsov D.M., Troyan V.N., Roslov Yu V. and Zerilli A. 2013. Joint inversion of seismic traveltimes and magnetotelluric data with a directed structural constraint. *Geophysical Prospecting* **21**, 1218–1228.

Moorkamp M., Lelièvre P.G., Linde N. and Khan A. 2016. *Integrated Imaging of the Earth: Theory and Applications*. American Geophysical Union.

Ogunbo J.N. 2019. Mono-model parameter joint inversion by Gramian constraints: EM methods examples. *Earth and Space Science* **6**, 741–751.

Rittgers J.B., Revil A., Mooney M.A., Karaoulis M., Wodajo L. and Hickey C.J. 2016. Time-lapse joint inversion of geophysical data with automatic joint constraints and dynamic attributes. *Geophysical Journal International* **207**, 1401–1419.

Sirgue L. and Pratt R.G. 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies. *Geophysics* **69**, 231–248.

Um E.S., Commer M. and Newman G. 2014. A strategy for coupled 3D imaging of large-scale seismic and electromagnetic data sets: application to subsalt imaging. *Geophysics* **79**, ID1–ID13.

Zacharov I., Arslanov R., Gunin M., Stefonishin D., Bykov A., Pavlov S. *et al*. 2019. "Zhores" — Petaflops supercomputer for data-driven modeling, machine learning and artificial intelligence installed in Skolkovo Institute of Science and Technology. *Open Engineering* **9**, 512–520.

Zhang R., Li T., Deng X., Huang X. and Pak Y. 2019. Two-dimensional data-space joint inversion of magnetotelluric, gravity, magnetic and seismic data with cross-gradient constraints. *Geophysical Prospecting*. https://doi.org/10.1111/1365-2478.12858

Zhdanov M. and Cai H. 2017. Joint Inversion of Gravity and Magnetotelluric Data for the Depth-to-Basement Estimation. *IEEE Geoscience and Remote Sensing Letters* **14**, 1228–1232.

Zhdanov M.S. 2002. *Geophysical Inverse Theory and Regularization Problems*. Elsevier.

Zhdanov M.S. 2015. *Inverse Theory and Applications in Geophysics*. Elsevier.

Zhdanov M.S., Endo M., Čuma M., Sunwall D., Malmberg J.-A., McKay A. *et al*. 2016. Large-scale seismically guided anisotropic inversion of towed-streamer EM data in the Barents Sea. *First Break* **34**, 75–79.

Zhdanov M.S., Gribenko A. and Wilson G. 2012. Generalized joint inversion of multimodal geophysical data using Gramian constraints. *Geophysical Research Letters* **39**, L09301.

Zhu Y., Čuma M., Kinakin Y. and Zhdanov M.S. 2015. Joint inversion of airborne gravity, gradiometry and magnetic data from the Lac de Gras region of the Northwest Territories of Canada. 84th SEG Annual International meeting, Denver, CO.

Zhu Y. and Zhdanov M.S. 2015. Joint inversion of airborne magnetic and electromagnetic data: case study in the Northwest Territories of Canada. 85th SEG Annual International meeting, New Orleans, LA.

APPENDIX: REPRESENTATION OF THE GRAMIAN STABILIZING FUNCTIONAL AS A NON-NEGATIVE QUADRATIC FORM

The main advantage of the Gramian constraint (6) over some other approaches is that the stabilizing functional ψ, defined by equation (5), is convex given transformation T is linear. The convexity of $\gamma(\mathbf{m}, \mathbf{s})$ is a consequence of the following proposition (in which we assume T is the identity transform for simplicity).

Proposition. Given a real vector space equipped with Euclidian norm, E^M, and $\mathbf{c} \in E^M$, $\mathbf{c} \neq 0$, the Gramian term $\gamma(\mathbf{x}, \mathbf{c})$ is a positive semi-definite quadratic form on E^M, which has the following eigenvalues: $\lambda_1 = 0$, $\lambda_2 = .. = \lambda_M = \|\mathbf{c}\|^2$.

Proof. We may think of γ as a map $E^M \to \mathbb{R}$, since the second argument is fixed:

$$\gamma(\mathbf{x}) = \gamma(\mathbf{x}, \mathbf{c}) = \|\mathbf{x}\|^2 \|\mathbf{c}\|^2 - \left(\mathbf{x}^T \mathbf{c}\right)^2, \quad \mathbf{c} = \text{const}. \tag{A1}$$

Let us consider the following map $E^M \times E^M \to \mathbb{R}$:

$$B(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y} \|\mathbf{c}\|^2 - \left(\mathbf{x}^T \mathbf{c}\right)\left(\mathbf{y}^T \mathbf{c}\right), \tag{A2}$$

where $\mathbf{x}, \mathbf{y}, \mathbf{c} \in E^M$, $\mathbf{c} = \text{const}$. Clearly, B is a symmetric bilinear form, since $B(a\mathbf{x_1} + b\mathbf{x_1}, \mathbf{y}) = aB(\mathbf{x_1}, \mathbf{y}) + bB(\mathbf{x_2}, \mathbf{y})$,

$B(\mathbf{x}, a\mathbf{y}_1 + b\mathbf{y}_2) = aB(\mathbf{x}, \mathbf{y}_1) + bB(\mathbf{x}, \mathbf{y}_2)$ and $B(\mathbf{x}, \mathbf{y}) = B(\mathbf{y}, \mathbf{x})$ for any $\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}, \mathbf{y}_1, \mathbf{y}_2 \in E^M$ and $a, b \in \mathbb{R}$. The Gramian term can be expressed as follows:

$$\gamma(\mathbf{x}) = B(\mathbf{x}, \mathbf{x}). \tag{A3}$$

Thus, $\gamma(\mathbf{x})$ is a quadratic form. It means it can be expressed as follows:

$$\gamma(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} \tag{A4}$$

for some matrix $\mathbf{A} \in \mathbb{R}^{M \times M}$. Specifically,

$$\mathbf{A} = \mathbf{I}\|\mathbf{c}\|^2 - \mathbf{c}\mathbf{c}^T. \tag{A5}$$

The first term is a diagonal matrix having M eigenvalues equal to $\|\mathbf{c}\|^2$. The second term is a rank-one matrix having one non-zero eigenvalue equal to $\|\mathbf{c}\|^2$. To see this, let us denote by $\boldsymbol{\xi}_1 = \mathbf{c}/\|\mathbf{c}\|$ the basis vector in the subspace spanned by $\mathbf{c}$. It is associated with eigenvalue $\lambda_1 = \|\mathbf{c}\|^2$, since $\mathbf{c}\mathbf{c}^T\boldsymbol{\xi}_1 = \mathbf{c}\|\mathbf{c}\| = \boldsymbol{\xi}_1\|\mathbf{c}\|^2$. Let us further denote by $\{\boldsymbol{\xi}_2, .., \boldsymbol{\xi}_N\}$ the orthonormal basis of the orthogonal complement of $\mathbf{c}$. Since $\mathbf{c}\mathbf{c}^T\boldsymbol{\xi}_i = 0$ for all $i = 2, \ldots, M$, eigenvectors $\{\boldsymbol{\xi}_2, .., \boldsymbol{\xi}_M\}$ are associated with $M - 1$ zero eigenvalues. Eigenvalues are invariant to the change of the basis, so $\mathbf{A}$ has one zero eigenvalue, and $(M - 1)$ positive eigenvalues equal to $\|\mathbf{c}\|^2$. ■

Thus, ψ is convex as a sum of three convex functionals with non-negative weights.