Inter-parameter tradeoff quantification and reduction in isotropic-elastic full-waveform inversion

Wenyong Pan, Kristopher A. Innanen, Department of Geoscience, CREWES Project, University of Calgary, Yanhua O. Yuan, Department of Geoscience, Princeton University

SUMMARY

Quantifying inter-parameter tradeoffs (i.e., cross-talk artifacts) and reducing the introduced uncertainties are becoming increasingly important for multiparameter full-waveform inversion (FWI). Parameter resolution studies based on scattering patterns may neglect important information for understanding inter-parameter tradeoffs. Our objective is to evaluate and reduce the inter-parameter tradeoff errors experienced by isotropic-elastic FWI using multiparameter Hessian. We introduce the inter-parameter contamination kernel as a tool to expose the origin and characters of these tradeoffs. By applying the multiparameter Hessian to various types of diagnostic vectors, we are able to quantify interparameter tradeoffs either locally or within the whole volume, i.e., considering finite-frequency effects in complex and heterogeneous models, etc. We observe that S-wave velocity perturbations produce strong contaminations into density and P-wave velocity updates, but suffer little contaminations from other parameters. Based on our findings, we subtract approximated contamination kernels, constructed by applying off-diagonal Hessian blocks to estimated model perturbation vectors, from the standard sensitivity kernels. Numerical examples show that this approach increases the reliability of certain cross-talk prone model estimations, for instance density in isotropic-elastic FWI.

INTRODUCTION

Simultaneously reconstructing multiple physical parameters in seismic FWI suffers from interparameter tradeoffs, which reflects the inability to distinguish the influence of one from another elastic property (Tarantola, 1986; Operto et al., 2013; Innanen, 2014). Perturbations in one physical parameter are therefore in danger of being mapped into the estimation of another. Evaluating inter-parameter tradeoffs and reducing the introduced uncertainties are essential for multiparameter FWI.

Researchers have devoted intensive efforts to parameter resolution studies, usually basing them on analytic solutions of Fréchet derivative wavefields for different parameter classes (Tarantola, 1986; Golghali et al., 2013; Alkhalifa and Plessix, 2014). Scattering patterns provide invaluable information for understanding the inter-parameter coupling effects, but at the same time they neglect some important types of information (Podgornova et al., 2015). Conclusions drawn from the overlapping of scattering patterns hold asymptotically (Operto et al., 2013), meaning inter-parameter tradeoffs originating outside of a ray-approximate framework are neglected.

One objective of this research is to quantify inter-parameter tradeoffs in isotropic-elastic FWI by analyzing the multiparameter Hessian, which provides complete characterization of these issues (Fichtner and van Leeuwen, 2015). We find that the products of the off-diagonal blocks of multiparameter Hessian with the model perturbation vectors, referred to as inter-parameter contamination kernels in this paper, account for the inter-parameter tradeoffs encountered in a given FWI problem. For large-scale inverse problems, characteristics of the Hessian can be inferred via matrix probing techniques, which are useful when explicit representation of a matrix is too expensive (Trampert et al., 2013). The product of the Hessian with a point-localized vector generates one Hessian column, which is referred to as the point spread function (Hu et al., 2001; Valenciano et al., 2006; Tang and Lee, 2015). Fichtner and Trampert (2011) evaluated local inter-parameter tradeoffs in FWI using multiparameter point spread functions (MPSFs). To evaluate the strengths of the inter-parameter tradeoffs within the whole volume, we adopt a stochastic probing strategy, in which the multiparameter Hessian operator is applied to random vectors. Expectation values of the correlation between these random vectors with their corresponding Hessian-vector products approximate the diagonals of multiparameter Hessian off-diagonal blocks, which measure parameter coupling strengths within the whole volume. To quantify inter-parameter tradeoffs with off-diagonal elements of the off-diagonal blocks, the multiparameter Hessian is further applied to vectors with regularly distributed spikes. The resulting MPSF volumes approximate the row summations of multiparameter Hessian.

Due to strong contaminations from velocity parameters, density structures are poorly constrained in isotropic-elastic FWI (Jeong et al., 2013; Alkhalifa and Plessix, 2014). Scattering patterns provide approximate framework are neglected.

METHODS

Isotropic-elastic FWI review

In seismic FWI a misfit function which measures the differences between seismic observations and synthetic data is defined:

$$
\Phi(m) = \sum_{j=1}^{S} \sum_{i=1}^{R} \int_{0}^{T} \| \Delta d(x_j, x_m, t; m) \|^2 dt,
$$

where $\Delta d(x_j, x_m, t; m)$ are the data residuals. Within the Newton optimization framework, the search direction $\Delta m$ is obtained by solving the Newton linear system: $H \Delta m = -V m \Phi_q$. Here, $V m \Phi$ and $H$ are the gradient and Hessian, which represent the first and second derivatives of the misfit function. In the adjoint-state method, explicit expressions for the $\kappa$, $\mu$ and $\rho$ sensitivity kernels are written as (Tromp et al., 2005; Yuan and Simons, 2014):

$$
K_\kappa(x) = -\langle x | V \cdot \nabla \cdot \mathbf{u} (x, x, T - t) \cdot V \cdot \mathbf{u} (x, x, t) \rangle,
$$

$$
K_\mu(x) = -2 \mu \langle x | D (x, x, T - t) : D (x, x, t) \rangle,
$$

$$
K_\rho(x) = -\langle \rho \cdot x | \delta^2 \delta \mathbf{u} (x, x, t) \rangle,
$$

where $\mathbf{u}$ represents the adjoint waveform, $D$ is the traceless strain deviator and $\langle \cdot \rangle$ represents summation over source, receiver and time. Sensitivity kernels for $\alpha$, $\beta$ and $\rho'$ are given by:

$$
K_\alpha = 2 \left(1 + \frac{4 \mu}{3 \kappa}\right) K_\kappa, \quad K_\beta = 2 \left(1 + \frac{4 \mu}{3 \kappa}\right) K_\kappa, \quad K_\rho' = K_\kappa + K_\beta.
$$

Inter-parameter contamination kernels

The sensitivity kernel $K_\alpha$ in the Newton system of isotropic-elastic FWI can be written as (Fichtner and van Leeuwen, 2015):

$$
K_\alpha = K_{\alpha,\alpha,\alpha} + K_{\beta,\beta,\alpha} + K_{\rho',\rho',\alpha} = -H_{\alpha,\alpha} \Delta m_{\alpha} - H_{\beta,\beta} \Delta m_{\beta} - H_{\rho',\rho'} \Delta m_{\rho'},
$$
where $K_{\alpha \rightarrow \alpha}$ is the correct update kernel for $\alpha$, and the terms $K_{\beta \rightarrow \alpha}$ and $K_{\rho \rightarrow \alpha}$ are defined as inter-parameter contamination kernels. The model perturbation vectors $\Delta \mathbf{m}_\alpha$ and $\Delta \mathbf{m}_\rho$, blurred by off-diagonal blocks $H_{\alpha \alpha}$ and $H_{\alpha \rho}$, are mapped into the update for parameter $\alpha$. Similarly, sensitivity kernels $K_{\beta \rho}$ and $K_{\rho \beta \rightarrow \rho}$ can be written as:

$$
K_{\beta \rho} = K_{\beta \rightarrow \rho} + K_{\beta \rightarrow \rho} + K_{\rho \rightarrow \rho} = -H_{\beta\rho} \Delta \mathbf{m}_\beta - H_{\beta\rho} \Delta \mathbf{m}_\rho,
$$

$$
K_{\rho \beta} = K_{\rho \rightarrow \beta} + K_{\rho \rightarrow \beta} + K_{\beta \rightarrow \beta} = -H_{\rho\beta} \Delta \mathbf{m}_\rho - H_{\rho\beta} \Delta \mathbf{m}_\beta,
$$

where $K_{\beta \rightarrow \rho}$ and $K_{\rho \rightarrow \beta}$ are the correct update kernels for $\beta$ and $\rho$, $K_{\beta \rightarrow \beta}$ and $K_{\rho \rightarrow \rho}$ involving off-diagonal blocks $H_{\alpha \beta}$ and $H_{\beta \rho}$, measure contaminations from $\alpha$ and $\beta$ into $\rho$, $\beta$ into $\alpha$, and $\beta$ into $\rho$.

Quantifying inter-parameter tradeoffs

Local inter-parameter tradeoffs

Selecting model perturbations $\Delta \mathbf{m}_0 = 0$, $\Delta \mathbf{m}_\rho = 0$ and $\Delta \mathbf{m}_\beta$, the gradient update $K_{\beta}$ in equation (7) becomes:

$$
K_{\beta} (x) = -A_{\beta} \int_{\Omega(x)} H_{\beta\beta} (x,x') \delta (x' - z) \, dx' = -A_{\beta} H_{\beta\beta} (x,x),
$$

where $K_{\beta} (x)$ represents an conservative estimate of spike model perturbation blurred by diagonal block $H_{\beta\beta}$. The sensitivity kernels $K_{\beta \rho}$ and $K_{\rho \beta}$, due to this point-localized perturbation can likewise be expressed as:

$$
K_{\rho} (x) = -A_{\rho} H_{\rho\beta} (x,x), 
$$

$$
K_{\beta \rho} (x) = -A_{\beta} H_{\beta\rho} (x,x).
$$

The preserved multiparameter Hessianian $H_{\beta\rho} (x,x)$ is named as multiparameter point spread function (MPSF) following the convention in exploration geophysics. Applying the Hessian to spike model perturbations $\Delta \mathbf{m}_0 = A_{\alpha} \delta (x - z)$ or $\Delta \mathbf{m}_\rho = A_{\rho} \delta (x - z)$ allows us to calculate the MPSFs $H_{\beta\alpha} (x,x)$, $H_{\rho\alpha} (x,x)$, $H_{\alpha\rho} (x,x)$, and $H_{\beta\rho} (x,x)$, which describe the local contaminations from $\alpha$ to $\beta$ and $\rho$ and the contaminations from $\rho$ to $\alpha$ and $\beta$.

Stochastic probing and MPSF volumes

If the Hessian operator is diagonally dominant, expectation values of the correlations between a zero-mean random vector $\mathbf{v}$ and its Hessian-vector product $\mathbf{H} \mathbf{v}$ approximates the Hessian diagonals (Sacchi et al., 2007). In a multiparameter inverse problem, the random vector $\mathbf{v}$ can be partitioned into $N_p$ subvectors, and the multiparameter Hessian can be likewise divided into $N_p \times N_p$ subblock matrices. Applying the multiparameter Hessian to the random vector products gives $N_p$ Hessian-vector subproducts. Diagonals of the Hessian subblock matrices can then be estimated by:

$$
H_{\alpha\beta} = E [v_\alpha \otimes \delta_\beta] = E [v_\alpha \otimes \sum_{p=1}^{N_p} H_{\alpha p} v_p],
$$

where $\otimes$ means element-wise multiplication, $E$ denotes the expectation operation, $p$ and $q$ are indices for subvectors representing different physical parameters, and $\delta_\beta$ is the sub-Hessian-vector product. Assuming independent zero-mean random vectors $v_\alpha$ and $v_\beta$ for two different physical parameters, $E [v_\alpha (x) v_\beta (x')] = 0$, equation (11) becomes:

$$
H_{\alpha\beta} = \sum_{p=1}^{N_p} H_{\alpha p} E [v_p v_p] = H_{\alpha p} E [v_p v_p].
$$

For example, in isotropic-elastic FWI, diagonal elements of the off-diagonal block $H_{\alpha \rho}$ can be estimated by:

$$
H_{\alpha \rho} = \sum_{m=1}^{NR} v_{\alpha, m} \otimes H_{\alpha \rho, m} v_{\rho, m} + \sum_{m=1}^{NR} v_{\beta, m} \otimes v_{\rho, m},
$$

where $\otimes$ means element-wise division, $nr = [1, ..., NR]$ indicates the random vector index. If the subblocks of multiparameter Hessian are highly diagonally-dominant, a small number of random probes are needed. To measure the inter-parameter tradeoffs with off-diagonal elements in off-diagonal blocks, the multiparameter Hessian is applied to vectors with discrete spikes regularly distributed in the whole volume. The resulting MPSF volumes approximate row summations of multiparameter Hessian.

Reducing cross-talks with approximated contamination kernels

In equation (8), the sensitivity kernel $K_{\rho}$ is a linear summation of the correct update kernel $K_{\beta \rightarrow \rho}$ with the contamination kernels. We have observed, during parameter tradeoff analysis in the numerical modelling section, that density suffers contaminations primarily from S-wave velocity, and S-wave velocity is less contaminated by other parameters. This is suggestive that we can first update the S-wave velocity iteratively for a finite number of $k'$ iterations, which provides an estimated model $\mathbf{m}_0'$. The inversion is then started from initial models by simultaneously updating the three parameters. At the $k$th iteration, approximated contamination kernels are constructed:

$$
\tilde{K}_{\beta \rightarrow \rho} (x) = -\int_{\Omega(x')} H_{\beta\rho} (x,x') \Delta \tilde{m}_\rho' (x') \, dx' = \int_{\Omega(x')} H_{\beta\rho} (x,x') \Delta \tilde{m}_\rho' (x') \, dx',
$$

$$
\tilde{K}_{\rho \beta} (x) = -\int_{\Omega(x')} H_{\rho\beta} (x,x') \Delta \tilde{m}_\rho' (x') \, dx' = \int_{\Omega(x')} H_{\rho\beta} (x,x') \Delta \tilde{m}_\rho' (x') \, dx',
$$

where $\Delta \tilde{m}_\rho = \mathbf{m}_0' - \mathbf{m}_k'$ is the approximated model perturbation vector. The new update kernels for $\alpha$, $\beta$ and $\rho$ are given by:

$$
\tilde{K}_{\alpha \beta} = K_{\alpha \beta} - \tilde{K}_{\beta \rightarrow \rho} \tilde{K}_{\rho \beta}^{-1} \tilde{K}_{\beta \rightarrow \rho} - \tilde{K}_{\beta \rightarrow \rho} \tilde{K}_{\beta \rightarrow \rho}^{-1} \tilde{K}_{\beta \rightarrow \rho}.
$$

Subtracting the approximated contamination kernels from the sensitivity kernels, $\tilde{K}_{\alpha \beta}$ and $\tilde{K}_{\rho \beta}$, suppresses the contaminations partially. Better approximations of the model perturbation vector $\Delta \tilde{m}_\rho$ would remove contaminations more completely, but would be more computationally expensive.

NUMERICAL EXAMPLES

Multiparameter point spread functions

We use a 2D, 1km x 1km homogeneous and isotropic-elastic model to test the idea of quantifying parameter tradeoffs with multiparameter point spread functions (MPSFs). P-wave, S-wave velocity and density are 2000 m/s, 1400 m/s and 1.2 kg/m$^3$, and a P-SV source with a Ricker wavelet ($f_{dom} = 8$ Hz) is used for modelling. A total of 60 sources...
Interparameter tradeoffs quantification and reduction

Figure 2: Panels (a-c) show the true P-wave velocity, S-wave velocity and density of the Gaussian anomaly models: $m_{p}^{\text{true}}$, $m_{s}^{\text{true}}$ and $m_{\rho}^{\text{true}}$.

Figure 3: Panels (a-c) sensitivity kernels $K_{\alpha}$, $K_{\beta}$ and $K_{\rho}^{\prime}$; Panels (d-f) show estimated model vector $\Delta \mathbf{m}_{\alpha}$ and approximated contamination kernels $\tilde{K}_{\beta \rightarrow \alpha}$ and $\tilde{K}_{\beta \rightarrow \rho'}$; Panels (g-i) new update kernels $\tilde{K}_{\alpha}$, $\tilde{K}_{\beta}$ and $\tilde{K}_{\rho'}$.

and 200 receivers are arranged regularly along all boundaries of the model. We first apply a positive spike perturbation in P-wave velocity at position $x$ (center of the model): $\Delta \mathbf{m}_{\alpha}(x) = 100$ m/s. MPSFs $H_{\alpha\beta}(x,z)$, $H_{\beta\rho}(x,z)$, and $H_{\beta\rho'}(x,z)$ are calculated, where $H_{\beta\rho}(x,z)$ and $H_{\beta\rho'}(x,z)$ describe the mappings from $\alpha$ to $\beta$ and $\rho'$. Then, spike perturbations $\Delta \mathbf{m}_{\alpha}(x) = 100$ m/s and $\Delta \mathbf{m}_{\rho'}(x) = 100$ kg/m$^3$ are applied respectively. MPSFs $H_{\alpha\beta}(x,z)$, $H_{\beta\rho}(x,z)$, $H_{\beta\rho'}(x,z)$, $H_{\beta\rho'}(x,z)$, $H_{\beta\rho'}(x,z)$, $H_{\beta\rho'}(x,z)$ are obtained.

These MPSFs are arranged in a block structure which is consistent with their positions in the multiparameter Hessian, as shown in Figure 1. Magnitudes of the MPSFs differ significantly. Contaminations from $\alpha$ to $\beta$ and $\rho'$ appear to be relatively weak. An S-wave velocity $\beta$ perturbation produces strong negative mapping to $\alpha$ and positive mapping to $\rho'$. A density $\rho'$ perturbation also produces unwanted artifacts in $\alpha$ and $\beta$. S-wave velocity $\beta$ experiences the least contaminations of the three parameters. These contaminations may result in density being highly under- or overestimated. Figures 2a, 2b, and 2c show the true model perturbations.

Figure 4: Panels (a-c) estimated models using a standard simultaneous inversion strategy; Panels (d-f) estimated models with the cross-talk suppression strategy.

Figure 5: Panel (a-c) show true P-wave velocity, S-wave velocity and density models; Panels (d-f) show the corresponding initial models; Panels (g-i) show the true model perturbations.

The P-wave velocity and density however have strong cross-talk artifacts leaking from the S-wave velocity. Figures 4d, 4e and 4f plot the inverted models created using the new inversion strategy with approximated contamination kernels. As indicated by the arrows, the contaminations due S-wave velocity in the inverted P-wave and density models have been significantly suppressed.

Marmousi model example

In Figure 5, the true P-wave and S-wave velocity and density models for a more complex Marmousi model are illustrated. The second row and third row in Figure 5 are the corresponding initial models and model perturbations. We deploy sources and receivers regularly along the top surface of the model.

The stochastic probing approach is applied to estimate the diagonals of the subblock matrices in the multiparameter Hessian with two independent random vectors. The first row in Figure 6 shows the diagonals of the diagonal blocks $H_{\alpha\alpha}^{\beta\beta}$, $H_{\beta\beta}^{\beta\beta}$ and $H_{\rho\rho}^{\beta\beta}$ which can be used as preconditioners in the inversion process (Pan et al., 2014, 2015). The second row in Figure 6 illustrates the diagonals of the off-diagonal blocks $H_{\alpha\beta}^{\beta\rho}$, $H_{\beta\rho}^{\beta\rho}$ and $H_{\rho\rho}^{\beta\rho}$ which indicate the coupling strengths of the isotropic-elastic parameters within the whole volume. Similarity of $H_{\alpha\beta}^{\beta\rho}$ and $H_{\beta\rho}^{\beta\rho}$ indicates that contaminations from $\beta$ to $\rho'$ are strong.

A vector $\mathbf{v}$ consisting of densely distributed spikes with a constant magnitude of 100 is designed. The corresponding multiparameter Hessian-vector products are shown in Figure 7. The strengths of inter-parameter contaminations generally match our predictions with multiparameter Hessian diagonals. Comparing strengths of the off-diagonal Hessian-vector products (i.e., $O_{\alpha\beta}$) with those of the diagonal Hessian-vector products (i.e., $O_{\alpha\alpha}$), we conclude that contaminations from $\alpha$ to $\beta$ and $\rho'$ are relatively weak and can be ignored. Contaminations from $\rho'$ to $\alpha$ and $\beta$ are also relatively weak. However, contamination from $\beta$ to $\alpha$ could potentially decrease the $\alpha$ update by as much as 20%, and contamination from $\beta$ to $\rho'$ may increase the $\rho'$ update by as much as 2.8 times.
Interparameter tradeoffs quantification and reduction

To verify our predictions, we calculate the true contamination kernels by applying the multiparameter Hessian to the true model perturbation vectors $\Delta m_\alpha$, $\Delta m_\beta$ and $\Delta m_\rho$. Figure 8 illustrates the sensitivity kernels, the correct update kernels, and the inter-parameter contamination kernels respectively. The Hessian diagonals shown in Figure 6 and the Hessian-vector products shown in Figure 7 generally reproduce (i.e., predict) the energy distribution and relative strengths of the inter-parameter tradeoffs. The correct update kernel $K_{\beta\rightarrow\rho}$ will be decreased by the S-wave contamination kernel. The correct update $K_{\beta\rightarrow\rho}$ exhibits limited contaminations from $\alpha$ and $\rho'$. The contamination kernel $K_{\rho'\rightarrow\rho}$ is $\approx 1.7$ times stronger than the correct update kernel $K_{\rho'\rightarrow\rho}$ for density. The estimated density structures are dominated by contamination from S-wave velocity.

To mitigate the contaminations of S-wave velocity to other parameters, we first update the S-wave velocity with 15 iterations, which gives an estimated model perturbation $\Delta m_\beta$, as shown in Figure 9a. The approximated contamination kernels $\tilde{K}_{\beta\rightarrow\alpha}$ and $\tilde{K}_{\beta\rightarrow\rho}$ constructed with the estimated model perturbation vector are very close to the true contamination kernels $K_{\beta\rightarrow\alpha}$ and $K_{\beta\rightarrow\rho}$. The second row in Figure 9 show the new update kernels $\tilde{K}_\alpha$, $\tilde{K}_\beta$ and $\tilde{K}_{\rho'}$ respectively. The new update kernel $\tilde{K}_\rho'$ is very close to the correct update kernel $K_{\rho'\rightarrow\rho}$ shown in Figure 8. Figure 10a, 10b and 10c show the inverted P-wave, S-wave and density models using traditional simultaneous inversion strategy. The inverted density model is highly under- or overestimated due to the leakages from S-wave velocity, as indicated by the arrows in Figure 10c. Figure 10d, 10e and 10f show the inverted P-wave, S-wave and density models with approximated contamination kernels. Figure 11 is a well log comparison of the inverted models. The contaminations in inverted density model are significantly suppressed.

CONCLUSIONS

We present strategies for quantifying inter-parameter tradeoffs in isotropic-elastic FWI with multiparameter Hessian-vector products. We propose a strategy to reduce cross-talk artifacts with approximated contamination kernels, by applying multiparameter Hessian off-diagonal blocks to estimated model perturbation vectors; this proves effective in removing contaminations in estimated density structures.

ACKNOWLEDGMENTS

This research was supported by CREWES Consortium and National Science and Engineering Research Council of Canada (NSERC, CRDPJ 461179-13). Wenyong Pan is also supported by SEGChevron scholarship and Eyes High International Doctoral Scholarship. Thanks also to Tiger cluster (Princeton University) and Lattice cluster (Compute Canada) for providing parallel computing environment. Thanks greatly to Frederik Simons, Andreas Fichtner, Samual Gray, Daniel Trad, Yu Geng, Junxiao Li, Hassan Khaniani, Youyi Ruan, Wenjie Lei, and Ryan Modrak for valuable discussions.