A new method for constructing the Hessian in elastic full waveform inversion

Vegard Stenhjem Hagen and Børge Arntsen, Norwegian University of Science and Technology;*
Espen Birger Raknes, Aker BP and Norwegian University of Science and Technology

SUMMARY

We propose a method for constructing the Hessian in Elastic Full Waveform Inversion (EFWI) as a series of Hessian-vector products with model perturbations. The assembled Hessian of both simple and complex synthetic models are studied to draw conclusions regarding spatial resolution and parameter retrieval. The results indicate a strong cross-talk between the different model parameters of a common density-velocity parametrisation in EFWI. Furthermore we are able to comment on the influence of offset by comparing zero- and intermediate-offset results.

INTRODUCTION

The ill-posedness of EFWI introduces local minima which must be avoided to perform successful recovery of subsurface parameters (Operto et al., 2013). Commonly, the gradient update in EFWI is estimated by linearising the problem and approximating the Hessian with a scalar as a way of eliminating some computational complexity, but at the cost of information loss. We propose calculating the action of the Hessian matrix on a subset of the model (Hessian-vector products) as a way of estimating the resolution and parameter cross-talk in a defined region, helping us give an estimate of the accuracy of the inversion.

By utilising second-order events we are able to construct the Hessian kernel (Marquering et al., 1999; Fichtner and Trampert, 2011b). We can then analyse the Hessian kernel to help us better understand the inversion results obtained from EFWI. For instance the Hessian contain information about parameter cross-talk and sensitivity to model changes (Sager et al., 2017), indicating how well different areas of the model is resolved in terms of uncertainty and resolution.

We have calculated the Hessian-vector product for a vertical slice of the model for different model perturbations along the slice. The products have then been assembled to construct the Hessian of the model slice which we use for studying the resolution and uncertainty of the model slice.

The Hessian can also be directly used in a Newton inversion scheme (Pratt et al., 1998; Epanomeritakis et al., 2008), but these methods will not be discussed here.

THEORY

FWI is a technique for iteratively recovering model parameters using the entire recorded waveform (Tarantola, 1984; Mora, 1987). For a comprehensive overview of modern FWI we refer to Virieux and Operto (2009), but we will state a brief overview of the basic in elastic media using the adjoint formulation in the time-domain (Fichtner et al., 2006).

The elastic displacement field $\mathbf{u}(\mathbf{x}, t)$ in a model $\mathbf{m}(\mathbf{x}, t)$ with space-coordinates $\mathbf{x} \in G \subset \mathbb{R}^3$ and time $t \in [0, T] \subset \mathbb{R}$ can be described by the wave-operator $\mathbf{L}(\mathbf{u}, \mathbf{m})$ defined as

$$\mathbf{L}(\mathbf{u}, \mathbf{m}) = \rho(\mathbf{x})\ddot{\mathbf{u}}(\mathbf{x}, t) - \nabla\boldsymbol{\sigma}(\mathbf{x}, t) = \mathbf{f}(\mathbf{x}, t), \tag{1}$$

where $\mathbf{f}(\mathbf{x}, t)$ is the driving force, density is denoted by ρ, and the stress tensor $\boldsymbol{\sigma}(\mathbf{x}, t)$ given by the 4th order stiffness tensor $\mathbf{C}$ as

$$\sigma_{ij} = C_{ijkl}\partial_k u_l, \tag{2}$$

using the Einstein summation convention.

Next we define a misfit function

$$\Psi = \Psi(\mathbf{u}(\mathbf{m}, \mathbf{x}_r), \mathbf{d}_0) \tag{3}$$

as a measure of how good a fit our modelled recording $\mathbf{u}(\mathbf{x}_r, t)$ is to the true recorded data $\mathbf{d}_0(\mathbf{x}_r, t)$ at recording locations $\mathbf{x}_r$. By calculating the gradient of the misfit with respect to the model parameters we can then find a model update that will decrease the misfit. This is done by introducing the Jacobian as

$$\mathbf{J}(\mathbf{m} + \delta\mathbf{m}) = \nabla_m\Psi(\mathbf{m} + \delta\mathbf{m}), \tag{4}$$

and linearising around $\mathbf{m}$ resulting in

$$\mathbf{J}(\mathbf{m} + \delta\mathbf{m}) \simeq \mathbf{J}(\mathbf{m}) + \nabla_m\mathbf{J}(\mathbf{m})\delta\mathbf{m} = 0. \tag{5}$$

Thus introducing the Hessian

$$\mathbb{H}(\mathbf{m}) = \nabla_m\mathbf{J}(\mathbf{m}) = \nabla_m\nabla_m\Psi(\mathbf{m}). \tag{6}$$

The model update $\delta\mathbf{m}$ can then be obtained from equation (5) by solving

$$\mathbb{H}(\mathbf{m})\delta\mathbf{m} = -\mathbf{J}(\mathbf{m}) \tag{7}$$

for $\delta\mathbf{m}$.

To compute the full Hessian one would need to calculate the second partial derivative of every model point which is computationally prohibitive for reasonable large models. A common adequate approximation is substituting the Hessian with a scalar and performing a line search for the optimal $\alpha \in \mathbb{R}$ to instead solve (Raknes, 2014)

$$\delta\mathbf{m} \simeq \alpha\mathbf{J}, \tag{8}$$

thus losing 2nd order information in the process. In this work we will be calculating the action of the Hessian on model perturbations (Hessian-vector products) using an adjoint approach proposed by Fichtner and Trampert (2011a).

We can calculate the Jacobian in equation (4) by introducing the adjoint wavefield $\mathbf{u}^\dagger$. This wavefield can be obtained by time-reversing the kernel χ of the misfit function (3) as

$$\Psi = \int_T \int_G \chi\left[\mathbf{u}(\mathbf{m}; \mathbf{x}_r, t), \mathbf{d}_0\right] \mathrm{d}t\mathrm{d}G, \tag{9}$$

10.1190/segam2018-2998233.1

and inserting it at the receiver locations $\mathbf{x}_r$. The Jacobian is then computed using the Frechét derivative as

$$\mathbf{J} = \mathbf{J}(\mathbf{u}^\dagger, \mathbf{u}) = \int_T \mathbf{u}^\dagger \nabla_m \mathbf{L}(\mathbf{u}, \mathbf{m}) \mathrm{d}t, \tag{10}$$

which can be used to find a model update according to the linear approximation in equation (8). To solve the more general model update in equation (7) we need to also find the Hessian matrix.

Introducing the perturbed forward field $\delta\mathbf{u}$ as

$$\delta\mathbf{u} = \lim_{\nu\to 0} \frac{1}{\nu}\Big[\mathbf{u}(\mathbf{m} + \nu\delta\mathbf{m}) - \mathbf{u}(\mathbf{m})\Big], \tag{11}$$

and the perturbed adjoint field $\delta\mathbf{u}^\dagger$ as

$$\delta\mathbf{u}^\dagger = \lim_{\nu\to 0} \frac{1}{\nu}\Big[\mathbf{u}^\dagger(\mathbf{m} + \nu\delta\mathbf{m}) - \mathbf{u}^\dagger(\mathbf{m})\Big], \tag{12}$$

the Hessian acting on a model perturbation $\delta\mathbf{m}$ can be split up into three components (Fichtner and Trampert, 2011a)

$$\mathbb{H}\delta\mathbf{m} = \mathbf{H}_1(\mathbf{u}^\dagger, \delta\mathbf{u}) + \mathbf{H}_2(\delta\mathbf{u}^\dagger, \mathbf{u}) + \mathbf{H}_3(\mathbf{u}^\dagger, \mathbf{u}). \tag{13}$$

On the condition that the wave-operator $\mathbf{L}$ is linear with respect to the displacement $\mathbf{u}$, these components can be written out as

$$\mathbf{H}_1(\mathbf{u}^\dagger, \delta\mathbf{u}) = \int_T \mathbf{u}^\dagger \nabla_m \mathbf{L}(\delta\mathbf{u}, \mathbf{m}) \mathrm{d}t, \tag{14}$$

$$\mathbf{H}_2(\delta\mathbf{u}^\dagger, \mathbf{u}) = \int_T \delta\mathbf{u}^\dagger \nabla_m \mathbf{L}(\mathbf{u}, \mathbf{m}) \mathrm{d}t, \tag{15}$$

$$\mathbf{H}_3(\mathbf{u}^\dagger, \mathbf{u}) = \int_T \mathbf{u}^\dagger \nabla_m \nabla_m \mathbf{L}(\mathbf{u}, \mathbf{m})(\delta\mathbf{m}) \mathrm{d}t. \tag{16}$$

Comparing equations (14) and (15) with equation (10) we observe that we need only replace the appropriate field with the perturbed one to calculate $\mathbf{H}_1$ and $\mathbf{H}_2$, whereas $\mathbf{H}_3$ is dependent on the given parametrisation. Note that the $\mathbf{H}_3$ term will vanish when the wave-operator is linear with respect to the model parameters, e.g. using the Lamé parametrisation. In the density (ρ), pressure-wave velocity (v_p), and shear-wave velocity (v_s) parametrisation $\mathbf{H}_3$ will not vanish, but we can calculate it by recycling the Frechét kernels directly obtained from equation (10) and the model perturbation $\delta\mathbf{m} = \left[\delta\rho, \delta v_p, \delta v_s\right]^T$ as

$$\mathbf{H}_3 = \begin{bmatrix} 0 & \rho^{-1} J_{v_p} & \rho^{-1} J_{v_s} \\ \rho^{-1} J_{v_p} & v_p^{-1} J_{v_p} & 0 \\ \rho^{-1} J_{v_s} & 0 & v_s^{-1} J_{v_s} \end{bmatrix} \begin{bmatrix} \delta\rho \\ \delta v_p \\ \delta v_s \end{bmatrix} \tag{17}$$

The Hessian kernel at position $\mathbf{x}_i$ due to a perturbation $\delta\mathbf{m}_j$ can be constructed as a matrix with the element ij expressed as

$$\mathbb{H}(\mathbf{m}, \mathbf{x}_i)\delta\mathbf{m}_j = H_i^n \delta_j^m, \tag{18}$$

with $n, m \in \mathbf{m}$ looping over the parametrisation parameters the action of the Hessian on a model perturbation can be written on the block form

$$\mathbb{H}\delta\mathbf{m} = \left[\begin{array}{c|c|c} H_i^{n_1}\delta_j^{n_1} & H_i^{n_1}\delta_j^{n_2} & H_i^{n_1}\delta_j^{n_3} \\ \hline H_i^{n_2}\delta_j^{n_1} & H_i^{n_2}\delta_j^{n_2} & H_i^{n_2}\delta_j^{n_3} \\ \hline H_i^{n_3}\delta_j^{n_1} & H_i^{n_3}\delta_j^{n_2} & H_i^{n_3}\delta_j^{n_3} \end{array}\right] \tag{19}$$

METHOD

We will primarily focus on the ρ, v_p, v_s parametrisation of EFWI using a time-domain finite-difference scheme. In order to calculate the Hessian for a given model we start by generating a top-recording of the true model. Next we introduce grid-size perturbations to a background model in order to calculate the action of the Hessian in a model slice as shown in Figure 2 to 7.

The different sub-matrices of the Hessian can then tell us about the interaction between a given model perturbation and the neighbouring area. A perfectly resolved model with no parameter cross-talk is represented by a scaled identity matrix in our construction of the Hessian. In this case the scalar substitution in equation (8) would be valid.

Off-diagonal elements are indicative of a bandwidth limited response, or de-focusing of the signal in the slice direction, showing how a perturbation at a given point interferes with its neighbourhood. Parameter cross-talk is evident by signal in the off-diagonal sub-matrices showing how much a given parameter affects a different parameter.

We have calculated the Hessian for a vertical slice in both simple (homogeneous and linear gradient) and complex models

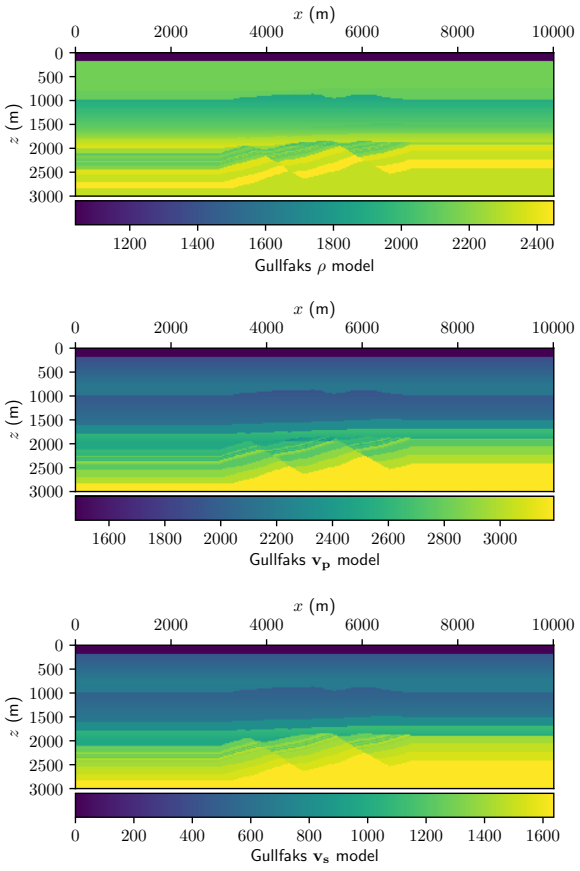

Figure 1: The Gullfaks density-velocity model

10.1190/segam2018-2998233.1

(Gullfaks model, Figure 1) in order to estimate how good we can resolve the models at a given depth with a predetermined source-receiver geometry.

As a source we have used a single Ricker wavelet with a dominant frequency of 10 Hz at offsets of both zero and 2 km to the left of the centre of the models. Receivers used to back-propagate the misfit kernels are placed along the full 10 km width at the top of the models. The simulations are performed using a fully resolved background model which differs from the target model by only the introduced perturbation. Attention is focused on the centre of the model at depths between 900 m and 2100 m where we construct the Hessian.

The obtained Hessian matrices are normalised to ± 1 and displayed with overexposure in order to emphasise structure.

RESULTS

In the obtained Hessian matrices (Figure 2 to 7) we observe strong cross-talk between all parameters as a signal in the off-diagonal sub-matrices. Ignoring the cross-talk sub-matrix elements, we have diagonally dominant matrix in the $\mathbf{H}^{v_P}\delta v_P$ and $\mathbf{H}^{v_S}\delta v_S$ elements, indicating a good spatial resolution for the velocities. The relative weaker signal in the $\mathbf{H}^{\rho}\delta\rho$ sub-matrix is indicative of a comparably lower recovery of the density parameter.

For the simple model zero-offset experiments (Figure 2 and 3) most of the de-focusing appears above the diagonal in the Hessian sub-matrices, with most of the de-focusing occurring at shallow depths. This is indicative of a lack of information about the placement of the perturbation. Moving further away from the centre we observe that the de-focusing is mostly gone in the intermediate-offset experiments (Figure 5 and 6), leaving us instead with a strong diagonal signal which slightly widens as we move deeper into the model.

Looking at the Hessian for the more complex Gullfaks model in the zero-offset case (Figure 4) we observe a more focused diagonal. This can be explained by internal reflections giving us more data as to where the perturbation is placed. The missing bands in the Hessian are thought to be caused by destructive interference due to the same internal reflections. Moving to the intermediate-offset Hessian (Figure 7), we notice less focused sub-matrix diagonals, especially in the $\mathbf{H}^{v_P}\delta v_P$ sub-matrix.

CONCLUSIONS

We have successfully constructed the Hessian matrix for a subset of a given model using Hessian-vector products based on simple model perturbations.

Despite added complexity the method of calculating the Hessian presented here repurposes many of the routines used in conventional FWI, making it fairly simple to implement in an existent solver. The main added computational cost comes from simulating twice as many fields, though some of this cost can be alleviated by focusing on a subset of the model.

Preliminary results suggest that there is a strong cross-talk between the different parameters in the ρ, v_P, v_S parametrisation without sampling the data. This results in a highly coupled response in the model update, making it difficult to distinguish the influence of different parameters. This is especially true when trying to invert for density as our results suggest density updates are more affected by changes in velocity than density itself.

Our results confirms that longer offsets help focus model inversion, but that the effect is not as prominent in a more complex model. A better resolution of the v_S signal lends us to propose sampling the recorded signal to construct the Hessian using mostly S-wave data.

ACKNOWLEDGMENTS

The authors would like to thank NTNU and Aker BP ASA for making the code available through the Codeshare project.

10.1190/segam2018-2998233.1

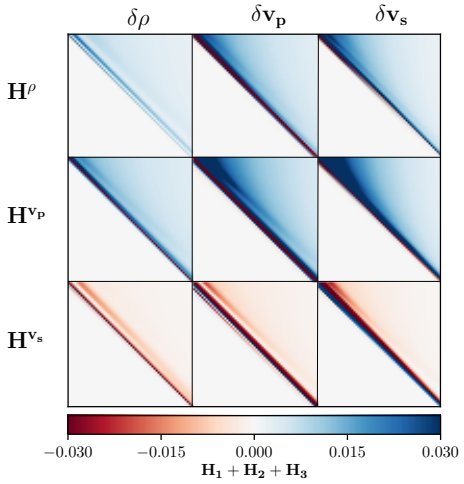

Figure 2: Zero-offset Hessian from a uniform model.

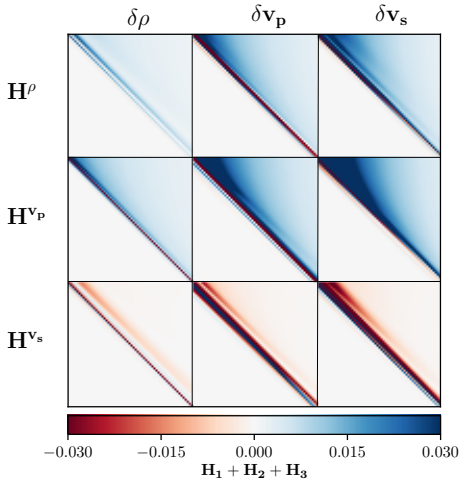

Figure 3: Zero-offset Hessian from a gradient model.

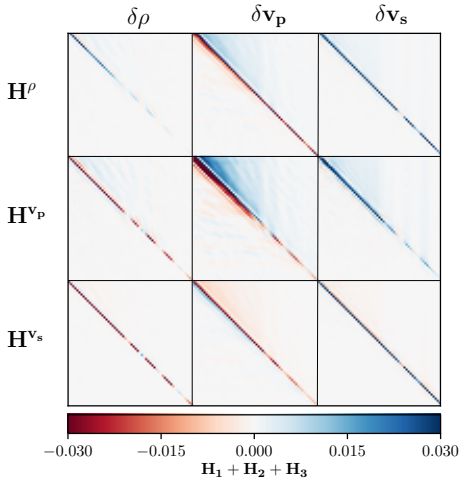

Figure 4: Zero-offset Hessian from the Gullfaks model.

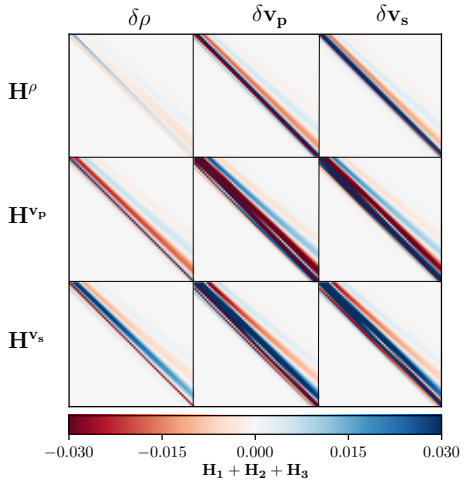

Figure 5: Indermediate-offset Hessian from a uniform model.

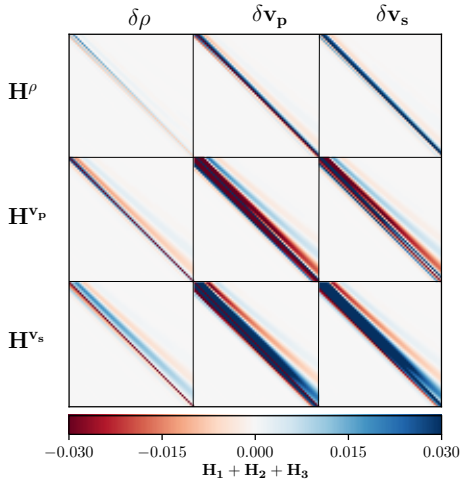

Figure 6: Intermediate-offset Hessian from a gradient model.

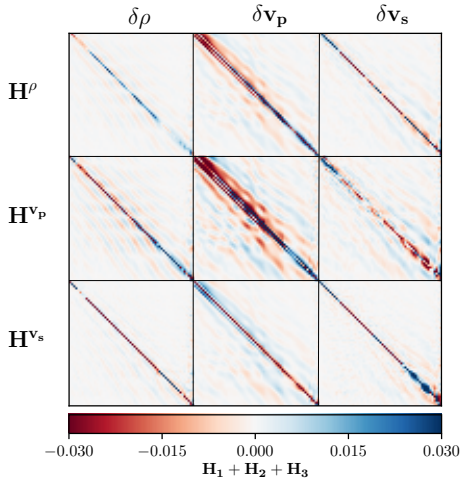

Figure 7: Intermediate-offset Hessian from the Gullfaks model.

10.1190/segam2018-2998233.1

REFERENCES

Epanomeritakis, I., V. Akçelik, O. Ghattas, and J. Bielak, 2008, A Newton-CG method for large-scale three-dimensional elastic full-waveform seismic inversion: Inverse Problem, **24**, 034015, https://doi.org/10.1088/0266-5611/24/3/034015.

Fichtner, A., H.-P. Bunge, and H. Igel, 2006, The adjoint method in seismology: Physics of the Earth and Planetary Interiors, **157**, 86–104, https://doi.org/10.1016/j.pepi.2006.03.016.

Fichtner, A., and J. Trampert, 2011a, Hessian kernels of seismic data functionals based upon adjoint techniques: Geophysical Journal International, **185**, 775–798.

Fichtner, A., and J. Trampert, 2011b, Resolution analysis in full waveform inversion: Geophysical Journal International, **187**, 1604–1624, https://doi.org/10.1111/j.1365-246X.2011.05218.x.

Marquering, H., F. A. Dahlen, and G. Nolet, 1999, Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox: Geophysical Journal International, **137**, 805–815.

Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **52**, 1211–1228, https://doi.org/10.1190/1.1442384.

Operto, S., Y. Gholami, V. Prieux, A. Ribodetti, R. Brossier, L. Metivier, and J. Virieux, 2013, A guided tour of multiparameter full-waveform inversion with multicomponent data. From theory to practice: The Leading Edge, **32**, 1040–1054, https://doi.org/10.1190/tle32091040.1.

Pratt, R. G., C. Shin, and G. J. Hicks, 1998, Gauss-Newton and full Newton methods in frequency domain seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Raknes, E. B., 2014, Time-lapse full-waveform inversion of limited-offset seismic data using a local migration regularization: Geophysics, **79**, no. 3, WA117–WA128, https://doi.org/10.1190/geo2013-0369.1.

Sager, K., L. Ermert, C. Boehm, and A. Fichtner, 2017, Towards full waveform ambient noise inversion: Geophysical Journal International, 566–590.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266, https://doi.org/10.1190/1.1441754.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration Geophysics: Geophysics, **74**, no. 6, WCC1–WCC26, https://doi.org/10.1190/1.3238367.