S23A-0779

Uncertainty Estimation in Elastic Full Waveform Inversion by Utilising the Hessian Matrix

Vegard Stenhjem Hagen*[1] <vegard.hagen@ntnu.no>, Børge Arntsen[1] <borge.arntsen@ntnu>, Espen B. Raknes[2] <espen.birger.raknes@akerbp.no>
[1]Norwegian University of Science and Technology, [2]Aker BP ASA

1 Introduction

Elastic Full Waveform Inversion (EFWI) is a computationally intensive iterative method for estimating elastic model parameters. A key element of EFWI is the numerical solution of the elastic wave equation which lies as a foundation to quantify the mismatch between synthetic (modelled) and true (target) measured seismic data. The misfit between the modelled and target receiver data is used to update the parameter model to yield a better fit between the modelled and true receiver signal. A common approach to the EFWI model update problem is to use a conjugate gradient search method. In our approach we suggest that the resolution and cross-coupling for the estimated parameter update can be found by computing the Hessian matrix. Both a homogeneous model and the Gullfaks model (fig. 1) are used to illustrate the influence of offset on parameter resolution and cross-coupling.

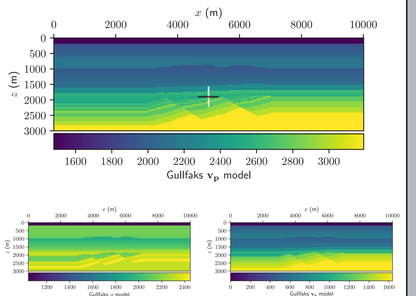


Figure 1: The Gullfaks model. The red diamond indicates a small inclusion δv_p in the target v_p model that is not present in the background model. The black and white lines indicate the position of the test perturbations used to calculate the Hessian kernel.

2 Aim

The main goal of this research is to get a better understanding of the wavefield interactions in seismic surveys and how to exploit this information in order to produce more accurate subsurface models, and with a higher confidence.

We are using the known adjoint-state approach to calculate a model update based on the difference in surface recorded field between the target model and an initial model as described by P.Mora and A.Tarantola[1, 2], and expanded by A.Fichtner[3, 4] to calculate the Hessian acting on a model perturbation. Here we have calculated the Hessian acting on several small perturbations in order to construct the Hessian for a defined section of the model. We then use the Hessian matrix to infer parameter resolution and cross-coupling for different acquisition geometries.

3 Elastic wave equation

Given position $\mathbf{x} \in G \subset \mathbb{R}^2$, time t, and model parameters $\mathbf{m} = (\rho, \mathbf{C})$, where ρ is density and $\mathbf{C}$ is the elastic tensor, the 2D Elastic wave equation operator $\mathbf{L}$ with driving force $\mathbf{f}$ can be written as[3]

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

where $\mathbf{u}$ is the displacement field, and $\boldsymbol{\sigma}$ is the stress tensor which can be expressed as

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

in a perfectly elastic medium. Note that the wave operator $\mathbf{L}$ is linear in $\mathbf{u}$.

Assuming an isotropic medium we can express the elastic tensor $\mathbf{C}$ using the Lamé parameters λ and μ as

$$C_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{3}$$

with δ_{ij} being the Kronecker delta.

Imposing the free surface boundary condition results in the normal components of the stress tensor $\boldsymbol{\sigma}$ vanishing at the surface ∂G of the Earth model, i.e.

$$\boldsymbol{\sigma}\mathbf{n}|_{x\in\partial G} = 0, \tag{4}$$

given the surface normal $\mathbf{n}$ on ∂G. Further, both the displacement field $\mathbf{u}$ and the velocity field $\dot{\mathbf{u}}$ are required to satisfy the initial condition

$$\mathbf{u}|_{t\leq 0} = \dot{\mathbf{u}}|_{t\leq 0} = 0. \tag{5}$$

4 Model update

Using a misfit functional $\Psi\left(\mathbf{d}_r(\mathbf{m}), \mathbf{d}_0\right)$, we can compare the target model recorded data $\mathbf{d}_0$ with the modelled data $\mathbf{d}_r = \mathbf{u}(\mathbf{m}, \mathbf{x}_r)$, where $\mathbf{x}_r$ is the position of the receivers. The goal of this method is to find a model update $\delta\mathbf{m}_k$ that decreases the misfit such that

$$\Psi(\mathbf{m}_{k+1} = \mathbf{m}_k + \delta\mathbf{m}_k) < \Psi(\mathbf{m}_k). \tag{6}$$

This can be done by calculating the Jacobian of the misfit as

$$\mathbf{J}(\mathbf{m}_k + \delta\mathbf{m}_k) = \nabla_m\Psi(\mathbf{m}_k + \delta\mathbf{m}_k), \tag{7}$$

and linearising it around $\mathbf{m}_k$

$$\mathbf{J}(\mathbf{m}_k + \delta\mathbf{m}_k) \simeq \mathbf{J}(\mathbf{m}_k) + \nabla_m\mathbf{J}(\mathbf{m}_k)\delta\mathbf{m}_k = \mathbf{0}. \tag{8}$$

With the Hessian expressed as

$$\mathbb{H}(\mathbf{m}_k) = \nabla_m\mathbf{J}(\mathbf{m}_k) = \nabla_m\nabla_m\Psi(\mathbf{m}_k) \tag{9}$$

we can rearrange eq. (8)

$$\mathbb{H}(\mathbf{m}_k)\delta\mathbf{m}_k = -\mathbf{J}(\mathbf{m}_k), \tag{10}$$

assuming the Hessian is invertible we can solve for $\delta\mathbf{m}_k$ and get

$$\delta\mathbf{m}_k = -\mathbb{H}^{-1}\mathbf{J}. \tag{11}$$

Since the full Hessian is computationally expensive, eq. (11) is commonly approximated[5] by substituting the inverse Hessian with a scalar $\alpha_k \in \mathbb{R}$ resulting in

$$\delta\mathbf{m}_k \simeq \alpha_k\mathbf{J}. \tag{12}$$

A model update can then be found via a linear search for the best step length α_k, but at the loss of information regarding parameter resolution and cross-talk. Here we will calculate the Hessian acting on a model perturbation as $\mathbb{H}\delta\mathbf{m}_k$ and use the result to study the effect of different perturbations $\delta\mathbf{m}_k$.

5 Fréchet kernels

We introduce the adjoint field $\mathbf{u}^\dagger$ shown in figure 2 as the backpropagation of the difference in the recorded field $\mathbf{d}_r$ and target recording $\mathbf{d}_0$ dependent on a norm χ which forms the kernel of the misfit functional Ψ.

By cross-correlating the adjoint field with the Fréchet derivative of the wave operator we can calculate the Jacobian as[4]

$$\mathbf{J} = \mathbf{F}_m(\mathbf{u}^\dagger, \mathbf{u}) = \int_T \mathbf{u}^\dagger\nabla_m\mathbf{L}(\mathbf{u}, \mathbf{m})\,\mathrm{d}t, \tag{13}$$

where $\mathbf{F}_m$ is the Fréchet kernel defined as the volumetric density of the Fréchet derivative.

In the perfectly elastic isotropic case using the $\mathbf{m}^0 = (\rho, \mu, \lambda)$ parametrisation, the Fréchet kernels can be written as

$$\mathbf{F}^0_\rho(\mathbf{x}) = -\int_T \dot{u}^\dagger_i\dot{u}_i\,\mathrm{d}t, \tag{14}$$

$$\mathbf{F}^0_\lambda(\mathbf{x}) = +\int_T \partial_l u^\dagger_i\partial_l u_j\,\mathrm{d}t, \tag{15}$$

$$\mathbf{F}^0_\mu(\mathbf{x}) = \frac{1}{2}\int_T \left(\partial_j u^\dagger_i + \partial_i u^\dagger_j\right)\left(\partial_j u_i + \partial_i u_j\right)\mathrm{d}t, \tag{16}$$

which can then be used to calculate model updates. The Fréchet kernels of different parametrisations can easily be obtained as a linear combination of the $\mathbf{m}^0$ kernels, e.g. in the $\mathbf{m} = (\rho, v_p, v_s)$ parametrisation we get

$$\mathbf{F}_\rho = \mathbf{F}^0_\rho + (v_p^2 - 2v_s^2)\mathbf{F}^0_\lambda + v_s^2\mathbf{F}^0_\mu, \tag{17}$$

$$\mathbf{F}_{v_p} = 2\rho v_p\mathbf{F}^0_\lambda, \tag{18}$$

$$\mathbf{F}_{v_s} = 2\rho v_s\mathbf{F}^0_\mu - 4\rho v_s\mathbf{F}^0_\lambda. \tag{19}$$

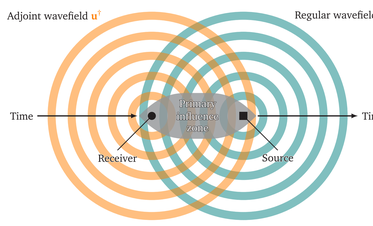


Figure 2: Interaction between the regular forward wavefield $\mathbf{u}$ and the adjoint wavefield $\mathbf{u}^\dagger$ takes place in the primary influence zone indicated by a grey area.

6 Hessian kernels

The Hessian $\mathbb{H}$ acting on a model perturbation $\delta\mathbf{m}$ can be expressed with the Hessian kernel $\mathbf{H}_m$ as[4]

$$\mathbb{H}\delta\mathbf{m} = \mathbf{H}_m. \tag{20}$$

The Hessian kernel can in turn be split up into three different terms

$$\mathbf{H}_m = \mathbf{H}^1_m(\mathbf{u}^\dagger, \delta\mathbf{u}) + \mathbf{H}^2_m(\delta\mathbf{u}^\dagger, \mathbf{u}) + \mathbf{H}^3_m(\mathbf{u}^\dagger, \mathbf{u}) \tag{21}$$

by introducing the perturbed forward and adjoint fields $\delta\mathbf{u}$ and $\delta\mathbf{u}^\dagger$ respectively as

$$\delta\mathbf{u} = \lim_{\nu\to 0}\frac{1}{\nu}\Big[\mathbf{u}(\mathbf{m} + \nu\delta\mathbf{m}) - \mathbf{u}(\mathbf{m})\Big], \tag{22}$$

$$\delta\mathbf{u}^\dagger = \lim_{\nu\to 0}\frac{1}{\nu}\Big[\mathbf{u}^\dagger(\mathbf{m} + \nu\delta\mathbf{m}) - \mathbf{u}^\dagger(\mathbf{m})\Big]. \tag{23}$$

It can be shown that the first two Hessian terms can be expressed using the same formula as for the Fréchet kernels, but substituting the wavefields[4]

$$\mathbf{H}^1_m = \int_T \mathbf{u}^\dagger\nabla_m\mathbf{L}(\delta\mathbf{u}, \mathbf{m})\,\mathrm{d}t = \mathbf{F}_m(\mathbf{u}^\dagger, \delta\mathbf{u}), \tag{24}$$

$$\mathbf{H}^2_m = \int_T \delta\mathbf{u}^\dagger\nabla_m\mathbf{L}(\mathbf{u}, \mathbf{m})\,\mathrm{d}t = \mathbf{F}_m(\delta\mathbf{u}^\dagger, \mathbf{u}.) \tag{25}$$

The third Hessian term can be expressed as

$$\mathbf{H}^3_m = \int_T \mathbf{u}^\dagger\nabla_m\nabla_m\mathbf{L}(\mathbf{u}, \mathbf{m})(\delta\mathbf{m})\,\mathrm{d}t \tag{26}$$

which vanishes in the ρ, μ, λ parametrisation. In the ρ, v_p, v_s parametrisation $\mathbf{H}^3_m$ is non-zero and can written using the Fréchet kernels as

$$\mathbf{H}^3_m = \begin{bmatrix}\mathbf{H}^3_\rho \\ \mathbf{H}^3_{v_p} \\ \mathbf{H}^3_{v_s}\end{bmatrix} = \begin{bmatrix}0 & \rho^{-1}\mathbf{F}_{v_p} & \rho^{-1}\mathbf{F}_{v_s} \\ \rho^{-1}\mathbf{F}_{v_p} & v_p^{-1}\mathbf{F}_{v_p} & 0 \\ \rho^{-1}\mathbf{F}_{v_s} & 0 & v_s^{-1}\mathbf{F}_{v_s}\end{bmatrix}\begin{bmatrix}\delta\rho \\ \delta v_p \\ \delta v_s\end{bmatrix}. \tag{27}$$

Note that the $\mathbf{H}^3_m$ element is explicitly dependent on the model perturbation.

We can construct a matrix representing the Hessian $\mathbb{H}$ at position $\mathbf{x}_i$ due to a model perturbation $\delta\mathbf{m}$ at position $\mathbf{x}_j$

$$\mathbb{H}\delta\mathbf{m} = \left[\begin{array}{ccc|ccc|ccc} H^\rho_0\delta^\rho_0 & H^\rho_0\delta^\rho_1 & \cdots & H^\rho_0\delta^{v_p}_0 & \cdots & H^\rho_0\delta^{v_s}_0 & \cdots \\ H^\rho_1\delta^\rho_0 & H^\rho_1\delta^\rho_1 & \ddots & \vdots & \ddots & \vdots & \ddots \\ H^\rho_2\delta^\rho_0 & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots \\ \hline H^{v_p}_0\delta^\rho_0 & \cdots & \cdots & H^{v_p}_0\delta^{v_p}_0 & \cdots & H^{v_p}_0\delta^{v_s}_0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots \\ \hline H^{v_s}_0\delta^\rho_0 & \cdots & \cdots & H^{v_s}_0\delta^{v_p}_0 & \cdots & H^{v_s}_0\delta^{v_s}_0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots \end{array}\right]$$

using the shorthand notation

$$\mathbb{H}(\mathbf{x}_i)\delta\mathbf{m}(\mathbf{x}_j) = H^n_i\delta^m_j. \tag{28}$$

7 Method

The models used are 10 km wide by 3 km deep with a grid spacing of 10 m. We have introduced a 30 m × 30 m inclusion in the v_p parameter centred at (5 km, 1.9 km) in the target model which we study the effect of by looking at the Hessian kernels. Two source-receiver configurations are considered, a zero-offset configuration where the sources are directly above the inclusion, and a long-offset configuration where the sources are placed at the edges of the domain. In both cases the sources are placed at a 30 m depth as 8 Hz Ricker wavelet stress monopoles, and a OBC-type receiver array running along the x-direction also at a 30 m depth.

Both models are probed using perturbations in all of the model parameters sequentially at distances up to 300 m from the centre of the inclusion in the x- and z-directions as indicated in figure 1 by the black and white lines.

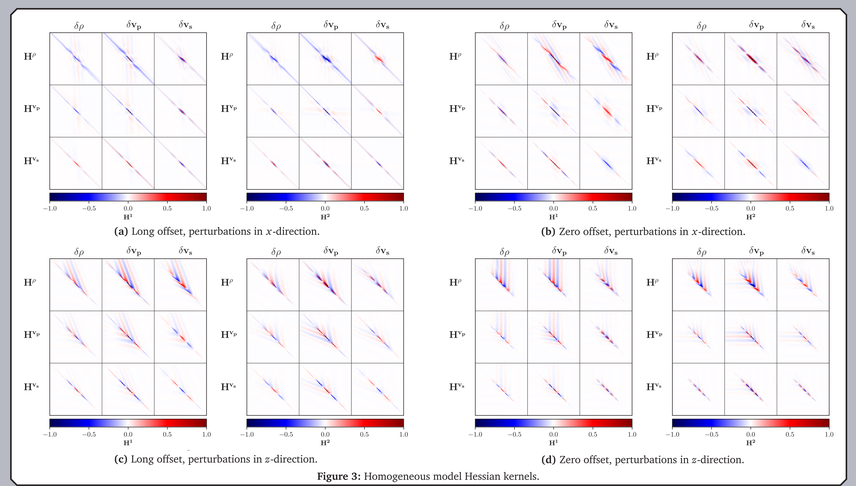
We first study the homogeneous model which has background velocities $v_p = 2000\,\mathrm{m/s}$, $v_s = 1000\,\mathrm{m/s}$, and density $\rho = 1500 kg/m^3$. The resulting $\mathbf{H}^1$ and $\mathbf{H}^2$ kernels are shown in figure 3.

Using the Gullfaks model illustrated in figure 1 results in the $\mathbf{H}^1$ and $\mathbf{H}^2$ Hessian kernels shown in figure 4.

The $\mathbf{H}^3$ kernels of both models are shown in figure (5). Due to the localised perturbations used in this investigation, only the diagonal elements of the $\mathbf{H}^3$ kernel are non-zero.

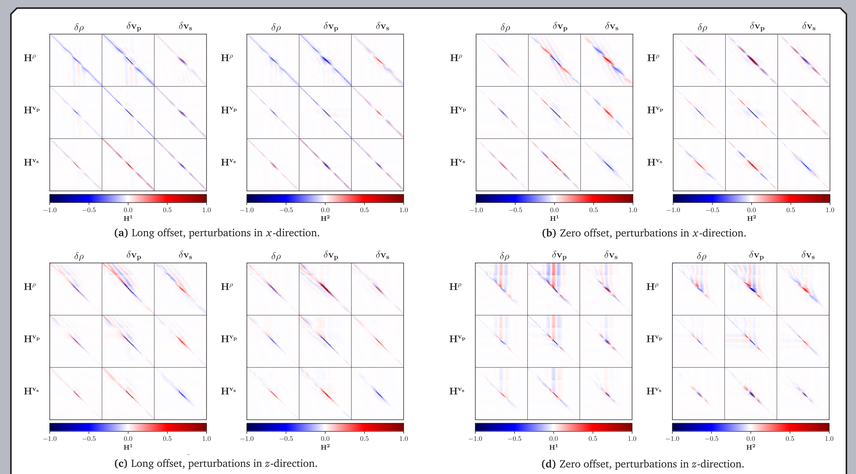
By comparing the different patterns of the Hessian kernels in a homogeneous medium we posit the possibility of inferring the resolution of different parameters and parameter coupling. Further comparing the homogeneous case with a more complex model we wish to comment on the influence of heterogeneities and internal scattering.

We will mainly focus on the $\mathbf{H}^1$ and $\mathbf{H}^1$ kernels in the common ρ, v_p, v_s parametrisation.



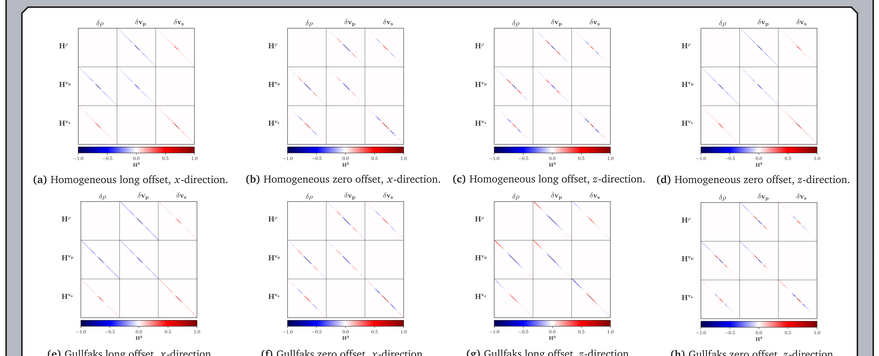
(a) Long offset, perturbations in x-direction. **(b)** Zero offset, perturbations in x-direction. **(c)** Long offset, perturbations in z-direction. **(d)** Zero offset, perturbations in z-direction.

Figure 3: Homogeneous model Hessian kernels.



(a) Long offset, perturbations in x-direction. **(b)** Zero offset, perturbations in x-direction. **(c)** Long offset, perturbations in z-direction. **(d)** Zero offset, perturbations in z-direction.

Figure 4: Gullfaks model Hessian kernels.



(a) Homogeneous long offset, x-direction. **(b)** Homogeneous zero offset, x-direction. **(c)** Homogeneous long offset, z-direction. **(d)** Homogeneous zero offset, z-direction. **(e)** Gullfaks long offset, x-direction. **(f)** Gullfaks zero offset, x-direction. **(g)** Gullfaks long offset, z-direction. **(h)** Gullfaks zero offset, z-direction.

Figure 5: $\mathbf{H}^3$ Hessian kernels.

8 Results and Discussion

Studying the Hessian kernels shown in figure 3-5 we observe a strong signal along the diagonal in all elements as an imprint of the perturbation. Especially in the $\mathbf{H}^3$ elements where only the diagonal is non-zero for some of the elements.

We believe the wave pattern to be caused by the finite frequency content and possibly linked to the choice of misfit functional Ψ and the cycle skipping problem[5].

Looking at the homogeneous case in figure 3 we notice a comparatively low level of smearing in the x-direction in both the long and zero offset cases compared with the z-direction. The best focusing is obtained at long offsets in both directions as expected.

The pattern in the top row of the horizontal direction kernels suggest that the $\mathbf{H}^\rho$-kernel is chiefly affected by interactions above the perturbation, which is also suggested by geometrical spreading studies[6]. The relatively smeared density kernel compared to the other elements is also in line with the notion that it is hard to invert for density[5]. There is also a strong coupling between the density and velocity parameters which is evident from the strong pattern from the $\delta\mathbf{v}_p$ and $\delta\mathbf{v}_s$ columns in the $\mathbf{H}^\rho$ row, especially in the z-direction.

It can appear that the v_s parameter is best resolved, though by closer inspection we see a arguably stronger coupling with the δv_p perturbation than the expected δv_s parameter.

All in all there is a significant coupling between all three parameters in the first two kernel terms having to do with the perturbed wavefields $\delta\mathbf{u}$ and $\delta\mathbf{u}^\dagger$. The pairwise strong coupling of parameters in the remaining $\mathbf{H}^3$ term might be able to alleviate some of the ambiguity by use of different perturbations.

Looking at the Hessian kernels obtained from the more complex Gullfaks model we can observe the same trends as in the homogeneous case. The resulting Hessian kernels for the heterogeneous case can be said to be better focused than in the homogeneous case, especially at the long offset. It should be noted that this effect can be due to a haphazardly better acquisition geometry and other models should be tested.

The kernels agree with previous observations and show promise to for applicability in uncertainty analysis for seismic exploration.

9 Future work

The presented Hessian kernels show interesting properties. Different strategies for calculating them should be tested, including different frequencies and misfit functionals to explore the effect it might have one the phase wave pattern, if any.

When properly implemented we suggest to incorporate the Hessian kernels in a iterative Full Waveform Inversion (FWI) scheme based on the Newton method[7, 8] to improve the inversion results and use the information contained in the kernels to provide uncertainty estimates of the result.

We also suggest to calculate the Hessian kernels for different and more complex parametrisations (e.g. VTI, orthorhombic) in the pursuit for an optimal parametrisation to use in seismic exploration.

References

[1] A. Tarantola. 'Inversion of seismic reflection data in the acoustic approximation'. In: *Geophysics* 49.8 (1984), p. 1259. DOI: 10.1190/1.1441754.

[2] P. Mora. 'Nonlinear two-dimensional elastic inversion of multioffset seismic data'. In: *Geophysics* 52.9 (1987), p. 1211. DOI: 10.1190/1.1442384.

[3] A. Fichtner, H. P. Bunge and H. Igel. 'The adjoint method in seismology. I. Theory'. In: *Physics of the Earth and Planetary Interiors* 157.1-2 (2006), pp. 86–104. DOI: 10.1016/j.pepi.2006.03.016.

[4] A. Fichtner. 'Full Seismic Waveform Modelling and Inversion'. In: *Advances* 1981 (2011), pp. 83–88. DOI: 10.1007/978-3-642-15807-0.

[5] J. Virieux and S. Operto. 'An overview of full-waveform inversion in exploration geophysics'. In: *Geophysics* 74.6 (2009), WCC1. DOI: 10.1190/1.3238367.

[6] B. Ursin. 'Offset-dependent geometrical spreading in a layered medium'. In: *GEOPHYSICS* 55.4 (1990), pp. 492–496. DOI: 10.1190/1.1442860. eprint: https://doi.org/10.1190/1.1442860.

[7] L Métivier, R Brossier, J Virieux and S Operto. 'The truncated Newton method for Full Waveform Inversion'. In: *Journal of Physics: Conference Series* 386.2 (2012), p. 012013. DOI: 10.1088/1742-6596/386/1/012013.

[8] I Epanomeritakis, V Akçelik, O Ghattas and J Bielak. 'A Newton-CG method for large-scale three-dimensional elastic full-waveform seismic inversion'. In: *Inverse Problems* 24.3 (2008), p. 034015. DOI: 10.1088/0266-5611/24/3/034015.

AGU American Geophysical Union

NTNU