Parameter resolution and cross-talk for Elastic Full Wavefrom Inversion

Vegard S. Hagen[†], Børge Arntsen[†] & Espen B. Raknes[‡]

[†]NTNU - Department of Geoscience and Petroleum, [‡]AkerBP

Outline

- Brief introduction to Full Waveform Inversion
  - Focus on the gradient
- Present the Frechét derivative
- Adjoint theory for calculating the Hessian
- Calculate the Hessian
  - Homogeneous media
  - Gullfaks model

Motivation

- Quantify parameter cross-talk
- Measure the resolution of FWI
- Investigate possibility of a Newton solver

Full Waveform inversion[1]

1. Guess a starting model (**m**) based on other information.

[1] Tarantola 1984; Mora 1987.

Full Waveform inversion[1]

1. Guess a starting model (**m**) based on other information.
2. Do a synthetic run.

[1] Tarantola 1984; Mora 1987.

Full Waveform inversion[1]

1. Guess a starting model (**m**) based on other information.
2. Do a synthetic run.
3. Compare the synthetic signal to the real.

[1] Tarantola 1984; Mora 1987.

Full Waveform inversion[1]

1. Guess a starting model (**m**) based on other information.
2. Do a synthetic run.
3. Compare the synthetic signal to the real.
4. Use this information to calculate a gradient update.

[1] Tarantola 1984; Mora 1987.

Full Waveform inversion[1]

1. Guess a starting model (**m**) based on other information.
2. Do a synthetic run.
3. Compare the synthetic signal to the real.
4. Use this information to calculate a gradient update.
5. Apply the model update.

[1] Tarantola 1984; Mora 1987.

Full Waveform inversion[1]

1. Guess a starting model ($\mathbf{m}$) based on other information.
2. Do a synthetic run.
3. Compare the synthetic signal to the real.
4. Use this information to calculate a gradient update.
5. Apply the model update.
6. Repeat.

[1] Tarantola 1984; Mora 1987.

Full Waveform inversion[1]

1. Guess a starting model (**m**) based on other information.
2. Do a synthetic run.
3. Compare the synthetic signal to the real.
4. Use this information to calculate a gradient update. ←
5. Apply the model update.
6. Repeat.

[1] Tarantola 1984; Mora 1987.

Iterative methods[2]

- Searching for a model $\mathbf{m}$ that describes the earth.
- Elastic wave equation

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

- Compare with true recorded data $\mathbf{d}_0$ using a misfit function

$$\Psi\left(\mathbf{u}(\mathbf{m},\mathbf{x}_r),\mathbf{d}_0\right).$$

- Iterative approach. Find a model update $\delta\mathbf{m}_k$ that decreases the misfit

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

[2] Tarantola 1984; Mora 1987; Fichtner et al. 2006; Fichtner 2011.

Iterative methods

- Calculate the gradient of the misfit

$$\mathbf{J}(\mathbf{m}+\delta\mathbf{m}) = \nabla_m \Psi(\mathbf{m}+\delta\mathbf{m}).$$

- Linearising the Jacobian results in

$$\mathbf{J}(\mathbf{m}+\delta\mathbf{m}) \simeq \mathbf{J}(\mathbf{m}) + \underbrace{\nabla_m \mathbf{J}(\mathbf{m})}_{\mathbf{H}(\mathbf{m})}\,\delta\mathbf{m} = \mathbf{0}.$$

Iterative methods

- Calculate the gradient of the misfit

$$\mathbf{J}(\mathbf{m} + \delta\mathbf{m}) = \nabla_m \Psi(\mathbf{m} + \delta\mathbf{m}).$$

- Linearising the Jacobian results in

$$\mathbf{J}(\mathbf{m} + \delta\mathbf{m}) \simeq \mathbf{J}(\mathbf{m}) + \underbrace{\nabla_m \mathbf{J}(\mathbf{m})}_{\mathbf{H}(\mathbf{m})} \delta\mathbf{m} = \mathbf{0}.$$

- The Hessian is given as

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

Iterative methods[3]

- Solving
$$\mathbf{H}(\mathbf{m})\delta\mathbf{m} = -\mathbf{J}(\mathbf{m})$$
for $\delta\mathbf{m}$ we find the next model update.
- Iff $\mathbf{H}$ is invertible we can "simply" solve
$$\delta\mathbf{m} = -\mathbf{H}^{-1}\mathbf{J}.$$

[3]Virieux and Operto 2009; Métivier et al. 2012; Epanomeritakis et al. 2008.

Iterative methods[3]

- Solving
$$\mathbf{H}(\mathbf{m})\delta\mathbf{m} = -\mathbf{J}(\mathbf{m})$$
for $\delta\mathbf{m}$ we find the next model update.
- Iff $\mathbf{H}$ is invertible we can “simply” solve
$$\delta\mathbf{m} = -\mathbf{H}^{-1}\mathbf{J}.$$
- A common approximation is
$$\delta\mathbf{m} \simeq \alpha\mathbf{J},$$
and a line search for the optimal $\alpha \in \mathbf{R}$.

[3] Virieux and Operto 2009; Métivier et al. 2012; Epanomeritakis et al. 2008.

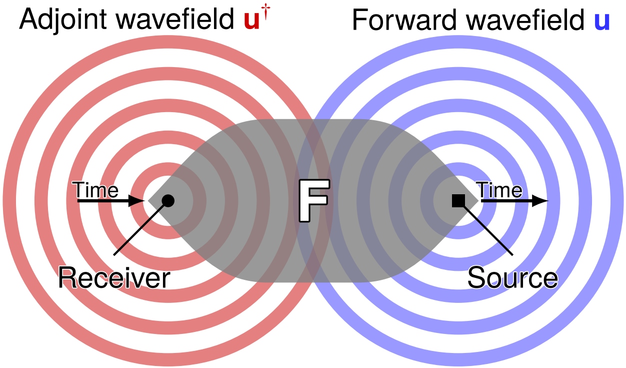
Gradient (Fréchet derivative)

- We can calculate the Jacobian by use of the Fréchet derivative

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

- Cross-correlate the adjoint field $\mathbf{u}^{\dagger}$ and the forward field $\mathbf{u}$.

Frechét kernel — $\mathbf{F}(\mathbf{u}^{\dagger}, \mathbf{u})$



Gradient

Fréchet kernel $\mathbf{F}(\mathbf{u}^{\dagger}, \mathbf{u})$

Background field $\mathbf{u}$

Adjoint field $\mathbf{u}^{\dagger}$

Hessian

The Hessian kernel $\mathbf{H}$ can be broken down into three parts[4]

$$\mathbf{H} = \mathbf{H}_1(\mathbf{u}^\dagger, \delta\mathbf{u}) + \mathbf{H}_2(\delta\mathbf{u}^\dagger, \mathbf{u}) + \mathbf{H}_3(\mathbf{u}^\dagger, \mathbf{u})$$

[4] Fichtner 2011.

Hessian

The Hessian kernel **H** can be broken down into three parts[4]

$$\mathbf{H} = \mathbf{H}_1(\mathbf{u}^\dagger, \delta\mathbf{u}) + \mathbf{H}_2(\delta\mathbf{u}^\dagger, \mathbf{u}) + \mathbf{H}_3(\mathbf{u}^\dagger, \mathbf{u})$$

$\mathbf{u}$ – Forward field.

$\mathbf{u}^\dagger$ – Adjoint field.

[4] Fichtner 2011.

Hessian

The Hessian kernel **H** can be broken down into three parts[4]

$$\mathbf{H} = \mathbf{H}_1(\mathbf{u}^\dagger, \delta\mathbf{u}) + \mathbf{H}_2(\delta\mathbf{u}^\dagger, \mathbf{u}) + \mathbf{H}_3(\mathbf{u}^\dagger, \mathbf{u})$$

$\mathbf{u}$ – Forward field.

$\mathbf{u}^\dagger$ – Adjoint field.

$\delta\mathbf{u}$ – Perturbed forward field.

[4] Fichtner 2011.

Hessian

The Hessian kernel **H** can be broken down into three parts[4]

$$\mathbf{H} = \mathbf{H}_1(\mathbf{u}^\dagger, \delta\mathbf{u}) + \mathbf{H}_2(\delta\mathbf{u}^\dagger, \mathbf{u}) + \mathbf{H}_3(\mathbf{u}^\dagger, \mathbf{u})$$

$\mathbf{u}$ – Forward field.

$\mathbf{u}^\dagger$ – Adjoint field.

$\delta\mathbf{u}$ – Perturbed forward field.

$\delta\mathbf{u}^\dagger$ – Perturbed adjoint field.

[4] Fichtner 2011.

Perturbed fields[5]

The perturbed forward field

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

[5] Fichtner and Trampert 2011b.

Perturbed fields[5]

The perturbed forward field

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

The perturbed adjoint field

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

[5] Fichtner and Trampert 2011b.

Hessian kernels

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

Hessian kernels

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

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

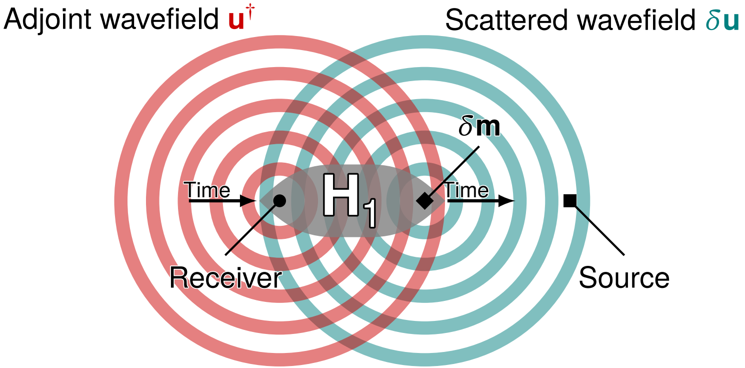
Hessian kernels

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

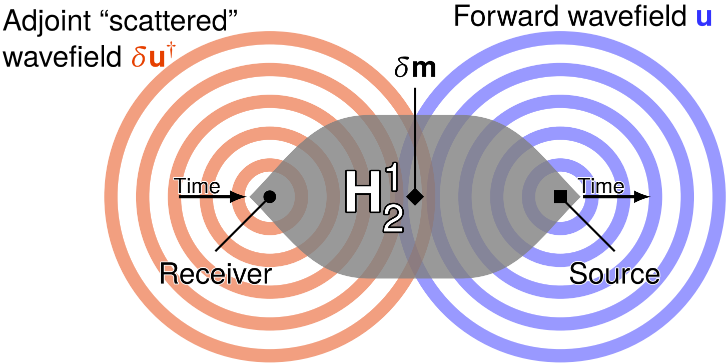
$$\mathbf{H}_2(\delta\mathbf{u}^\dagger, \mathbf{u}) = \int_T \delta\mathbf{u}^\dagger \nabla_m \mathbf{L}(\mathbf{u}, \mathbf{m})\,\mathrm{d}t \qquad = \mathbf{F}(\delta\mathbf{u}^\dagger, \mathbf{u})$$

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

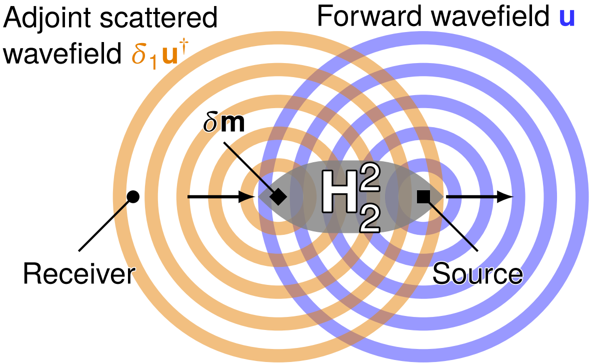
$\mathbf{H}_1(\mathbf{u}^\dagger, \delta\mathbf{u})$



$\mathbf{H}_2(\delta\mathbf{u}^\dagger, \mathbf{u})$ – 1st order scattering



$\mathbf{H}_2(\delta\mathbf{u}^\dagger, \mathbf{u})$ – 2nd order scattering



$\mathbf{H_3(u^\dagger, u)}$

Localised to the perturbation and dependent on the parametrisation

$\mathbf{H}_3(\mathbf{u}^\dagger, \mathbf{u})$

Localised to the perturbation and dependent on the parametrisation

- In the ρ, λ, μ parametrisation $\mathbf{H}_3 \equiv 0$ due to linearity w.r.t parameters.

$\mathbf{H}_3(\mathbf{u}^\dagger, \mathbf{u})$

Localised to the perturbation and dependent on the parametrisation

- In the ρ, λ, μ parametrisation $\mathbf{H}_3 \equiv 0$ due to linearity w.r.t parameters.
- In the ρ, v_p, v_s parametrisation $\mathbf{H}_3$ can be expressed using model parameters and Fréchet kernels.

$\mathbf{H}_3(\mathbf{u}^\dagger, \mathbf{u})$

Localised to the perturbation and dependent on the parametrisation

- In the ρ, λ, μ parametrisation $\mathbf{H}_3 \equiv 0$ due to linearity w.r.t parameters.
- In the ρ, v_p, v_s parametrisation $\mathbf{H}_3$ can be expressed using model parameters and Fréchet kernels.

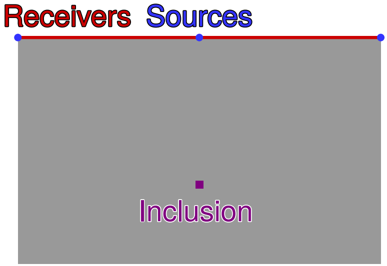
$$\mathbf{H}_3(\mathbf{u}^\dagger, \mathbf{u}) = \begin{bmatrix} H_3^\rho \\ H_3^{v_p} \\ H_3^{v_s} \end{bmatrix} = \begin{bmatrix} 0 & \rho^{-1} F_{v_p} & \rho^{-1} K_{v_s} \\ \rho^{-1} K_{v_p} & v_p^{-1} K_{v_p} & 0 \\ \rho^{-1} K_{v_s} & 0 & v_s^{-1} K_{v_s} \end{bmatrix} \begin{bmatrix} \delta \rho \\ \delta v_p \\ \delta v_s \end{bmatrix}$$

Anatomy of the Hessian

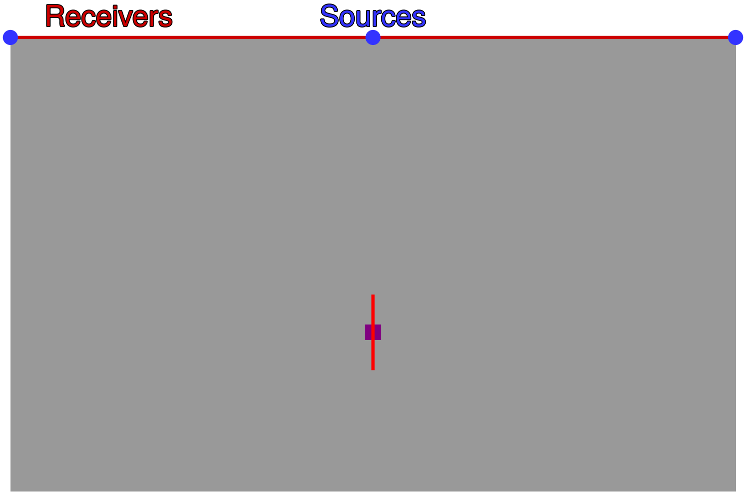
$$\mathbf{H}^m(x_i)\delta m(x_j) = \mathbf{H}_i^m \delta_j^m$$

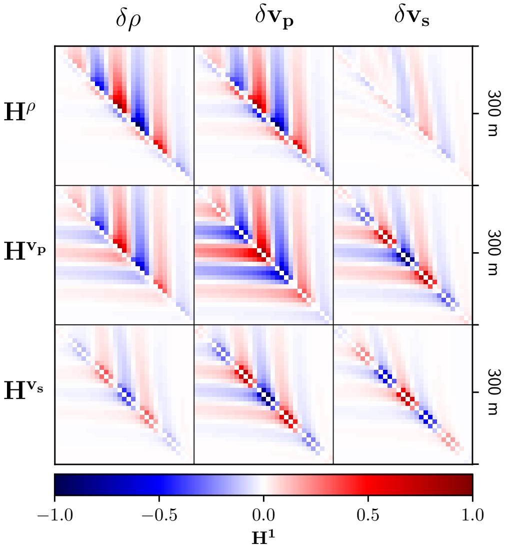
$$\left[\begin{array}{ccc|cc|cc}
\mathbf{H}_0^{\rho}\delta_0^{\rho} & \mathbf{H}_0^{\rho}\delta_1^{\rho} & \cdots & \mathbf{H}_0^{\rho}\delta_0^{v_p} & \cdots & \mathbf{H}_0^{\rho}\delta_0^{v_s} & \cdots \\
\mathbf{H}_1^{\rho}\delta_0^{\rho} & \mathbf{H}_1^{\rho}\delta_1^{\rho} & \ddots & \vdots & \ddots & \vdots & \ddots \\
\mathbf{H}_2^{\rho}\delta_0^{\rho} & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots \\
\hline
\mathbf{H}_0^{v_p}\delta_0^{\rho} & \cdots & \cdots & \mathbf{H}_0^{v_p}\delta_0^{v_p} & \cdots & \mathbf{H}_0^{v_p}\delta_0^{v_s} & \cdots \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots \\
\hline
\mathbf{H}_0^{v_s}\delta_0^{\rho} & \cdots & \cdots & \mathbf{H}_0^{v_s}\delta_0^{v_p} & \cdots & \mathbf{H}_0^{v_s}\delta_0^{v_s} & \cdots \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots
\end{array}\right]$$

Model

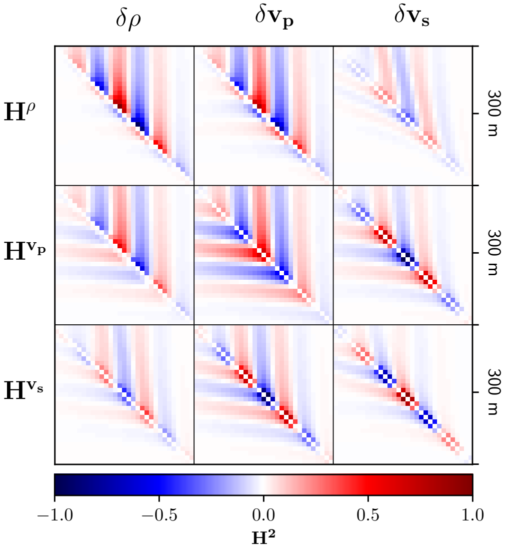


- Elastic 2-D
- 4010 m × 3000 m
- 10 m × 10 m grid cells
- 8 Hz and 32 Hz Ricker
- Background:
  $\rho = 1.5\,\mathrm{kg/m^3}$,
  $v_p = 2.0\,\mathrm{km/s}$,
  $v_s = 1.0\,\mathrm{km/s}$
- Inclusion:
  100 m/s, 30 m × 30 m

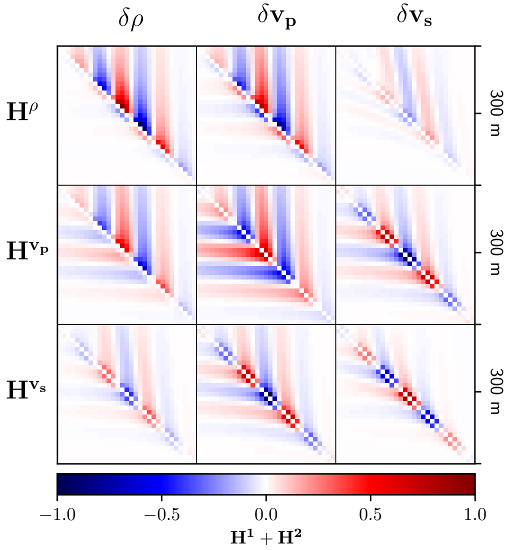

Receivers
Sources



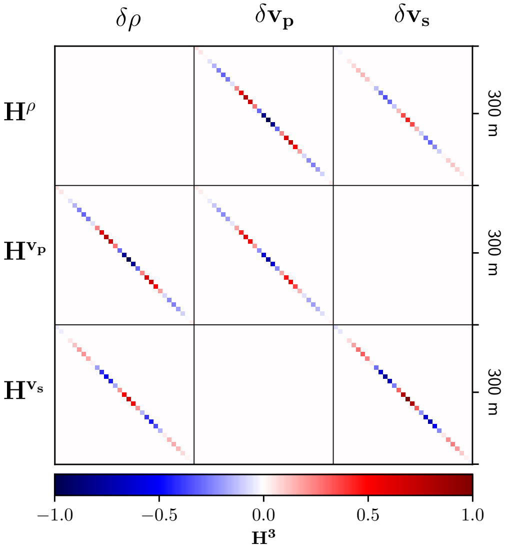
Homogeneous media, 8 Hz - $\mathbf{H}_1$



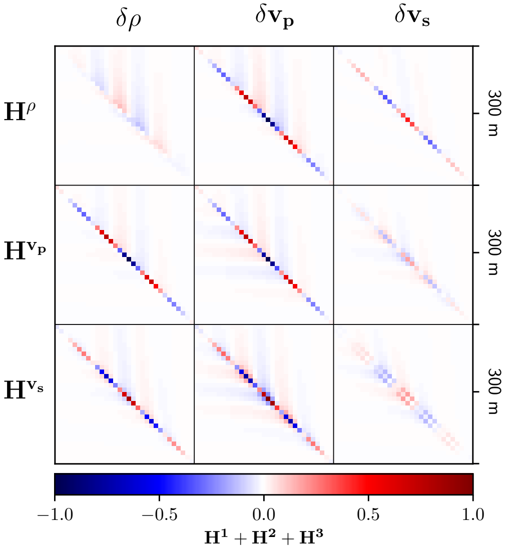
Homogeneous media, 8 Hz - $\mathbf{H}_2$



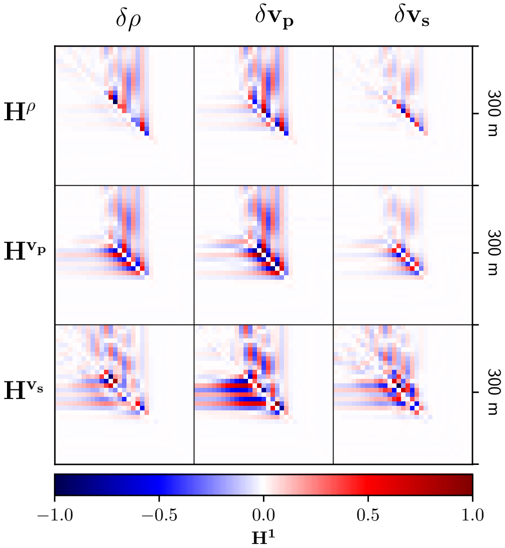
Homogeneous media, 8 Hz - $\mathbf{H}_1 + \mathbf{H}_2$



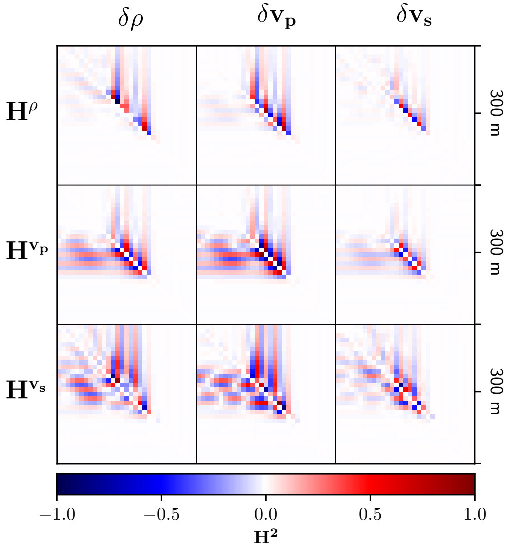
Homogeneous media, 8 Hz - $\mathbf{H}_3$



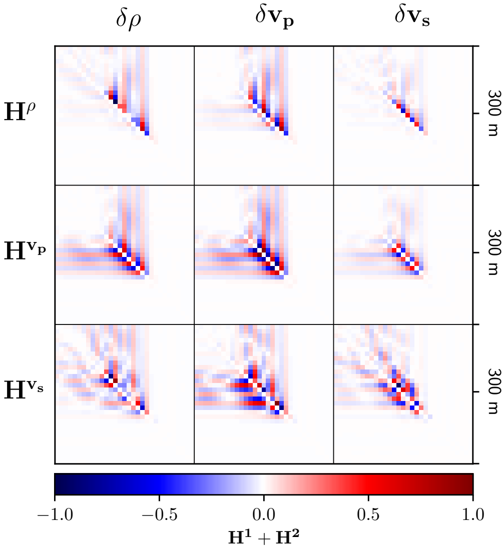
Homogeneous media, 8 Hz - $\mathbf{H}_1 + \mathbf{H}_2 + \mathbf{H}_3$



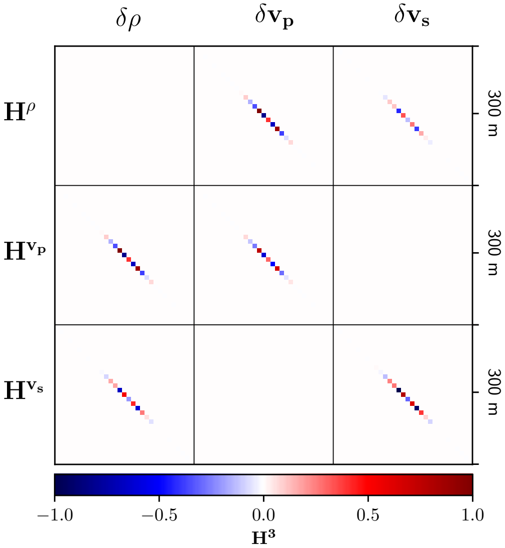
Homogeneous media, 32 Hz - $\mathbf{H}_1$



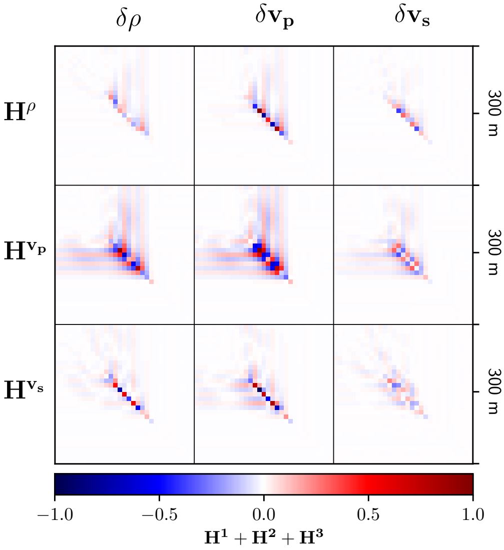
Homogeneous media, 32 Hz - $\mathbf{H}_2$



Homogeneous media, 32 Hz - $\mathbf{H}_1 + \mathbf{H}_2$

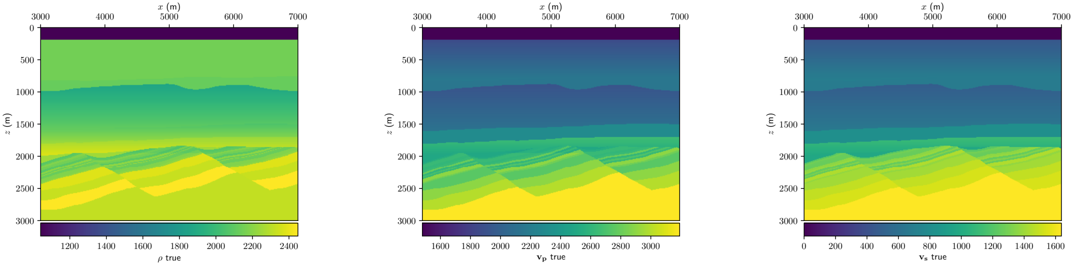


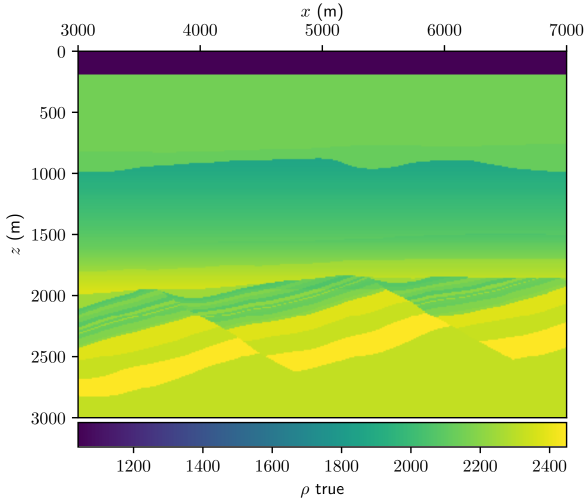
Homogeneous media, 32 Hz - $\mathbf{H}_3$



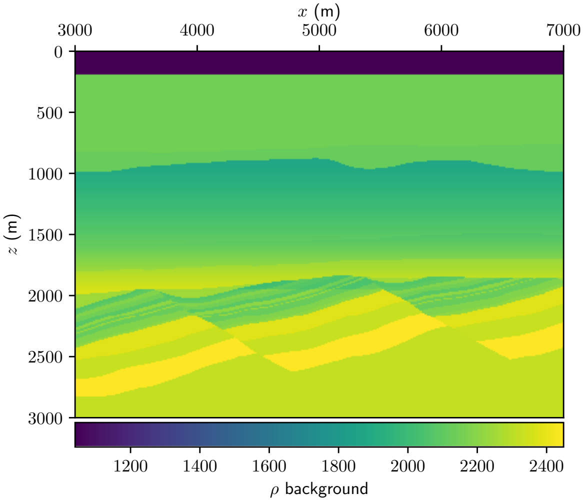
Homogeneous media, 32 Hz - $\mathbf{H}_1 + \mathbf{H}_2 + \mathbf{H}_3$

Gullfaks model

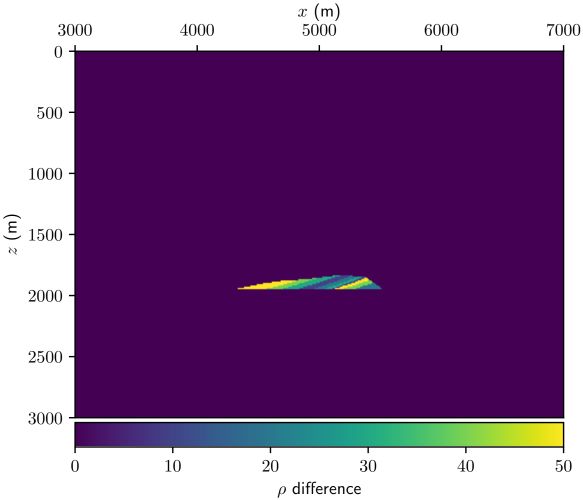




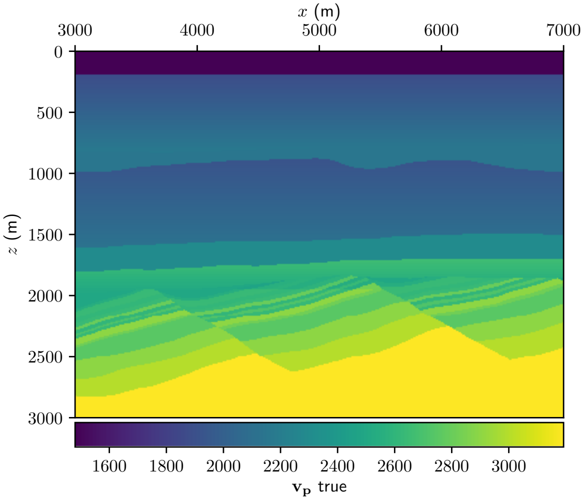
x (m)
3000
4000
5000
6000
7000
0
500
1000
1500
2000
2500
3000
z (m)
1200
1400
1600
1800
2000
2200
2400
ρ true



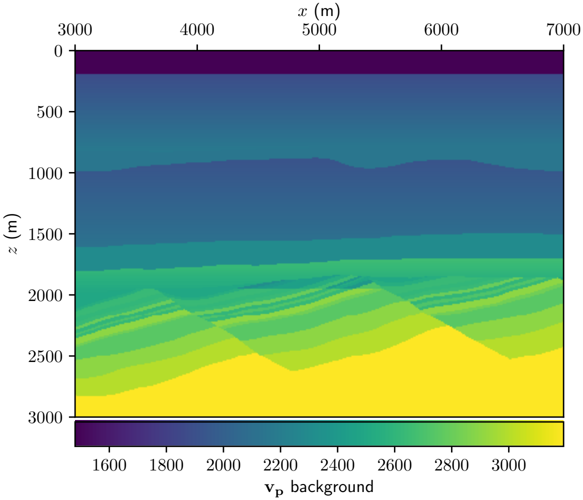
x (m)
3000
4000
5000
6000
7000
0
500
1000
1500
2000
2500
3000
z (m)
1200
1400
1600
1800
2000
2200
2400
ρ background



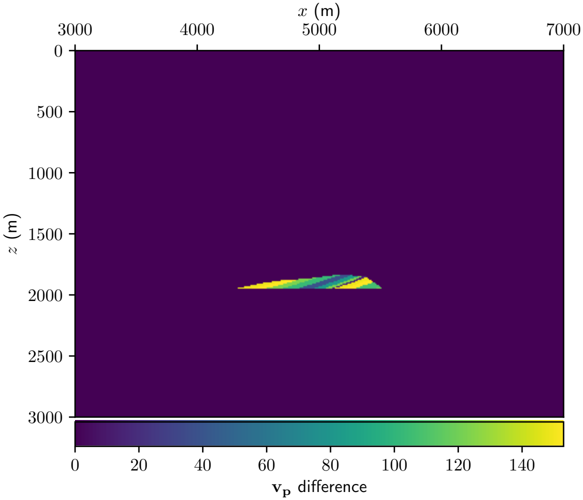
x (m)
3000
4000
5000
6000
7000
0
500
1000
1500
2000
2500
3000
z (m)
0
10
20
30
40
50
ρ difference



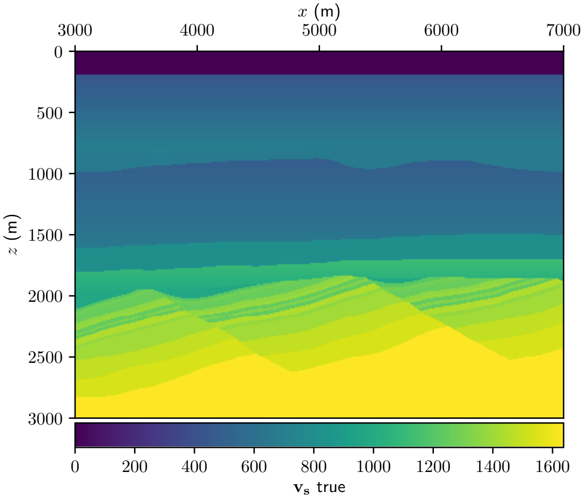
x (m)
3000
4000
5000
6000
7000
0
500
1000
1500
2000
2500
3000
z (m)
1600
1800
2000
2200
2400
2600
2800
3000
$\mathbf{v_p}$ true



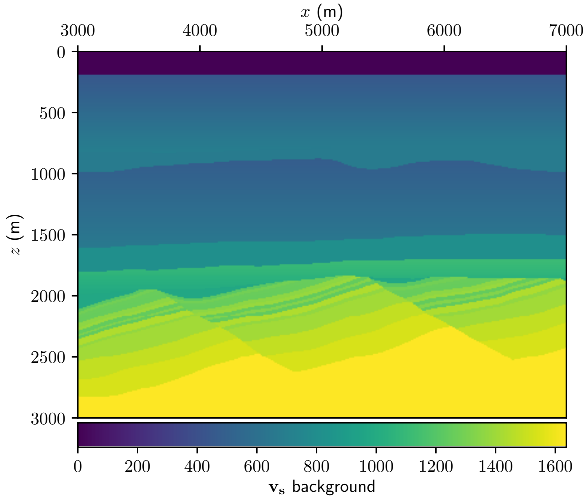
x (m)
3000
4000
5000
6000
7000
0
500
1000
1500
2000
2500
3000
z (m)
1600
1800
2000
2200
2400
2600
2800
3000
$\mathbf{v_p}$ background



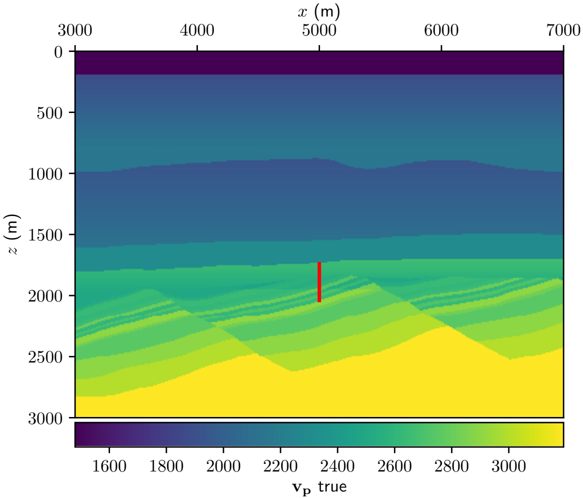
x (m)
3000
4000
5000
6000
7000
0
500
1000
1500
2000
2500
3000
z (m)
0
20
40
60
80
100
120
140
$\mathbf{v_p}$ difference



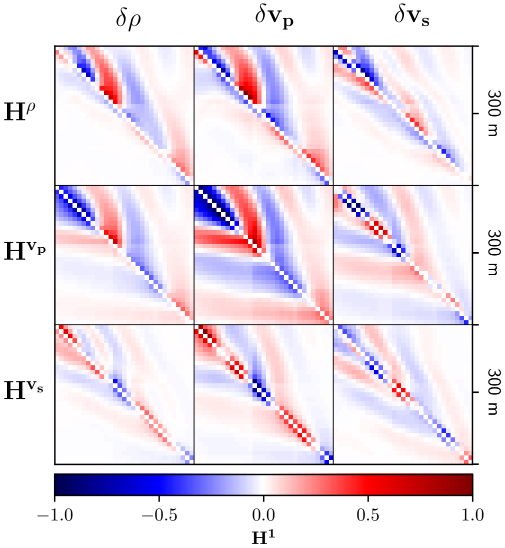
x (m)
3000
4000
5000
6000
7000
0
500
1000
1500
2000
2500
3000
z (m)
0
200
400
600
800
1000
1200
1400
1600
vs true



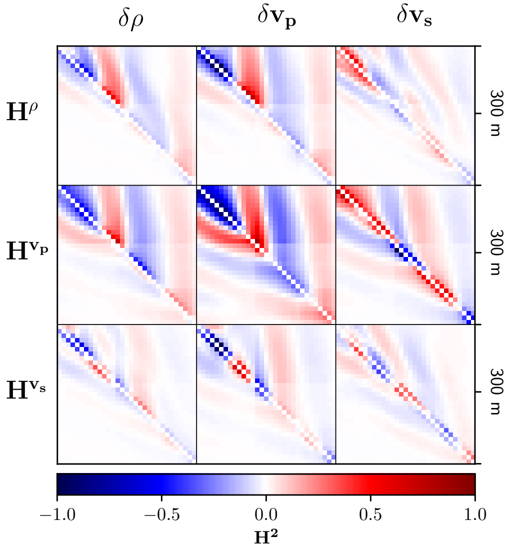
x (m)
3000
4000
5000
6000
7000
0
500
1000
1500
2000
2500
3000
z (m)
0
200
400
600
800
1000
1200
1400
1600
$\mathbf{v_s}$ background



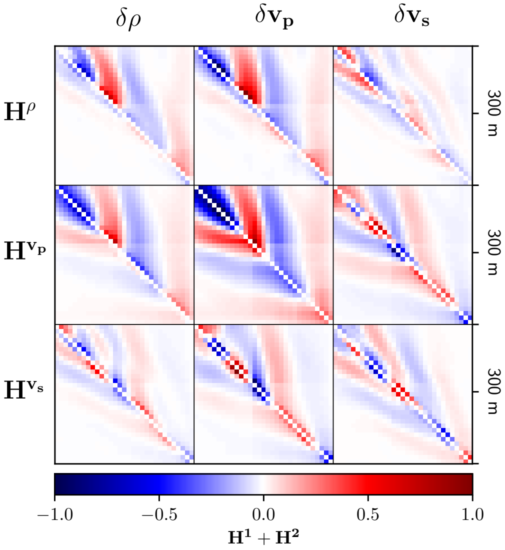
x (m)
3000
4000
5000
6000
7000
0
500
1000
1500
2000
2500
3000
z (m)
1600
1800
2000
2200
2400
2600
2800
3000
$\mathbf{v_p}$ true



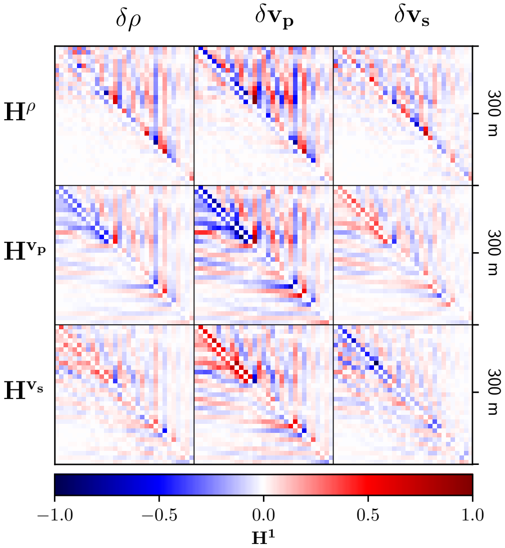
Homogeneous media, 8 Hz - $\mathbf{H}_1$



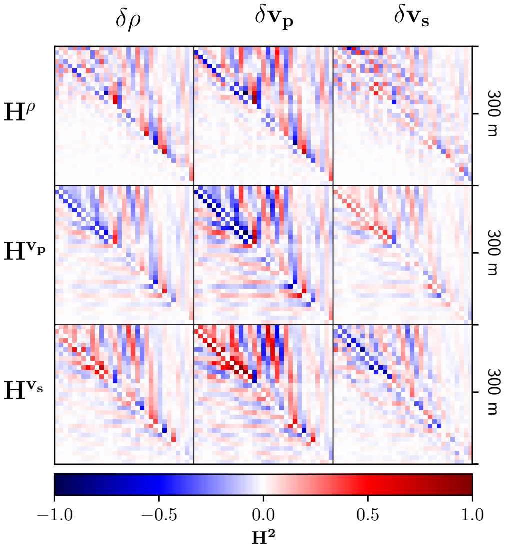
Homogeneous media, 8 Hz - $\mathbf{H}_2$



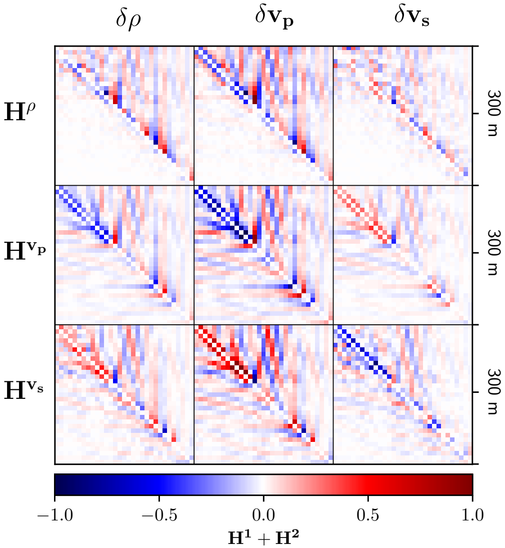
Homogeneous media, 8 Hz - $\mathbf{H}_1 + \mathbf{H}_2$



Homogeneous media, 32 Hz - $\mathbf{H}_1$



Homogeneous media, 32 Hz - $\mathbf{H}_2$



Homogeneous media, 32 Hz - $\mathbf{H}_1 + \mathbf{H}_2$

Conclusions

- Need twice the amount of fields. More computational expensive than linear methods.

Conclusions

- Need twice the amount of fields. More computational expensive than linear methods.
- We can compute the action of the Hessian on a model perturbation without the need to calculate the whole Hessian.

Conclusions

- Need twice the amount of fields. More computational expensive than linear methods.
- We can compute the action of the Hessian on a model perturbation without the need to calculate the whole Hessian.
- Much of the code can be recycled.

Conclusions

- Need twice the amount of fields. More computational expensive than linear methods.
- We can compute the action of the Hessian on a model perturbation without the need to calculate the whole Hessian.
- Much of the code can be recycled.
- Possible to perform accuracy and resolution analysis on the results.

Conclusions

- Need twice the amount of fields. More computational expensive than linear methods.
- We can compute the action of the Hessian on a model perturbation without the need to calculate the whole Hessian.
- Much of the code can be recycled.
- Possible to perform accuracy and resolution analysis on the results.
- Parameter cross-talk analysis.

Future work

- Write a focused full Newton inversion algorithm.
- Investigate cross-talk in different parametrisations.

References

- Tarantola, A. (1984) Inversion of seismic reflection data in the acoustic approximation in: *Geophysics*, **49**:8, 1259 doi: 10.1190/1.1441754
- Mora, P. (1987) Nonlinear two-dimensional elastic inversion of multioffset seismic data in: *Geophysics*, **52**:9, 1211 doi: 10.1190/1.1442384
- Fichtner, A., H. P. Bunge and H. Igel (2006) The adjoint method in seismology. I. Theory in: *Physics of the Earth and Planetary Interiors*, **157**:1-2, 86–104 doi: 10.1016/j.pepi.2006.03.016
- Fichtner, A. (2011) Full Seismic Waveform Modelling and Inversion in: *Advances*, 1981, 83–88 doi: 10.1007/978-3-642-15807-0
- Virieux, J. and S. Operto (2009) An overview of full-waveform inversion in exploration geophysics in: *Geophysics*, **74**:6, WCC1 doi: 10.1190/1.3238367
- Métivier, L, R Brossier, J Virieux and S Operto (2012) The truncated Newton method for Full Waveform Inversion in: *Journal of Physics: Conference Series*, **386**:2, 012013 doi: 10.1088/1742-6596/386/1/012013
- Epanomeritakis, I, V Akçelik, O Ghattas and J Bielak (2008) A Newton-CG method for large-scale three-dimensional elastic full-waveform seismic inversion in: *Inverse Problems*, **24**:3, 034015 doi: 10.1088/0266-5611/24/3/034015
- Pratt, R. G., C Shin and G. J. Hicks (1998) Gauss-Newton and full Newton methods in frequency domain seismic waveform inversion. in: *Geophysical Journal International*, **133**: 341–362
- Fichtner, A. and J. Trampert (2011b) Resolution analysis in full waveform inversion in: *Geophysical Journal International*, **187**:3, 1604–1624 doi: 10.1111/j.1365-246x.2011.05218.x
- Trampert, J., A. Fichtner and J. Ritsema (2013) Resolution tests revisited: The power of random numbers in: *Geophysical Journal International*, **192**: 676–680 doi: 10.1093/gji/ggs057
- Biondi, B., E. Biondi, M. Maharramov and Y. Ma (2015) Dissection of the full-waveform inversion Hessian in: *Unpublished*,
- Fichtner, A. and J. Trampert (2011a) Hessian kernels of seismic data functionals based upon adjoint techniques in: *Geophysical Journal International*, **185**: 775–798 doi: 10.1111/j.1365-246X.2011.04966.x

$\mathbf{H}_1$ kernel

$$\mathbf{H}_1(\mathbf{u}^\dagger, \delta\mathbf{u}) = \begin{bmatrix} H_1^\rho \\ H_1^\lambda \\ H_1^\mu \end{bmatrix} = \int_T \begin{bmatrix} -u_i^\dagger \cdot \delta u_i \\ \varepsilon_{ii}^\dagger \cdot \delta\varepsilon_{jj} \\ 2\varepsilon_{ij}^\dagger \cdot \delta\varepsilon_{ij} \end{bmatrix} \mathrm{d}t$$