A Comparative Study of Three Hessian Approximations in Wave-Equation Migration

Abolhassani, S.; Verschuur, E.

DOI
10.3997/2214-4609.202410492

Publication date
2024

Document Version
Final published version

Citation (APA)
Abolhassani, S., & Verschuur, E. (2024). *A Comparative Study of Three Hessian Approximations in Wave-Equation Migration*. Paper presented at 85th EAGE Annual Conference & Exhibition 2024, Oslo, Lillestrøm, Norway. https://doi.org/10.3997/2214-4609.202410492

2024 EAGE ANNUAL 85TH CONFERENCE & EXHIBITION OSLO | NORWAY

A Comparative Study of Three Hessian Approximations in Wave-Equation Migration

S. Abolhassani[1], E. Verschuur[1]

[1] TU Delft

Summary

Enhanced pre-stack depth migration, characterized by improved resolution and amplitudes, ensures a more accurate representation of the subsurface, proving essential for reducing the likelihood of geological misinterpretations and facilitating informed decision-making in seismic exploration. However, obtaining high-resolution images with preserved amplitudes through standard depth migration could face several hurdles known as migration artifacts. Iterative least-squares migration (LSM) was developed to address these migration artifacts. However, the convergence rate of LSM using a gradient descent approach tends to be slow. Several researchers have attempted to achieve computational efficiency in linearized LSM through gradient preconditioning. In the context of iterative least-squares wave-equation migration, this extended abstract compares three minimization approaches that differ in error functions and gradient preconditioning—including the depth-dependent Hessian approximation inverse—through two numerical examples, one with an inverse-crime scenario and the other with a non-inverse-crime scenario.

A Comparative Study of Three Hessian Approximations in Wave-Equation Migration

Introduction

Pre-stack depth migration in exploration seismology transforms seismic data into a subsurface depth image, aiming for accurate positioning of seismic amplitudes. However, obtaining high-resolution images with preserved amplitudes through standard pre-stack depth migration faces several hurdles associated with irregular acquisition geometry and organized noises, among others, degrading image accuracy, known as migration artifacts (Jones, 2018).

Iterative least-squares migration (LSM) represents an advanced variant of pre-stack depth migration, where the subsurface reflectivity model (image) receives iterative updates through an iterative migration-demigration cycle to converge towards an "ideal representation" of the reflectivity model, aiming to mitigate the migration artifacts (Nemeth et al., 1999). However, the convergence rate of LSM using a gradient descent approach tends to be rather slow. It may require quite a few iterations to get higher accuracy due to the band-limited nature of seismic data, making it a demanding computational effort.

Linearized LSM, where the forward problem is linearized through the Born approximation, aims to accelerate the convergence rate and achieve the highest possible accuracy within the resolution power of the acquired seismic data. It minimizes S, a quadratic function of $\delta\mathbf{r}^{\cup}$,

$$S(\delta\mathbf{r}^{\cup}) = \frac{1}{2}\int_{\omega} d\omega \sum_{s} \|\delta\mathbf{d}_{s,\omega} - \mathbf{J}_{s,\omega}\delta\mathbf{r}^{\cup}\|_2^2, \tag{1}$$

with the minimizer (Lines and Treitel, 1984),

$$\frac{\partial S}{\partial \delta\mathbf{r}^{\cup}} = 0 \;\;\rightarrow\;\; \delta\mathbf{r}^{\cup} = -\frac{\mathbf{Re}\left\{\int_{\omega} d\omega \sum_{s} \mathbf{J}_{s,\omega}^{\dagger}\delta\mathbf{d}_{s,\omega}\right\}}{\mathbf{Re}\left\{\int_{\omega} d\omega \sum_{s} \mathbf{J}_{s,\omega}^{\dagger}\mathbf{J}_{s,\omega}\right\}}, \tag{2}$$

in which $\delta\mathbf{r}^{\cup}$ is the upward reflectivity model perturbation vector, ω is an angular frequency component, s means a shot, $\delta\mathbf{d}$ denotes the data residual vector, $\dagger$ means adjoint, $\mathbf{J}$ stands for the partial derivatives of the modeled wavefield, known as the Jacobian matrix, and $\mathbf{J}^{\dagger}\mathbf{J}$ is known as the Hessian approximation matrix ($\mathbf{H}^{\mathrm{a}}$). The upward reflectivity model can be then updated iteratively via $\mathbf{r}_{k+1}^{\cup} = \mathbf{r}_{k}^{\cup} + \alpha_k\delta\mathbf{r}_{k}^{\cup}$, where k and α are the iteration number and the minimization step length, respectively. If one assumes that S is a quadratic function of the $\delta\mathbf{r}_{\omega}^{\cup}$ (frequency-dependent reflectivity model perturbation), the minimizer transforms to (Jang et al., 2009; Oh and Min, 2013),

$$\delta\mathbf{r}^{\cup} = -\int_{\omega} d\omega\, \delta\mathbf{r}_{\omega}^{\cup} = -\int_{\omega} d\omega \frac{\mathbf{Re}\left\{\sum_{s} \mathbf{J}_{s,\omega}^{\dagger}\delta\mathbf{d}_{s,\omega}\right\}}{\mathbf{Re}\left\{\sum_{s} \mathbf{J}_{s,\omega}^{\dagger}\mathbf{J}_{s,\omega}\right\}}, \tag{3}$$

in which, an effective deconvolution effect on the source signature is facilitated through the sequential application of the denominator (Hessian approximation) to the gradient at each frequency, compared to Equation 2.

Over time, the replacement of the denominator in Equations 2 and 3 (gradient preconditioning) with cost-effective alternatives has been pursued by several researchers to ensure the computational efficiency of linearized LSM in large-scale applications (e.g., Shin et al., 2001; Plessix and Mulder, 2004; Choi et al., 2008; Lu et al., 2018). In the context of space-frequency linearized least-squares one-way wave-equation migration, which uses angle-independent reflection/transmission coefficients and involves upward/downward wavefield extrapolation between virtual data levels within the medium (Berkhout, 2014a), Abolhassani and Verschuur (2022) introduced a depth-dependent $\mathbf{H}^{\mathrm{a}}$. They partition $\mathbf{H}^{\mathrm{a}}$ into smaller operators, each capturing the correlation of partial derivative wavefields at identical depth levels. Constructed and applied (its inverse) directly at each virtual data level within the medium, the depth-dependent Hessian approximation efficiently minimizes the associated computational expenses, as the number of model parameters per virtual data level represents only a small fraction of the total. Combined into a single matrix, these operators form a block-diagonal approximation for $\mathbf{H}^{\mathrm{a}}$.

Table 1: Investigated approaches

Approach	Error Function	Update Equation	Denominator in the Update Equation	Explanation
AP-c	$S(\delta\mathbf{r}^{\cup})$	2	"Equation 27" in Plessix and Mulder (2004) (i.e., diagonal approximation of $\mathbf{H}^{a}$)	–
AP-1	$S(\delta\mathbf{r}^{\cup}_{\omega})$	3	depth-dependent version of $\sum_s \mathbf{J}^{\dagger}_{s,\omega}\mathbf{J}_{s,\omega}$ (i.e., block-diagonal approximation of $\mathbf{H}^{a}$)	–
AP-2	$S(\delta\mathbf{r}^{\cup}_{\omega})$	3	depth-dependent version of $(\sum_s \mathbf{J}_{s,\omega})^{\dagger}\sum_s \mathbf{J}_{s,\omega}$ (i.e., block-diagonal approximation of $\mathbf{H}^{a}$)	including interference between sources in the denominator but mathematically cheaper compared to the AP-1 approach

In this extended abstract, we initially compare three minimization approaches, as listed in Table 1, using an "inverse-crime" scenario. Subsequently, given the similar performance of AP-1 and AP-2, but with AP-2 being cheaper, we further compare AP-c and AP-2 using a full-wavefield "non-inverse-crime" scenario. We end the paper with conclusions.

Theory

In LS-WEM, we use a space-frequency one-way wavefield extrapolation scheme called primary-wavefield modeling (PWMod) (Berkhout, 2014b). This scheme models angle-independent two-way primary reflection data, including reflection and transmission effects,

$$\mathbf{p}^{-}_{\text{mod}}(z_0,\omega) = \sum_{m=N}^{1} \overbrace{\left[\left[\prod_{n=1}^{m-1}\mathbf{W}^{-}_{z_{n-1};z_n}\mathbf{T}^{-}(z_n)\right]\mathbf{W}^{-}_{z_{m-1};z_m}\right]}^{\mathbf{L}^{-}_{z_0;z_m}} \left(\mathbf{r}^{\cup}(z_m)\circ\mathbf{p}^{+}_{\text{mod}}(z_m,\omega)\right) \tag{4}$$

$$\mathbf{p}^{+}_{\text{mod}}(z_m,\omega) = \left[\left[\prod_{n=m-1}^{1}\mathbf{W}^{+}_{z_{n+1};z_n}\mathbf{T}^{+}(z_n)\right]\mathbf{W}^{+}_{z_1;z_0}\right]\mathbf{s}^{+}(z_0,\omega), \tag{5}$$

in which $\mathbf{p}^{-}_{\text{mod}}(z_0,\omega)$ is the monochromatic upgoing wavefield modeled at z_0, $\mathbf{p}^{+}_{\text{mod}}(z_m,\omega_f)$ is the monochromatic downgoing wavefield modeled at z_m, $\mathbf{W}^{-}_{z_{n-1};z_n}$ is the upward phase-shift extrapolation matrix-operator from z_n to z_{n-1}, $\mathbf{W}^{+}_{z_{n+1};z_n}$ is the downward phase-shift extrapolation matrix-operator from z_n to z_{n+1}, $\mathbf{r}^{\cup}(z_m)$ represents the angle-independent upward reflectivity vector-operator at z_m, $\mathbf{T}^{\pm}(z_n)$ denotes the downward/upward transmission diagonal-matrix-operator $\left(\mathbf{T}^{\pm}(z_n)=\mathbf{I}\pm\text{diag}(\mathbf{r}^{\cup}(z_n))\right)$, $\mathbf{s}^{+}(z_0,\omega_f)$ shows the monochromatic downgoing physical source at the Earth's surface, N is the total number of virtual depth levels, and the symbol $\circ$ means element-wise product.

Compared to PWMod, FWMod (full-wavefield modeling) takes into account the multiple scattering by replacing Equation 5 with (Berkhout, 2014b),

$$\mathbf{P}^{+}_{\text{mod}}(z_m,\omega) = \sum_{m=1}^{N}\mathbf{L}^{+}_{z_m;z_0}\left(\mathbf{s}^{+}(z_0,\omega)+\mathbf{r}^{\cap}(z_m)\circ\mathbf{p}^{-}_{\text{mod}}(z_m,\omega)\right), \tag{6}$$

where $\mathbf{r}^{\cap}(z_m)$ is the angle-independent downward reflectivity vector-operator at z_m with $\mathbf{r}^{\cap}=-\mathbf{r}^{\cup}$. Now, running the recursive summation in Equation 4 for multiple iterations results in multiple scattering.

In the introduction, we outlined three minimization approaches for linearized least-squares one-way wave-equation migration, listed in Table 1, to be examined and compared in this study, namely AP-c, AP-1, and AP-2, where AP stands for amplitude preserved. The nominator in all update equations ("gradient") for a pair of source and frequency components and a given depth level (z_m) reads,

$$\mathbf{g}_s(z_m,\omega) = \mathbf{Re}\left\{\left(\mathbf{p}^{+}_{\text{mod,s}}(z_m,\omega)\right)^{*}\circ\overbrace{\left(\mathbf{L}^{-^{\dagger}}_{z_0;z_m}\delta\mathbf{d}^{-}_{s}(\omega)\right)}^{\text{back-propagated wavefield}}\right\}, \tag{7}$$

where $*$ denotes the complex conjugate transpose.

To calculate the denominators for AP-c, AP-1, and AP-2, the only absent component is the Jacobian matrix. Abolhassani and Verschuur (2022) presented the expression of $\mathbf{J}_s(z_m,\omega)$ as,

$$\mathbf{J}_s(z_m,\omega) = \left[p^{+}_{1\ \text{mod},s}(z_m,\omega_f)\mathbf{L}^{-}_{*,1_{z_0;z_m}} \quad p^{+}_{2\ \text{mod},s}(z_m,\omega_f)\mathbf{L}^{-}_{*,2_{z_0;z_m}} \quad \dots \quad p^{+}_{n_x\text{mod},s}(z_m,\omega_f)\mathbf{L}^{-}_{*,n_x z_0;z_m}\right], \tag{8}$$

in which $\mathbf{L}^{-}_{*,j_{z_0:z_m}}$ denotes the jth column of $\mathbf{L}^{-}_{z_0:z_m}$, j denotes the lateral location of a model parameter at z_m, n_x denotes the number of model parameters at z_m.

Examples

In the first example, our true model is a section of the BP 2004 model, including gas clouds. To enhance the model, we introduce three point diffractors to the model, as shown in Figure 1. The model consists of a grid with 351 × 326 points in both horizontal and vertical directions. On top of the model, we have positioned 51 shot points with a spacing of 175 meters. For each shot, 351 receivers spaced 25 meters apart record reflection data. The source function used in this example is a Ricker wavelet with a dominant frequency of 10 Hz. Our forward modeling tool for generating the observed dataset is PWMod, i.e., only primaries, applied on the true model ("inverse-crime" scenario), with a recording time of 4.092 seconds. Employing an accurate migration velocity model (without diffractors), we evaluate and compare AP-c, AP-1, and AP-2. Figures 2a, 2b, and 2c compare the results after 5 iterations.

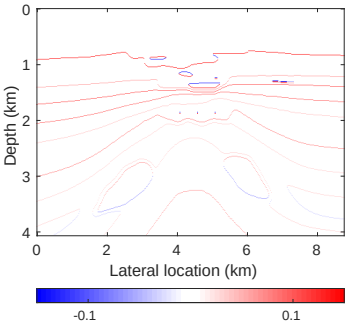
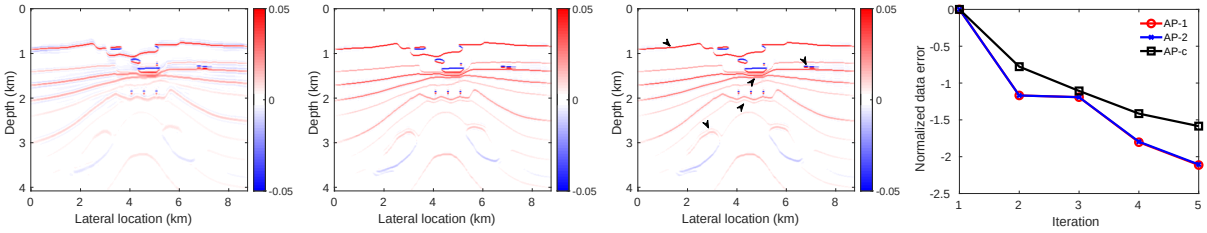


Figure 1: BP 2004 model: True image



(a) AP-c (b) AP-1 (c) AP-2 (d) Convergence history

Figure 2: BP 2004 model example: Estimated images after 5 iterations and convergence history

AP-1 and AP-2 outperform AP-c in resolution and amplitude preservation. Figure 2c highlights significant disparities. Notably, interference between neighboring mono-frequency sources in AP-2 doesn't degrade the image. Figure 2d confirms the faster convergence of AP-1 and AP-2 over AP-c. Given the similar performance of AP-1 and AP-2, the second example exclusively compares AP-2 with AP-c, as AP-2 is cheaper than AP-1 due to requiring fewer matrix-matrix multiplications.

In the second example, we use a density-driven wedge model (Figure 3) with alternating reflection coefficients ($r^{\cup} = \pm 0.1$), where the true velocity model is a homogeneous velocity of 2000 m/s. Acoustic finite-difference modeling with a 1 m grid and a 15 Hz Ricker wavelet generates observed reflection data. For 2.92 s, 401 receivers, planted on top of the model, at 10 m spacing record 31 shots at 120 m spacing from 200 m to 3800 m, triggered sequentially on top of the model. For inversion, with the same source-receiver setup and the true homogeneous velocity model, FWMod with a 10 m horizontal grid and a 4 m vertical grid models the reflection data in each iteration (non-inverse-crime scenario). With this experimental setup, we evaluate and compare AP-c and AP-2. Figure 4 displays the estimated images after 5 and 10 iterations.

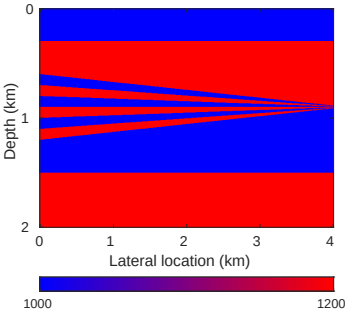
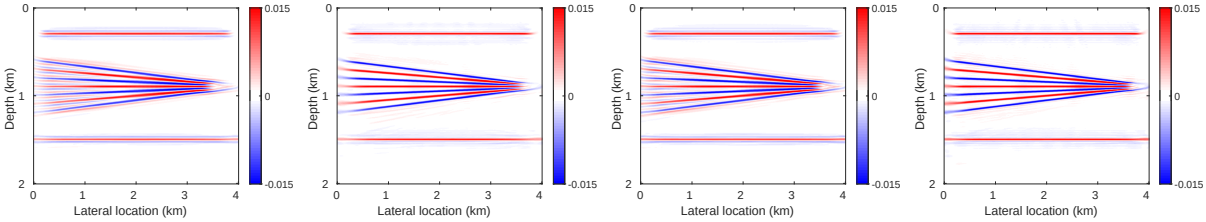
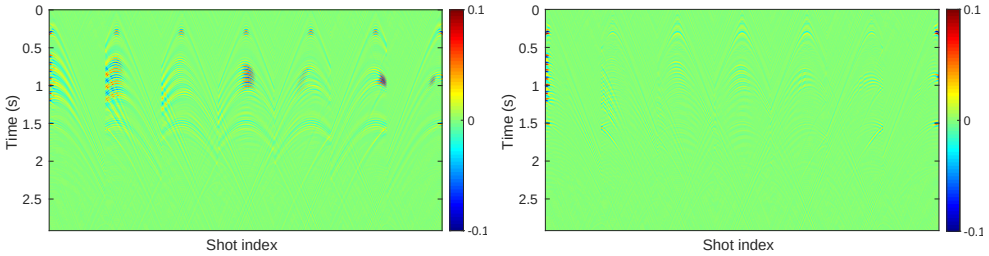


Figure 3: Wedge model: True density



(a) AP-c: Iteration 5 (b) AP-2: Iteration 5 (c) AP-c: Iteration 10 (d) AP-2: Iteration 10

Figure 4: Wedge model example: Estimated images after 5 and 10 iterations



(a) AP-c (b) AP-2

Figure 5: Wedge model example: Data residuals after 10 iterations for selected sources.

As seen in Figure 4, AP-2 exhibits superior resolution and significantly better-preserved amplitudes in half the iterations of AP-c. Once again, and now in a non-inverse-crime scenario, we observe that the interference between adjacent mono-frequency sources in AP-2 doesn't degrade the image. Figure 5 depicts the data residual of AP-c and AP-2 after 10 iterations for selected sources. It confirms that AP-2 reaches smaller data errors (faster convergence in terms of iterations).

Conclusions

This study highlights the notable impact of depth-dependent Hessian approximations in large-scale linearized least-squares migration, with AP-2 distinguishing itself in resolution power and amplitude perseverance. The example with "non-inverse-crime" scenario pinpoints the convergence benefit (in terms of iterations) of AP-2 over AP-c, achieving superior images and less data errors in half the iterations. The resistance of AP-2 to the interference between sources adds further to its appeal. These findings make clear the potential of depth-dependent Hessian approximations to enhance seismic imaging, fostering improved subsurface characterization and interpretation.

Acknowledgements

The authors appreciate Delphi consortium members for their support and engaging discussions.

References

Abolhassani, S. and Verschuur, E. [2022] Fast Gauss-Newton full-wavefield migration. SEG/AAPG, 2709–2713.

Berkhout, A.J. [2014a] Review paper: An outlook on the future of seismic imaging, part II: Full-wavefield migration. *Geophysical Prospecting*, **62**(5), 931–949.

Berkhout, A.J. [2014b] Review paper: An outlook on the future seismic imaging, part I: Forward and reverse modelling. *Geophysical Prospecting*, **62**(5), 911–930.

Choi, Y., Min, D.J. and Shin, C. [2008] Frequency-domain elastic full waveform inversion using the new pseudo-Hessian matrix: Experience of elastic Marmousi-2 synthetic data. *Bulletin of the Seismological Society of America*, **98**(5), 2402–2415.

Jang, U., Min, D.J. and Shin, C. [2009] Comparison of scaling methods for waveform inversion. *Geophysical Prospecting*, **57**(1), 49–59.

Jones, I.F. [2018] *Velocities, imaging, and waveform inversion*. EAGE.

Lines, L. and Treitel, S. [1984] A review of least-squares inversion and its application to geophysical problems. *Geophysical Prospecting*, **32**(2), 159–186.

Lu, S., Liu, F., Chemingui, N., Valenciano, A. and Long, A. [2018] Least-squares full-wavefield migration. *The Leading Edge*, **37**(1), 46–51.

Nemeth, T., Wu, C. and Schuster, G.T. [1999] Least-squares migration of incomplete reflection data. *Geophysics*, **64**(1), 208–221.

Oh, J.W. and Min, D.J. [2013] Weighting technique using backpropagated wavefields incited by deconvolved residuals for frequency-domain elastic full waveform inversion. *Geophysical Journal International*, **194**(1), 322–347.

Plessix, R.E. and Mulder, W.A. [2004] Frequency-domain finite-difference amplitude-preserving migration. *Geophysical Journal International*, **157**(3), 975–987.

Shin, C., Jang, C.S. and Min, D.J. [2001] Improved amplitude preservation for prestack depth migration by inverse scattering theory. *Geophysical Prospecting*, **49**(5), 592–606.