### Full waveform inversion using envelope-based global correlation norm

<table>
<thead>
<tr>
<th>Item Type</th>
<th>Article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors</td>
<td>Oh, Juwon; Alkhalifah, Tariq Ali</td>
</tr>
<tr>
<td>Eprint version</td>
<td>Post-print</td>
</tr>
<tr>
<td>DOI</td>
<td>10.1093/gji/ggy031</td>
</tr>
<tr>
<td>Publisher</td>
<td>Oxford University Press (OUP)</td>
</tr>
<tr>
<td>Journal</td>
<td>Geophysical Journal International</td>
</tr>
<tr>
<td>Rights</td>
<td>This is a pre-copyedited, author-produced PDF of an article accepted for publication in Geophysical Journal International following peer review. The version of record is available online at: <a href="https://academic.oup.com/gji/advance-article-abstract/doi/10.1093/gji/ggy034/4830112">https://academic.oup.com/gji/advance-article-abstract/doi/10.1093/gji/ggy034/4830112</a>.</td>
</tr>
<tr>
<td>Download date</td>
<td>2023-11-01 10:45:21</td>
</tr>
<tr>
<td>Link to Item</td>
<td><a href="http://hdl.handle.net/10754/627058">http://hdl.handle.net/10754/627058</a></td>
</tr>
</tbody>
</table>
Full waveform inversion using envelope-based global correlation norm

Ju-Won Oh\textsuperscript{1,2)*}, and Tariq Alkhalifah\textsuperscript{1)}

1) King Abdullah University of Science and Technology

Physical Science and Engineering Division,
4700, King Abdullah University of Science and Technology, Thuwal
23955-6900, Kingdom of Saudi Arabia

2) (Currently) Chonbuk National University

Department of Resources and Energy Engineering,
567, Baekje-daero, Deokjin-gu, Jeonju-si, Jeollabuk-do
54896, Republic of Korea

* Corresponding author

TEL: 82-63-270-2364

e-mail: juwon24@jbnu.ac.kr

Abstract

To increase the feasibility of full waveform inversion on real data, we suggest a new objective function, which is defined as the global correlation of the envelopes of modeled and observed data. The envelope-based global correlation norm has the advantage of the envelope inversion that generates artificial low-frequency information, which provides the possibility
to recover long-wavelength structure in an early stage. In addition, the envelope-based global correlation norm maintains the advantage of the global correlation norm, which reduces the sensitivity of the misfit to amplitude errors so that the performance of inversion on real data can be enhanced when the exact source wavelet is not available and more complex physics are ignored.

Through the synthetic example for 2D SEG/EAGE overthrust model with inaccurate source wavelet, we compare the performance of 4 different approaches, which are the least squares waveform inversion, least squares envelope inversion, global correlation norm and envelope-based global correlation norm. Finally, we apply the envelope-based global correlation norm on the 3D OBC data from the North Sea. The envelope-based global correlation norm captures the strong reflections from the high-velocity caprock and generates artificial low-frequency reflection energy that helps us recover long wavelength structure of the model domain in the early stages. From this long-wavelength model, the conventional global correlation norm is sequentially applied to invert for higher-resolution features of the model.

**Keywords:** full waveform inversion, objective function, envelope-based global correlation norm, artificial low frequency
Introduction

Full waveform inversion (FWI) is a promising seismic data interpretation tool to estimate subsurface model parameters through an inverse mathematical procedure. Recently, our attention in FWI has focused on improving its reliability in handling real data without low-frequency components as reliable low-frequency components are hard to acquire in real experiments. For this reason, without a good initial model that is reasonably kinematically accurate, FWI usually converges to local minima of the corresponding objective function (Virieux and Operto, 2009).

The common used approaches to extract long wavelength propagation velocities are traveltime tomography (Woodward et al., 2008) and migration velocity analysis (MVA; Symes, 2008). These two approaches have been widely applied to obtain initial models for FWI for real seismic data acquired for oil-exploration purposes (Ravaut et al., 2004; Operto et al., 2006; Malinowsky and Operto, 2008). However, drawbacks in these methods still exist. Using traveltime tomography, reliable picking of desired arrivals is an important component (Virieux and Operto, 2009) and MVA requires an extension in the image domain, which increases the computational cost (Symes, 2008), particularly in the 3D problem.

In the hope to extract low wavenumbers from inversion methods, many geophysicists have developed techniques that provide such low wavenumbers from an inversion implementation. Shin and Cha (2008) developed Laplace-domain inversion that estimates the gradient direction in the Laplace domain to build a smoothed velocity model from the data without low-frequency components. They extended it to Laplace-Fourier domain inversion to increase the resolution limit (Shin and Cha, 2009). Alkhalifah (2015) suggested a scattering-angle based filtering of the waveform inversion gradients. Using the scattering-angle
filtering, the wide scattering angle components that have long-wavelength updates can be extracted in the early stage of FWI (Miller et al., 1987). As another way to avoid cycle skipping, Warner and Guasch (2016) introduced adapted waveform inversion (AWI), in which least squares convolutional filters are designed to transform the modeled data into the observed data with a penalty applied to the filter. Reflected waveform inversion (RWI) is also a promising candidate to obtain a smooth initial velocity model (Xu et al, 2012; Wu and Alkhalifah, 2015) for FWI. In RWI, we obtain a smooth velocity model by inverting for wave path updates to the reflector (so-called ‘rabbit-ear’). Another interesting way to measure the misfit between observed and modeled data avoiding cycle skipping, Métilvier et al. (2016) suggested an objective function based on an optimal transport distance that allows us to account for the lateral coherency of events within the seismograms. From a different point of view, Li and Demanet (2016) introduced a way to obtain artificially low-frequency data extrapolated from the higher frequency data.

Among the many new approaches used to build a starting model for FWI, here, we focus on the envelope inversion (EI). Bozdağ et al. (2011) first introduced the envelope inversion for the global seismology problem by utilizing the envelopes of wavefields to measure the misfit between envelopes of observed and modeled data. By converting the wavefields to their envelopes, envelope inversion offers the possibility to obtain artificial low-frequency data. Utilizing this feature, Wu et al., (2014) and Chi et al., (2014) adopted the least squares envelope inversion in the context of exploration seismology. Following that, many worked on enhancing the feasibility of the envelope inversion. Huang et al. (2015) examined the possibility of applying the envelope inversion within an elastic FWI and Gao et al. (2016) combined frequency selection strategy with envelope inversion to improve its performance. Han et al. (2016) suggested a modified envelope inversion, so-called spline envelope
inversion. For a large-scale 3D problem, the feasibility of using the envelope inversion on synthetic elastic data (Borisov et al., 2016) and synthetic anisotropic data (Wang and Tsvankin, 2016) has been verified. However, the feasibility of applying the envelope inversion on real field data has not yet been examined.

In this study, we increase the reliability of the envelope inversion to handle real data by introducing the envelope-based global correlation norm that is defined as the global correlation of the envelopes of wavefields. The synthetic and real data examples demonstrate that the new objective function generates artificial low-frequency data, which provide long wavelength subsurface models, with less sensitivity to amplitude errors.
Least squares norm: FWI / EI

The objective function using the $l_2$ norm of residuals between modeled ($u$) and field ($d$) data in the time domain can be expressed as (Tarantola, 1986)

$$E_{FWI}^{L_2}(m) = \sum_s \sum_r \left[ u(s,r,m) - d(s,r) \right]^2,$$

where $r$ and $s$ denote the receiver and source, respectively; $m$ represents the model parameter. The gradient and the adjoint source ($b_{FWI}^{L_2}$) are expressed as

$$\frac{\partial E_{FWI}^{L_2}(m)}{\partial m_i} = \sum_s \sum_r \left[ \frac{\partial u(s,r,m)}{\partial m_i} \right] b_{FWI}^{L_2}(s,r,m)$$

and

$$b_{FWI}^{L_2}(s,r,m) = u(s,r,m) - d(s,r),$$

respectively. For least squares FWI, the adjoint source is the residual wavefield, which directly measures the difference between modeled and observed data. However, as illustrated by many (Virieux and Operto, 2009), the least squares FWI suffers from cycle skipping problem, which is caused when an event in the modeled data is more than a half-cycle away from its location in the observed data. As we use lower frequencies, FWI becomes more robust to cycle skipping as the half-cycle is larger. However, in general, reliable low-frequency energy is usually absent in field data. To solve this problem, the envelope inversion was proposed to produce artificial low frequency (Bozdaş et al., 2011; Wu et al., 2014; Chi et al., 2014). The objective function for the least squares EI can be expressed as

$$E_{EI}^{L_2}(m) = \sum_s \sum_r \left[ e^p(s,r,m) - e^p_d(s,r) \right]^2,$$
where $p$ is the power of envelope (Wu et al., 2014). Here, $e_u$ and $e_d$ are the envelopes of modeled and observed data, respectively, as follows:

$$e_u(s,r,m) = \sqrt{u^2(s,r,m) + u^2_H(s,r,m)}$$

and

$$e_d(s,r) = \sqrt{d^2(s,r) + d^2_H(s,r)},$$

where $u_H$ and $d_H$ are Hilbert-transformed modeled and observed wavefields. The gradient direction for a model parameter from EI (Wu et al., 2014) is expressed as

$$\frac{\partial E^{L2}_{EI}(m)}{\partial m_i} = \sum_s \sum_r \left[ \frac{\partial u(s,r,m)}{\partial m_i} \cdot b^{L2}_{EI}(s,r,m) \right]$$

with

$$b^{L2}_{EI}(s,r,m) = \left[ A^{L2}u(s,r,m)e_u^p(s,r,m) \right. \left. H \left\{ A^{L2}u_H(s,r,m)e_u^p(s,r,m) \right\} \right]$$

and

$$A^{L2} = e_u^p(s,r,m) e_d^p(s,r).$$

However, the least squares FWI, as well as the least squares EI, are sensitive to the amplitude misfit, which is an issue especially when the exact source wavelet is not available (Warner and Guasch, 2015).
Envelope-based global correlation norm

To make the envelope inversion more practical for real data, we suggest a new objective function for the envelope inversion by combining the original envelope inversion with the global correlation norm, which is less sensitive to amplitude errors.

Review of global correlation norm and its gradients

Before introducing the envelope-based global correlation norm (EGCN), we first review the gradient and adjoint source of the global correlation norm (GCN). The global correlation norm (Routh et al., 2011; Choi and Alkhalifah, 2012) is defined as the zero-lag cross-correlation between normalized modeled and observed wavefields, expressed by

\[
E_{\text{Global}}^{\text{FWI}}(m) = \left\{ \hat{u}(s,r,m) \hat{d}(s,r) \right\},
\]

where

\[
\hat{u}(s,r,m) = \frac{u(s,r,m)}{||u(s,r,m)||}
\]

and

\[
\hat{d}(s,r) = \frac{d(s,r)}{||d(s,r)||}.
\]

By taking the partial derivative with respect to the model parameter, we obtain the gradient direction from the global correlation norm as expressed by
\[
\frac{\partial E^{Global}_{FWI}(m)}{\partial m_i} = \sum_s \sum_r \left[ \frac{\partial u(s,r,m)}{\partial m_i} \cdot b^{Global}_{FWI}(s,r,m) \right]
\]  

(13)

with the adjoint source given by

\[
b^{Global}_{FWI}(s,r,m) = \frac{1}{\|u(s,r,m)\|} \left\{ \hat{u}(s,r,m) \left( \hat{u}(s,r,m) \times \hat{d}(s,r) \right) \right\} \hat{d}(s,r). \]

(14)

However, the conventional global correlation norm, as well as the least squares FWI, also suffers from the cycle skipping problem when we start FWI from an inaccurate initial model and reliable low-frequency data are not available.

**Envelope-based global correlation norm and its gradients**

The envelope-based global correlation norm is defined as the global correlation of normalized envelopes of the modeled and observed data as below:

\[
E^{Global}_{ELJ}(m) = \sum_s \sum_r \left\{ \hat{e}^p_u(s,r,m) \times \hat{e}^p_d(s,r) \right\}.
\]

(15)

where

\[
\hat{e}^p_u(s,r,m) = \frac{e^p_u(s,r,m)}{\|e^p_u(s,r,m)\|}
\]

(16)

and

\[
\hat{e}^p_d(s,r) = \frac{e^p_d(s,r)}{\|e^p_d(s,r)\|}.
\]

(17)

Like in the global correlation norm case, the gradient direction is expressed by
The adjoint source using the envelope-based global correlation norm has both features of the envelope inversion in eq. (8) and global correlation norm in eq. (14) as follows:

$$b_{EI}^{Global}(s,r,m) = p A^{Global} u(s,r,m) e_u^p (s,r,m) H(A^{Global} u_H(s,r,m) e_u^p (s,r,m))$$

(19)

with

$$A^{Global} = \frac{1}{||e_u^p (s,r,m)||} \left\{ \hat{e}_u^p (s,r,m) \hat{e}_d^p (s,r) \times \hat{e}_d^p (s,r) \right\}$$

(20)

As Wu et al. (2014) showed, the envelope of wavefields ($e_d$) has artificial low frequencies even though the original wavefields ($d$) may not contain low frequencies. For this reason, the adjoint source in eq. (19) includes artificial low frequencies generated by the conversion of wavefields to their envelopes. In addition, as shown in eq. (20), the adjoint source of EGCN still maintains the characteristics of the adjoint source of GCN in eq. (14), in which the global correlation operator increases the robustness in the presence of unexpected amplitude misfits. In the general case of FWI applied to real data, it is hard to trust the amplitude of wavefields because of the inaccurate source wavelet, attenuation, elastic effects, preprocessing to enhance the signal to noise ratio and so on. Thanks to the normalization feature of GCN and the artificial low frequencies in EI, the EGCN improves the performance of FWI for real data in the early iterations by providing artificial low-frequency data that are less sensitive to the amplitude misfit.
Multi-stage FWI with envelope-based global correlation norm

As Wu et al., (2014) suggested in the least squares FWI and EI, we also propose a two-stage inversion strategy by switching the inversion from EI with EGCN to FWI with GCN. In the 1st stage, the envelope inversion using EGCN is applied to capture long-wavelength features of the subsurface. Then, FWI with GCN is sequentially applied to the inverted model from EI with EGCN to add the high-resolution structures. For a synthetic model and real data from the North Sea, we show the feasibility of the two-stage inversion strategy.

Synthetic example: 2D SEG/EAGE overthrust model

For a synthetic 2D SEG/EAGE overthrust model (Figure 1a), we compare the performance of 4 approaches; FWI with $L_2$, EI with $L_2$, FWI with GCN and EI with EGCN. Total 150 pressure sources and 600 hydrophones are aligned regularly on the water surface with 80 m and 20 m intervals, respectively. The true source is a Ricker wavelet, whose peak and maximum frequencies are 5 Hz and 10 Hz, respectively. On the other hand, the source wavelet for the inversion is a first derivative Gaussian function, whose peak and maximum frequencies are 2.5 and 10 Hz, respectively. We also apply a damping function over time ($e^{-1.2t}$) to each trace in the observed data, which can be practically caused by weak gain control or errors in the 3D-to-2D amplitude corrections. This mismatch between the observed and modeled data also happens in general because of the incompleteness of the physics we assume to represent Earth and the preprocessing of raw data that affects the amplitude of wavefields. We filtered out the low-frequency data below 3 Hz. For this synthetic data set, we
perform the acoustic inversion starting from a smoothed initial P-wave velocity model (Figure 1b). This smoothed initial model is required because the artificial ultra-low frequency components (near 0 Hz) from envelope inversion are not reliable based on our experiences. The power of the envelope ($p$) is 1. As the power of envelope increases, EI focuses on more energetic waves in the data. Thus we can miss some weak reflections from deeper structures. The gradient direction is scaled by the diagonal of pseudo-Hessian matrix (Shin et al., 2001). We use a fixed step length (50 m/s).

Figure 2 shows the observed and modeled data with their envelopes. Because of the inaccurate source wavelet and damping effects, the direct wave in modeled data ($u$) looks quite different from that in observed data ($d$). In this case, the least squares FWI and EI that match the amplitude fall into local minima as Warner and Guasch (2015) showed. Figure 3 shows the adjoint sources using the least squares FWI and EI, FWI with GCN and EI with EGCN for different frequency bands. Because we muted out data with frequencies below 3 Hz, the adjoint sources from the FWI with $L_2$ or GCN does not contain reliable low-frequency information. On the other hand, the EI with $L_2$ or EGCN produces artificial low-frequency data, which are generated by converting the original wavefields to their envelopes in Figure 2. With the help of artificial low-frequency data, the EI with $L_2$ or EGCN produces long wavelength structure (Figure 4). Thus the gradient directions have low-frequency features. However, the EI with $L_2$, as well as FWI with $L_2$, suffers from strong artifacts in the shallow part because of the large amplitude error in direct waves (yellow arrows in Figure 3) courtesy of an inaccurate source wavelet (Figure 2). On the other hand, EGCN and GCN produce better gradient directions regardless of the source signature.

From these observations, we notice that multi-stage inversion from EI with EGCN to FWI with GCN helps us obtain long-wavelength structures in the initial stage using EGCN and...
then short-wavelength structures can be recovered by GCN. Because the least squares FWI and EI suffer from the potential amplitude mismatch that cannot be compensated in the modeling process, we ignore them in this comparison. In Figure 5, we compare the performances of least-squares FWI, a pure GCN and the proposed two-stage FWI. Because of the severe amplitude differences, the least squares FWI failed (Figure 5a) and, due to the lack of low-frequency data, GCN (after 30 iteration) only inverts sharp boundaries of the subsurface layers (Figure 5b). On the other hand, EI with EGCN (after 20 iteration), thanks to the artificial low-frequency information, recovers long wavelength structures (Figure 5c). In the 2nd stage (after 20 iteration), the high-resolution features can be recovered by FWI with GCN (Figure 5d).

Real data application: North Sea OBC data

We further verify the feasibility of envelope-based global correlation norm on real OBC data from North Sea (Szydlik et al., 2007; Alkhalifah et al., 2016). The initial P-wave velocity model (Figure 6b) is a smoothed version the layer-based tomographic model (Figure 6a; Jones, 2003; Szydlik et al., 2007) obtained by applying a smoothing operator of length 1 km. The background density is 1 g/cm$^3$. The model domain is 12.3 km $\times$ 6.8 km $\times$ 4.5 km. Total 12 ocean bottom cables with 400 m separation are used and each cable contains 240 geophones with 25 m interval. To perform the inversion on such a large 3D area, we solve the first-order acoustic wave equations using the staggered grid finite-difference method with domain decomposition (Bohlen, 2002). Among the many ways to avoid storing the huge forward (source) wavefields (Nguyen and McMechan, 2015) in 3D FWI, we choose the time-
frequency hybrid domain approach (Sirgue et al., 2008), in which the forward and backward wavefields are transformed to the frequency domain on the fly so that we can calculate the gradient direction in the frequency domain within the desired frequency band. To scale the gradient direction, the diagonal of pseudo-Hessian matrix is used (Shin et al., 2001). We use a fixed step length (50 m/s). The observed data correspond to the vertical component (particle acceleration) measured at the bottom of the sea and even PS waves are muted. For these reasons, the acoustic inversion with first-order acoustic wave equation should be good enough.

The direct waves are muted by the data-owner so least squares FWI and EI are not capable of handling the data set. In addition, due to the lack of direct waves that mainly preserve the source signature, the source wavelet estimation was also not successful. For this reason, least squares FWI and EI are not practical without elaborate muting on the direct arrival on the synthetic data at each iteration. This muting can be complicated when initial shallow velocity is too slow because we might still have direct waves in the synthetic data outside the muting window, particularly at far offset. Despite these limitations, GCN and EGCN show reasonable performance because the global correlation norm is capable of reducing the influence of the amplitude component of waves, as well as, reduce the contribution of data not predicted by the modeled data (Choi and Alkhalifah, 2012).

At first, we conduct the inversion for a low-frequency band of the data, specifically from 2.75 to 10 Hz using 300 sources, which are almost regularly placed on the water surface with about a 500 m spacing interval along both the inline and crossline directions. Figure 7 shows seismograms of the modeled and observed data, as well as their envelopes and adjoint sources from FWI with GCN and EI with EGCN, which correspond to a line given by the black dashed line from a source represented by the black dot in Figure 6b. At first, we
observe strong reflections in the observed data (red arrow in Figure 7a). The envelope of the observed data also captures this strong reflection. As we showed in the synthetic example, converting the wavefields to their envelopes generates artificial low-frequencies even though real low-frequencies were filtered out. While the adjoint source using FWI with GCN does not contain reliable low-frequency components (Figure 7c), the adjoint source using EI with EGCN includes artificial low frequencies that include the strong reflection (Figure 7d). Thanks to this artificial low-frequency information, the gradient direction from EGCN at 1\textsuperscript{st} iteration (Figure 8b) shows long-wavelength features compared to that from GCN (Figure 8a), particularly near the high-velocity caprock. On the other hand, FWI using GCN (after 30 iteration) might fall into a local minimum (Figure 9a) compared to the tomography result (Figure 6a). This is also clear in the vertical profile of the inverted velocity compared to the well-log velocity (Figure 10a). Due to the lack of reliable low-frequency information, GCN failed to detect the right depth information, particularly for the high-velocity caprock (black arrow). This is also related to the cycle skipping observed at the reflection events from this strong velocity contrast (red arrow in Figure 11b). On the other hand, the gradient using EGCN (Figure 8b) offers reasonable long-wavelength structures. As a result, EI with EGCN (after 20 iteration) successfully detects the smoothed background velocity providing an accurate location of high-velocity caprock as also supported by the tomography result (Figure 6a) and the well-log velocity (Figure 10b). In Figure 11c (red arrow), we also observe that the cycle skipping issue is generally resolved after EGCN. Using the inverted model by EGCN (Figure 9b) as an initial model, we further perform FWI with GCN to recover the high-frequency features. In the 2\textsuperscript{nd} stage, two different frequency bands are sequentially inverted from low (2.75 – 10 Hz with 300 sources) to high (2.75 – 20 Hz with 1200 sources, which are almost regularly placed on the water surface) frequencies. Figure 9c (after 20 iteration) shows the final result that contains some high-resolution features of the subsurface. In the depth
profiles, the inverted model compares well with well log velocity. The inverted model also produces quite reasonable reflections compared to the observed data (red arrows in Figure 11d). At the near offset (left red arrow), the reflected waves are matched well while, at far offset (right red arrow), the modelled reflected waves look slower. The possible reason could be our ignorance of the anisotropic parameters since the Volve area has significant anisotropy above the caprock (Alkhalifah et al., 2016). The angle domain common image gathers (yellow arrows in Figure 12; Sava and Fomel, 2003) are also improved after using the proposed two-stage inversion.

In summary, all the Quality Control (QC) measures listed above indicate that the proposed inversion strategy provides reasonable velocity structure near the high-velocity caprock. We think that the large velocity contrast near the caprock generates strong reflection events (red arrow in Figure 7). For this reason, EI was dominated by this strong reflection event in the early stage as is shown in the adjoint source of the artificial low frequency (Figure 7d). Thanks to the artificial low-frequency band, we obtained smoothed updates at this depth (near 2.5 km in depth; Figure 8b) and finally obtained a smoothed velocity model at the 1st stage (Figure 9b). For this reason, the updated model using EGCN (Figure 9b) is quite similar to the layer-based tomography result (Figure 6a), which is also based on the reflection events. The improvements in the angle gathers are also observed near the caprock.

Conclusions

We suggest the envelope-based global correlation norm, which improves the inversion performance on real data. The envelope-based global correlation norm is designed by combining the envelope attribute of signals and the global correlation norm to utilize their features in providing artificial low frequencies that are insensitive to amplitude misfits. In the
synthetic example for 2D SEG/EAGE overthrust model, we observed that the least squares envelope inversion, which provides artificial low-frequency data, suffers from the amplitude mismatch caused by the inaccurate source wavelet and the preprocessing applied to the raw data. On the other hand, the global correlation norm is reasonably immune from amplitude errors, but prone to the classic cycle skipping problem when the initial velocity is poor or low frequencies are not available. The envelope attribute provides artificial low frequency, and combining it with the global correlation norm, reduces its sensitivity to misfits in the amplitude that can be caused by many factors (i.e. an incomplete physics). For the 3D real OBC data from the North Sea, we observed that the envelope-based global correlation norm generated artificial low-frequency data to build the long-wavelength structure in the early stages. Then global correlation norm helps to get higher resolution updates of the area. In our experience, the ultra-low frequency band from the envelope inversion is still not reliable, and thus, the success of the envelope inversion still relies on the initial macro-velocity model. Further research is needed to mitigate this limitation.

**Acknowledgements**

Research reported in this publication was supported by competitive research funding from King Abdullah University of Science and Technology (KAUST) in Thuwal, Saudi Arabia. This research was also supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017R1C1B5077123). The authors would like to thank Statoil ASA and the Volve license partners ExxonMobil Exploration and Production Norway AS and Bayerngas Norge AS for the release of the Volve data. We would like to thank Marianne Houbiers from Statoil for
providing some helpful suggestions and corrections. For computer time, this research used the resources of the Supercomputing Laboratory in KAUST. We thank the members of Seismic Wave Analysis Group (SWAG) in KAUST for their helpful discussion. We would like to thank Lapo Boschi as the Associate Editor, Jean-Xavier Dessa and one anonymous reviewer for their helpful suggestions.

References


Figure 1 (a) The true SEG/EAGE overthrust model and (b) smoothed initial model.
**Figure 2** The observed data from the true model \(d\), modeled data from the initial model \(u\) and their envelopes \(e_d\) and \(e_u\).
Figure 3 The adjoint sources using FWI with $L_2$ ($h_{FWI}^{L2}$), FWI with GCN ($h_{FWI}^{Global}$), EI with $L_2$ ($h_{EI}^{L2}$) and EI with EGCN ($h_{EI}^{Global}$) from the frequency bands [0 to 10 Hz] and [0 to 3 Hz]. Note that both modeled and observed data are filtered in the frequency domain and then the adjoint source is calculated in the time domain after inverse Fourier transform. That’s why we have some numerical errors (caused by FFT) in the low-frequency band of $L_2$ FWI and GCN.
Figure 4 The gradient direction at the 1st iteration using FWI with L2, FWI with GCN, EI with L2 and EI with EGCN from each adjoint source in Figure 3. Note that the gradient directions are calculated in the frequency domain within the frequency bands shown in each figure after the Fourier transform of the source and receiver wavefields. Also, note that the
gradient directions of $L_2$ FWI and GCN for $[0 \text{ – } 3 \text{ Hz}]$ are very weak but they are enhanced after normalization.
**Figure 5** The inverted models using (a) least-squares FWI, (b) FWI with GCN, (c) EI with EGCN in 1st stage and (d) FWI with GCN in 2nd stage.
Figure 6 (a) The tomography result and (b) a smoothed version of it acting as the initial velocity model. Notice that different scale bars are used for horizontal and vertical slices.

Note that yellow dashed lines indicate the positions of vertical and horizontal slices. Red and blue dots indicate the coordinates of an oblique well at top and bottom (Figure 10). A black
A dot and a black dashed line denote the locations of a source and a receiver cable, which are used to compare the shot gathers (Figure 11) and angle gathers (Figure 12).
Figure 7 The seismograms of (a) observed and modeled wavefields, (b) their envelopes and adjoint sources using (c) GCN and (d) EGCN for original wavefields and low-pass filtered versions of them. The source is located on the black dot and receivers are aligned along the black-dashed line in Figure 6b. The red arrows indicate strong reflections from the high-velocity caprock. Note that both modeled and observed data are filtered in the frequency domain then the adjoint source is calculated in the time domain after inverse Fourier transform.
Figure 8 The gradient direction at 1st iteration using (a) GCN and (b) EGCN.
Figure 9 (a) The inverted velocity model using only FWI with GCN. The inverted velocity model (b) after 1st stage using EI with EGCN and (c) after 2nd stage using FWI with GCN.

Figure 10 The depth profiles along well path from a red dot on the top and a blue dot at the bottom in Figure 6b. The black arrows indicate the upper boundary of the high-velocity caprock. The well log velocities upper 1.6 km in depth are not available thus they are interpolated. Note that, to avoid a complexity caused by overlapping many curves in the same figure, (a) conventional approaches (FWI with only GCN and tomography) and (b) the proposed approach (1st stage EI with EGCN and 2nd stage FWI with GCN) are displayed in separate figures to compare them with well log and initial velocities.
Figure 11 Comparison of observed data to the modeled data along a black dashed line from a source at the black dot in Figure 6b: (a) initial model, (b) inverted model using FWI with GCN and inverted models using (c) EI with EGCN in 1st stage and (d) FWI with GCN in 2nd stage.
Figure 12 Angle domain common image gathers along the black dashed line in Figure 6b from (a) initial model and (b) final inverted model.