

The value of constrained conjugate-gradient least-squares migration in seismic inversion: Application to a fractured-basement play, Texas Panhandle

Thang Ha¹ and Kurt Marfurt¹

Abstract

Seismic inversion has become almost routine in quantitative 3D seismic interpretation. To ensure the quality of the seismic inversion, the input seismic data need to have a high signal-to-noise ratio. With the current low oil price environment, seismic reprocessing is often preferred over reacquisition to improve data quality. Common filter pairs include forward and inverse f - k and Radon transforms. Forward and inverse migrations (i.e., migration and demigration) are a more recently introduced transform pair that, when used together in an iterative workflow, results in a least-squares migration algorithm. Least-squares migration compensates for surface variation in data density and, when combined with a filter applied to prestack migrated images, suppresses the operator and data aliasing. We apply a least-squares migration workflow to a fractured-basement data set from the Texas Panhandle to demonstrate the enhancement in signal-to-noise ratio, the reduction in acquisition footprint and migration artifacts, and the improvement in the P-impedance inversion result.

Introduction

Thanks to improved algorithms and simple user interfaces, seismic impedance inversion has become a routine means of incorporating well data to estimate lithologic properties in 3D quantitative interpretation. However, the quality of seismic inversion depends on the quality of the seismic amplitude data. To improve seismic data quality, we either acquire and process new data, or reprocess old data to obtain a better, more amplitude-friendly image. Even though reacquiring data using a denser, wider azimuth survey is more likely to produce better images, such acquisition is costly and time consuming. Given the current low oil price, reprocessing might be the only feasible choice for many operators.

Reprocessing old seismic data involves many different tasks, such as surface-consistent residual statics correction, velocity model refinement, coherent noise suppression, trace balancing, 5D interpolation, prestack time/depth migration, and other forms of data conditioning. Each of these tasks contributes to the final image improvement. In this paper, we focus on the last two tasks — migration and data conditioning through the use of a constrained conjugate-gradient least-squares migration method.

Nemeth et al. (1999) were perhaps the first to use migration and demigration (least-squares migration) as a seismic processing pair with Nemeth et al. (2000) showing how one can separate signal from noise. Most sub-

sequent least-squares migration has focused on marine data, in which the acquisition, while aliased, is still regular, resulting in Yu et al.'s. (2006) migration deconvolution algorithm. Since then, the use of least-squares migration has further advanced, with modern implementations by Zeng et al. (2014) inverting for impedance changes rather than reflectivity at offset. Applications of the least-squares migration of 3D land data have been much more limited, with Guo et al. (2016) applying constrained least-squares migration to Mississippi Lime plays in Ness Co., Kansas, and Osage Co., Oklahoma, and Verma et al. (2016) applying constrained least-squares migration to a Mississippi Lime survey in north Texas. Our work builds on these last two publications.

We begin our paper with a short review of the least-squares migration workflow, with mathematical and workflow implementation relegated to the appendices. We then describe the data quality and processing challenges for a 3D data volume acquired over fractured basement in the Texas Panhandle. Although this is a modern wide-azimuth survey acquired in 2013 with a nominal bin size of 82.5×82.5 ft, the shallow target at 2500 ft results in a strong operator aliasing and acquisition footprint. We apply conventional migration and constrained least-squares migration to this data volume and compare the results. Finally, we validate our findings through seismic attributes, improved well ties, and P-wave impedance inversion.

¹The University of Oklahoma, Norman, Oklahoma, USA. E-mail: thang.n.ha-1@ou.edu; kmarfurt@ou.edu.

Manuscript received by the Editor 14 December 2016; revised manuscript received 7 March 2017; published online 15 May 2017. This paper appears in *Interpretation*, Vol. 5, No. 3 (August 2017); p. SN13–SN23, 11 FIGS.

<http://dx.doi.org/10.1190/INT-2016-0240.1>. © 2017 Society of Exploration Geophysicists. All rights reserved.

Method

Migration is central to seismic imaging. The main idea behind migration is to broadcast each amplitude sample onto an ellipsoidal pattern computed from the velocity model and then stack all the resulting ellipsoids (Figure 1). If adequately sampled on the surface and given an accurate velocity model, areas of constructive interference result in reflectors and diffractors. Other areas of destructive interference result in low reflectivity. Compared with simple stacked images of normal moveout-corrected gathers, migration collapses diffractors and moves (or “migrates”) dipping reflectors to steeper dips in the updip direction.

If the surface data are insufficiently sampled, the image suffers from two types of aliasing. First, the ellipsoids associated with reflected and diffracted energy may not destructively interfere at steep dips, giving rise to operator aliasing (Figure 1). Second, undersampled noise, such as ground roll and shallow diffractions, which (if properly sampled) should exhibit short apparent wavelength and strong moveout and should be filtered out by migration, are instead aliased to longer apparent wavelength and gentler moveout and will be passed and subsequently imaged by migration (Dev and McMechan, 2009; Pramik, 2011). Biondi (2001) describes a common workflow to suppress operator aliasing that simply filters out higher frequencies when migrating to steeper dips.

Thus, the key objectives of our least-squares migration workflow are to (1) suppress aliased signal and aliased noise in the final image, (2) preserve rather than reject the higher frequency information of steeply dipping reflectors (Zeng et al., 2017), and (3) compensate for irregular surface sampling that gives rise to acquisition footprint and other amplitude artifacts.

To understand constrained conjugate-gradient least-squares migration, let us break down the method into four elements: migration, least squares, conjugate gradient, and constraint. The meaning of each element is

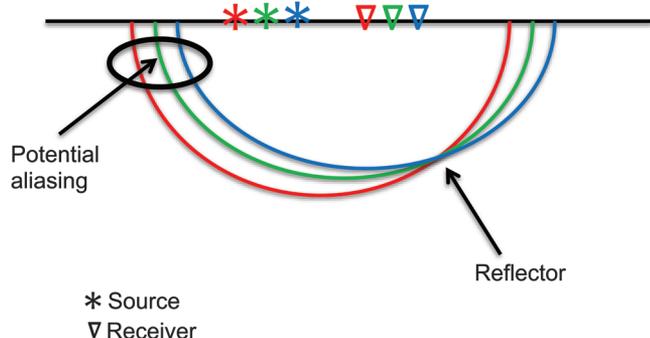


Figure 1. Schematic showing the migration process. The main idea is to copy each amplitude value along the ellipsoids and then stack all the subsequent images together. If the data are sufficiently sampled, some areas constructively interfere into reflectors, whereas others destructively interfere into zero-data zones. If the data are insufficiently sampled, there may be only partial destructive interference, resulting in aliasing.

discussed below, in that order, with details given in the appendices.

Migration

The migration operator can be understood as a filter implemented as a matrix operator, applied to the prestack data to produce a reflectivity model (i.e., the migrated images). The reverse operator is demigration (more commonly known as the forward-modeling operator). This demigration operator is usually denoted as \mathbf{G} (for Green’s function), which is a filter that, when applied to a reflectivity model \mathbf{m} , would produce the prestack raw seismic data \mathbf{d} :

$$\mathbf{d} = \mathbf{G}\mathbf{m}. \quad (1)$$

We want to solve for the reflectivity model \mathbf{m} . To do so, we need to invert operator \mathbf{G} :

$$\mathbf{m} = \mathbf{G}^{-1}\mathbf{d}. \quad (2)$$

However, in almost all cases, the reflectivity model \mathbf{m} and the prestack raw seismic data \mathbf{d} do not have the same dimensional configuration, and therefore, the demigration operator \mathbf{G} is not only a nonsquare matrix but may also be “rank deficient,” which physically means that some areas of the subsurface are poorly illuminated (such as the shadow zone in Kirchhoff migration). Such a matrix \mathbf{G} cannot be directly inverted. Fortunately, we can approximate the inverse of \mathbf{G} by its transpose, \mathbf{G}^T (Nemeth et al., 2000). Hence, migration is the transpose of demigration:

$$\mathbf{m} = \mathbf{G}^T\mathbf{d}. \quad (3)$$

Using \mathbf{G}^T as an approximation of \mathbf{G}^{-1} typically produces adequate image quality for dense data. However, for old

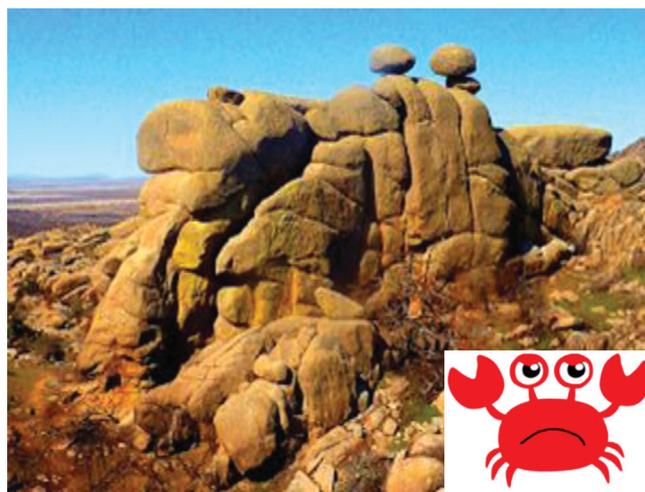


Figure 2. Crab-eye rock at Charon’s Garden, Wichita Mountains (see how it looks like a frowny crab?). The rock is composed of fractured granite with multiple sets of joints that are several tens of feet apart. This is an outcrop analog to the fractured basement 2500 ft below the ground.

data with low fold or modern data like our application with a shallow target, this approximation no longer holds, resulting in the annoying crosscutting migration artifacts seen on the final image.

Least squares

The goal of “least squares” is to find a solution \mathbf{m} such that the *sum of the squares* of all elements of the error ($\mathbf{d} - \mathbf{Gm}$) is the *least*. That is, instead of finding \mathbf{m} such that

$$\mathbf{d} - \mathbf{Gm} = \mathbf{0}, \quad (4)$$

we want to find \mathbf{m} such that the objective function,

$$J = (\mathbf{d} - \mathbf{Gm})^T (\mathbf{d} - \mathbf{Gm}) \quad (5)$$

is minimum.

Conjugate gradient

The conjugate-gradient method is first developed by [Hestenes and Stiefel \(1952\)](#). For a least-squares problem in the form

$$\mathbf{Ax} = \mathbf{b}, \quad (6)$$

the conjugate-gradient method involves finding a set of \mathbf{A} -conjugate directional vectors $\mathbf{S} = \{\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_{n-1}\}$ and incorporating them to the solution \mathbf{x} iteratively. [Appendix B](#) explains the conjugate-gradient method in detail.

The key advantage that differentiates the conjugate-gradient method from the other first-order iterative approaches to solve least-squares problems, including QR-decomposition, singular-value decomposition, and steepest descent method, is that theoretically it guarantees to converge after n iterations for an $n \times n$ matrix \mathbf{A} , whereas other methods cannot guarantee to converge after a finite number of iterations ([Lewis et al., 2006](#)). Therefore, the conjugate-gradient method generally has the fastest convergence rate — the rate at which we approach the true solution iteratively.

Constraints

Finally, we need to be aware that the data consist of signal and noise. Because of this, if we focus only on solving the exact solution for the least-squares problem, we are in fact solving for a reflectivity model that is responsible for signal and noise. We do not want our seismic image to represent the noise in the raw gathers. Therefore, we need to constrain the reflectivity model to only represent the signal portion of the data. The constraint we use in this paper is prestack structure-oriented filtering (SOF). The workflow of SOF involves calculating the structural dip from the stacked volume, computing the coherence attribute along the structural

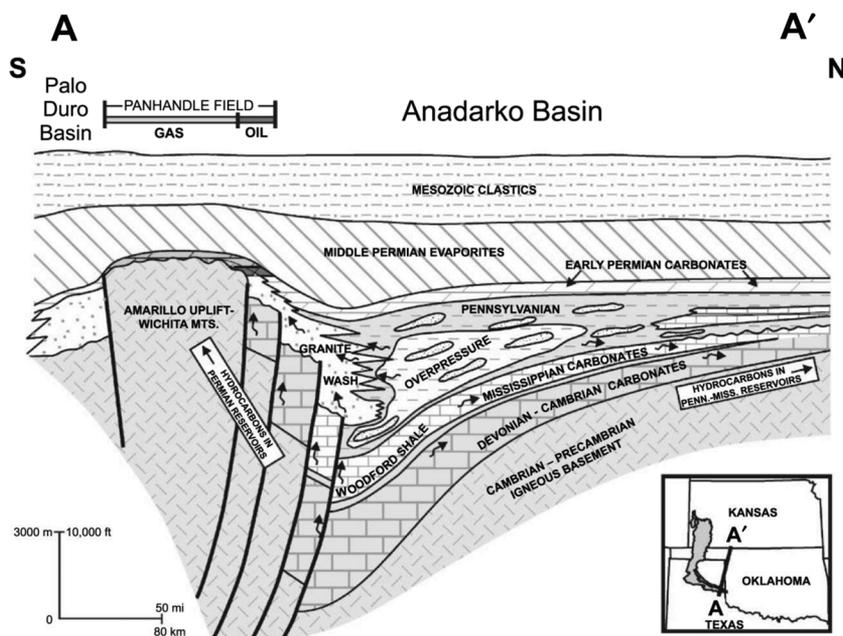


Figure 3. Regional geologic cross section through the Panhandle field ([Sorenson, 2005](#)). Source rocks are located in the deeper part of the Anadarko Basin and have an age range from Ordovician to Pennsylvanian, including the Mississippian Woodford Shale. The most common reservoir rocks are the early Permian carbonate and the Granite Wash. Oil also fills joints and fractures that formed in the previously exposed basement highs. Above the reservoir rocks, middle Permian evaporites act as a seal. Such a thick, high-velocity layer of evaporite is the cause of strong head waves in seismic data.

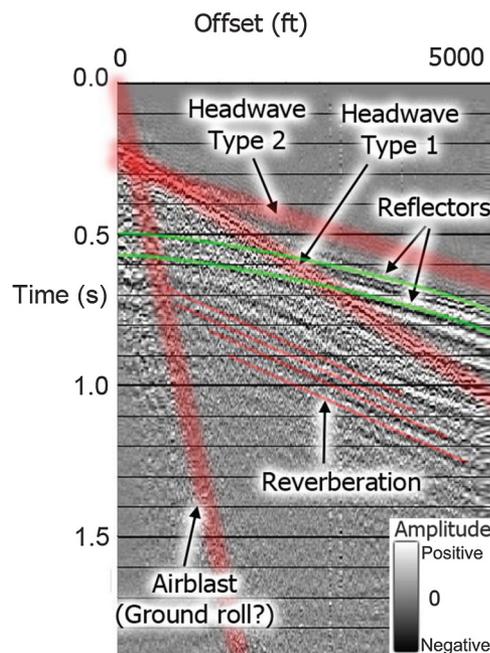


Figure 4. A representative 3D shot gather sorted by offset with interpreted events, including head waves, reflections, air blast (or ground roll?), and reverberations. At the target, top basement depth ($t = 0.57$ s), critical refraction occurs at offset $h = 3200$ ft. Beyond this point, the signals are highly contaminated by coherent, moderate-bandwidth head waves.

dip, and finally applying a lower-upper-middle filter to the input data based on coherency. The data are filtered where the coherency is high (e.g., strong reflectors) and are kept the same where the coherency is low (e.g., faults and discontinuities). Therefore, SOF suppresses random noise and the operator and data aliasing that cuts across the dominant reflectors, while preserving edges (Zhang et al., 2016). Our workflow follows that of Guo et al. (2016), whereas Appendix E provides details on our implementation of these constraints within the conjugate-gradient solution framework.

Application and results

We applied the constrained conjugate-gradient least-squares migration to a Texas Panhandle data set. The Panhandle-Hugoton field, of Texas, Oklahoma, and Kansas, is a giant oil field and the largest conventional gas field in North America, with an estimated ultimate recovery (EUR) of 1400 million barrels of oil and 75 trillion cubic feet of gas (Sorenson, 2005). Although the field has been extensively produced, previously untapped local hydrocarbon accumulations are still encountered. Recent drilling activity indicates that some wells produce directly from basement fractures, suggesting a shallow “buried-hill” reservoir type (Figure 2). Our main objec-

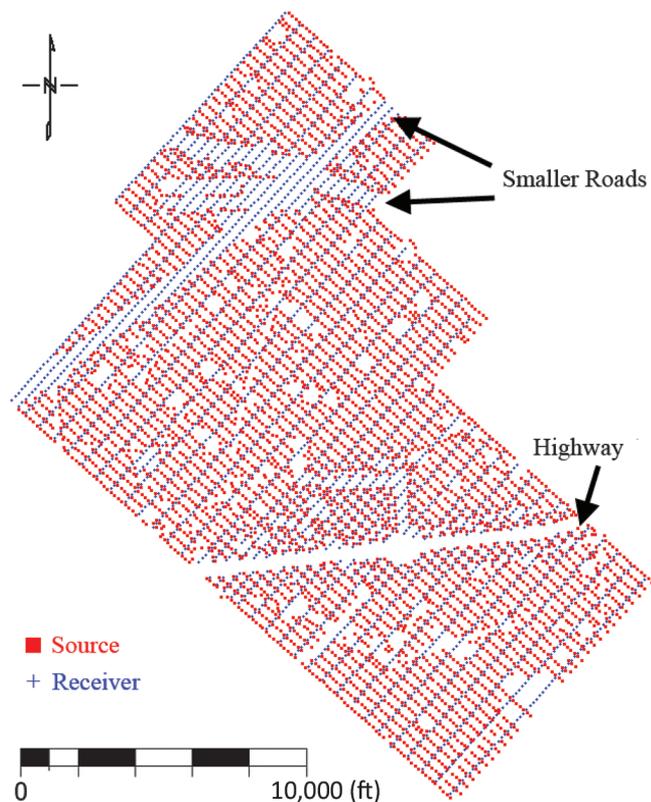


Figure 5. The source (red squares) and receiver (blue crosses) geometry of the seismic survey. Linear gaps in the source and receiver locations are associated with roads. Other smaller, circular gaps are areas inaccessible to vibroseis trucks. These gaps, together with the rectangular gridding geometry, generate acquisition footprint in seismic data, especially at a shallow target depth.

tive is to use seismic attributes and inversion results to identify open fracture zones that are potentially filled with hydrocarbon.

The top basement is very shallow (approximately 2500 ft deep, equivalent to approximately 600 ms two-way-traveltime), giving rise to some processing challenges. Overlaying the top of the basement is a thick Permian evaporite layer, causing strong head waves (Figure 3). Reflection signals are overprinted by strong coherent noise, including ground roll and reverberating refractions (Figure 4). Due to the shallow target, some gaps in the source-receiver geometry, and the nature of the orthogonal shot and receiver line acquisition program (Figure 5), the seismic data suffer from the acquisition footprint.

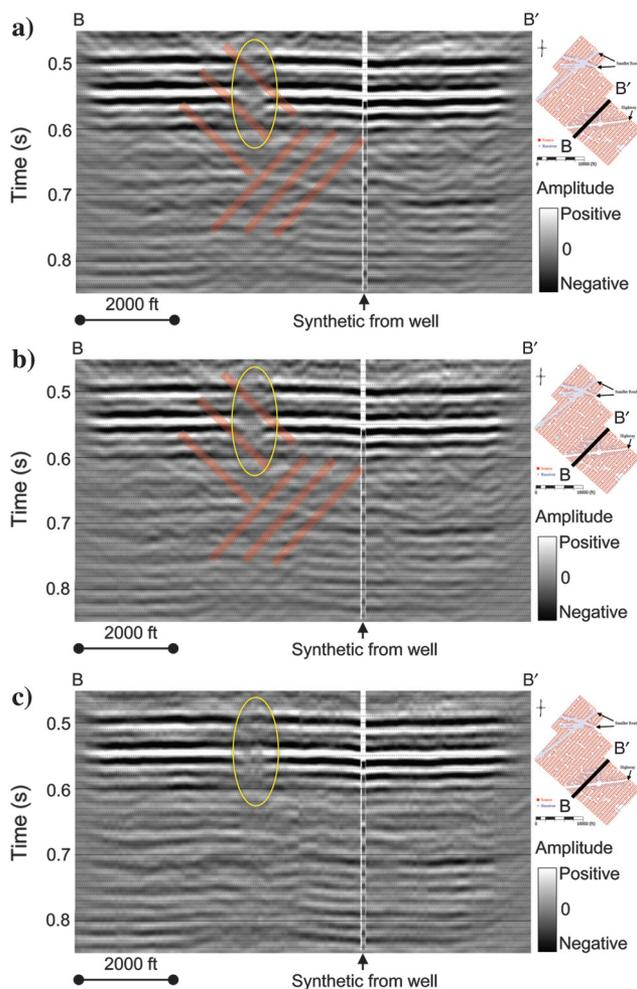


Figure 6. Vertical slice through the seismic amplitude volume generated from (a) conventional Kirchhoff prestack time migration, (b) unconstrained conjugate gradient least-squares prestack time migration, and (c) constrained conjugate gradient least-squares prestack time migration. The same migration algorithm is used in all cases. Note the crosscutting migration artifacts (red lines) in the conventional migrated image, which are enhanced in the least-squares migrated image without the constraint. Also, note a sudden amplitude decrease (yellow ellipses) where a highway intersects the profile. The constrained conjugate gradient least-squares migration suppresses those artifacts and provides better amplitude balancing.

After applying coherent noise-suppression techniques described by Verma et al. (2016), we evaluated both conventional Kirchhoff prestack time migration and constrained conjugate-gradient least-squares migration. We then computed geometric attributes and impedance inversion from both volumes in order to quantify any improvement from constrained conjugate-gradient least-squares migration.

Figure 6 shows a vertical slice through the seismic amplitude volumes generated by conventional Kirchhoff, unconstrained conjugate-gradient least-squares, and constrained conjugate-gradient least-squares prestack time migration. The conventional Kirchhoff migrated result exhibits strong steeply dipping migration artifacts due to operator aliasing, even though we used only the low-frequency components to image steep dips as described by Biondi (2001). This aliasing gives rise to acquisition footprint in subsequent attribute and inversion results. Without the constraint, least-squares migration image enhances those artifacts because the least-squares element alone preserves noise components that have leaked through migration. The constrained conjugate-gradient least-squares migration result increases the signal-to-noise ratio, enhances reflection clarity, and fills in the illumination gaps caused by highways. The SOF constraint and least-squares migration contribute to a higher signal-to-noise ratio and better amplitude balancing. The SOF constraint takes an initial reflectivity model and rejects components that are inconsistent with its neighbors and with the local dips. Least-squares migration then adjusts this model further to better fit the (sparse) surface data.

Figure 7 shows coherence time slices below the top basement generated by conventional Kirchhoff and constrained conjugate gradient least-squares prestack time migration. Most of the grid-like (hash) artifacts are suppressed on the coherence time slice computed from the constrained conjugate-gradient least-squares migration result.

The same effect can be observed in the near-offset-stack P-impedance inversion results and inversion misfit error maps (Figures 8 and 9). The acquisition footprint is greatly reduced in the constrained conjugate-gradient least-squares migration result, making low-impedance zones of interest smoother and easier to identify. These low-impedance zones

contain producing well locations in the survey area and correspond to open fractures filled with hydrocarbons. The impact on prestack azimuthal anisotropy analysis is also significant (T. Ha and K. J. Marfurt, personal communication, 2017). Unfortunately, the absence of an S-wave sonic log prevented prestack inversion.

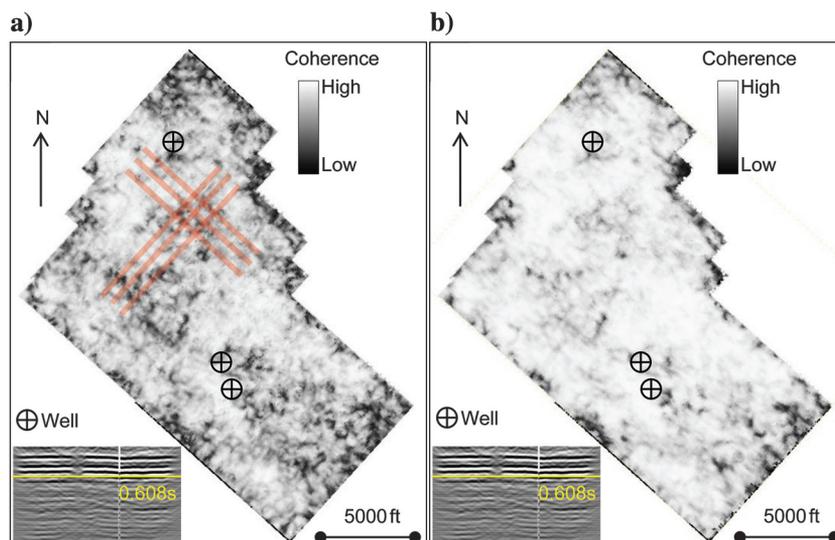


Figure 7. Coherence time slice at $t = 0.608$ s (close to top basement) generated from (a) conventional Kirchhoff prestack time migration and (b) constrained conjugate gradient least-squares prestack time migration. Most of the grid-like low-coherence hash pattern (red lines) seen in the conventional migrated coherence map is suppressed in the constrained conjugate gradient least-squares migrated coherence map.

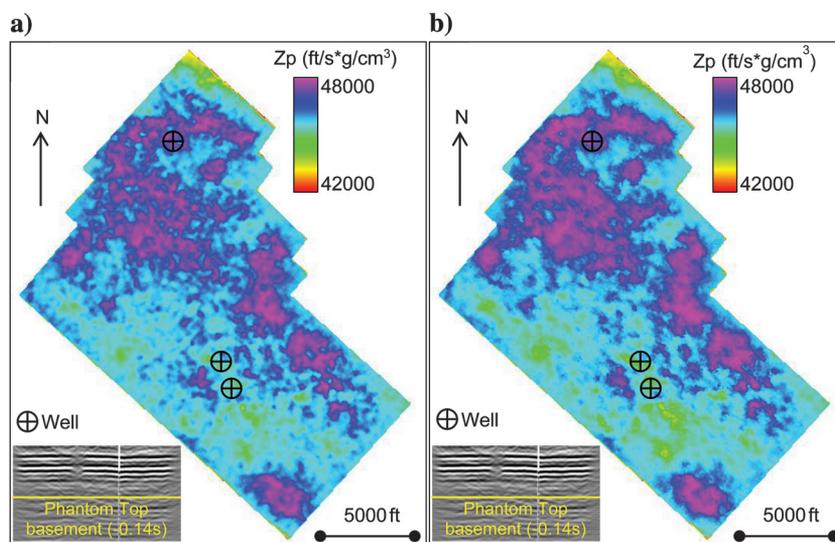


Figure 8. Phantom horizon map 0.14 s below the top basement through P-impedance volumes generated from (a) conventional Kirchhoff prestack time migration and (b) constrained conjugate gradient least-squares prestack time migration. The impedance map created by constrained conjugate gradient least-squares migration exhibits less hash-pattern noise, making it easier to isolate zones of low impedance corresponding to potential open fractures filled with hydrocarbons.

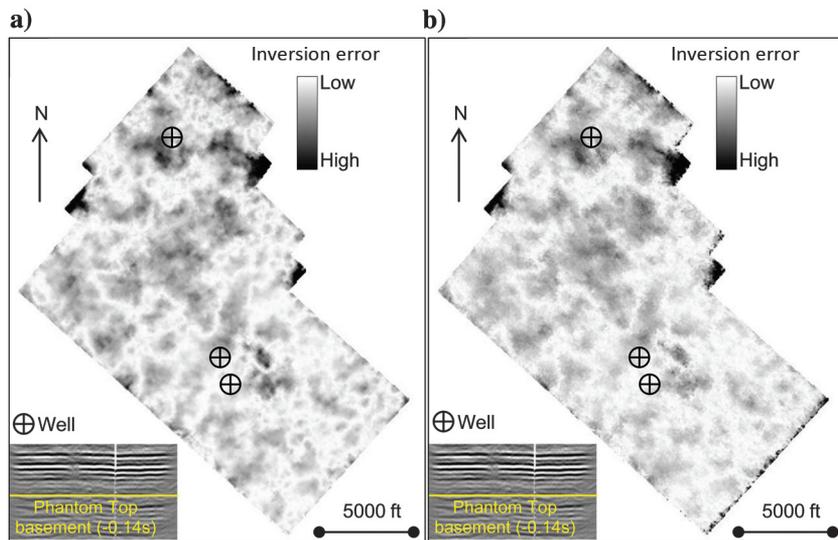


Figure 9. Phantom horizon map 0.14 s below the top basement through inversion misfit error volume generated from (a) conventional Kirchhoff prestack time migration and (b) constrained conjugate gradient least-squares prestack time migration. The constrained conjugate gradient least-squares migration errors are smoother and contain less hash-pattern artifacts than those created by conventional Kirchhoff migration.

Conclusions

Evaluation of the results of the constrained conjugate-gradient least-squares migration on a Texas Panhandle data set shows a significant improvement in seismic data quality by reducing migration artifacts, suppressing acquisition footprint, and enhancing reflection clarity. Zones of low-impedance, hydrocarbon-filled open fractures are better delineated using inversion results from the constrained conjugate-gradient least-squares migration. Three iterations of the conjugate-gradient solution required three prestack migrations and three prestack demigrations, with a computation cost six times that of conventional migration. This increase in computation cost is small compared with the velocity analysis cost associated with conventional migration followed by residual velocity analysis in a Derogowski loop. In contrast, the same velocity is used for each iteration of migration and demigration. Our caveat is that interpreters must recognize that even this sophisticated workflow needs careful data processing. Specifically, a good velocity model, surface-consistent residual statics corrections, and coherent noise suppression, are equally important. Only together with a good velocity model and higher signal-to-noise data can the constrained conjugate-gradient least-squares migration exhibit its full potential.

Acknowledgments

We thank Schlumberger for a license to their Vista software package used in preprocessing the data and CGG Veritas for Hampson-Russell used in inversion, and all of the sponsors of our AASPI consortium. Our thanks go to CIMAREX Energy Company for providing us a license to their seismic data volume for use in research and education. Financial support was through

the industry sponsors of the OU Attribute-Assisted Seismic Processing and Interpretation (AASPI) consortium.

Appendix A

Least-squares migration

In this section, we provide the mathematical background for least-squares migration.

We first start with the well-known linear equation:

$$\mathbf{d} = \mathbf{G}\mathbf{m}, \quad (\text{A-1})$$

where \mathbf{d} is the original prestack data measured on the earth's surface; \mathbf{G} is the forward Kirchhoff modeling operator (i. e., the “demigration” operator); and \mathbf{m} is the true migrated result, which is what we seek.

In general, \mathbf{G} is not a square matrix, because \mathbf{d} and \mathbf{m} may have different lengths. Therefore, we cannot simply invert \mathbf{G} and put it on the other side of the equation. The transpose of \mathbf{G} (demigration) operator \mathbf{G}^T is the Kirchhoff migration operator. We usually consider

$$\mathbf{m}' = \mathbf{G}^T\mathbf{d} \quad (\text{A-2})$$

as the migrated result, which is an accurate assumption if the surface data are regularly and densely sampled. In general, the scale of the migrated result is not equivalent to the scale of the original data, although relative amplitude changes, such as the amplitude variation with offset and amplitude variation with azimuth (AVAz) effects, are preserved. We also assume that \mathbf{G}^T (migration operator) is a good approximation to the inverse of \mathbf{G} (demigration operator). This assumption is not valid under the mathematical lens, particularly for undersampled or irregularly sampled data.

Instead, we solve for \mathbf{m} that minimizes the following objective function:

$$J = \|\mathbf{d} - \mathbf{G}\mathbf{m}\|^2. \quad (\text{A-3})$$

Expanding J , we obtain

$$\begin{aligned} J &= \|\mathbf{d} - \mathbf{G}\mathbf{m}\|^2 = (\mathbf{d} - \mathbf{G}\mathbf{m})^T(\mathbf{d} - \mathbf{G}\mathbf{m}) \\ &= \mathbf{d}^T\mathbf{d} - (\mathbf{G}\mathbf{m})^T\mathbf{d} - \mathbf{d}^T(\mathbf{G}\mathbf{m}) + (\mathbf{G}\mathbf{m})^T(\mathbf{G}\mathbf{m}) \\ &= \mathbf{d}^T\mathbf{d} - \mathbf{m}^T\mathbf{G}^T\mathbf{d} - \mathbf{d}^T\mathbf{G}\mathbf{m} + \mathbf{m}^T(\mathbf{G}^T\mathbf{G})\mathbf{m} \\ &= \mathbf{d}^T\mathbf{d} - 2\mathbf{d}^T\mathbf{G}\mathbf{m} + \mathbf{m}^T(\mathbf{G}^T\mathbf{G})\mathbf{m}, \end{aligned} \quad (\text{A-4})$$

where the scalars

$$\mathbf{m}^T\mathbf{G}^T\mathbf{d} = \mathbf{d}^T\mathbf{G}\mathbf{m}. \quad (\text{A-5})$$

To find the minimum of J , we take the gradient of $J(\nabla J)$ and set it to zero:

$$\nabla J = -2\mathbf{G}^T\mathbf{d} + 2\mathbf{G}^T\mathbf{G}\mathbf{m} = \mathbf{0}, \quad (\text{A-6})$$

or

$$\mathbf{G}^T\mathbf{G}\mathbf{m} = \mathbf{G}^T\mathbf{d}. \quad (\text{A-7})$$

An alternative way to write equation A-7 is to turn equation A-1 into what we call the “normal” equation by multiplying both sides with the transpose of \mathbf{G} (i.e., \mathbf{G}^T).

From here, we can multiply both sides by the inverse of $\mathbf{G}^T\mathbf{G}$ (i.e., $(\mathbf{G}^T\mathbf{G})^{-1}$) to obtain

$$(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{G}\mathbf{m} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{d} \quad (\text{A-8})$$

such that the left of \mathbf{m} becomes an identity matrix, giving

$$\mathbf{m} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{d}. \quad (\text{A-9})$$

To avoid instability, the least-squares solution (such as in deconvolution) often introduce a prewhitening factor $\epsilon\mathbf{I}$, which is a fraction of the diagonal of $\mathbf{G}^T\mathbf{G}$ to obtain

$$\mathbf{m} = (\mathbf{G}^T\mathbf{G} + \epsilon\mathbf{I})^{-1}\mathbf{G}^T\mathbf{d}, \quad (\text{A-10})$$

thereby favoring the solution with the minimum reflectivity energy. In our application, we will use a constraint (SOF) that favors piecewise continuous (i.e., edge-preserving) solutions over all others.

Appendix B

Conjugate-gradient least-squares migration

The linear system involved in the conjugate gradient method has the form

$$\mathbf{b} = \mathbf{A}\mathbf{x}, \quad (\text{B-1})$$

where

$$\mathbf{A} = \mathbf{G}^T\mathbf{G}, \quad (\text{B-2})$$

$$\mathbf{x} = \mathbf{m}, \quad (\text{B-3})$$

and

$$\mathbf{b} = \mathbf{G}^T\mathbf{d}. \quad (\text{B-4})$$

To understand the conjugate *gradient* method, we first examine the conjugate *direction* method. A set of vectors $\mathbf{S} = \{\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_{n-1}\}$ is said to be \mathbf{A} -conjugate if $\mathbf{p}_i^T\mathbf{A}\mathbf{p}_j = \mathbf{0}$ for $i \neq j (i, j \in [0, n-1])$. If such a set of vectors is provided, the conjugate direction method

is guaranteed to converge after n iterations. A summary of the conjugate direction method is as follows:

Step 0. choose a starting point \mathbf{x}_0 and compute $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$.

For $k = 0$ to $(n-1)$ do:

Step (1) $\alpha_k = \mathbf{p}_k^T\mathbf{r}_k / \mathbf{p}_k^T\mathbf{A}\mathbf{p}_k$

Step (2) $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{p}_k$

Step (3) $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\mathbf{A}\mathbf{p}_k$

Step (4) If the residual $\|\mathbf{r}_k + \mathbf{1}\| < \epsilon$, a convergence testing threshold, then $\mathbf{x}_k + \mathbf{1}$ is the solution, and one quits the loop. Otherwise, continue.

Normally, the set of vector \mathbf{p}_i is not known beforehand. Instead, we will generate a specific set of vectors \mathbf{p}_i while iterating through the conjugate *gradient* method:

Step 0. choose a starting point \mathbf{x}_0 and compute $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$. Let $\mathbf{p}_0 = \mathbf{r}_0$

For $k = 0$ to $(n-1)$ do:

Step (1) $\alpha_k = \mathbf{r}_k^T\mathbf{r}_k / \mathbf{p}_k^T\mathbf{A}\mathbf{p}_k$

Step (2) $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{p}_k$

Step (3) $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\mathbf{A}\mathbf{p}_k$

Step (4) If the residual $\|\mathbf{r}_k + \mathbf{1}\| < \epsilon$, a convergence testing threshold, then $\mathbf{x}_k + \mathbf{1}$ is the solution, and one quits the loop. Otherwise, continue.

Step (5) $\beta_k = \mathbf{r}_{k+1}^T\mathbf{r}_{k+1} / \mathbf{r}_k^T\mathbf{r}_k$

Step (6) $\mathbf{p}_{k+1} = \mathbf{r}_{k+1} + \beta_k\mathbf{p}_k$.

Among the various methods to solve the least-squares problem, the conjugate gradient method usually yields the fastest rate of convergence (the algorithm does not require external input and guarantees to converge after n iterations). For these reasons, it is often the preferred method in many fields of study, including migration in geophysics.

For consistency with the notation used in previous geophysical research, we now replace \mathbf{p} with \mathbf{h} , and \mathbf{r} with \mathbf{g} :

Step 0. choose a starting solution \mathbf{m}_0 and compute $\mathbf{g}_0 = \mathbf{G}^T\mathbf{d} - \mathbf{G}^T\mathbf{G}\mathbf{m}_0$. Let $\mathbf{h}_0 = \mathbf{g}_0$ be the conjugate direction.

For $k = 0$ to $(n-1)$ do:

Step (1) $\alpha_k = \mathbf{g}_k^T\mathbf{g}_k / \mathbf{h}_k^T(\mathbf{G}^T\mathbf{G})\mathbf{h}_k = \mathbf{g}_k^T\mathbf{g}_k / (\mathbf{G}\mathbf{h}_k)^T\mathbf{G}\mathbf{h}_k$

Step (2) $\mathbf{m}_{k+1} = \mathbf{m}_k + \alpha_k\mathbf{h}_k$

Step (3) $\mathbf{g}_{k+1} = \mathbf{g}_k - \alpha_k\mathbf{G}^T\mathbf{G}\mathbf{h}_k$

Step (4) If $\|\mathbf{g}_{k+1}\| < \epsilon$, a convergence testing threshold, then \mathbf{m}_{k+1} is the solution, and one quits the loop. Otherwise, continue.

Step (5) $\beta_k = \mathbf{g}_{k+1}^T\mathbf{g}_{k+1} / \mathbf{g}_k^T\mathbf{g}_k$

Step (6) $\mathbf{p}_{k+1} = \mathbf{r}_{k+1} + \beta_k\mathbf{p}_k$.

Because migration (\mathbf{G}^T is the migration operator) and demigration (\mathbf{G} is the demigration operator) are computationally intensive, we want to minimize the number of times we apply them. Therefore, we introduce \mathbf{r} as the residual in the demigrated domain (i.e., the domain of the original data \mathbf{d}), in contrast to \mathbf{g} , which is the residual in the model domain (i.e., the domain of the solution \mathbf{m}).

To do so, in step 0, set

$$\mathbf{r}_0 = \mathbf{d} - \mathbf{G}\mathbf{m}_0, \quad (\text{B-5})$$

and

$$\mathbf{g}_0 = \mathbf{G}^T \mathbf{r}_0 = \mathbf{G}^T (\mathbf{d} - \mathbf{G}\mathbf{m}_0) = \mathbf{G}^T \mathbf{d} - \mathbf{G}^T \mathbf{G}\mathbf{m}_0. \quad (\text{B-6})$$

Also, to make the result unbiased, we choose

$$\mathbf{m}_0 = \mathbf{0}. \quad (\text{B-7})$$

Thus,

$$\mathbf{r}_0 = \mathbf{d} \quad (\text{B-8})$$

and

$$\mathbf{g}_0 = \mathbf{G}^T \mathbf{d}. \quad (\text{B-9})$$

In step 1, we note how $\mathbf{A} = \mathbf{G}^T \mathbf{G}$ changes the denominator to a simple dot product. We only need to apply demigration \mathbf{G} to the conjugate gradient \mathbf{h}_k and save one migration operator \mathbf{G}^T . In step 3, we note that \mathbf{g}_{k+1} is the migrated \mathbf{r}_{k+1} , and \mathbf{g}_k is the migrated residual \mathbf{r}_k (i. e., $\mathbf{g}_{k+1} = \mathbf{G}^T \mathbf{r}_{k+1}$, and $\mathbf{g}_k = \mathbf{G}^T \mathbf{r}_k$). Therefore, we can separate step 3 into two parts:

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{G}\mathbf{h}_k, \quad (\text{B-10})$$

and

$$\mathbf{g}_{k+1} = \mathbf{G}^T \mathbf{r}_{k+1}. \quad (\text{B-11})$$

Because step 1 (calculating α_k) involves reading data from \mathbf{g}_k and $\mathbf{G}\mathbf{h}_k$, it would be more efficient to merge step 1 with $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{G}\mathbf{h}_k$. Similarly, we can merge step 5 with step 6 for the same reason.

Applying the above modifications, we can rewrite the conjugate gradient method for migration as follows:

Step 0. choose a starting solution $\mathbf{m}_0 = \mathbf{0}$. Set $\mathbf{r}_0 = \mathbf{d}$. Compute $\mathbf{g}_0 = \mathbf{G}^T \mathbf{d}$. Let $\mathbf{h}_0 = \mathbf{g}_0$ be the conjugate direction.

For $k = 0$ to $(n - 1)$, do:

- Step (1) $\alpha_k = (\mathbf{g}_k^T \mathbf{g}_k) / (\mathbf{G}\mathbf{h}_k)^T \mathbf{G}\mathbf{h}_k$, $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{G}\mathbf{h}_k$
- Step (2) $\mathbf{m}_{k+1} = \mathbf{m}_k + \alpha_k \mathbf{h}_k$
- Step (3) $\mathbf{g}_{k+1} = \mathbf{G}^T \mathbf{r}_{k+1}$
- Step (4) If $\|\mathbf{g}_{k+1} + \mathbf{1}\| < \varepsilon$, a convergence testing threshold, then \mathbf{m}_{k+1} is the solution, and one quits the loop. Otherwise, continue.
- Step (5) $\beta_k = \mathbf{g}_{(k+1)}^T \mathbf{g}_{(k+1)} / \mathbf{g}_k^T \mathbf{g}_k$, and $\mathbf{h}_{k+1} = \mathbf{g}_{k+1} + \beta_k \mathbf{h}_k$

The above suite of equations should be familiar to most people working with the conjugate gradient least-squares migration. What we have done here is nothing more than a “translation” between mathematical papers and geophysical migration, by replacing notations of variables.

Appendix C

Constrained least-squares migration

Now that we have reviewed the workflow of conjugate gradient least-squares migration, let us move on to add a constraint to the workflow. The purpose of the constraint is to increase the signal-to-noise ratio and reduce the migration aliasing artifacts. A constraint is simply a filter \mathbf{F} applied to the solution \mathbf{m} — a means to “bend” the result to our will:

$$\bar{\mathbf{m}}_{k+1} = \mathbf{F}(\mathbf{m}_{k+1}) = \mathbf{F}(\mathbf{m}_k + \alpha_k \mathbf{h}_k). \quad (\text{C-1})$$

To improve the conjugate direction, we also need to apply filter \mathbf{F} to \mathbf{h}_k : $\bar{\mathbf{h}}_k = \mathbf{F}(\mathbf{h}_k)$.

Because filter \mathbf{F} might be computationally intensive, we assume \mathbf{F} to be a relatively linear operator. Therefore, we can reduce computational cost by setting $\bar{\mathbf{h}}_k = \bar{\mathbf{m}}_{k+1} - \mathbf{m}_k / \alpha_k$.

As we introduce the constraint to the workflow, we expect the residual \mathbf{r}_k to be as close to the noise portion of the input data as possible. That is, the magnitude of the residual $\|\mathbf{r}_k\|$ should become smaller and then stabilize, but should never reach zero. Because we do not know the magnitude of the noise portion in the original data, a constant threshold ε of the residual’s magnitude is no longer valid as a stopping criterion. We need a new condition to quit the loop.

In all of our experiments, the result is considered sufficiently improved for interpretation purposes when the change of the residual is within 10% of the residual in the current iteration (typically after three iterations). Therefore, we define a new stopping criterion for our workflow as when $\|\mathbf{r}_{k+1} - \mathbf{r}_k\| / \|\mathbf{r}_{k+1}\| < 0.1$.

The constrained conjugate gradient least-squares migration workflow is as follows:

Step (0). choose a starting solution $\mathbf{m}_0 = \mathbf{0}$, set $\mathbf{r}_0 = \mathbf{d}$, compute $\mathbf{g}_0 = \mathbf{G}^T \mathbf{d}$, and let $\mathbf{h}_0 = \mathbf{g}_0$ be the conjugate direction.

For $k = 0$ to $(n - 1)$ do:

- Step (1) $\alpha_k = \mathbf{g}_k^T \mathbf{g}_k / (\mathbf{G}\mathbf{h}_k)^T \mathbf{G}\mathbf{h}_k$, and $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{G}\mathbf{h}_k$.
- Step (2) $\mathbf{m}_{k+1} = \mathbf{m}_k + \alpha_k \mathbf{h}_k$
- Step (3) $\bar{\mathbf{m}}_{k+1} = \mathbf{F}(\mathbf{m}_{k+1})$
- Step (4) If $\|\mathbf{r}_{k+1} - \mathbf{r}_k\| / \|\mathbf{r}_{k+1}\| < 0.1$, then $\bar{\mathbf{m}}_{k+1}$ is the solution, and one quits the loop. Otherwise, continue.
- Step (5) $\bar{\mathbf{h}}_k = \bar{\mathbf{m}}_{k+1} - \mathbf{m}_k / \alpha_k$
- Step (6) $\bar{\mathbf{h}}_k = \bar{\mathbf{m}}_{k+1} - \mathbf{m}_k / \alpha_k$
- Step (7) $\beta_k = \mathbf{g}_{(k+1)}^T \mathbf{g}_{(k+1)} / \mathbf{g}_k^T \mathbf{g}_k$, and $\mathbf{h}_{k+1} = \mathbf{g}_{k+1} + \beta_k \bar{\mathbf{h}}_k$.

Appendix D

Modification for weighted least-squares migration

Seismic data often contain undesirable linear noise, such as head waves. Ideally, such noise should be muted before migrating the data. Similarly, the demigration process may include some data-truncation artifacts due to the nature of inverse Fourier transform used to apply the $i\omega$ operator, causing potential wrap-around or

other artifacts after migration (Figure D-1). Such artifacts in demigrated data must also be muted.

Muting can be understood as a weighting operator, in which the points representing head waves and wrap-around artifacts are set with weights = 0, whereas the points representing useful data are set with weights = 1. Tapering gives weights between 0 and 1. We need a solid mathematical ground for the weighted least-squares problem.

The objective function that we want to minimize now becomes

$$J = \|\mathbf{W}(\mathbf{d} - \mathbf{G}\mathbf{m})\|^2. \quad (\text{D-1})$$

Expanding J , we have

$$\begin{aligned} J &= (\mathbf{W}(\mathbf{d} - \mathbf{G}\mathbf{m}))^T \mathbf{W}(\mathbf{d} - \mathbf{G}\mathbf{m}) \\ &= (\mathbf{d} - \mathbf{G}\mathbf{m})^T \mathbf{W}^T \mathbf{W}(\mathbf{d} - \mathbf{G}\mathbf{m}) \\ &= (\mathbf{d}^T - \mathbf{m}^T \mathbf{G}^T) \mathbf{W}^T \mathbf{W}(\mathbf{d} - \mathbf{G}\mathbf{m}) \\ &= \mathbf{d}^T \mathbf{W}^T \mathbf{W} \mathbf{d} - \mathbf{m}^T \mathbf{G}^T \mathbf{W}^T \mathbf{W} \mathbf{d} \\ &\quad - \mathbf{d}^T \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{m} + \mathbf{m}^T \mathbf{G}^T \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{m}. \end{aligned} \quad (\text{D-2})$$

The two middle terms are basically the transpose of each other and thus are equal to each other, giving

$$\begin{aligned} J &= \mathbf{d}^T \mathbf{W}^T \mathbf{W} \mathbf{d} - 2\mathbf{d}^T \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{m} \\ &\quad + \mathbf{m}^T \mathbf{G}^T \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{m}. \end{aligned} \quad (\text{D-3})$$

To minimize J , we need to find where $\nabla J = \mathbf{0}$ with respect to \mathbf{m} :

$$\begin{aligned} \nabla J &= -2(\mathbf{d}^T \mathbf{W}^T \mathbf{W} \mathbf{G})^T \\ &\quad + 2\mathbf{G}^T \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{m} = \mathbf{0}, \end{aligned} \quad (\text{D-4})$$

or

$$\mathbf{G}^T \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{m} = \mathbf{G}^T \mathbf{W}^T \mathbf{W} \mathbf{d}. \quad (\text{D-5})$$

Updating the workflow, we have

Step (0). choose a starting solution $\mathbf{m}_0 = \mathbf{0}$, set $\mathbf{r}_0 = \mathbf{d}$, compute $\mathbf{g}_0 = \mathbf{G}^T \mathbf{W}^T \mathbf{W} \mathbf{d}$, and let $\mathbf{h}_0 = \mathbf{g}_0$ be the conjugate direction.

For $k = 1$ to n do:

Step (1) $\alpha_k = \mathbf{g}_{k-1}^T \mathbf{g}_{k-1} / (\mathbf{G} \mathbf{h}_{k-1})^T \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_{k-1}$, and $\mathbf{r}_{k+1} = \mathbf{r}_{k-1} - \alpha_k \mathbf{G} \mathbf{h}_{k-1}$

Step (2) $\mathbf{m}_k = \mathbf{m}_{k-1} + \alpha_k \mathbf{h}_{k-1}$

Step (3) $\mathbf{h}_k = \mathbf{F}(\mathbf{m}_k)$

Step (4) If, $\|\mathbf{r}_k - \mathbf{r}_{k-1}\| / \|\mathbf{r}_{k+1}\| < 0.1$, then \mathbf{m}_k is the solution, and one quits the loop. Otherwise, continue.

Step (5) $\mathbf{g}_k = \mathbf{G}^T \mathbf{W}^T \mathbf{W} \mathbf{r}_k$

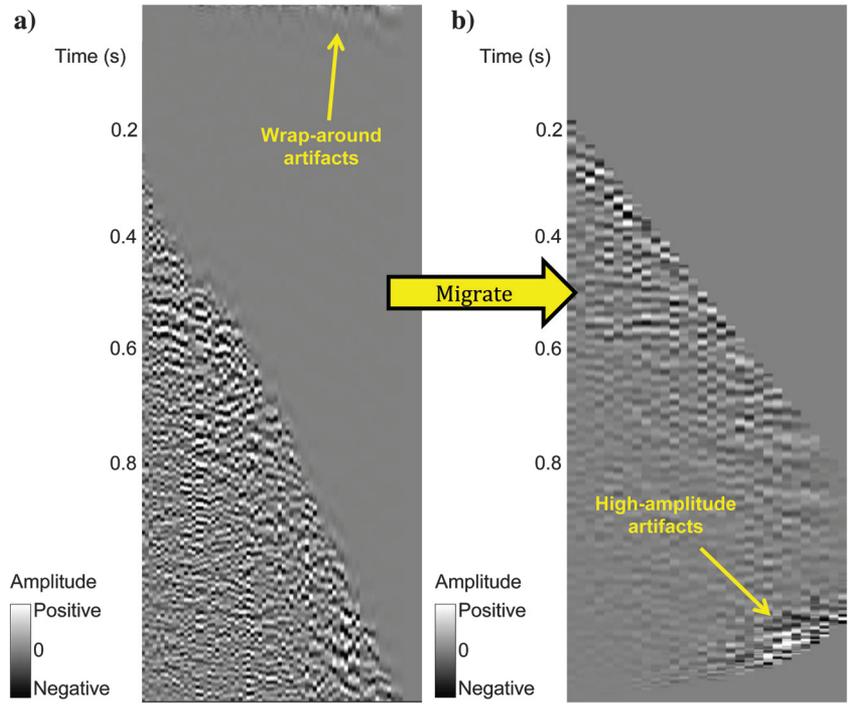


Figure D-1. (a) A demigrated CDP gather showing wrap-around artifacts close to the time zero. These artifacts are inherent to the forward modeling (i.e., demigration) operator, due to the nature of the inverse Fourier transform. (b) Migrated result of demigrated data in panel (a). The high-amplitude artifacts at the bottom of the migrated CDP gather are caused by such wrap-around artifacts. Thus, it is important to mute (or alternatively, sufficiently pad) the top part of the demigrated gathers.

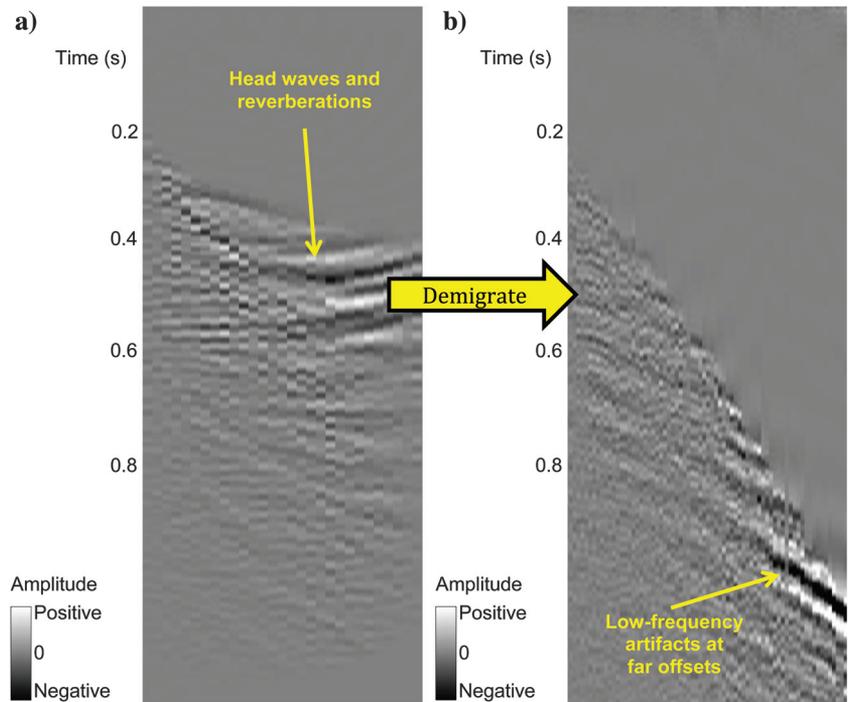


Figure D-2. (a) A migrated CDP gather showing head waves and reverberations. (b) Demigrated result of the migrated data in panel (a). The low-frequency artifacts at far offset in demigrated gather are caused by inadequate suppression of head waves and reverberations. Thus, if the muting and noise suppression applied on the original data is insufficient to remove the head waves and reverberations, we need to mute the migrated result as well.

Step (6) $\bar{\mathbf{h}}_k = \bar{\mathbf{m}}_k - \mathbf{m}_{k-1}/\alpha_k$

Step (7) $\beta_k = \mathbf{g}_k^T \mathbf{g}_k / \mathbf{g}_{k-1}^T \mathbf{g}_{k-1}$ and $\mathbf{h}_k = \mathbf{g}_k + \beta_k \bar{\mathbf{h}}_k$.

In practice, applying such weights is equivalent to muting the data in the unmigrated domain (i.e., the domain of the original data and the demigrated data). Sometimes, muting is also applied in the migrated data domain, in case the demigrated data exhibit enhanced low-frequency artifacts at far-offset traces (Figure D-2).

Appendix E

Step-by-step workflow

Based on the mathematical description of the constrained conjugate gradient least-squares migration, we can see that it is quite difficult to represent the workflow by a simple flowchart. Even the first iteration requires many small steps and generates many different outputs, each to be used in either the same iteration or the next one. Therefore, we choose to represent the workflow with a step-by-step guide for each iteration.

A step-by-step workflow for a maximum of n iterations of constrained conjugate gradient least-squares migration is as follows:

Iteration #0 (i.e., preparation iteration):

- 1) Mute the original data to avoid head wave contamination. This is basically multiplying $\mathbf{W}^T \mathbf{W}$ with data \mathbf{d} : $\mathbf{W}^T \mathbf{W} \mathbf{d}$.
- 2) Compute the fold and offset map. This step is a part of Kirchhoff migration procedure, but it only needs to be done once.
- 3) Migrate the data with antialiasing enabled, using the muted original data and the calculated fold and offset map. An antialiasing feature in migration reduces migration artifacts. Applying the migration operator \mathbf{G}^T to $\mathbf{W}^T \mathbf{W} \mathbf{d}$: $\mathbf{g}_0 = \mathbf{G}^T \mathbf{W}^T \mathbf{W} \mathbf{d}$.
- 4) Apply SOF to \mathbf{g}_0 : $\mathbf{h}_0 = \mathbf{F}(\mathbf{g}_0)$ and perform muting if needed. The reason we apply SOF in iteration #0 is to further constrain the result to improve the signal-to-noise ratio. Plus, running SOF in this iteration means we do not need to run SOF for iteration #1, thereby saving us one step.
- 5) Demigrate the SOF-migrated result. This is equivalent to applying the forward modeling operator \mathbf{G} to \mathbf{h}_0 : $\mathbf{G} \mathbf{h}_0$.
- 6) Mute the demigrated result. Again, multiplying weight $\mathbf{W}^T \mathbf{W}$ with $\mathbf{G} \mathbf{h}_0$: $\mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_0$.

Iteration #1:

- 1) Update the residual: calculate α_1 and \mathbf{r}_1 :

$$\alpha_1 = \mathbf{g}_0^T \mathbf{g}_0 / (\mathbf{G} \mathbf{h}_0)^T \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_0 \quad \text{and} \\ \mathbf{W}^T \mathbf{W} \mathbf{r}_1 = \mathbf{W}^T \mathbf{W} \mathbf{d} - \alpha_1 \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_0. \quad (\text{E-1})$$

The denominator of α_1 is basically the square of muted demigrated result $\mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_0$. Because \mathbf{W} consists of 0 and 1 only (i.e., muted: $\mathbf{w} = \mathbf{0}$, nonmuted: $\mathbf{w} = \mathbf{1}$), $\mathbf{W}^T \mathbf{W} = \mathbf{W}$, and thus $(\mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_0)^T (\mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_0) = (\mathbf{W} \mathbf{G} \mathbf{h}_0)^T (\mathbf{W} \mathbf{G} \mathbf{h}_0) = (\mathbf{G} \mathbf{h}_0)^T \mathbf{W}^T \mathbf{W} (\mathbf{G} \mathbf{h}_0)$.

Note that in the residual calculation, we use the muted result of original data and demigrated data, and thus the updated residual is muted and we can skip the muting step later on.

- 2) Update the model: $\mathbf{m}_1 = \mathbf{m}_0 + \alpha_1 \mathbf{h}_0 = \mathbf{0} + \alpha_1 \mathbf{h}_0 = \alpha_1 \mathbf{h}_0$ (because we assume $\mathbf{m}_0 = \mathbf{0}$ as the starting solution). Because \mathbf{h}_0 is the SOF-applied result, we can skip the constrain step.
- 3) Migrate the muted updated residual with antialias disabled (because we do not want the residual to be too smoothed): $\mathbf{g}_1 = \mathbf{G}^T \mathbf{W}^T \mathbf{W} \mathbf{r}_1$.
- 4) Mute the migrated residual, if needed (optional).
- 5) Update the conjugate gradient: calculate β_1 and \mathbf{h}_1 .
- 6) Demigrate the updated conjugate gradient: $\mathbf{G} \mathbf{h}_1$.
- 7) Mute the demigrated result: $\mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_1$.

Iteration #k ($2 \leq k < n$):

- 1) update the residual: calculate α_k and \mathbf{r}_k :

$$\alpha_k = \frac{\mathbf{g}_{(k-1)}^T \mathbf{g}_{(k-1)}}{\mathbf{G} \mathbf{h}_{(k-1)}^T \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_{(k-1)}} \quad \text{and} \quad \mathbf{W}^T \mathbf{W} \mathbf{r}_k \\ = \mathbf{W}^T \mathbf{W} \mathbf{r}_{k-1} - \alpha_k \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_{k-1}. \quad (\text{E-2})$$

- 2) Update the model: $\mathbf{m}_k = \mathbf{m}_{k-1} + \alpha_k \mathbf{h}_{k-1}$.
- 3) Apply SOF on \mathbf{m}_k : $\bar{\mathbf{m}}_k = \mathbf{F}(\mathbf{m}_k)$.
- 4) Update the directional vector: $\bar{\mathbf{h}}_k = \bar{\mathbf{m}}_k - \mathbf{m}_{k-1}/\alpha_k$.
- 5) Migrate the muted updated residual with anti-alias disabled: $\mathbf{g}_k = \mathbf{G}^T \mathbf{W}^T \mathbf{W} \mathbf{r}_k$.
- 6) Mute the migrated residual, if needed (optional).
- 7) Update the conjugate gradient: calculate β_k and \mathbf{h}_k :

$$\beta_k = \mathbf{g}_k^T \mathbf{g}_k / \mathbf{g}_{(k-1)}^T \mathbf{g}_{(k-1)}, \quad \text{and} \quad \mathbf{h}_k = \mathbf{g}_k + \beta_k \bar{\mathbf{h}}_k. \quad (\text{E-3})$$

- 8) Demigrate the updated conjugate gradient: $\mathbf{G} \mathbf{h}_k$.
- 9) Mute the demigrated result: $\mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_k$.

Iteration #n (last iteration):

- 1) Update the residual: calculate α_n and \mathbf{r}_n :

$$\alpha_n = \mathbf{g}_{(n-1)}^T \mathbf{g}_{(n-1)} / (\mathbf{G} \mathbf{h}_{(n-1)})^T \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_{(n-1)} \\ \text{and} \quad \mathbf{W}^T \mathbf{W} \mathbf{r}_n = \mathbf{W}^T \mathbf{W} \mathbf{r}_{n-1} - \alpha_n \mathbf{W}^T \mathbf{W} \mathbf{G} \mathbf{h}_{n-1} \quad (\text{E-4})$$

- 2) Update the model: $\mathbf{m}_n = \mathbf{m}_{n-1} + \alpha_n \mathbf{h}_{n-1}$.
- 3) Apply the SOF on \mathbf{m}_n : $\bar{\mathbf{m}}_n = \mathbf{F}(\mathbf{m}_n)$.

References

- Biondi, B., 2001, Kirchhoff imaging beyond aliasing: Geophysics, **66**, 654–666, doi: [10.1190/1.1444956](https://doi.org/10.1190/1.1444956).
- Dev, A., and G. A. McMechan, 2009, Spatial antialias filtering in the slowness-frequency domain: Geophysics, **74**, no. 2, V35–V42, doi: [10.1190/1.3052115](https://doi.org/10.1190/1.3052115).
- Guo, S., B. Zhang, Q. Wang, A. Cabrales-Vargas, and K. J. Marfurt, 2016, Noise suppression using preconditioned least-squares prestack time migration: Application to the

- Mississippian Limestone: *Journal of Geophysics and Engineering*, **13**, 441–453, doi: [10.1088/1742-2132/13/4/441](https://doi.org/10.1088/1742-2132/13/4/441).
- Hestenes, M., and E. Stiefel, 1952, Methods of conjugate gradient for solving linear systems: *Journal of Research of the National Bureau of Standards*, **49** 409–436, doi: [10.6028/jres.049.044](https://doi.org/10.6028/jres.049.044).
- Lewis, J. M., S. Lakshmivarahan, and S. K. Dhall, 2006, *Dynamic data assimilation: A least squares approach*: Cambridge University Press.
- Nemeth, T., H. Sun, and G. T. Schuster, 2000, Separation of signal and coherent noise by migration filtering: *Geophysics*, **65** 574–583, doi: [10.1190/1.1444753](https://doi.org/10.1190/1.1444753).
- Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: *Geophysics*, **64** 208–221, doi: [10.1190/1.1444517](https://doi.org/10.1190/1.1444517).
- Pramik, B., 2011, Broadband land acquisition — Survey design issues: 81st Annual International Meeting, SEG, Expanded Abstracts, 4344–4348.
- Sorenson, R., 2005, A dynamic model for the Permian Panhandle and Hugoton fields, western Anadarko basin: *AAPG Bulletin*, **89**, 921–938, doi: [10.1306/03010504045](https://doi.org/10.1306/03010504045).
- Verma, S., S. Guo, T. Ha, and K. J. Marfurt, 2016, Highly aliased ground-roll suppression using a 3D multiwindow Karhunen-Loeve filter: Application to a legacy Mississippi Lime survey: *Geophysics*, **81**, no. 1, V79–V88, doi: [10.1190/geo2014-0442.1](https://doi.org/10.1190/geo2014-0442.1).
- Yu, J., J. Hu, G. T. Schuster, and R. Estill, 2006, Prestack migration deconvolution: *Geophysics*, **71**, no. 2, S53–S62, doi: [10.1190/1.2187783](https://doi.org/10.1190/1.2187783).
- Zeng, C., S. Dong, and B. Wang, 2014, Least-squares reverse time migration: Inversion-based imaging toward true reflectivity: *The Leading Edge*, **33**, 962–968, doi: [10.1190/tle33090962.1](https://doi.org/10.1190/tle33090962.1).
- Zeng, C., S. Dong, and B. Wang, 2017, A guide to least-squares RTM for subsalt imaging: Challenges and solutions: *Interpretation*, **5**, this issue, doi: [10.1190/int-2016-0196.1](https://doi.org/10.1190/int-2016-0196.1).
- Zhang, B., T. Lin, S. Guo, O. E. Davogustto, and K. J. Marfurt, 2016, Noise suppression of time-migrated gathers

using prestack structure-oriented filtering: *Interpretation*, **4**, no. 2, SG19–SG29, doi: [10.1190/INT-2015-0146.1](https://doi.org/10.1190/INT-2015-0146.1).



Thang Ha received bachelor's and master's degrees in geophysics from the University of Oklahoma, where he is a Ph.D. student in geophysics advised by Kurt Marfurt. He is a member of SEG, AAPG, EAGE, and the Geophysical Society of Oklahoma City. His research interests include seismic processing, 3D seismic attributes, prestack data conditioning, least-squares migration, and automatic classification of geological facies.



Kurt Marfurt began his geophysical career teaching geophysics and contributing to an industry-supported consortium on migration, inversion, and scattering (project MIDAS) at Columbia University's Henry Krumb School of Mines in New York City. In 1981, He joined Amoco's Tulsa Research Center and spent the next 18 years leading research efforts in modeling, migration, signal analysis, basin analysis, seismic attribute analysis, reflection tomography, seismic inversion, and multicomponent data analysis. In 1999, he joined the University of Houston as a professor in the department of geosciences and as a director of the Allied Geophysics Laboratories. He is currently a member of the Geophysical Societies of Tulsa and Houston, SEG, EAGE, AAPG, AGU, and SIAM, and he serves as an assistant editor for *GEOPHYSICS*. His current research activity includes prestack imaging, velocity analysis and inversion of converted waves, computer-assisted pattern recognition of geologic features on 3D seismic data, and interpreter-driven seismic processing. His research interests are in seismic signal analysis, 3D seismic attributes, seismic velocity analysis, subsurface imaging, and multicomponent data analysis.