GEOPHYSICS, VOL. 80, NO. 5 (SEPTEMBER-OCTOBER 2015); P. A95–A101, 5 FIGS.
10.1190/GEO2015-0070.1

Block row recursive least-squares migration

Nasser Kazemi[1] and Mauricio D. Sacchi[1]

ABSTRACT

Recursive estimates of large systems of equations in the context of least-squares fitting is a common practice in different fields of study. For example, recursive adaptive filtering is extensively used in signal processing and control applications. The necessity of solving least-squares problems recursively stems from the need for fast real-time signal processing strategies. The computational cost of least-squares solvers can also limit the applicability of this technique in geophysical problems. We have considered a recursive least-squares solution for least-squares wave equation migration with sliding windows involving several fixed rank downdating and updating computations. This technique can be applied for dynamic and stationary processes. If we use enough data in each windowed setup, the spectrum of the preconditioned system is clustered around one and the method will converge superlinearly with probability one. Numerical experiments were performed to test the effectiveness of the technique for least-squares migration.

INTRODUCTION

Seismic migration aims to produce true structural and stratigraphical images of the subsurface. There are different types of migration in the geophysical literature. Migration methods can be divided into two main categories. The first category is a ray-tracing-based approach (e.g., Kirchhoff migration; Schneider, 1978). The ray-tracing method is computationally efficient, and it is easily adaptable to irregular acquisition geometry. Another category is wave equation methods. Although wave equation methods are computationally expensive, they provide accurate wavefield extrapolation and high-quality images for complex areas. In this approach, we solve for one- or two-way wave equations (Gazdag, 1978; Stolt, 1978; Baysal et al., 1983; McMechan, 1983; Whitmore, 1983; Gazdag and Sguazzero, 1984; Stoffa et al., 1990). The action of these migration operators will be calculated on the fly, and in some cases, it is even impossible to have a very straight forward explicit matrix format for these operators. Moreover, it is worth mentioning that each one of the migration techniques has its own artifacts. These artifacts arise from approximations that were made at the time of designing migration operators and also from incompleteness of the data.

These artifacts can be removed by adopting data regularization techniques prior to imaging (Fomel and Guitton, 2006). Artifacts can also be attenuated by implementing least-squares migration (Chavent and Plessix, 1999; Nemeth et al., 1999; Duijndam et al., 2000; Kuhl and Sacchi, 2003; Plessix and Mulder, 2004; Symes, 2008; Kaplan et al., 2010a, 2010b). In least-squares migration, we try to fit the data by inverting the demigration operator, and in general, we adopt constraints to minimize the artifacts produced by data incompleteness. However, the computational cost of the least-squares approach is high, and with present computer resources, it is really difficult to implement this method in its full potential for industrial applications. In other words, the direct inverse computation of the Hessian is expensive and we need to approximate its inverse (Hu et al., 2001; Etgen, 2002; Guitton, 2004; Yu et al., 2006; Lecomte, 2008; Toxopeus et al., 2008; Naoshi and Schuster, 2009; Kazemi and Sacchi, 2014). Another way of reducing the computational cost of least-squares migration is by adopting encoding methods (Wei et al., 2010; Dai et al., 2011). The other way to look at the problem is to apply least-squares migration in a recursive fashion. In this approach, we can implement the technique on the memory-limited machine and in a fast way.

Recursive least-squares algorithms are extensively used for a dynamic system of equations (e.g., radar data) and also for adaptive filtering in the case of dynamic and stationary environments. The main idea is to implement infinite memory algorithms by solving the problem via introducing one data point at a time to the system of equations (e.g., adding one row to the data matrix or in our case the demigration operator). By *infinite memory algorithm*, we mean a recursive method that operates at a given time on a small segment of data but without forgetting the influence of previous segments of data in the current solution. However, very little effort has been

Manuscript received by the Editor 30 January 2015; revised manuscript received 27 April 2015; published online 6 July 2015.
[1]University of Alberta, Department of Physics, Alberta, Canada. E-mail: kazemino@ualberta.ca; msacchi@ualberta.ca.

made in the geophysical community to adapt this technique for different practical applications. This is mainly because recursive least-squares algorithms require explicit matrices. In large-scale problems (e.g., migration), everything will be done on the fly, and adding or removing one row from the system of equations (e.g., demigration operator) has no physical meaning. Moreover, the one-data-point update scheme in the recursive least-squares approach causes stability issues. In other words, it is difficult to track the stability of the process. To tackle these shortcomings, one can update the system of equations in a block fashion by introducing more than one data point at each step, making the algorithm faster and more stable. These kinds of algorithms belong to the family of block row or block column recursive least squares. However, in this family, we need to explore some structures in the data matrix to make the algorithms faster and also update the solution in a way that in the end we solve for the original least-squares problem with the accuracy in the bounded error limit. However, in most of geophysical problems, we do not have explicit matrices for data matrix and algorithms will be implemented on the fly. Of course, in some cases such as autoregressive (AR) or moving average (MA) or AR with MA applications, one can explore special structures in the data matrix and use directly the common recursive algorithms. For example, Naghizadeh and Sacchi (2009) use rank one update of the recursive least-squares fitting with some exponentially weighted forgetting factor for f-x adaptive filtering in the context of seismic interpolation. It is obvious that this technique cannot be applied to least-squares migration.

In this paper, we propose a block row recursive method to solve the least-squares migration. Our method, in essence, is closest to the work of Ng and Plemmons (1996). We consider a recursive least-squares solution of the wave equation migration with sliding windows involving several rank K downdating and updating computations. The least-squares estimator can be found by solving the partial least-squares settings in each step, recursively. We apply preconditioned conjugate gradient (CG) method with proper preconditioners that cluster the eigenvalues of the partial Hessian operators. From practical point of view, to name a few, the blocks can be a group of shot gathers, group of frequency slices, or group of offset classes.

The outline of the paper is as follows. First, least-squares migration will be explained. Then, we will introduce a block row recursive algorithm. Moreover, we will discuss some of the practical considerations for the proposed method and compare the computational cost of the method with full least-squares migration and examine the efficiency of the proposed method on a simple toy example and the Marmousi model. Finally, the main conclusions will be summarized.

LEAST-SQUARES MIGRATION

Data generating model under the action of demigration operator $\mathbf{A}$ can be written as

$$\mathbf{d} = \mathbf{A}\mathbf{m} + \mathbf{n}, \tag{1}$$

where $\mathbf{d}$ is $N \times 1$ vectorized version of the recorded data at the surface, $\mathbf{m}$ is the migrated model with $M \times 1$, and $\mathbf{n}$ is the noise content and, often, also a term to absorb waves not modeled by the demigration operator. Using the adjoint operator of demigration, one can estimate the migrated model as

$$\hat{\mathbf{m}} = \mathbf{A}^T\mathbf{d}, \tag{2}$$

where $\hat{\mathbf{m}}$ is the adjoint estimated image of the earth, and $\mathbf{A}^T$ is the migration operator. Although $\hat{\mathbf{m}}$ can capture the main structures of the true geologic model $\mathbf{m}$, the produced model does not honor the data. In other words, application of the demigration operator on the migrated image yields poor data prediction. In addition, the migrated image contains blurring and sampling artifacts. These artifacts come from the fact that migration and demigration operators are not orthogonal, and the energy of the signal in the complimentary image space of the operator will be zeroed out. To tackle the problem, we will solve

$$\tilde{\mathbf{m}} = \underset{\mathbf{m}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{m} - \mathbf{d}\|_2^2 + \lambda\|\mathbf{m}\|_2^2, \tag{3}$$

where $\mathbf{m}$ is a desired model, $\mathbf{A}$ is a demigration operator, $\mathbf{d}$ is a recorded data, and λ is a regularization parameter (Nemeth et al., 1999; Kuhl and Sacchi, 2003; Kaplan et al., 2010b). The cost function of equation 3 is convex and has a closed-form solution

$$\tilde{\mathbf{m}} = (\mathbf{A}^T\mathbf{A} + \lambda\mathbf{I})^{-1}\mathbf{A}^T\mathbf{d}. \tag{4}$$

However, the computational cost of the least-squares approach is high, and to make the algorithm faster, one can approximate the Hessian (i.e., $\mathbf{A}^T\mathbf{A}$) inverse (Hu et al., 2001; Kazemi and Sacchi, 2014). In the next section, we will propose a sliding window scheme for solving equation 3 in recursive fashion. The main motivations behind the method are reducing the computational cost and, at the same time, producing migrated images that honor the recorded data using memory-limited resources.

BLOCK ROW RECURSIVE LEAST-SQUARES MIGRATION

In this section, we will follow the recursive least-squares solution via the rank K updating and rank K downdating procedure introduced by Ng and Plemmons (1996). However, there are some differences between the proposed method in Ng and Plemmons (1996) with our technique. We are not considering a near-Toeplitz structure for data matrix, and in the updating procedure, we just use the previous solution of the block row setup as an initial solution for the next sliding window. This way, we will give up fast implementation of the technique in favor of not considering special structures in the demigration operator. However, one can show that in the case of a laterally invariant medium, the data matrix can be considered as blockwise Toeplitz.

To explain the block row recursive least-squares method, let us consider again the problem of equation 3. In recursive least-squares computations, it is required to calculate $\mathbf{m}$ while observations are successively added to or deleted from the system of equations. Suppose we have estimated the model with the first set of measurement points $\mathbf{d}_0$ in a least-squares sense,

$$\mathbf{m}_0 = (\mathbf{A}_0^T\mathbf{A}_0)^{-1}\mathbf{A}_0^T\mathbf{d}_0. \tag{5}$$

Now, the question we should ask is that by introducing new data points to the system of equations, can the best estimate for the combined system $\mathbf{A}_0\mathbf{m} = \mathbf{d}_0$ and $\mathbf{A}_1\mathbf{m} = \mathbf{d}_1$ be estimated using only $\mathbf{m}_0$ and $\mathbf{d}_1$?

To do that, we define the new matrix as

$$\mathbf{P}_1^{-1} = \mathbf{A}_0^T\mathbf{A}_0 + \mathbf{A}_1^T\mathbf{A}_1, \tag{6}$$

then we have

$$\mathbf{m}_1 = \mathbf{P}_1(\mathbf{A}_0^T\mathbf{d}_0 + \mathbf{A}_1^T\mathbf{d}_1), \tag{7}$$

and note that $\mathbf{m}_1$ is the best model for combined system of equations. At this point, we need to eliminate $\mathbf{d}_0$ term from equation 7. Let us rewrite equation 6 as

$$\mathbf{P}_1^{-1} = \mathbf{P}_0^{-1} + \mathbf{A}_1^T\mathbf{A}_1, \tag{8}$$

and after a few algebraic manipulations, one can show that equation 7 changes to

$$\mathbf{m}_1 = \mathbf{m}_0 + \mathbf{P}_1\mathbf{A}_1^T(\mathbf{d}_1 - \mathbf{A}_1\mathbf{m}_0), \tag{9}$$

where $\mathbf{P}_1\mathbf{A}_1^T$ is the gaining factor. In the case of one data point update, we can use the matrix inversion lemma (MIL) and calculate the gaining factor without direct inversion, and in the case of a blockwise update (more than one data point update), to apply fast calculations of the gaining matrix, we need to explore MIL with QR decomposition. However, none of these approaches are applicable to least-squares migration. It is mainly because in migration, everything will be done on the fly and there are no simple explicit matrices for migration.

However, it is easy to show that if the system of equations satisfies some assumptions, we can relax the gaining factor term of equation 9 and solve the problem recursively on the overlapping windows using the CG method. Let us explain the method step by step. The least-squares estimator at step i can be found by solving for the $M \times 1$ vector $\mathbf{m}(b_i)$ in

$$\begin{aligned}\tilde{\mathbf{m}}(b_i) = \underset{\mathbf{m}(b_i)}{\operatorname{argmin}} \|\mathbf{A}(b_i)\mathbf{m}(b_i) - \mathbf{d}(b_i)\|_2^2 \\ + \lambda\|\mathbf{m}(b_i) - \tilde{\mathbf{m}}(b_{i-1})\|_2^2,\end{aligned} \tag{10}$$

where $\tilde{\mathbf{m}}(b_i)$ is the least-squares solution of the model till block i, and $\mathbf{d}(b_i)$ is the recorded data corresponding to block i with size $Q \times 1$, $\mathbf{A}(b_i)$ is the $Q \times M$ data matrix of block i, and Q is the length of sliding window. Figure 1 shows the schematic representation of the setup. To update the solution recursively, we will add K data points to the system of equations and remove K data points from the beginning of the previous data vector. We call this step *rank* K *updating and downdating* (see Figure 1 for more information). For this new configuration, one can use

$$\begin{aligned}\tilde{\mathbf{m}}(b_{i+1}) = \tilde{\mathbf{m}}(b_i) + (\mathbf{A}^T(b_{i+1})\mathbf{A}(b_{i+1}) + \lambda\mathbf{I})^{-1}\mathbf{A}^T(b_{i+1}) \\ \times [\mathbf{d}(b_{i+1}) - \mathbf{A}(b_{i+1})\tilde{\mathbf{m}}(b_i)],\end{aligned} \tag{11}$$

to estimate the best model in a least-squares sense that fits the whole system of equations till step $i + 1$. The second term in equation 11 can be interpreted as the least-squares solution of the unpredicted part of the new data set by the solution of the previous step. It is worth mentioning that we use CG to solve equation 11 and we never calculate the inverse of the Hessian (i.e., $(\mathbf{A}^T(b_{i+1})\mathbf{A}(b_{i+1}) + \lambda\mathbf{I})^{-1}$). Using the previous solution as an initial model for the next block has some advantages. This "warm start" for the new sliding window will result in fast convergence of the CG method. This is mainly because the nearby data points are highly correlated in seismic acquisition. It is also good to point out that in each setup, we must use proper preconditioners and regularization term to cluster the eigenvalues of the partial Hessians around one. Ng and Plemmons (1996) prove the convergence of this recursive least-squares technique in probabilistic terms. They show that the method will converge superlinearly with probability of one, provided the underlying process satisfies some assumptions. First of all, the input discrete-time stochastic process should be stationary. Second, the autocovariances of the kernels in each step should be absolutely summable. This in turn will assure the invertibility of the processes in each windowed setup. Third, the variances between the autocovariances of the kernels between different setups should be bounded. Finally, the stationary process has zero mean. All of these assumptions are valid for many time series analysis problems, but it is not clear if they are fully applicable to migration. However, in some cases, one can show that these assumptions are more or less valid for migration (e.g., the acquisition system has a full aperture and sufficiently fine sampling in time and spatial directions). Moreover, the physical properties of the medium such as slowness should be spatially invariant (Stolk, 2000).

For this idealized situation following the works by Gelius et al. (2002), Sjoeberg et al. (2003), and Lecomte (2008), it is easy to show that the action of the Hessian can be approximated via convolutional operators. Note that this relationship is only valid for a

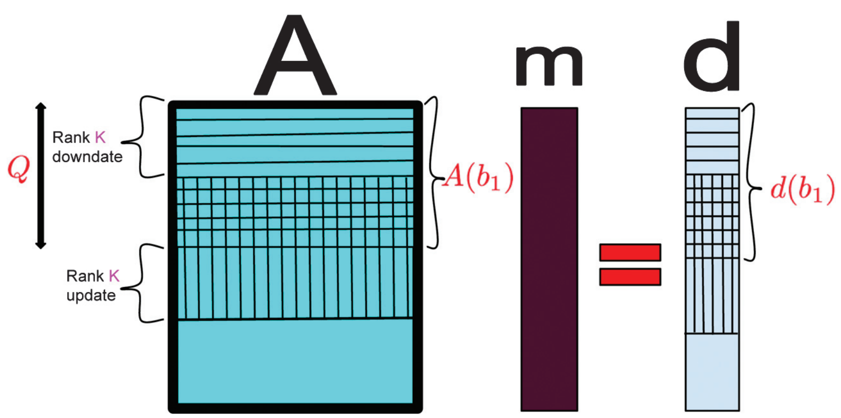

Figure 1. Schematic representation of the block row recursive least-squares algorithm.

spatially invariant medium (i.e., constant slowness in spatial directions).

Nevertheless, the ideas of Ng and Plemmons (1996) can be explored for migration applications and we will show that this method works quite well even for complex medium. In the next section, we will show the efficiency of the proposed method using a simple toy example and the Marmousi model.

PRACTICAL CONSIDERATIONS

It is worth mentioning that in this paper, we assume that the velocity field is known and there is no need for residual velocity analysis. From an application point of view, the proposed method can be implemented via different configurations. The key element to keep in mind is the fact that we need some overlapping between consecutive blocks to assure smooth changes in the Hessians. This method can be applied on a subset of shot gathers in the context of shot-profile migration or in the frequency domain using a subset of frequency slices with overlapping. The subset of shots can be chosen with regular patterns (e.g., in two dimensions from left to right or right to left). We cannot use a pure stochastic gradient method with a random selection of shot gathers because we need overlapping between the blocks. In the frequency domain, one could start from low frequencies and move to high frequencies. Application of the proposed technique using near offsets (near angles) and moving to far offsets (high angles) is also straightforward. A block row recursive least-squares method can also be extended to 3D imaging. One has the option of using subset of shots or offset classes as a block, provided that a proper migration operator is chosen.

COMPUTATIONAL COST ANALYSIS

To compare the computational cost of the full least-squares migration and block row recursive method, let us assume that the number of grid points in one coordinate direction is n; hence, in 2D coordinates, there are $O(n^2)$ points. We will borrow some of the complexity estimates from Marfurt and Shin (1989), Mulder and Plessix (2004a), and Mulder and Plessix (2004b). In general, the one-way wave equation has complexity of $O(n_s n_\omega n^2)$ for a 2D acquisition system, where n_s is the total number of shots and n_ω is the number of frequency realizations in data space. Considering the scheme described in Collino and Joly (1995), the complexity of the one-way wave equation in three dimensions increases by the factor of n (i.e., the complexity is $O(n_s n_\omega n^3)$). Moreover, in the CG method, in each iteration, we need to call migration and demigration operators. Let us say that CG will converge in N iterations. So, the computational cost of the CG method for the 2D case is roughly $O(2N n_s n_\omega n^2)$. On the other hand, the cost of block row recursive least-squares migration in the frequency domain is $O(2\bar{N} n_s n_b n_\omega^b n^2)$, where $\bar{N}$ is the average number of iterations per block, n_b is the number of blocks, and n_ω^b is the number of frequency realizations in each block. In the case of a subset of shot gathers as a block, the cost will be $O(2\bar{N} n_b n_s^b n_\omega n^2)$, where n_b is the number of shot gather blocks and n_s^b is the number of shot gath-

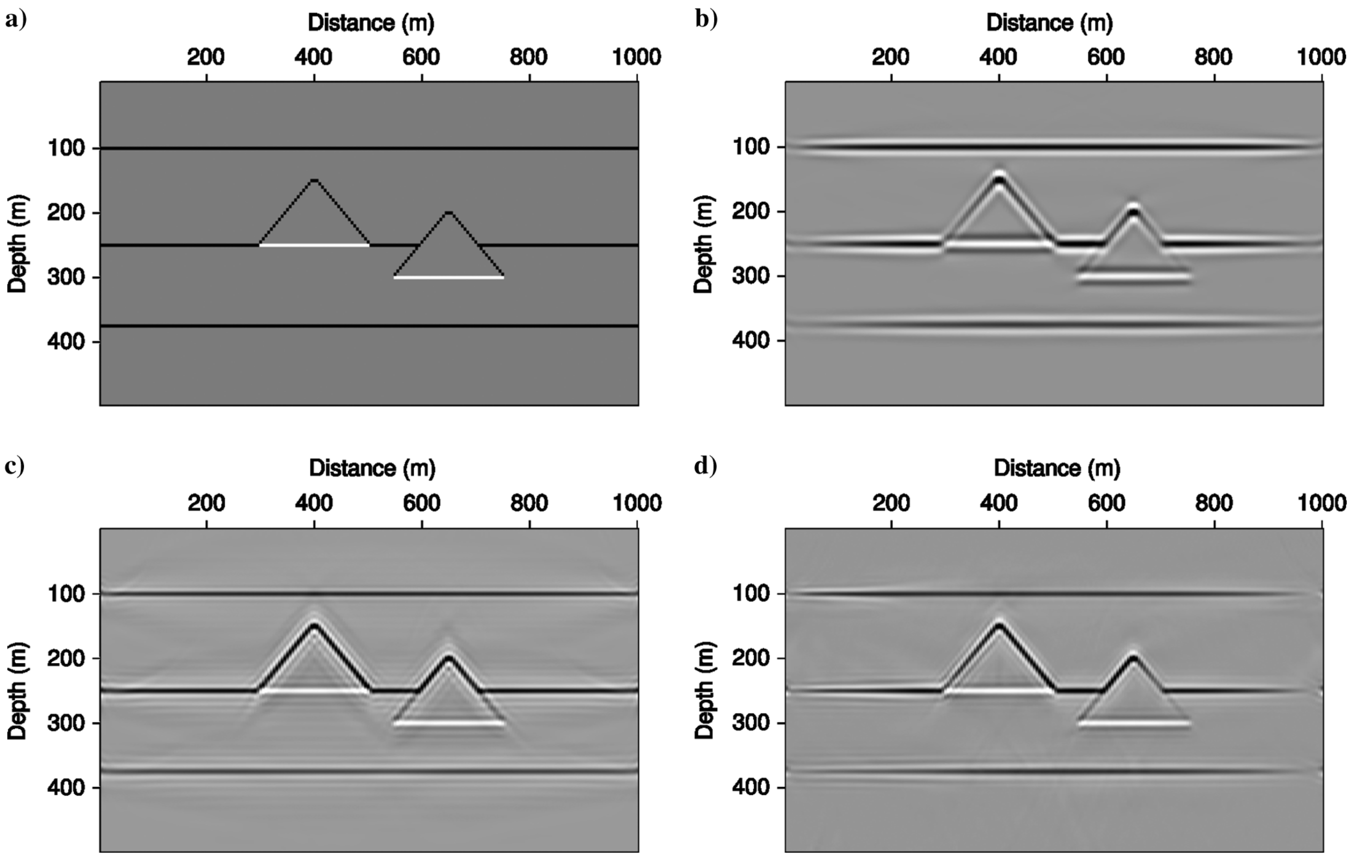

Figure 2. True and migrated reflectivity models. (a) True reflectivity. (b) Adjoint migrated model. (c) Least-squares migrated model. (d) Block row recursive least-squares migrated model.

ers per block. Moreover, to make the algorithm more efficient, one can use a subset of the model for each block (i.e., there are no illuminations outside the coverage zone of subset of shots) and reduce the cost even more. Hence, the new cost will be $O(2\bar{N}n_b n_s^b n_\omega n n_m)$ where n_m is the number of grid points of model in the horizontal direction that covers the illumination aperture corresponding to the geometry of subset of shots.

In the case of 3D imaging the I/O cost can also be drastically decreased. In the block row recursive approach, we need to load the data for each subset of shot gathers only once, rather than repeatedly as is the case for CG applied to the full least-squares migration problem when the data size exceeds core storage, which is routinely the case in 3D imaging.

EXAMPLES

In the following examples, we used shot profile wave equation adjoint and forward operators with a split step correction in the context of adjoint, least-squares, and block row recursive least-squares algorithms.

To test the performance of the proposed method, we generate a 2D reflectivity model (Figure 2a) and we use demigration operator to produce the data set with a 10-m shot interval and a 5-m receiver interval. We use a Ricker wavelet with dominant frequency of 30 Hz, and the receivers are active for all of the shots. Then, we apply the adjoint operator to migrate the data set (Figure 2b). Figure 2c and 2d shows the least-squares and the block row recursive approach migrated models, respectively. In the case of the block row recursive approach, we use 10 consecutive shot gathers in each group and we delete five shot gathers from the beginnings of the previous windowed setup and add five new shot gathers to the end of the new windowed setup. The stopping criterion, for both methods, is set to be equal to the drop of the data residual of the first iteration by the order of 10^6, and the regularization parameter is set to $\lambda = 100$. Please note that in the case of block row recursive method, the stopping criterion is defined for each block separately. The full least-squares approach converged after 13 iterations, and in the case of the block row recursive method, the average iteration number per block was approximately $\bar{N} = 3$. Hence, the speed up for this experiment was by a factor of two. It is worth mentioning that for this experiment, all the receivers were active and we cannot use the second cost criterion (e.g., $O(2\bar{N}n_b n_s^b n_\omega n n_m)$) for the block row recursive method. The block row recursive approach did a good job in recovering the true reflectivity model, and the result is comparable to that of the least squares. To show how they honor the recorded wavefield, we use the recovered reflectivity models to predict the data set. Figure 3a shows the true near-offset section of the data set, and Figure 3b is the predicted near-offset section using the adjoint migrated model. Finally, Figure 3c and 3d shows the predicted near-offset sections using the least-squares and block row

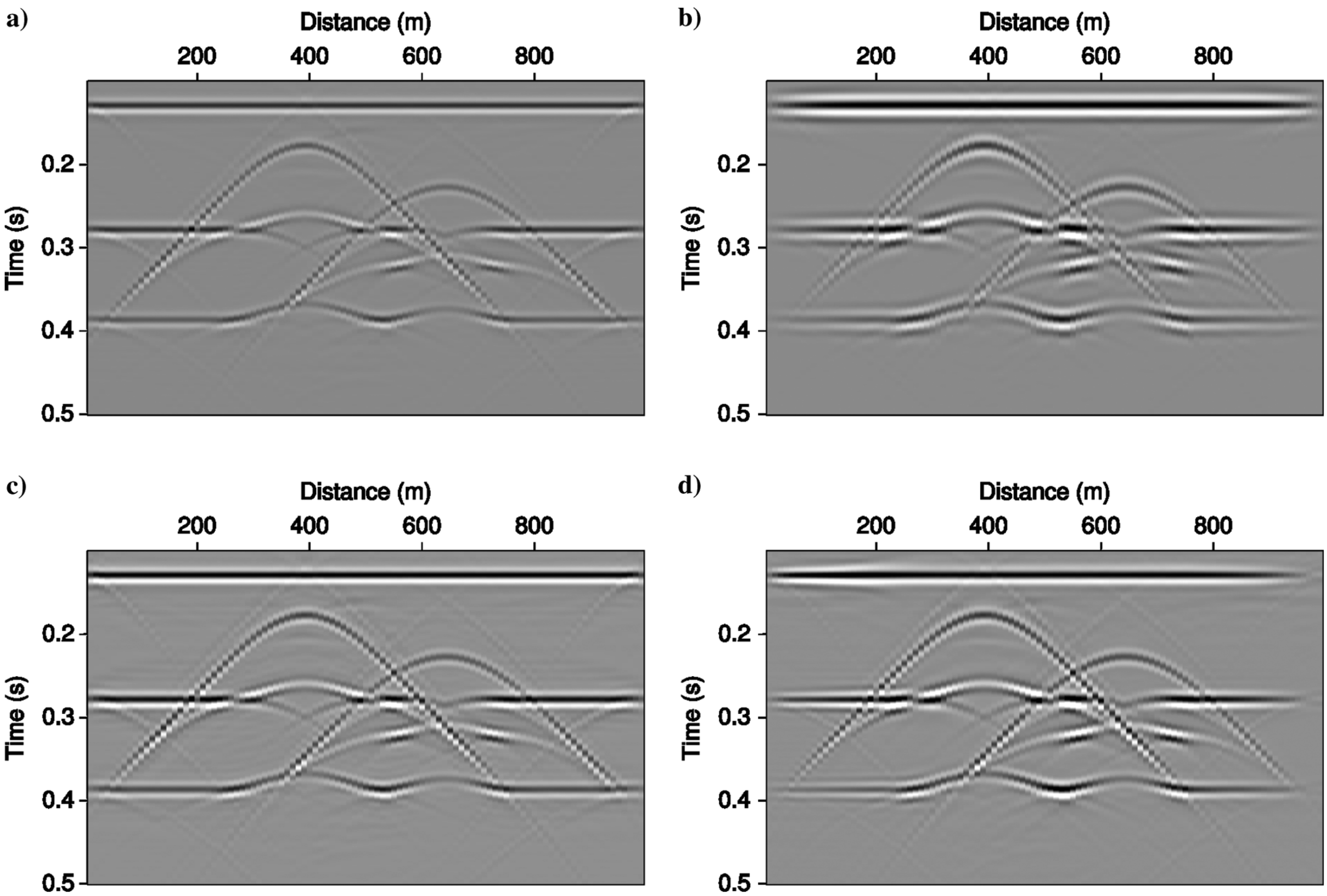

Figure 3. True and predicted near-offset sections. (a) True near-offset section. (b) Adjoint predicted section. (c) Least-squares predicted section. (d) Block row recursive least-squares predicted section.

recursive approach models, respectively. It is clear that the predicted near-offset sections via the least-squares approach and the block row recursive methods are in good accordance with the true near-offset section. Moreover, in both cases, the amplitudes are well preserved and also the detailed features are better predicted than the adjoint predicted data set.

Next, we apply the method on the Marmousi model. It is worth mentioning that the data set is generated by the finite-difference method with a Ricker wavelet with dominant frequency of 20 Hz. The data set consists of 240 shot gathers with 25-m shot interval that are modeled with an off-end survey with receivers to the left of the source being pulled toward the right. Each shot gather consists of 96 traces, with the smallest offset being 200 m, and the receiver intervals are 25 m. The nonsmooth velocity model of the Marmousi is shown in Figure 4a. Figure 4b shows the adjoint-migrated image of the Marmousi data set using shot profile wave equation migration with split-step correction. Figure 5a shows the least-squares-migrated image of the Marmousi data set after 15 iterations using split-step Fourier migration and demigration operators as an adjoint and forward operators required by the CG algorithm. The stopping criterion for full least squares is set to be equal to the drop of the data residual of the first iteration by the order of 10^6, and the regularization parameter is set to $\lambda = 1$. Finally, Figure 5b shows the migrated image produced by the block row recursive approach. In the case of the block row recursive approach, we test different configurations and finally we use five consecutive shot gathers in each group, and we delete three shot gathers from the beginnings of the previous windowed setup and add three new shot gathers to the end of new windowed setup. The stopping criterion for the block row recursive approach is set to be equal to the drop of the data residual of the first iteration by the order of 10^6, and the regularization parameter is set to $\lambda = 1$. Please note that the stopping criterion is defined for each block separately. The average iteration number per block is approximately $\bar{N} = 4$, and we choose n_m to be one-third of the model size in the horizontal direction. Hence, the speed up for this experiment is approximately by a factor of six. It is worth mentioning that we get the same results using different configurations. The block row recursive approach does a good job in recovering the true reflectivity model, and the result is comparable to that of least squares. Comparing the proposed method's result with adjoint, the block row recursive approach did a reasonable job in preserving the amplitude of the reflectors and removed some of the defocusing problems inherent in the adjoint migrated image.

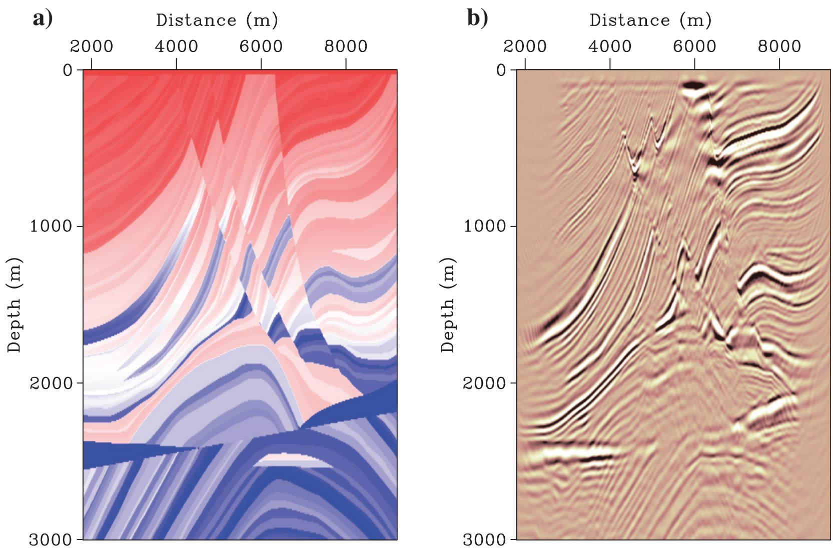

Figure 4. Velocity field of the Marmousi model and the corresponding shot-profile wave-equation-migrated image. (a) True velocity field. (b) Adjoint migrated model.

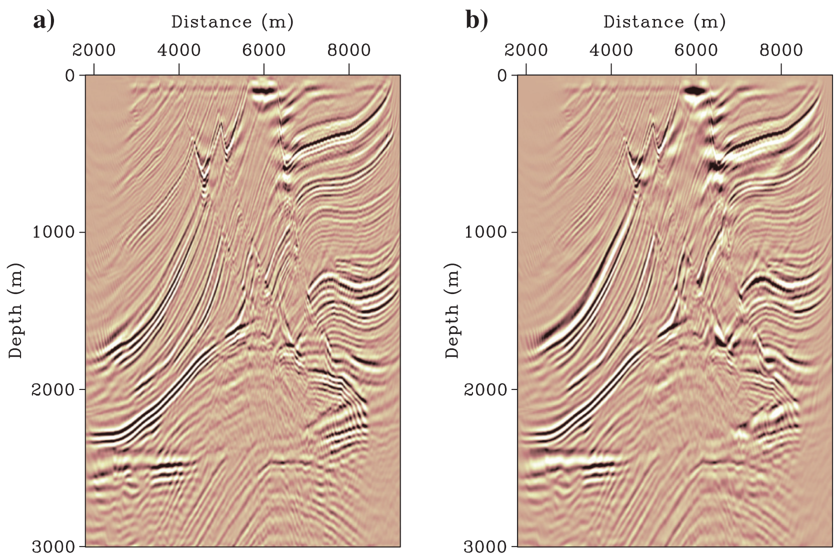

Figure 5. Comparison of the results of full least-squares migration and the proposed method on the Marmousi model. (a) Full least-squares-migrated model. (b) Block row recursive least-squares-migrated model using blocks of shot gathers.

CONCLUSION

We have presented a block row recursive least-squares migration method. This method uses a blockwise update of the demigration operator via rank K update and downdate in each setup while the new data points are successively added to the data vector. In each windowed setup, the CG algorithm was used to solve the system of equations in least-squares sense. To have fast convergence, the previous solution of the method is implemented as an initial solution for the next block. This warm start will result in fast convergence of the CG algorithm. This is supported by the fact that nearby blocks are highly correlated. The results of applying this technique on a simple toy and the Marmousi model convinced us that the block row recursive method can be a practical tool for improving the spatial resolution of the migrated images.

ACKNOWLEDGMENTS

We thank the sponsors of the Signal Analysis and Imaging Group (SAIG) at the University of Alberta. We also thank J. Rickett, K. Van Wijk, J. Thorbecke, J. Etgen, and another anonymous reviewer for their constructive comments and suggestions.

REFERENCES

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524, doi: 10.1190/1.1441434.

Chavent, G., and R. E. Plessix, 1999, An optimal true-amplitude least squares prestack depth-migration operator: Geophysics, **64**, 508–515, doi: 10.1190/1.1444557.

Collino, F., and P. Joly, 1995, Splitting of operators, alternate directions, and paraxial approximations for the three-dimensional wave equation: SIAM Journal on Scientific Computing, **16**, 1019–1048, doi: 10.1137/0916059.

Dai, W., X. Wang, and G. T. Schuster, 2011, Least-squares migration of multisource data with a deblurring filter: Geophysics, **76**, no. 5, R135–R146, doi: 10.1190/geo2010-0159.1.

Duijndam, A. J. W., A. W. F. Volker, and P. M. Zwartjes, 2000, Reconstruction as efficient alternative for least squares migration: 70th Annual International Meeting, SEG, Expanded Abstracts, 1012–1015.

Etgen, J., 2002, Recent advances and the road ahead: An invited talk presented at 72nd Annual International Meeting, SEG, .

Fomel, S., and A. Guitton, 2006, Regularizing seismic inverse problems by model reparameterization using plane-wave construction: Geophysics, **71**, no. 5, A43–A47, doi: 10.1190/1.2335609.

Gazdag, J., 1978, Wave equation migration with the phase-shift method: Geophysics, **43**, 1342–1351, doi: 10.1190/1.1440899.

Gazdag, J., and P. Sguazzero, 1984, Migration of seismic data by phase shift plus interpolation: Geophysics, **49**, 124–131, doi: 10.1190/1.1441643.

Gelius, L.-J., I. Lecomte, and H. Tabti, 2002, Analysis of the resolution function in seismic prestack depth imaging: Geophysical Prospecting, **50**, 505–515, doi: 10.1046/j.1365-2478.2002.00331.x.

Guitton, A., 2004, Amplitude and kinematic corrections of migrated images for nonunitary imaging operators: Geophysics, **69**, 1017–1024, doi: 10.1190/1.1778244.

Hu, J., G. T. Schuster, and P. A. Valasek, 2001, Poststack migration deconvolution: Geophysics, **66**, 939–952, doi: 10.1190/1.1444984.

Kaplan, S. T., M. Naghizadeh, and M. D. Sacchi, 2010a, Data reconstruction with shot–profile least squares migration: Geophysics, **75**, no. 6, WB121–WB136, doi: 10.1190/1.3478375.

Kaplan, S. T., P. S. Routh, and M. D. Sacchi, 2010b, Derivation of forward and adjoint operators for least-squares shot-profile split-step migration: Geophysics, **75**, no. 6, S225–S235, doi: 10.1190/1.3506146.

Kazemi, N., and M. Sacchi, 2014, Filter-based least squares wave equation shot profile migration: 76th Annual International Conference and Exhibition, EAGE, Extended Abstracts, We P02 03, doi: 10.3997/2214-4609.20141225.

Kuhl, H., and M. D. Sacchi, 2003, Least-squares wave-equation migration for AVP/AVA inversion: Geophysics, **68**, 262–273, doi: 10.1190/1.1543212.

Lecomte, I., 2008, Resolution and illumination analyses in PSDM: A ray-based approach: The Leading Edge, **27**, 650–663, doi: 10.1190/1.2919584.

Marfurt, K., and C. Shin, 1989, The future of iterative modeling in geophysical exploration, *in* E. Eisner, ed., Handbook of geophysical exploration: Pergamon Press, Supercomputers in Seismic Exploration 21, 203–228.

McMechan, G. A., 1983, Migration by extrapolation of time-dependent boundary values: Geophysical Prospecting, **31**, 413–420, doi: 10.1111/j.1365-2478.1983.tb01060.x.

Mulder, W. A., and R. E. Plessix, 2004b, How to choose a subset of frequencies in frequency domain finite-difference migration: Geophysical Journal International, **158**, 801–812, doi: 10.1111/j.1365-246X.2004.02336.x.

Mulder, W. A., and R. E. Plessix, 2004a, A comparison between one-way and two-way wave-equation migration: Geophysics, **69**, 1491–1504, doi: 10.1190/1.1836822.

Naghizadeh, M., and M. Sacchi, 2009, f-x adaptive seismic-trace interpolation: Geophysics, **74**, no. 1, V9–V16, doi: 10.1190/1.3008547.

Naoshi, A., and G. T. Schuster, 2009, Fast least-squares migration with a deblurring filter: Geophysics, **74**, no. 6, WCA83–WCA93, doi: 10.1190/1.3155162.

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.

Ng, M., and R. Plemmons, 1996, Fast recursive least squares adaptive filtering by fast Fourier transform-based conjugate gradient iterations: SIAM Journal on Scientific Computing, **17**, 920–941, doi: 10.1137/0917060.

Plessix, R. E., and W. A. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophysical Journal International, **157**, 975–987, doi: 10.1111/j.1365-246X.2004.02282.x.

Schneider, W. A., 1978, Integral formulation for migration in two and three dimensions: Geophysics, **43**, 49–76, doi: 10.1190/1.1440828.

Sjoeberg, T. A., L. Gelius, and I. Lecomte, 2003, 2D deconvolution of seismic image blur: 73rd Annual International Meeting, SEG Expanded Abstracts, 1055–1058.

Stoffa, P. L., J. T. Fokkema, M. de Luna Freir, and W. P. Kessinger, 1990, Split-step Fourier migration: Geophysics, **55**, 410–421, doi: 10.1190/1.1442850.

Stolk, C. C., 2000, Microlocal analysis of a seismic linearized inverse problem: Wave Motion, **32**, 267–290, doi: 10.1016/S0165-2125(00)00043-3.

Stolt, R. H., 1978, Migration by Fourier transform: Geophysics, **43**, 23–48, doi: 10.1190/1.1440826.

Symes, W. W., 2008, Approximate linearized inversion by optimal scaling of prestack depth migration: Geophysics, **73**, no. 2, R23–R35, doi: 10.1190/1.2836323.

Toxopeus, G., J. Thorbecke, K. Wapenaar, S. Petersen, E. Slob, and J. Fokkema, 2008, Simulating migrated and inverted seismic data by filtering a geologic model: Geophysics, **73**, no. 2, T1–T10, doi: 10.1190/1.2827875.

Wei, D., C. Boonyasiriwat, and G. T. Schuster, 2010, 3D multisource least-squares reverse time migration, 3120–3124.

Whitmore, N. D., 1983, Iterative depth migration by backward time propagation: 53rd Annual International Meeting, SEG, Expanded Abstracts, 382–385.

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.