GEOPHYSICS, VOL. 82, NO. 4 (JULY-AUGUST 2017); P. A31–A35, 4 FIGS., 2 TABLES.
10.1190/GEO2016-0552.1

Efficient anisotropic dip filtering via inverse correlation functions

Paul Wellington[1], Romain Brossier[1], Okba Hamitou[2], Phuong-Thu Trinh[1], and Jean Virieux[1]

ABSTRACT

We have developed a computational framework that allows an efficient, spatially variant correlation filter for anisotropic dip filtering. The approach is based on the Laplace correlation function, for which there exists analytical expressions for the correlation kernel and its inverse kernel in the 1D case. An extension to higher dimensions by adding orthogonal 1D inverse functions provides a linear equation whose solution is identical to applying a Bessel filter. We have found that a good approximation of the Laplace filter function is obtained by applying a cascade of Bessel filters. We implement such an inverse operator on regular grids with a finite-difference stencil, giving a sparse matrix for the linear equation, which can be solved efficiently through a conjugate-gradient algorithm. Computing such a correlation operation by solving the linear system involving the inverse operator is significantly faster than applying the correlation function via convolution or windowed convolution: Solving the linear system is as fast as applying often-used tensorized 1D convolutions.

INTRODUCTION

Seismic imaging techniques build representations of the subsurface from acquired seismic data. Strategies for this imaging process are numerous and are constantly improving. One key aspect is improving the resolution either by altering how kinematic information is used or by incorporating information that is more sensitive to amplitude. The two key families of imaging techniques can be summarized as (1) migration approaches based on kinematic projections mitigating amplitude influence, from classic migration (Claerbout, 1971, 1976) to least-squares migration (Lailly, 1983; Nemeth et al., 1999), and (2) migration approaches based on iterative inversion of waveform misfit increasing the influence of true seismic amplitude, such as full-waveform inversion (FWI) (Lailly, 1983; Tarantola, 1984, for the seminal references). Both approaches rely on imaging conditions that are sensitive to several undesired factors, such as acquisition footprints, imaging artifacts, spatial aliasing, and coherent and incoherent noise. Many of these artifacts may have orientations conflicting with the underlying true geologic structure. They can be effectively attenuated by smoothing along the orientation of the expected local geologic dip.

There are several strategies that can be applied to attempt to smooth along specific structural dips of a seismic image. Some approaches window over a few samples parallel to the dip, filtering out unwanted features using simple statistical methods, such as median, mean, or alpha-trimmed mean (Lee and Lassa, 1985; Longbotham and Bovak, 1989). Other strategies include those that look to attenuate wavenumbers in a particular orientation, such as the wave-kill filter (Claerbout, 1992), the plane-wave annihilation filter (Fomel, 2002), or those based on anisotropic diffusion kernels (Hale, 2007) and structure-oriented filtering (Fehmers and Höcker, 2003; Hale, 2011). The inverse of these notch filters is often used to "steer" along the geologic dip by selecting a proper diffusive-like tensor. For example, such an approach has been applied as part of FWI (Guitton et al., 2012).

In this paper, we discuss the calculation for an image value at a given point by summing over adjacent points using a monotonically decreasing weighting function. The weight of this function decreases as the distance increases between the point of interest and adjacent points. This weighting function acts as a low-pass spatial filter, and it is referred to as the correlation kernel. The rate of decay of the correlation function depends on the shape of the weight, and it is often characterized by a correlation length: long correlation lengths smooth more than shorter ones. Designing weights and, in fact, local correlation lengths gives an intuitive link with the resolution one expects in the image at a given point. By explicitly rotating orthogonal correlation lengths to be aligned along the dip, we argue that the short and long correlation directions and lengths will be correctly represented, which should prevent undesired "shaping" of the image. We believe that this behavior is particularly important especially when we have used this filter as a preconditioner in FWI (Wellington et al., 2016).

Manuscript received by the Editor 28 October 2016; revised manuscript received 22 December 2016; published online 05 May 2017.
[1]Université Grenoble Alpes, ISTerre, Grenoble, France. E-mail: paul.wellington@univ-grenoble-alpes.fr; romain.brossier@univ-grenoble-alpes.fr; phuong-thu.trinh@univ-grenoble-alpes.fr; jean.virieux@univ-grenoble-alpes.fr.
[2]Université Grenoble Alpes, ISTerre and Laboratoire Jean Kuntzmann, Grenoble, France. E-mail: okba.hamitou@univ-grenoble-alpes.fr.

Because geology varies spatially, there is a strong desire for the correlation characteristics to be defined locally at each point. This factor precludes the option of efficiently applying correlation functions in the wavenumber-frequency domain. In addition, we typically want to use a shorter correlation length orthogonal to the geologic dip to minimize filtering of key details across contrasts, whereas a longer correlation length is more suitable parallel to the dip (for how the dip can be approximated locally, see Fomel, 2002).

In this work, we introduce an efficient tool to perform the convolution by a correlation function of the vector describing the image to be filtered. The application of this correlation function is shown to be efficient in 1D, 2D, and 3D. Moreover, it can handle dip fields that vary smoothly in space (associated with anisotropic correlation lengths) and spatially varying correlation lengths. In the first part, the application of a correlation function based on convolution is discussed, prior to addressing the interest of using an inverse filter operator in the second part. In these two parts, the theoretical computational cost is evaluated. These costs are then compared on realistically sized 2D and 3D images from field data, prior to conclusions.

CORRELATION FUNCTIONS

The convolution by a correlation function to an image or volume vector $\mathbf{u}$ is defined as

$$\tilde{\mathbf{u}}(\mathbf{x}) = \int \mathcal{C}(\mathbf{x}, \mathbf{x}')\mathbf{u}(\mathbf{x}')d\mathbf{x}', \qquad (1)$$

where $\tilde{\mathbf{u}}$ is the output vector and $\mathcal{C}(\mathbf{x}, \mathbf{x}')$ is the convolution kernel that describes how the image $\mathbf{u}$ at location $\mathbf{x}$ is influenced by the value of the image at an offset location $\mathbf{x}'$. This expression can be applied in 1D, 2D, or 3D (where in 3D, $\mathbf{x} = (x, y, z)$ and in 2D, $\mathbf{x} = (x, z)$). When applied in a discrete world, the convolutional integral becomes a summation over all elements of the finite-size image or volume. Such a computation can be intensive because the convolution requires a loop over the whole input vector for each element of the output vector. Table 1 provides the computational complexity of this procedure in the third column.

One simple approach to decrease the computational complexity of the explicit convolution would be to move to the Fourier domain in which the convolution of two functions becomes the product of the Fourier transforms of the two functions. We would then benefit from the efficiency of the forward and inverse fast Fourier transforms. However, such Fourier domain approaches apply only to stationary filters and therefore do not allow for spatially variant properties.

In practice, the kernel of most correlation filters decays with physical distance and we can approximate the filtered image by

$$\tilde{\mathbf{u}}(\mathbf{x}) \approx \int_{\Omega_l} \mathcal{C}(\mathbf{x}, \mathbf{x}')\mathbf{u}(\mathbf{x}')d\mathbf{x}', \qquad (2)$$

where Ω_l is a region of radius l, where the correlation weights are above a certain threshold. We consider the convolution kernel to be zero outside this domain. The windowed version of the convolution provided in equation 2 requires an area integral in 2D and a volume integral in 3D. As shown in the fourth column of Table 1, this windowed convolution significantly reduces the computational complexity, when the radius l is much smaller than the image dimensions.

A 2D example of a correlation function used as the convolution kernel is shown in Figure 1, for the case in which the correlation length L_x is longer in one orientation than in the other L_z. In such a case, the region Ω_l forms an ellipse described by the window lengths l_x and l_z, outside of which the correlation function is considered to be zero.

An often applied approach to accelerate the filtering process involves using multidimensional correlation functions formed via tensorized application of 1D correlation functions. A common example of such a correlation function is the Gaussian correlation function $\mathcal{G}$ that can be tensorized such that $\mathcal{G}(x, z, x', z') = \mathcal{G}_x(x, x')\mathcal{G}_z(z, z')$. In 2D, for example, we can recast the 2D integral as the successive application of 1D functions, yielding

$$\tilde{\mathbf{u}}(x, z) \approx \int_{-l_z}^{l_z} \mathcal{G}_z(z; z') \left[\int_{-l_x}^{l_x} \mathcal{G}_x(x; x')\mathbf{u}(x', z')dx' \right] dz'. \quad (3)$$

Table 1. Theoretical computational complexity for the different filtering approaches, assuming a square (2D) or cube (3D) shape with n discrete points in each spatial direction.

Dimension	Domain size	Convolution	Windowed convolution	Tensorized convolution	Inverse operator
2D	n^2	n^4	$a^2n^2 \approx n^2$	$2an^2 \approx n^2$	$2b_{2D}n^2N_{\text{iter}} \approx n^2$
3D	n^3	n^6	$a^3n^3 \approx n^3$	$3an^3 \approx n^3$	$3b_{3D}n^3N_{\text{iter}} \approx n^3$

Note: The term a is the number of operations for computing a 1D windowed correlation function to a single point (a constant if $a \ll n$), b is the number of nonzero entries in each row of the $(C^{-1}_{X_{\text{RES}}})$ matrix (a constant), N_{iter} is the number of conjugate gradient iterations to reach convergence (in practice, numerical tests show that N_{iter} does not scale with n and can be considered as a constant). Note that some constants, which do not affect the leading order of the complexities, have been omitted for clarity.

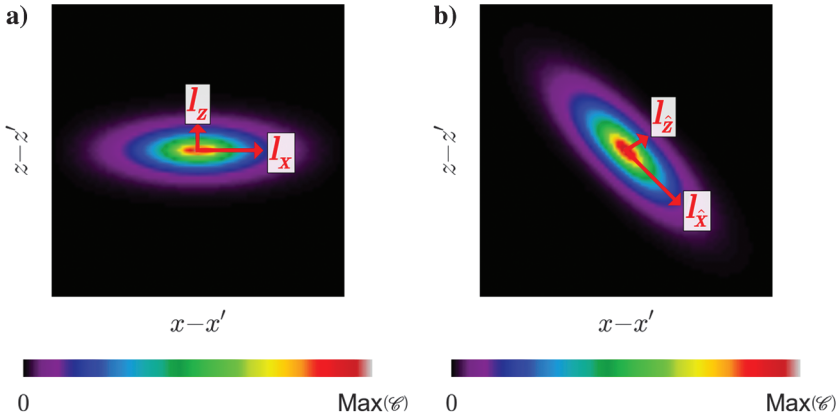

Figure 1. (a) An anisotropic 2D correlation function defined with correlation directions l_x and l_z. (b) An anisotropic correlation function not aligned with the x- and z-directions.

In such a tensorized formulation, the computation cost is reduced (Table 1, fifth column) because we only need to perform several 1D integrations instead of one multidimensional integral.

A shortcoming of the tensorized approach exists if one needs to align the anisotropic filter to the spatially varying orientation of geologic dip. In such a case, it is no longer possible to use the tensorized property available to some multidimensional correlation functions such as the Gaussian function. The importance of aligning the correlation functions along the dip is highlighted in an example (Figure 2) taken from an offshore region of the Netherlands. The 2D image shows reflectors dipping from left to right, which are contaminated by weak noise. This example is filtered with a 2D Laplace function, and three different correlation strategies are applied. The first strategy involves isotropic filtering in which the correlation weight decreases isotropically in space. It results not only in an attenuation of the noise content but also of the reflector signal. Anisotropic filtering is applied in the second case, in which the correlation function decays more slowly in the horizontal direction. We see that we not only have preserved some of the vertical high-wavenumber detail, but we have also added artifacts caused by horizontal smearing across the reflectors. The best result is obtained by aligning the true local orientation of the reflectors through a change of the local coordinate system $(\hat{x}, \hat{z})$ by a rotation angle corresponding to the expected local dip. In other words, accounting for this dip information requires an integration over the ellipse defined by the rotated window lengths $l_{\hat{x}}$ and $l_{\hat{z}}$ using equation 2 in the new coordinate system.

CORRELATION FILTERING VIA INVERSE CORRELATION FUNCTIONS

For some correlation functions, such as the 1D Laplace function,

$$\mathcal{C}_{1\mathrm{D}_{\mathrm{LAP}}}(x;x') = \frac{1}{2L} e^{-\frac{|x-x'|}{L}}, \tag{4}$$

an analytical inverse correlation function exists. The Laplace correlation function is commonly used as a low-pass filter, in which the rate of decay is controlled by the correlation length L. Tarantola (2005) highlights that the inverse of the 1D Laplace correlation function is expressed as

$$\mathcal{C}^{-1}_{1\mathrm{D}_{\mathrm{LAP}}}(x;x') = L\left(\frac{1}{L}\delta(x-x') - L\delta^2(x-x')\right), \tag{5}$$

where $\delta(x)$ is a delta function in space and $\delta^2(x)$ is its second derivative with respect to space. On a regular grid, this 1D inverse correlation function can be discretized using finite-difference techniques to build the discrete inverse correlation matrix C^{-1}_{LAP}, using the Neumann boundary conditions. Using a second-order centered finite-difference stencil, we obtain a sparse matrix with only three nonzero entries per row. Here, we propose using this roughening operator to perform the correlation filtering by solving the sparse linear system:

$$C^{-1}_{1\mathrm{D}_{\mathrm{LAP}}}\tilde{\mathbf{u}} = \mathbf{u}, \tag{6}$$

where $\mathbf{u}$ is the input raw vector and $\tilde{\mathbf{u}}$ is the output filtered vector, as opposed to a standard approach in which $C_{1\mathrm{D}_{\mathrm{LAP}}}$ is used for discrete convolution. Equation 6 can be efficiently solved using a conjugate gradient algorithm, providing the same solution as equation 1. It is important to note that our inverse correlation operator is stable as long as the correlation parameter L is greater than the used finite-difference grid increment. We find that the required number of iterations to solve equation 6 increases linearly with L. These observations also hold for the 2D and 3D cases expressed later allowing an effective and efficient strategy to perform filtering.

By adding two 1D inverse correlation functions, we build a 2D inverse correlation function:

$$\mathcal{C}^{-1}_{2\mathrm{D}_{\mathrm{BES}}}(x,z;x',z') = \frac{L_x}{2}\left(\frac{1}{L_x}\delta(x-x') - L_x\delta^2(x-x')\right) + \frac{L_z}{2}\left(\frac{1}{L_z}\delta(z-z') - L_z\delta^2(z-z')\right), \tag{7}$$

where L_x and L_z are the correlation lengths in the x- and z-directions. The 2D corresponding correlation function is the elementary solution of the partial differential equation (PDE):

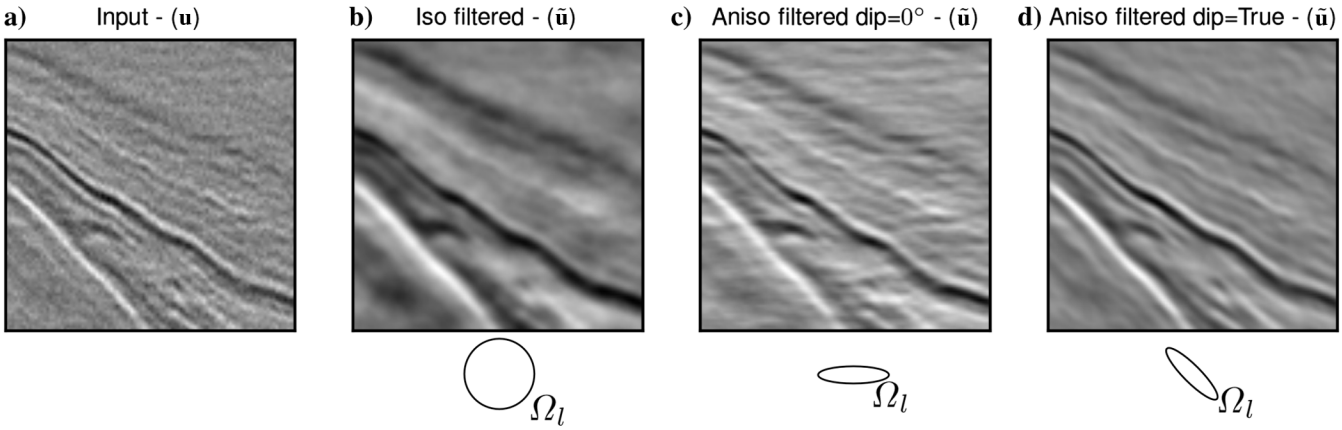

Figure 2. (a) An input seismic image, (b) seismic image filtered with isotropic correlation lengths, (c) filtered with anisotropic correlation lengths along Cartesian coordinates, and (d) filtered with anisotropic correlation lengths aligned with the geologic dip. The approximate region of the filter is shown by Ω_l.

$$C_{2D_{BES}}(x,z)-\frac{1}{2}\left(L_x^2\frac{\partial^2}{\partial x^2}+L_z^2\frac{\partial^2}{\partial z^2}\right)C_{2D_{BES}}(x,z)=\delta(x,z), \quad (8)$$

which corresponds to a modified Helmholtz equation when the correlation lengths are constant. The solution of this PDE is the modified Bessel function of the second kind (Abramowitz and Stegun, 1972; Polyanin and Nazaikinskii, 2002). Therefore, through the specific construction of the corresponding inverse operator, we propose a convolution as a Bessel filter. In practice, the function of equation 7 can be discretized using the second-order centered finite-difference technique, yielding a matrix with only five nonzero entries per row. This 2D discrete scheme can also be rotated in space to honor dip, totaling nine entries per row. We can even approximate locally this discretization of the inverse operator with local values of correlation lengths. The modified Helmholtz is no longer honored, but smooth variations of correlation lengths lead to small artifacts in the output solution.

The Bessel function does not exhibit the same decay as the Laplace function; however, we find numerically the following relation:

$$C_{2D_{BES}}(C_{2D_{BES}}\mathbf{u}) \approx C_{2D_{LAP}}\mathbf{u}. \quad (9)$$

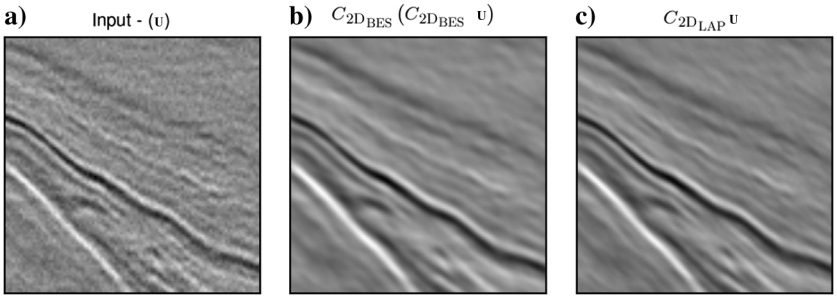

Figure 3. (a) The input noisy seismic record. (b) Result filtered twice by solving the linear system associated with the inverse Bessel function. (c) The input filtered with the analytical Laplace correlation function.

A filtering comparison of the left-hand side (LHS) and right-hand side (RHS) in the above equation is shown in Figure 3.

If the number of iterations of the conjugate gradient solver remains small, it appears that filtering with the LHS of equation 9, through two linear system resolutions involving the discrete matrix $C_{2D_{BES}}^{-1}$, appears competitive with the RHS, which relies on windowed convolution (equation 2), as shown in Table 1.

The same procedure can be applied in 3D, leading to

$$\begin{aligned}C_{3D_{BES}}^{-1}(x,y,z;x',y',z') = &\frac{L_x}{3}\left(\frac{1}{L_x}\delta(x-x')-L_x\delta^2(x-x')\right)\\ &+\frac{L_y}{3}\left(\frac{1}{L_y}\delta(y-y')-L_y\delta^2(y-y')\right)\\ &+\frac{L_z}{3}\left(\frac{1}{L_z}\delta(z-z')-L_z\delta^2(z-z')\right). \quad (10)\end{aligned}$$

It can be shown that the correlation function associated with this inverse function is a 3D spherical Bessel function of the second kind when correlation lengths are constant. The inverse function (equation 11) can be discretized using a centered second-order finite-difference technique, requiring seven nonzero entries per row. As in the 2D case, rotation can be applied to honor any dip geometry, leading to 19 nonzero entries per row in the linear system to be solved.

Again, we numerically find an approximate relation:

$$C_{3D_{BES}}[C_{3D_{BES}}(C_{3D_{BES}}\mathbf{u})] \approx (C_{3D_{LAP}})\mathbf{u}, \quad (11)$$

where $C_{3D_{LAP}}$ is the 3D Laplace correlation function. The demonstration of equations 9 and 11 is covered numerically (Wellington, 2016) and alluded to analytically (see Tarantola, 2005, section 7.23) but not discussed in detail in this paper.

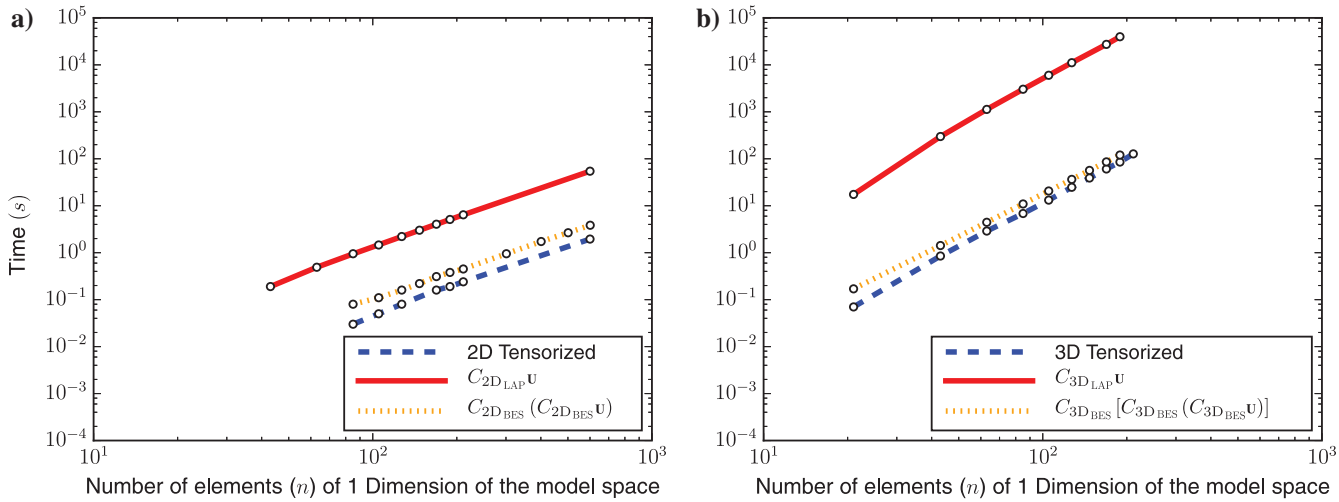

Figure 4. Computational cost comparison of filtering an arbitrary vector **u** of increasing size n. Different strategies are applied (blue, tensorized filtering; red, truncated convolution; orange, using the inverse operator based approach) for (a) 2D and (b) 3D.

DISCUSSION ON COMPUTATIONAL COST

The theoretical computational complexity of the different approaches is summarized in Table 1. From this table, we conclude that all approaches except the standard convolution scale equivalently from the theoretical point of view. Note that, from the most efficient approaches, only the windowed convolution and the inverse-operator-based approach allow for spatially varying dips and correlation lengths.

To compare the practical efficiency of the approaches, we illustrate one application of the three most efficient methods using a relatively small constant correlation length (equivalent to the distance between two points) in all orientations (isotropic filtering), and without dip to apply a tensorized Gaussian filter. The size of the model domain n is increased to assess the practical complexities. The results of the computational performance are highlighted in Figure 4. One first observation is that results of the practical application match very well the theoretical complexity: The slopes in the log-log scale match their expected dimension exponents. Second, using linear system resolution implying our inverse operator of the Bessel function allows us to filter the vector $\mathbf{u}$ several orders of magnitude faster than the truncated convolution approach of equation 2. Indeed, numerical tests show that the number of iterations for the convergence of the linear system scales approximately linearly with the correlation length as the operator becomes less diagonally dominant, meaning that the N_{iter} factor remains relatively small. Therefore, it implies a significantly lower coefficient for the leading complexity term in the inverse operator strategy. This makes the strategy attractive for large data sets and, in the presence of long correlation length, with the added advantage of being able to include spatially variant dip, unlike the similarly efficient tensorized approach.

CONCLUSION

Many practical applications of filtering can benefit from an anisotropic correlation operator aligned with the geologic dip. We present an efficient means of wavenumber filtering of images using correlation functions. The 1D inverse Laplace correlation function is used at the core of our design of sparse 2D and 3D inverse correlation functions. The associated correlation function to these inverse 2D and 3D functions is shown to be modified Bessel and spherical Bessel functions of the second kind. We find that, when applied twice and three times in cascade, they accurately mimic the 2D and 3D Laplace filters, respectively. The application of the discretized inverse correlation operator to a vector can be achieved efficiently by solving a well-conditioned sparse linear system with a conjugate gradient algorithm. Despite sharing the same leading order for their computational complexities, using the inverse filter through linear system resolution is far more efficient than conventional approaches of windowed convolution, and it scales favorably to long correlation lengths, large numbers of cells, and 3D applications. The practical application of this filtering operator to FWI and further seismic examples is the focus of our ongoing work.

ACKNOWLEDGMENTS

This study was partially funded by the SEISCOPE consortium, sponsored by CGG, Chevron, Exxon-Mobil, JGI, Shell, Sinopec, Statoil, and Total and Woodside. This study was granted access to the HPC resources of CIMENT infrastructure and CINES/IDRIS under the allocation 046091 made by GENCI. The authors would like to acknowledge fruitful discussions made and assistance provided by L. Métivier. The authors would also like to acknowledge the Open Seismic Repository and dGB Earth Sciences for providing access to the data.

REFERENCES

Abramowitz, M., and I. A. Stegun, 1972, Handbook of mathematical functions with formulas, graphs, and mathematical tables: Dover publications Inc.

Claerbout, J., 1971, Towards a unified theory of reflector mapping: Geophysics, **36**, 467–481, doi: 10.1190/1.1440185.

Claerbout, J. F., 1976, Fundamentals of geophysical data processing: McGraw-Hill Book Co.

Claerbout, J. F., 1992, Earth sounding analysis: Blackwell Scientific Publications.

Fehmers, G., and C. Höcker, 2003, Fast structural interpretation with structure-oriented filtering: Geophysics, **68**, 1286–1293, doi: 10.1190/1.1598121.

Fomel, S., 2002, Applications of plane-wave destruction filters: Geophysics, **67**, 1946–1960, doi: 10.1190/1.1527095.

Guitton, A., G. Ayeni, and E. Díaz, 2012, Constrained full-waveform inversion by model reparameterization: Geophysics, **77**, no. 2, R117–R127, doi: 10.1190/geo2011-0196.1.

Hale, D., 2007, Local dip filtering with directional Laplacians: Technical report, Centre for Wave Phenomena, Colorado School of Mines.

Hale, D., 2011, Structure-oriented bilateral filtering of seismic images: 81st Annual International Meeting, SEG, Expanded Abstracts, 3596–3600.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering, Theory and Application, SIAM, Expanded Abstracts, 206–220.

Lee, Y. H., and S. A. Lassa, 1985, Generalized median filtering and related nonlinear filtering techniques: IEEE Transactions on Acoustics Speech and Signal Processing, **33**, 627–683.

Longbotham, H. G., and A. C. Bovak, 1989, Theory of order statistics filters and their relationship to linear fir filters: IEEE Transactions on Acoustics Speech and Signal Processing, **37**, 275–287.

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.

Polyanin, A. D., and V. E. Nazaikinskii, 2002, Handbook of linear partial differential equations for engineers and scientists: CRC Press.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266, doi: 10.1190/1.1441754.

Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: SIAM.

Wellington, P., 2016, Efficient 1D, 2D and 3D geostatistical constraints and their application to full waveform inversion: Ph.D. thesis, University of Grenoble Alpes.

Wellington, P., R. Brossier, O. Hamitou, S. Garambois, and J. Virieux, 2016, Preconditioning FWI with the sparse additive inverse Laplacian correlation function: 86th Annual International Meeting, SEG, Expanded Abstracts, 1462–1466.