2025 EAGE ANNUAL
86TH CONFERENCE & EXHIBITION
TOULOUSE | FRANCE

Total-Variation regularised image-domain least-squares migration

Matteo Ravasi, Lorenzo Casasanta, Volkan Akcelik, Ivan Vasconcelos
Shearwater GeoServices

Image-domain least-squares migration (i-LSM) is a powerful imaging tool that estimates a full-bandwidth reflectivity (or perturbation) model from a seismic image by means of a spatially non-stationary deconvolution process with so-called point-spread functions. However, neither a reflectivity nor a perturbation model are a direct representation of any subsurface reservoir property. In this work, we propose to recast i-LSM to retrieve full-bandwidthmedium properties (e.g., velocity); this is achieved by applying a Total-Variation (TV) regularisation term to the sum of the background velocity model (used to propagate wavefields in the Born modeling operator) and the velocity perturbation (used by the Born modeling operator to construct single-scattering seismic data). By leveraging the property of pre-composition of proximal operators, the proposed TV-regularised inverse problem can be successfully solved using the Primal-Dual algorithm with minimal additional cost to standard least-squares regularised inversion (or sparsity-promoting inversion when inverting for reflectivity). Numerical results on both synthetic and field datasets are presented to verify the effectiveness of the proposed approach.

Total-Variation regularised image-domain least-squares migration

Introduction

Least-squares migration (LSM) is an advanced seismic imaging tool that provides enhanced resolution and amplitude fidelity compared to conventional (adjoint-based) migration methods. In image-domain least-squares migration (i-LSM), the goal is typically to invert pre-stack seismic data —such as common shot, receiver, or offset gathers— to obtain a reflectivity model or a medium property perturbation model (e.g., velocity or impedance) relative to a known background model. In the literature, the former case is often associated with Kirchhoff modelling/migration, whilst the latter is referred to Born modelling/migration (Yang and Zhang, 2019), regardless of whether these operators rely on asymptotic approximations or band-limited wavefield modeling engines.

The primary objective of this work is to establish a framework to invert pre-stack seismic data for full-bandwidthmedium properties. This is accomplished by applying Total Variation (TV) regularisation to the medium property of interest (e.g., velocity) while inverting for its perturbation component. This approach ensures an optimal blending of the frequency content between the two components, aligning with subsurface expectations, such as layered structures with sharp discontinuities between geological units. Numerical results on both synthetic and field datasets are presented to demonstrate the effectiveness of the proposed approach.

Theory

Let us begin by deriving the Born modelling operator, which will be utilised in the numerical examples. The variable-density, variable-velocity acoustic wave equation can be written in the time domain as:

$$\mathcal{L}(p; v, Z) \coloneqq \nabla \cdot \frac{v(\boldsymbol{x})}{Z(\boldsymbol{x})} \nabla p(\boldsymbol{x}, t) + \frac{1}{Z(\boldsymbol{x}) v(\boldsymbol{x})} \frac{\partial p(\boldsymbol{x}, t)}{\partial t^2} = -\frac{v(\boldsymbol{x})}{Z(\boldsymbol{x})} s(\boldsymbol{x}, t) \tag{1}$$

where $v(\boldsymbol{x})$ is the velocity model, $Z(\boldsymbol{x}) = \rho(\boldsymbol{x}) v(\boldsymbol{x})$ is the acoustic impedance model, $p(\boldsymbol{x}, t)$ is the acoustic wavefield, $s(\boldsymbol{x}, t)$ is the source wavefield, and ∇ and $\nabla \cdot$ are the gradient and divergence operators. We define a background model $v_0(\boldsymbol{x})$ for velocity and $Z_0(\boldsymbol{x})$ for impedance such that $v(\boldsymbol{x}) = v_0(\boldsymbol{x}) + \delta v(\boldsymbol{x})$ and $Z(\boldsymbol{x}) = Z_0(\boldsymbol{x}) + \delta Z(\boldsymbol{x})$. We then subtract $\mathcal{L}(p_0; v_0, Z_0)$ from $\mathcal{L}(p; v, Z)$ and perform a first-order perturbation of the terms $1/(Z_0 + \delta Z)$ and $1/(Z_0 + \delta Z)(v_0 + \delta v)$ that results in the linearised version of the wave equation or Born modeling engine:

$$\begin{aligned} \mathcal{L}(p_0; v_0, Z_0) &\coloneqq \nabla \cdot \frac{v_0(\boldsymbol{x})}{Z_0(\boldsymbol{x})} \nabla p_0(\boldsymbol{x}, t) + \frac{1}{Z_0(\boldsymbol{x}) v_0(\boldsymbol{x})} \frac{\partial p_0(\boldsymbol{x}, t)}{\partial t^2} = -\frac{v_0(\boldsymbol{x})}{Z_0(\boldsymbol{x})} s(\boldsymbol{x}, t) \\ \delta\mathcal{L}(\delta p; v_0, Z_0) &\coloneqq \nabla \cdot \frac{v_0(\boldsymbol{x})}{Z_0(\boldsymbol{x})} \nabla \delta p(\boldsymbol{x}, t) + \frac{1}{Z_0(\boldsymbol{x}) v_0(\boldsymbol{x})} \frac{\partial \delta p(\boldsymbol{x}, t)}{\partial t^2} = \underbrace{-\frac{2}{v_0(\boldsymbol{x}) Z_0(\boldsymbol{x})} \left(cos^2\gamma \frac{\delta v(\boldsymbol{x})}{v_0(\boldsymbol{x})} + sin^2\gamma \frac{\delta Z(\boldsymbol{x})}{Z_0(\boldsymbol{x})} \right)}_{\mathcal{V}(\boldsymbol{x})} \frac{\partial p_0(\boldsymbol{x}, t)}{\partial t^2} \end{aligned} \tag{2}$$

The wave equation $\mathcal{L}(p_0; v_0, Z_0)$ in the background medium is solved to compute the incident wavefield $p_0(\boldsymbol{x}, t)$ that, once multiplied with the scattering potential $\mathcal{V}(\boldsymbol{x})$ in the right-hand side of second wave equation, leads to the secondary source of the Born modelling operator. Whilst this setting is general, in the following we will only consider the velocity perturbation δv (i.e., assuming constant density), which leads to a simplified scattering potential $\mathcal{V}(\boldsymbol{x}) = -2\delta v / v_0^3$.

Since Born modelling is a linear mapping between the perturbation $\boldsymbol{\delta v}$ and the single-scattering seismic waveform, from here onwards we use a discrete matrix-vector notation to represent such an operator:

$$\boldsymbol{\delta p} = \mathbf{L} \boldsymbol{\delta v} \tag{3}$$

where $\mathbf{L}$ is the modeling operator (that includes also any scaling in the scattering potential and a spatial sampling of the wavefield at the receiver locations to produce the recorded data $\boldsymbol{\delta p}$). Finding the perturbation that matches the observed data entails solving an inverse problem, usually referred to as data-domain least-squares migration. Alternatively, one can pre-multiply each side of equation 2 by $\mathbf{L}^{\mathrm{H}}$ to obtain:

$$\boldsymbol{\delta \tilde{v}} = \mathbf{H} \boldsymbol{\delta v} \tag{4}$$

where $\boldsymbol{\delta\tilde{\mathbf{v}}} = \mathbf{L}^{\mathrm{H}}\boldsymbol{\delta\mathbf{p}}$ represents the migrated image, and $\mathbf{H} = \mathbf{L}^{\mathrm{H}}\mathbf{L}$ is the so-called Hessian of the demigration/migration operator's chain. In industry practice, $\mathbf{H}$ is usually approximated via a non-stationary convolution operator where the filters are the impulse response or point-spread functions (PSFs) of the demigration-migration process due to Dirac-like velocity perturbations $\boldsymbol{\delta\mathbf{v}}=\boldsymbol{\delta}(\boldsymbol{x}-\boldsymbol{x_0})$ at an ensemble of arbitrarily spaced locations $\boldsymbol{x_0}$. The process of solving equation 3 for $\boldsymbol{\delta\mathbf{v}}$ is usually referred to as image-domain least-squares migration. Despite its apparent simplicity, image-domain least-squares migration is a highly ill-posed inverse problem due to the band-limited nature of seismic data and the fact that the Born modelling operator can only approximately synthesize the observed data. First, to simplify the solution of this inverse problem and streamline the selection of the regularisation parameters for the inversion algorithms used in this work, the maximum eigenvalue of $\mathbf{H}^{\mathrm{H}}\mathbf{H}$ (i.e., the Lipschitz constant of $1/2 \parallel \boldsymbol{\delta\tilde{\mathbf{v}}} - \mathbf{H}\boldsymbol{\delta\mathbf{v}} \parallel_2^2$ is computed and the problem is scaled by the square root of its inverse. Second, to improve the convergence of the inversion process, a spatially dependent weighting is computed by interpolating the central values of each PSF over the grid of interest, and applying its inverse via a diagonal (or element-wise) operator $\mathbf{K}$ to both the image-domain data and modeling operator. Moreover, given the ill-posed nature of the problem, it is of paramount importance to identify a robust, ad-hoc regularisation strategy. Whilst it may be complicated to identify a suitable regularisation for the velocity perturbation itself (given its nature as shown Figure 1e), a more natural approach is to impose a-priori knowledge on the full-bandwidth velocity model ($\mathbf{v} = \mathbf{v}_0 + \boldsymbol{\delta\mathbf{v}}$) with $\mathbf{v}_0$ representing the known background velocity model used to propagate the background wavefield p_0 in equation 2. By doing so, we can ensure that the frequency content of the background model and the inverted one seamlessly blend together in a optimal fashion (or at least optimal with respect to our prior knowledge injected by the regulariser). We therefore propose to solve the following optimization problem:

$$\min_{\boldsymbol{\delta\mathbf{v}}} \; 1/2 \parallel \mathbf{K}\boldsymbol{\delta\tilde{\mathbf{v}}} - \mathbf{K}\mathbf{H}\boldsymbol{\delta\mathbf{v}} \parallel_2^2 + \epsilon \parallel \nabla(\boldsymbol{\delta\mathbf{v}} + \mathbf{v}_0) \parallel_{2,1} \tag{5}$$

where ∇ is the gradient operator that takes a vector $\boldsymbol{\delta\mathbf{v}} \in \mathcal{R}^n$ and returns a matrix $\nabla\boldsymbol{\delta\mathbf{v}} = [\nabla_{\mathbf{x}}\boldsymbol{\delta\mathbf{v}}, \nabla_{\mathbf{z}}\boldsymbol{\delta\mathbf{v}}]^T \in \mathcal{R}^{2\times n}$ (for simplicity, we consider here the 2-dimensional case). The regularisation term represents the isotropic TV norm of the full-bandwidthvelocity model and is used to promote blocky structures (i.e., sharp discontinuities between layers). Due to the non-smoothness of this term, the overall objective function cannot be minimised using gradient-based solvers; instead, we must rely on proximal algorithms such as the Alternating Direction Method of Multipliers (ADMM) on the Primal-Dual (PD) solvers. In this work, we decide to use the latter: such choice is motivated by numerical evidence that PD outperforms ADMM in various seismic inversion tasks (e.g., Ravasi and Birnie, 2023; Romero et al., 2023). The PD solver (Chambolle and Pock, 2011) is a type of proximal solver that works by recasting any problem of this kind:

$$\min_{\boldsymbol{\delta\mathbf{v}}} \; f(\boldsymbol{\delta\mathbf{v}}) + g(\mathbf{G}\boldsymbol{\delta\mathbf{v}}) \tag{6}$$

into its primal-dual equivalent (i.e., a saddle-point problem), where f and g are convex (and possibly non-smooth) functionals and $\boldsymbol{G}$ is a linear operator. We can easily see that equation 5 is a special case of equation 6 where $f(\boldsymbol{\delta\mathbf{v}}) = 1/2 \parallel \mathbf{K}\boldsymbol{\delta\tilde{\mathbf{v}}} - \mathbf{K}\mathbf{H}\boldsymbol{\delta\mathbf{v}} \parallel_2^2$, $\boldsymbol{G} = \nabla$, and $g(\cdot) = \epsilon \parallel \cdot + \nabla\mathbf{v}_0 \parallel_{2,1}$. A series of iterations is then introduced to obtain convergence to the saddle point:

$$\begin{cases} \mathbf{Y}^{\mathrm{k}+1} = prox_{\mu g^*}(\mathbf{Y}^{\mathrm{k}} + \mu\nabla\overline{\boldsymbol{\delta\mathbf{v}}}^{\mathrm{k}}) \\ \boldsymbol{\delta\mathbf{v}}^{\mathrm{k}+1} = prox_{\tau f}(\boldsymbol{\delta\mathbf{v}}^{\mathrm{k}} - \tau\nabla^{\mathrm{H}}\mathbf{Y}^{\mathrm{k}+1}) \\ \overline{\boldsymbol{\delta\mathbf{v}}}^{\mathrm{k}+1} = 2\boldsymbol{\delta\mathbf{v}}^{\mathrm{k}+1} - \boldsymbol{\delta\mathbf{v}}^{\mathrm{k}} \end{cases} \tag{7}$$

where $\mathbf{Y} \in \mathcal{R}^{2\times n}$ is an auxiliary variable, and τ and μ represent the step-lengths of the two sub-gradients that must be chosen such that to $\tau\mu < 1/\lambda_{max}(\nabla^{\mathrm{H}}\nabla)^2$ (where $\lambda_{max}(\nabla^{\mathrm{H}}\nabla)$ equals to 8 in 2-dimensions and 12 in 3-dimensions). The proximal operator f is evaluated by inexactly solving the following optimization with a gradient-based iterative solver (e.g., LSQR):

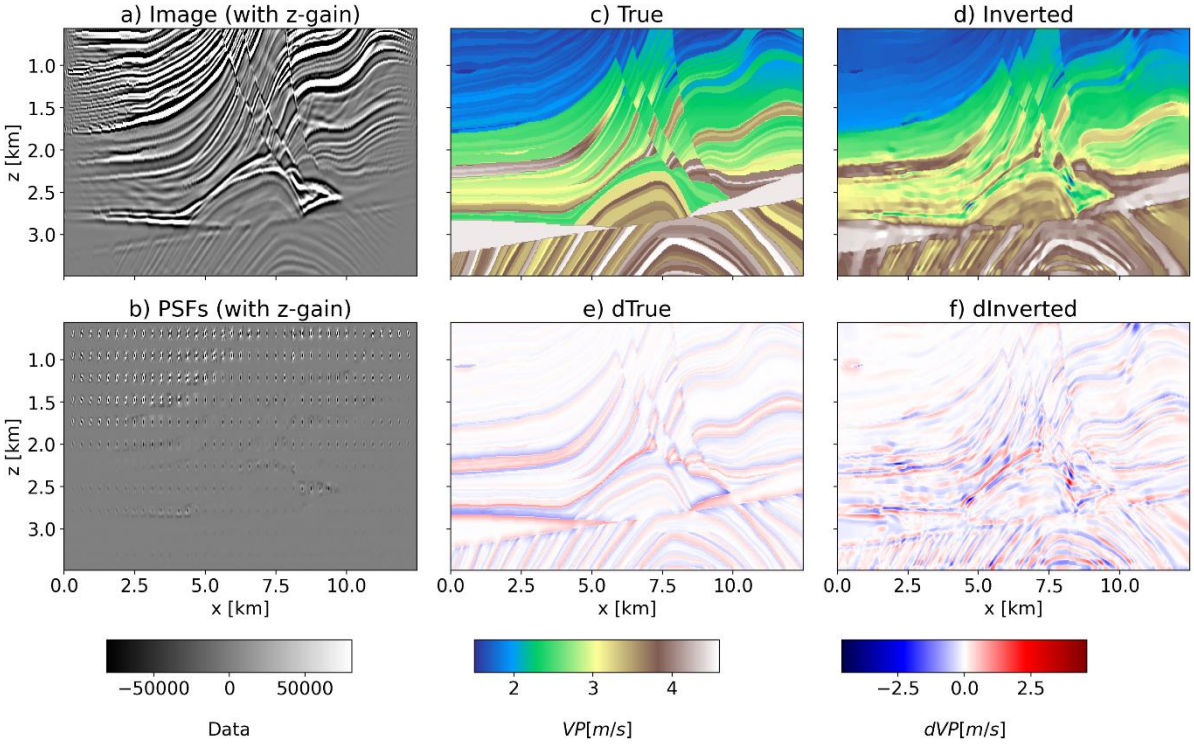

Figure 1 *Marmousi model example. a) Migrated image, b) point-spread functions, c) true velocity model, and d) TV-regularised inverted model. e-f) Corresponding velocity perturbations.*

$$prox_{\tau f}(\mathbf{x}) = \left(\mathbf{I} + \tau\mathbf{H}^{\mathrm{H}}\mathbf{K}^{\mathrm{H}}\mathbf{K}\mathbf{H}\right)^{-1}(\mathbf{x} + \tau\mathbf{H}^{\mathrm{H}}\mathbf{K}^{\mathrm{H}}\mathbf{K}\boldsymbol{\delta}\tilde{\mathbf{v}}) \quad (8)$$

To evaluate the proximal operator of g, we instead leverage the property of pre-composition of a proximal operator (Parikh, 2013), together with the observation that the $\|\cdot\|_{2,1}$ norm of a matrix is the sum of Euclidian norms of its columns (i.e., separable functions):

$$prox_{\mu\epsilon\|\cdot+\nabla\mathbf{v}_0\|_{2,1}}(\mathbf{X}) = prox_{\mu\epsilon\|\cdot\|_{2,1}}(\mathbf{X} + \nabla\mathbf{v}_0) - \nabla\mathbf{v}_0 = \left[prox_{\mu\epsilon\|\cdot\|_2}(\mathbf{X}_i + \nabla\mathbf{v}_{0,i}) - \nabla\mathbf{v}_{0,i} \text{ for } i = 1, \dots, n\right] \quad (9)$$

where $\mathbf{X_i}$ represents the i-th column of $\mathbf{X}$ and $prox_{\mu\epsilon\|\|_2}(\mathbf{x}) = (1 - \mu\epsilon/max\{||\mathbf{x}||_2, \mu\})\mathbf{x}$. Finally, the Moreau identity is used to compute the dual proximal (i.e., proximal of the convex conjugate g^*): $prox_{\mu\epsilon\|\|_2^*} = \mathbf{x} - \mu prox_{\epsilon\|\|_2/\mu}(\mathbf{x}/\mu)$.

Numerical Results

We apply the proposed i-LSM method to a synthetic dataset constructed from the Marmousi model and a 2D line of the Volve OBC field dataset. In both cases, we mimic an ocean-bottom acquisition setting and use band-limited two-way propagators in the Born migration/demigration operators. Starting from the Marmousi model, we create a dataset composed of 301 sources (41m spacing) and 101 receivers (122m spacing) with a 15Hz Ricker wavelet. To adhere to a single-scattering assumption, absorbing boundaries are added on all four sides of the model. Figure 1a displays the reverse time migrated (RTM) image obtained by applying the adjoint of the modeling operator in equation 2 with a background velocity model v_0 obtained by applying a smoothing with a 20-sample boxcar filter to the true model (Figure 1c). Note that because of the presence of the $-2/v_0^3$ scaling factor, a depth-dependent gain is applied for visualization purposes. Similarly, PSFs are modeled by placing a grid of 30 x 7 equally spaced point-scatterers (387.5m in both directions) and shown in Figure 1b. In both cases a mask is applied to remove artifacts from the water column. Figure 1d shows the velocity model obtained by solving the optimization problem in equation 5. Finally, Figures 1e-f show the velocity perturbations for the true and inverted models. We observe that our approach can produce a satisfactory velocity model, and TV regularisation contributes to recovering more details and sharper discontinuities between

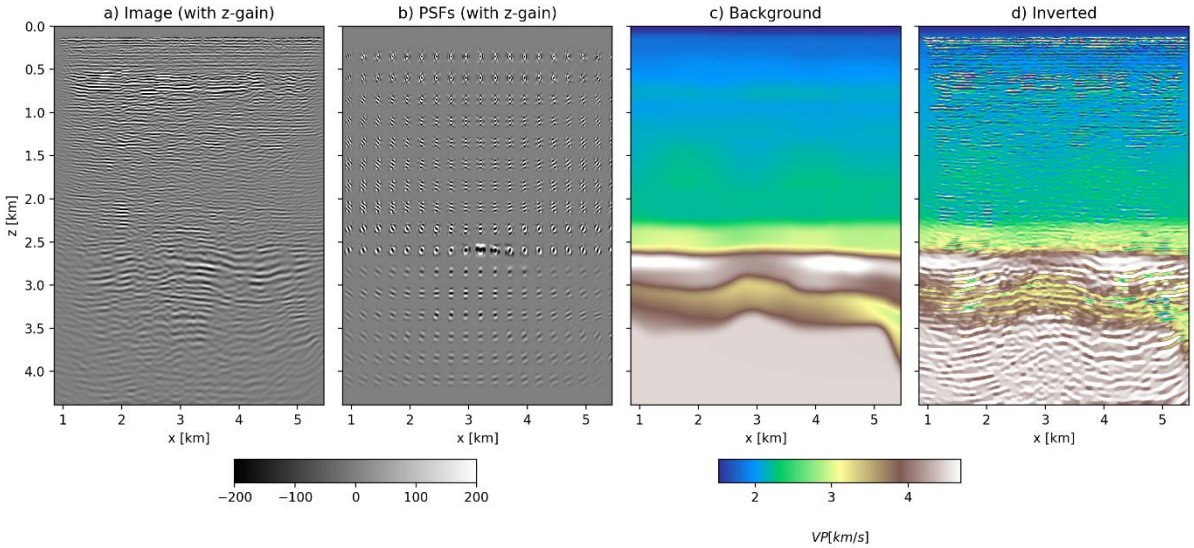

Figure 2 *Volve field data example. a) Migrated image, b) point-spread functions, c) background velocity model, and d) TV-regularised inverted model.*

layers. A similar procedure is now applied to a 2D line of the Volve dataset, which is pre-processed to remove free-surface effects — see Ravasi et al. (2022). The processed shot gathers are migrated to produce the image in Figure 2a with the background velocity model in Figure 2c. The velocity model obtained by solving equation 5 is shown in Figure 2d. Like the synthetic case, we can observe how the high-frequency component of the seismic image is satisfactorily mapped into a velocity perturbation that nicely blends in with the background model.

Conclusions

We have presented a new formulation for image-domain least-squares migration that aims to reconstruct a full-bandwidthvelocity model by combining the known background model (used to propagate the wavefields in the Born modeling operator) and the velocity perturbation in the seismic image. By applying TV-regularisation to the velocity model, we ensure optimal mixing of the frequency-content of these two components with respect to our expectations of the subsurface. This is practically achieved by formulating the i-LSM problem in a way amenable to proximal algorithms, particularly to the Primal-Dual solver. The presented framework is generic and can be easily extended to invert for alternative parametrizations (e.g., impedance plus velocity using equations 1 and 2 for modelling/migration, or elastic parameters using elastic Born modelling/migration engines); this represents an alternative to the traditional migration-plus-inversion industry-standard workflows, which fully accounts for and deconvolves the 3D effect of the seismic wavelet together with survey-design imprints/illumination, offering a more comprehensive solution beyond 1D convolutional models.

Acknowledgements

The authors thank Shearwater GeoServices for supporting this research.

References

[1] Chambolle, A. And Pock, T. [2011] A first-order primal-dual algorithm for convex problems with applications to imaging, Journal of Mathematical Imaging and Vision, 40, 120–145.
[2] Parikh, N. [2013] Proximal Algorithms, Foundations and Trends in Optimization.
[3] Ravasi, M. and Birnie, C. [2022] A Joint Inversion-Segmentation approach to Assisted Seismic Interpretation: Geophysical Journal International, 228 (2), 893-912.
[4] Ravasi, M., Selvan, T. and Luiken, N. [2022] Stochastic multi-dimensional deconvolution M Ravasi, IEEE Transactions on Geoscience and Remote Sensing 60, 1-14.
[5] Romero, J. and Luiken, N. and Ravasi, M. [2023] Seeing through the CO2 plume: Joint inversion-segmentation of the Sleipner 4D seismic data set The Leading Edge 42 (7), 457-464
[6] Yang, K. and Zhang, J. [2019] Comparison between Born and Kirchhoff operators for least-squares reverse time migration and the constraint of the propagation of the background wavefield: Geophysics, 84, R725–R739.