A fast least-squares migration with ultra-wide-band phase-space beam summation method

Ram Tuvi, Zeyu Zhao and Mrinal K Sen
Institute for Geophysics, John A. and Katherine G. Jackson School of Geosciences, The University of Texas at Austin, Austin, TX 78758, United States of America
E-mail: ram@ig.utexas.edu

Keywords: wave propagation, gaussian beams, least-squares migration, inverse theory

Abstract
Least-squares migration (LSM) problems are often formulated as iterative schemes. At each iteration, traditional LSM methods require solving the wave equation several times to compute the gradient and single scattering (Born) predicted data, which makes LSM computationally expensive. We use a beam approach to overcome this problem, based on the ultra-wide-band phase space beam summation (UWB-PS-BS) method. We use the beams to expand and propagate the measured data. The beams calculation is performed via a small number of Green’s functions (GFs). These GFs are calculated once and then stored in the computer’s disk memory. Thus, wave equation calculations are avoided within each iteration. We use a beam transformation of shot gathers which can be considered as the data obtained by a pair of source and receiver beams. This data has a physical interpretation. The beam amplitudes extract the medium reflectivity generated from a localized region in subspace at a specific angle (local Snell’s law). This data calculation is performed before the iterative scheme. Thus, we may a-priori threshold beams with low amplitude beams, as they contain less essential information. Using the beam propagators spectral locality, we may further reduce the number of beams used to construct the image. In this work, we establish a new local beam-based data domain LSM in the frequency domain (FD). We use the properties described above to reduce the computational complexity of the LSM problem. We demonstrate the advantages of the proposed LSM algorithm via numerical examples.

1. Introduction
Migration techniques are important tools to recover subsurface structures from observed seismic data [1]. These formulations approximate the adjoint of the forward modeling operation. An image obtained by the migration operation contains errors due to inaccuracy of the estimated velocity model, data limitation, and noise [2]. The resulting image can be verified by comparing the predicted data obtained from the migrated image, to the observed data. The error caused by the migration operation can be reduced by formulating the migration problem as a least-squares inverse problem [3, 4]. Given a sufficiently accurate velocity model, this problem is often formulated as least-squares migration (LSM) [5].

LSM approaches are linearized inversion methods which minimize the misfit between the observed data and estimated data. Early implementation of LSM has been demonstrated in [5], using crosswell data. In [2], this approach has been applied to surface data, where Kirchhoff operators have been used to propagate the wavefield. It has been shown that the use of the LSM method improves the resolution, and reduces migration artifacts. The method has then been extended to a one-way wave equation [6, 7]. To enhance the image quality, two-way wave equation and reverse-time LSM methods have been explored [6, 8, 9–11].

Generally, LSM can be implemented in either image domain or data domain. The image domain LSM is based on explicitly constructing the Hessian matrix [12–15]. The LSM result is obtained by applying the inverse Hessian on a pre-computed migration image (also known as the gradient of the misfit function). Constructing
the full Hessian matrix is computationally expensive for realistic problems. Data domain LSM utilizes iterative schemes to reduce the misfit function, where the explicit Hessian calculation is avoided. In this paper, we will employ the data domain LSM scheme.

In data domain LSM formulations, \( L_2 \) norm misfit function is often minimized by utilizing gradient-based optimization methods. The data domain LSM approach comprises of two steps: a migration part where the gradient of the misfit function is calculated via an adjoint-state method, and a Born modeling (demigration) part that generates estimated data with Born linear forward mapping [12]. These two steps make the implementation of LSM computationally expensive. In a time domain LSM implementation, wavefield propagations are necessary to compute Born data and the gradient. The number of wavefield propagations can be large when there is a large number of shot records. The computational cost of time domain LSM formulations often greatly increases with an increase in the amount of data. To reduce the computational cost, phase encoding formulations have been suggested in [16–20].

Frequency domain (FD) LSM methods have been studied by [21–26]. It has been shown that only a subset of the entire frequency components is needed to successfully recover subsurface images. The use of a frequency-dependent grid has been proposed in [27]. Using this approach the computational complexity of the LSM problem is reduced. In [25] it has been shown that the Born data and the gradient can be obtained by using pre-computed source and receiver Green’s functions (GFs) that are stored in the computer’s disk or memory. While this scheme works well for small datasets, implementing the scheme for large datasets becomes infeasible since the number of needed GFs would be tremendously large. Therefore, the adjoint-state method is often implemented for shot domain frequency LSM. In [26], a FD double plane wave LSM method has been proposed to improve the computational efficiency of the FD LSM problem. This method is based on the double plane wave transform of [28, 29] (see also [30] and [31]) for the problem of reverse time migration. They have shown that only a small number of plane wave GFs are needed to perform LSM, which substantially reduces the wavefield calculations in the iterative updating process. The method works well for large datasets as well.

Beam summation (BS) methods play an important role in the study of wave propagation. In these formulations, the propagated field is represented as a superposition of beams emanating from a source, propagating along certain directions in space. The beams can be tracked locally in an inhomogeneous medium. In addition, these formulations are localized in space, since only beams that pass near a subsurface point need to be taken into account in the field calculation [32]. Due to these advantages beam-based migration techniques have been successfully applied [33–38]. To increase the image resolution and reduce the computational cost of LSM formulations, beam-based LSM approaches have been suggested in [39–42]. These beam-based formulations are generally based on point source expansions methods [43–45], where the source and receiver GFs are expanded using superposition of beams. Due to the spectral locality of beam propagators, only beams that contribute to an imaging point need to be used in the LSM, which makes it efficient for target-oriented LSM approaches.

There is another category of beam summation method based on extended (distributed) sources 1. In the extended-source BS formulations, the field is expressed as a discrete superposition of shifted and tilted beams that emanate from a discrete set of points with initial ray parameters over the surface (see [46], figure 1(b)). The method has been structured initially upon a Gabor-series expansion of surface sources [47–49]. It has then been extended by using Gabor frames in the context of radiation from surface field distribution [50–53] and from volume sources [54, 55]. For the seismic survey, the field is measured over a wide aperture. For this reason, we use the extended sources approach. Specifically, we use the ultra-wide-band phase space beam summation (UWB-PS-BS) method [51]. In this method, one utilizes windowed Fourier transform (WFT) frame decomposition of a surface field distribution to propagate the surface field. The frame elements are Gaussian windows, centered in the spatial domain around surface points, and the spectral domain, around a set of ray parameters. The propagated field is given as a superposition of weighted-beams, which emanate from the surface. In this theory, the beam amplitudes are calculated by projecting the surface field distribution on a set of dual frame functions. Due to the frame functions’ structure, the projection operation extracts the local radiation spectrum (windowed Fourier transform) of the surface field near a surface point and specific direction (ray-parameter). In [46] the UWB-PS-BS method has been used for seismic migration applications. The image is given as the sum of the source and receiver pair of beams and beam transformation of the shot gathers. Using the physical interpretation of the beam amplitudes (local radiation spectrum), it has been shown that the number of beams used for migration can be \textit{a priori} reduced, and the migration method can be used to image specific targets over the subsurface.

In this research, we formulate the UWB-PS-BS migration method into LSM scheme. In our approach the beam propagators are solutions for the wave equation. They are calculated using numerical integration with

---

1 A detailed comparison between both of these approaches has been carried out in [46]. Here for simplicity, we only summarize the key points of these methods.
subsurface GFs. The number of GFs needed to construct the beam propagators is relatively small. One can pre-compute FD GFs and store those in the computer’s disk or even memory. Thus, in the iterative updating process, solving the two-way wave equation is no longer necessary, pre-computed GFs can be repeatedly used to construct the beam propagators and hence update images. Comparing with shot domain LSM methods, we significantly reduce the computational cost of the LSM implementation by eliminating the time-consuming wave-equation solving process. Additionally, since original shot gathers are decomposed into spatial and spectral localized beam domain data, one may filter out low amplitude beams that correspond to insignificant scattering data. With reduced input beam data, we should obtain an image close to the one received by migrating all the beam data but with reduced computational effort. Comparing with traditional shot domain LSM methods, the proposed beam-based LSM greatly improves the computational efficiency.

2. Theory of least-squares migration

2.1. Least-squares migration—Data domain

We begin with a brief review of data domain LSM. Under the Born approximation, the seismic data, \( \hat{d} \), collected at the surface is related to the subsurface model parameter \( m_r \) via a propagation operator \( \hat{L} \),

\[
\hat{d} = \hat{L} m_r.
\]

In the LSM inverse problem, the goal is to estimate the model parameter, reflectivity, from the data. In general, this can be obtained by,

\[
m_r = (\hat{L}^\dagger \hat{L})^{-1} \hat{L}^\dagger \hat{d},
\]

with the superscript \( \dagger \) being the conjugate transpose operator.

In large-scale LSM, since directly constructing \( (\hat{L}^\dagger \hat{L})^{-1} \) is computationally infeasible, we often formulate the inverse problem as an iterative scheme. We can iteratively minimize the difference between observed data and predicted data via misfit or objective function

\[
F(m_r) = (1/2) \Delta \hat{d}^\dagger \Delta \hat{d},
\]

with \( F \) is the misfit function, the superscript \( \dagger \) represents the conjugate transpose operator, and the superscript \( \ast \) denotes FD quantities. In equation (3), \( \Delta \hat{d} = \hat{d}^{\text{obs}} - \hat{d}^{\text{pre}} \) is the difference between the observed (measured) data \( \hat{d}^{\text{obs}} \) and the predicted data \( \hat{d}^{\text{pre}} \) calculated from the model parameter \( m_r \), via equation (1). In the data domain LSM problem, we iteratively update \( m_r \), via gradient-based optimization method with \( \nabla_m F(m_r) = g \) represents the gradient of the misfit function. The best set of parameters is obtained when the misfit function reaches its minimum.

2.2. Frequency domain data under the Born approximation

Considering a 3D coordinate frame \( r = (x, z) = (x_1, x_2, z) \) with velocity \( v(r) \), the FD acoustic wave equation due to a point source illumination at \( x = x_s \) with frequency signature of \( \hat{f} \) is given by
\[(\nabla^2 + \omega^2/v^2(r))\hat{u}(r) = -\hat{f}(x - x_s).\]  
(4)

Here \(\omega\) is the angular frequency, \(\hat{u}\) is the propagated field and \(\delta\) denotes the Kronecker delta function.

Given a sufficiently accurate background velocity model, \(v_B\), the velocity and field can be written as \(v = v_B + \Delta_v\) and \(\hat{u}(r) = \hat{u}_B(r) + \hat{u}^*(r)\). Here \(\Delta_v\) is the velocity perturbation while \(\hat{u}_B(r)\) and \(\hat{u}^*(r)\) are the background and scattered fields, respectively. The background field is given by

\[\hat{u}_B(r, x_i) = \hat{f} \hat{G}_B(r, x_i),\]
(5)

where \(\hat{G}_B\) is the background GF. Since \(|\Delta_v| \ll |v_B|\), we have \(|\hat{u}^*(r)| \ll |\hat{u}_B(r)|\). For this case, the scattering data collected at \(x = x_i\) can be written as

\[\hat{d}^i(x_i, x_i) = \hat{u}^*(r, x_i)|_{r = (x_i, 0)} \simeq \omega^2 \int d^3r \ m_r(r) \hat{G}_B(r, x_i) \hat{G}_B^*(r, x_i),\]
(6)

with \(m_r(r) = -2\delta_r/v_0^2(r)\). Thus, the weak scattering approximation of the propagation operator \(\hat{L}\) of equation (1) is given by

\[\hat{L}(r) = \hat{f} \omega^2 \hat{G}_B(r; x_i) \hat{G}_B^*(r; x_i).\]
(7)

The gradient of the misfit function is given by [56]

\[\nabla_m F(m_r) = g = \text{Re} (\hat{L}^* \Delta \hat{d}),\]
(8)

with \(\nabla_m = [\partial_{m_1} \partial_{m_2} \ldots \partial_{m_n}]\) being the gradient of the misfit function with respect to the medium reflectivity. Note that the gradient \(g\) is in the length of number of model points \(r\). Writing it explicitly in terms of equation (7),

\[g(r) = \text{Re} \left( \sum \sum \sum \omega^2 \hat{f} \hat{G}_B^*(r; x_i) \hat{G}_B(r; x_i) \Delta \hat{d} \right).\]
(9)

Typically a preconditioner which is usually the approximation of the Hessian matrix is applied to the gradient to improve the convergence of the iterative scheme. Following [57], the approximation of the main diagonal is chosen as

\[k(r) = \text{Re} \left( \sum \omega \hat{f} \left| \sum \hat{G}_B(r; x_i) \right|^2 \sum \left| \hat{G}_B(r; x_i) \right|^2 \right)^{-1} \].
(10)

This preconditioner is referred to as the illumination compensation.

3. Data domain least-squares migration using WFT frame expansions

In this section, we introduce our beam-based LSM method. The formulation is based on the UWB-PS-BS method of [51]. The main details for UWB-PS-BS theory are summarized in appendix.

3.1. Beam representation of the scattered data

Given measured or predicted shot gather data \(\hat{d}_0(x_s, x_i)\), the beam domain representation of the data is given by [46], equation (22)

\[\hat{A}_{\mu, \mu'} = \int d^2x_s \ \hat{\varphi}_{\mu'}(x_s) \int d^2x_r \ \hat{\varphi}_{\mu}^*(x_r) \hat{d}_0(x_r, x_s).\]
(11)

Here \(\mu = (x_m, p_m)\) and \(\mu' = (x_m', p_m')\) are the phase-space indices of equation (A.1), which corresponds to the source and receiver arrays, respectively.

The Born approximation of the scattering data is obtained by first multiplying both sides of (6) by \(\hat{\varphi}_\mu(x_s)\) and \(\hat{\varphi}_\mu^*(x_s)\), and then changing the order of integration,

\[\hat{A}_{\mu, \mu'} = \hat{f} \omega^2 \int d^3r m_r(r) \int d^2x_s \ \hat{\varphi}_\mu(x_s) \hat{G}_B(r; x_s) \int d^2x_r \ \hat{\varphi}_{\mu'}^*(x_r) \hat{G}_B^*(r; x_r),\]
(12a)

\[= \hat{f} \frac{\nu^4}{||\hat{\psi}(x)||^4} \omega^2 \int d^3r m_r(r) \int d^2x_s \ \hat{\varphi}_\mu(x_s) \hat{G}_B(r; x_s) \int d^2x_r \ \hat{\varphi}_{\mu'}^*(x_r) \hat{G}_B^*(r; x_r),\]
(12b)

\[= \hat{f} \frac{\nu^4}{||\hat{\psi}(x)||^4} \omega^2 \int d^3r m_r(r) \hat{\psi}_{\mu'}(r) \hat{\psi}_{\mu'}^*(r),\]
(12c)

\[= \hat{f} \frac{\nu^4}{||\hat{\psi}(x)||^4} \omega^2 \int d^3r m_r(r) \hat{\psi}_{\mu'}^*(r).\]
(12d)

Equation (12b) is obtained by using the frame relation of (A.7). In equation (12c) \(\hat{\psi}_{\mu'}(r)\) are the beam propagators of equation (A.5), while \(\hat{\psi}_{\mu'}^*(r)\) are the backpropagated beams, obtained by the same expression with
replacing $\hat{\psi}_\mu$ by its conjugate complex. The final expression in equation (12d) is obtained by using the imaging kernels of [46], equation (21), $\hat{\Lambda}_{\mu,\mu'} = \hat{\Psi}_\mu^* \hat{\Psi}_{\mu'}^b$.

The operation in equation (11) projects the shot gathers data on localized spatial-spectral window functions. Following equation (A.6) this operation identified as WFT of the data near the phase-space region $(x_m - x_m', p_n - p_n')$. The Born approximation of this data is given by equation (12). Following the discussion after [46], equation (22) the beam representation of the data has an analytic model. It is given by the multiplication of the imaging kernels and the reflectivity of the medium. Each beam propagator propagates along a certain direction in space, and decays away from it. The multiplication $\hat{\Psi}_\mu(r) \hat{\Psi}_{\mu'}^b(r)$ is localized around a scattering cell (see also [46], figure 2). In addition, since each propagator propagates along a certain direction defined by the initial ray-parameter, each imaging kernel is oriented along the direction defined by the incident and scattered beams ray parameters. It follows from this relation that the Born approximation of the beam domain data is given by the local reflectivity of the subsurface, in an angle defined by the incident and backpropagated beams ray parameters.

3.2. Phase-space representation of the data misfit and the adjoint operator

In this work, we use a conjugate gradient algorithm to update the model. To use the conjugate gradient method, we first define a normalized phase-space data $d^{ps} = \hat{\Lambda}_{\mu,\mu'}^s \frac{\| \hat{\psi}(x) \|^2}{\mu^2}$ given by

$$d^{ps} = \left\| \hat{\psi}(x) \right\|^2 \int d^2x_r \hat{\varphi}_\mu(x_s) \int d^2x_r \hat{\varphi}_{\mu'}^b(x_r) \hat{d}_0(x_r, x_s) = \hat{D}_{\mu,\mu'} \hat{d}_0(x_r, x_s),$$

Figure 2. Marmousi model. (a) Image generated by the standard beam-based migration of [46]. (b) The image generated by the proposed beam-based LSM 30 iterations. (c) Convergence of the misfit function. (d) Comparison over line. The result obtained by the migration algorithm is normalized to compare it with other results.
with
\[ \hat{D}_{\mu, \mu'} = \frac{||\hat{\psi}(x)||^2}{\nu^2} \hat{\psi}_{\mu'}(x) \hat{\psi}_{\mu'}(x). \]  

(14)

With this normalization, the Born data of equation (12) can be written as
\[ d^{\mu'} = \hat{D}_{\mu, \mu'} \hat{L} \hat{m} = \hat{L}_{\mu, \mu'} \hat{m}, \quad \hat{L}_{\mu, \mu'} = \int \hat{\psi}(x) \hat{\psi}_{\mu'}(r) \, dr, \]

(15)

while the migrated image is now given by applying the adjoint of the \( \hat{L}_{\mu, \mu'} \) operator on the normalized data.

The gradient is obtained by applying \( \hat{L} \) on the beam transformation of the data misfit \( \Delta \hat{d}_{\mu, \mu'} \)
\[ \Delta \hat{d}_{\mu, \mu'} = \hat{D}_{\mu, \mu'} \Delta \hat{d}(x, x_s) \]
\[ g(r) = \hat{L}_{\mu, \mu'} \Delta \hat{d}_{\mu, \mu'}. \]

To minimize the misfit function, we apply the preconditioned conjugate gradient method with the preconditioner defined by
\[ k_{\mu, \mu'}(r) = \left( \frac{\int \hat{\psi}^2 \, d^2 \omega \sum_{\mu, \mu'} |\hat{\psi}_{\mu}(r)|^2 |\hat{\psi}_{\mu'}(r)|^2}{\nu^2} \right)^{-1}. \]

(17)

The representation of equations (16)–(17) is directly in the beam domain. The beam propagators are calculated via equation (A.5) using pre-computed GFs, and no wave propagation is needed at each iteration. Given a subsurface point, only pairs of beams that highly overlap near this point are taken into account in the propagation operator. In addition, only the beams with high amplitudes in equation (16) will contribute to the propagation operator. Thus we reduce the computational complexity of the LSM problem at each iteration. Next, we demonstrate the properties of the algorithm through numerical examples.

4. Numerical example

In this section we demonstrate the proposed LSM algorithm with the 2D Marmousi model.

4.1. The physical configuration

The model contains 460 and 150 grid points in the horizontal and vertical directions, respectively, with 0.02 km grid spacing in both directions. The velocity model is plotted in figure 1(a). We subtract a slightly smoothed velocity model from the true velocity model to obtain the reflectivity model plotted in figure 1(b).

4.2. Choosing the beam parameters

The first step in selecting the beam parameters is to set the beam frequency independent parameter, \( b \). As demonstrated in [51], this parameter is chosen such that we have collimated beam propagators through the propagation domain, while also maintaining \( \omega_{\text{min}} b / v_0 \ll 1 \). Following [46] we set \( b = 20 \) km which leads to collimated beams for the subsurface. Note that the beam width is determined by [34]. We use the frequency band \( f = [5, 25] \) Hz, with 0.25 Hz frequency interval for imaging. Following equation (A.10) we obtain the phase space lattice dimensions \( \epsilon \approx 0.6 \) km, \( \bar{\epsilon} = 0.02 \) s/km. We also obtain 16 beam initiation points with 43 propagation directions, giving the ray-parameter range \( p_x \approx [-0.42, 0.42] \) s/km. As illustrated below, we do not have to use all the beams for the proposed LSM to achieve good results.

4.3. Image results

4.3.1. Imaging with full data set and all GFs

We begin by exploring the proposed LSM algorithm using Born data as the observed data, which is calculated using equation (15). For forward modeling, we use a Ricker wavelet with central frequency \( f = 15 \) Hz. We use the frequency range \( f = [5, 25] \) Hz, with \( \Delta f = 0.25 \) Hz. The beam propagators are calculated using equation (A.5). We use here numerical integration with pre-computed GFs. We use the frequency domain finite difference method to generate 92 GFs, with a horizontal spacing of 0.1 km. Note that the GFs spacing does not satisfy the horizontal spacing Nyquist sampling rate over the surface for frequency \( f > 7.5 \) Hz. This relatively small number of GFs allows us to calculate the GFs only once and store them in the computer’s memory. Hence no wave propagation is needed at each iteration.

The image obtained by the standard beam-based migration algorithm of [46] is plotted in figure 2(a). As can be seen, most of the subsurface is reconstructed. The sides of the image, especially at the bottom parts are not clearly recovered. In figure 2(b) we plot the image obtained by the proposed LSM algorithm after 30 iterations.
As can be seen, migration artifacts are reduced, and the image resolution is substantially improved. Figure 2(c) shows the convergence of the misfit function. We also plot the reflectivity profiles at $x = 4.6$ km in Figure 2(d), to get a better comparison. We also plot the migrated image of [46] over this line. The proposed LSM algorithm successfully retrieves the true reflectivity value for most parts of the model.

In Figure 3 we compare the input data and the estimated data obtained by the estimated reflectivity model after 30 iterations. We plot the data obtained by a source located at $x = 4.6$ km. For clarity all plots are scaled with the same colorbar. The data obtained by the proposed method show a good agreement with the input data.

4.3.2. Imaging with a reduced number of GFs

The number of GFs used for constructing the beam propagators can be reduced. In Figure 4(a) and 4(c) we plot the migrated image and LSM image obtained using 46 GFs, with a spacing 0.2 km. As can be seen, the iterative LSM scheme helps to suppress migration artifacts and balance the reflectivity value for the entire model even though with a small number of GFs. In Figure 4(b) we plot the image obtained by the beam-based migration algorithm, where we use 23 GFs to construct the beam propagators. For this case, the GFs spacing over the surface is $\Delta x = 0.4$ km while the Nyquist sampling rate is 0.15 km for $f = f_{\text{min}}$. Therefore, the image obtained by the beam-based migration algorithm produces an image with poor resolution and spatial artifacts. In Figure 4(d), we plot the image obtained by our LSM algorithm. Using the iterative scheme we manage to reconstruct most parts of the subsurface. The spatial resolution is greatly improved with a very small number of GFs. On the other hand, the image quality with 23 GFs is much degraded compared with that of 46 GFs mainly due to the fact that the Nyquist theorem is severely violated. In practice, one might want to use a reasonable size of GFs to reduce the computation cost in the meantime have most of the frequency components satisfy the Nyquist theorem.

4.3.3. Imaging with sparse data

The use of the UWB-PS-BS method also allows us to a-priori reduce the computational cost by using only the data that significantly contribute to the subsurface image. Following the discussion after equation (12), the beam
domain data is governed by the subsurface local Snell’s law reflection. Thus, we may neglect imaging kernels with low amplitudes as they do not contain significant scattering events. In [46] this relation was extended for the problem of scattering under the Born approximation, where it was shown that a pair of beams with high amplitudes corresponds to reflection events over the subsurface (see discussion after equation (12) and [46], equation (20)). Here, we use this property to reduce the number of elements that construct the propagation operator.

In figure 5(a) we plot the absolute value of the beam domain data generated with equation (15) at frequency \( f = 10 \) Hz. The data is plotted in the source and receiver beams domain. Each axis is arranged as a concatenated vector with each initiation point \( x_m \) and all the propagation directions \( p_n \). The data is tagged in colors from the brightest to the darkest. The data set is sparse, and we can use only the pair of beams with high amplitudes, as explained above, to carry out the proposed LSM algorithm. Note that the beam amplitudes calculation is performed over the surface field, as a pre-processing phase, before the beam propagators computation. Thus, one may a-priori reduce the number of elements used to construct the propagation operator in the LSM scheme.

We use this property in the imaging phase. For each frequency we threshold the highest 20% beam amplitudes, and use only the imaging kernels corresponding to these \( \mathbf{\mu}, \mathbf{\mu}' \) values to construct the propagation operator. For comparison, we plot the corresponding data at \( f = 10 \) Hz in figure 5(b). One can see that the two data sets agree with each other. The images obtained by the sparse data set are plotted in figure 6(c). In figure 6(a) we plot the image obtained by the beam-based migration of [46], using 46 GFs. (b) Image obtained by the standard beam-based migration of [46], using 23 GFs. (c) LSM image obtained by using 46 GFs. (d) LSM image obtained by using 23 GFs.

Figure 4. Image results with a reduced number of GFs. (a) Image obtained by the standard beam-based migration of [46], using 46 GFs. (b) Image obtained by the standard beam-based migration of [46], using 23 GFs. (c) LSM image obtained by using 46 GFs. (d) LSM image obtained by using 23 GFs.
that is comparable with that obtained with the full data set, while reducing the computational complexity at each iteration.

As above, we compare the estimated and input data in figure 7 for a source located at $x = 4.6$ km. For clarity all plots are scaled with the same colorbar. Note that although we used a sparse data set the final result matches the input data.

**Figure 5.** Marmousi model—Beam domain data set. The figure shows the source-receiver beam data corresponds to the amplitudes of equation (15) for frequency $f = 10$ Hz. The magnitudes are tagged by colors, from the brightest to the darkest. (a) The full data set, (b) Sparse data set. As can be seen, the data set is spares. Only dominant coefficients should be used to reconstruct the image.

**Figure 6.** Marmousi model—images generated by sparse data set. (a) Image generated by the standard beam-based migration of [46]. (b) The image generated by the proposed beam-based LSM 30 iterations. (c) Comparison over line. We also add the LSM result obtained from the full data set in order to compare the results. (d) The final reflectivity model difference—full and sparse data.
4.3.4. Imaging with a full wavefield synthetic data

After exploring the algorithm via a Born data set, we explore the LSM algorithm via a full wavefield synthetic data set. The data set is generated using the time domain rapid expansion method [58]. The physical configuration is the same as the Born data set.

There are 460 shots, with 460 receivers for each shot. Shot-profile methods require solving and storing 460 GFs at each frequency. Here, as above, we use only 92 GFs to construct the beams.

The image results are plotted in figure 8. As can be seen, the proposed LSM algorithm substantially improves the subsurface image resolution, image amplitudes are balanced from the shallow to the deep part of the model. The dipping events at the bottom of the model, which are not visible in the regular migration image, can be clearly identified in the LSM result. Note that the full wave data doesn’t fully obey the Born approximation. A different propagation operator is used to form the full wave data than for imaging. For this reason we see less accurate results than the image in figure 2(c).

To increase the accuracy of the full wavefield synthetic data LSM image, we use 460 GFs to construct the beam propagators. The image is plotted in figure 9(a). For comparison we also added the data obtained from a source located at $x = 4.6$ km. As can be seen, the two data sets agree.

5. Discussion

We introduced here a data domain LSM algorithm using beam propagators. The proposed approach is based on the migration operation of [46]. With this migration algorithm, we use WFT frame expansions to transform the LSM problem directly to the beam domain. We establish the data transformation (Born modeling) and migration (image reconstruction) operators and show that they are conjugate operators. This property allows us to use a gradient scheme to solve the LSM problem.

Data domain beam-based LSM algorithms are iterative schemes that may be implemented in the time-domain or FD. In a traditional shot-profile time-domain LSM formulation, the adjoint-state method is used to
compute the gradient. These time-domain implementations require solving the wave equation twice for every shot—one for the forward and one for the backward propagation. Denoting by $n$ as the number of iterations, $n_s$ the number of shots, the number of wave equations that needs to be solved is $2 \times n_s \times n$.

The current implementation is formulated using FD. In FD LSM implementations, we may calculate the gradient in two ways. The first is to use the adjoint-state method. In this approach, we multiply the inverse of the left hand side of Helmoltz equation to the right hand side twice for each shot over the entire frequency range. When the number of shots and iterations is large, the multiplication of the inverse of the left hand side to the

Figure 8. LSM image using synthetic data. Here we use 92 GFs to construct the beam propagators. (a) Image generated by the standard beam-based migration of [46]. (b) The image generated by the proposed beam-based LSM.

Figure 9. LSM image using synthetic data. Here we use 460 GFs to construct the beam propagators (a) The image generated by the proposed beam-based LSM. (b) Input data. (c) Estimated Born data using the final reflectivity model of figure 9(a).
right hand side can becomes time-consuming. An alternative method is to use the scattering integral approach. There, we save GFs for all source and receiver in the computers memory or disk. Constructing the forward and backward wavefields repeatedly at each iteration is avoided. However, this implementation is storage-consuming and thus limited, especially when the number of sources and receivers is large, storing the GFs for every shot and receiver becomes prohibitive.

In the proposed method, we only need to compute a certain small number of GFs even with large numbers of shots and receivers. The GFs are used to numerically construct the beam propagators (see equation (A.5)). The number of GFs is determined to obtain a sufficient sampling interval over the surface. This sampling interval is generally determined by the Nyquist sampling interval ($\Delta_s < \lambda/2$). Note that this number only gives a bound to the number of GFs that have to be used. As shown in the numerical examples, we can use fewer GFs to obtain the image. The number of GFs is relatively small. Note that the same GFs are used to represent the source and receiver beams. We use here the same GFs throughout iterations. Thus there is no need to solve the wave equation or construct the forward and backward wavefield in each iteration. These points greatly reduce the computational cost of the LSM implementation.

The proposed method is different from the conventional GB LSM algorithms discussed in the introduction. Most of these formulations are based on the localized sources approach [33, 34]. These methods are based on expanding each point source GF as a superposition of GBs. In the localized sources GB approaches, one constructs the beam propagators using a paraxial wave equation near the beam axis. These are ray-based methods that require solving a Riccati equation along the axis. On the other hand, as solutions to the wave equation, beams are insensitive to caustics and reduce multipathing. Our approach is based on a full wavefield solution, which improves the accuracy of wave propagation, hence improve the accuracy for the reflectivity estimation.

The use of the WFT expansions allows to a-priori reduce the number of propagators that construct the propagation operator $\hat{L}$ in the LSM problem. With the WFT expansions, we decompose the shot gathers into a beam to beam scattering amplitudes (beam domain data). Only pairs of beams with high amplitudes have to be taken into account when constructing $\hat{L}$. This filtering step is performed as a pre-processing phase before the iterative scheme, and a priori reduces the number of beam cross-correlation operations used to form the image. Comparing to localized source GB methods, the steps described above reduce the computational complexity of the iterative scheme.

The UWB-PS-BS utilizes an overcomplete, frequency independent, beam lattice. This use allows us to transform the theory directly to the time-domain [59], while maintaining the physical properties of the beam amplitudes. On the other hand, due to the overcompleteness of the frame, the UWB-PS-BS method suffers from low overcompleteness at low frequencies (see [51] and also comment after (A.3)). This problem arises mainly in applications with a very wide frequency range (see [60]). Note that similar results for the oversampling criteria of the beams over the surface obtained in [33]. Thus the same overcompleteness arises in formulations based on the latter theory. To overcome this problem we use here a sparse data set, and reduce the number of imaging kernels used for imaging (see section 4.3.3). Alternatively, one may reduce the overcompleteness at low frequencies by using the multiband UWB-PS-BS algorithm introduced in [60]. In this approach, the frequency band is divided into sub-bands, and the UWB-PS-BS algorithm is implemented in each sub-band. As a result, we obtain fewer beam elements at the lower frequency bands while keeping the overcompleteness in all the bands below a given level.

6. Conclusions

We introduced here a new approach of beam-based data domain LSM using the UWB-PS-BS method. This method is based on WFT frame expansion of a surface field. The beams are used to decompose the surface field and then to propagate it. In the current implementation, we use this method to reduce the computational complexity of the data-domain LSM problem.

In the proposed method, the beams are calculated using the medium GFs. We use here a relatively small number of GFs to construct the beams, which can be calculated only once and store at the computer memory. Thus, no wave propagation needs to be performed at each iteration. Compared to conventional shot domain LSM implementations, we reduce the computational complexity of the problem. For a given subsurface point, only beams passing near it need to be considered. Our approach also allows us to further reduce the computational complexity of the LSM problem, in contrary to the localized sources beam-based LSM formulations. The beam amplitudes correspond to scattering events over the subsurface. Thus, we may filter out pairs of beams with relatively low magnitudes, which reduces the data size and the number of wave fields used to reconstruct the subsurface, while obtaining a good agreement with the true model. Since the beam amplitudes are calculated before the migration phase, this approach a-priori reduces the computational complexity of the
LSM implementation. Finally, since the beam propagators are calculated via GFs obtained by solving the two-way wave equation, and not solving ray-based equation, we also enhance the image accuracy comparing with ray-based LSM methods. These unique properties are utilized here to improve the computational efficiency for the LSM problem.

Data availability statement

The data generated and/or analysed during the current study are not publicly available for legal/ethical reasons but are available from the corresponding author on reasonable request.

Appendix A. The ultra-wide-band phase-space beam summation method

In this work we are apply a beam approach for data-domain LSM, utilizing the UWB-PS-BS method. The key features of the method are summarized below. Further details of the method, with application to seismic wave propagation are given in [46].

A.1. Windowed Fourier transform (WFT) frame expansion

The UWB-PS-BS method has been introduced in [51], in the context of time harmonic radiation of aperture (surface) field distribution to \( z > 0 \) half space, with frequency band \( \Omega = [\omega_{\text{min}}, \omega_{\text{max}}] \).

The method is structured on WFT frame expansion of aperture field distribution. The frame set defined over the \( z = 0 \) surface is given by

\[
\hat{\psi}_\mu(x) = \hat{\psi}(x - x_m)e^{-i\pi p_n^x (x - x_m)},
\mu = (m, n) = ((m_1, m_2), (n_1, n_2))
\]

with \( \hat{\psi}(x) \) is localized window function, typically obtained by multiplying 2 1D window functions. The frame elements are localized around the phase-space grid of points, defined by the 4-index \( \mu \) via (see also [46], figure 1)

\[
x_m = mx, p_n = n\beta.
\]

In order that the set in equation (A.1) to constitute a frame, it is required that the grid spacing provides an overcomplete coverage of the phase-space, i.e.,

\[
\omega \beta \beta = 2\pi \nu, \quad \nu \leq 1.
\]

Here \( \nu \) being the overcompleteness parameter. Note that the case \( \nu = 1 \), corresponds to the Gabor limit, where the frame is critically complete. For UWB applications we choose \( \nu = \nu_{\text{max}} \omega / \omega_{\text{max}} \), so that the lattice is frequency independent.

The WFT frame can be used to construct the propagated field into the half space \( z > 0 \), due to surface field distribution, \( u_0(x) \) over the \( z = 0 \) plane,

\[
\hat{u}(r) = \sum_\mu \hat{a}_\mu \hat{\psi}_\mu(r).
\]

Here \( \hat{\psi}_\mu(r) \) are the beam propagators. They are given by propagating each frame element into the half space \( z > 0 \),

\[
\hat{\psi}_\mu(r) = \int d^2r' \hat{\psi}_\mu(r') 2\partial_{r'} \hat{G}(r, r'),
\]

with \( \hat{G} \) being the medium GF and \( r' = (x', 0) \) represents a coordinate along the \( z = 0 \) surface. In equation (A.4) \( \hat{a}_\mu \) is the set of beam amplitudes. Since the set \( \{\hat{\psi}_\mu(x)\} \) is overcomplete, the set of coefficients is not unique. Following [51], a unique set of coefficients is given by using the dual frame set \( \{\hat{\psi}_\mu(x)\} \). This set has the same structure as \( \{\hat{\psi}_\mu(x)\} \), with replacing the window functions \( \hat{\psi}(x) \) by \( \hat{\varphi}(x) \). This particular set of coefficients is given by

\[
\hat{a}_\mu = \langle \hat{u}_0(x), \hat{\psi}_\mu(x) \rangle,
\]

with \( \langle \cdot, \cdot \rangle \) denotes the \( L_2 \) inner product. The operation in equation (A.6) projects \( \hat{u}_0(x) \) on window functions which are localized around the points \( x_m, p_n \). This operation extracts the radiation spectrum of this field distribution around the point \( x_m \) in the direction defined \( p_n \).

Generally \( \hat{\varphi} \) should be calculated numerically for each frequency in the frequency band. For the case of highly overcomplete regime (\( \nu < 0.3 \)), \( \hat{\varphi} \) can be evaluated directly from the frames via

\[
\hat{\varphi}(x) \approx (\nu |\hat{\psi}(x)|^2 \hat{\psi}(x).
\]
which simplifies the calculations. This choice also leads to stable and localized coefficients ([51]). Equation (A.4) represents the radiated field as superposition of beam propagators, emanate from the points \( \mathbf{x} = \mathbf{x}_m \), and propagate along a direction defined by the initial ray parameter \( \mathbf{p} = \mathbf{p}_n \) (see [46], figure 1), and decay away from it. The direction of propagation also known as the beam axis. The beam amplitudes are obtained by projecting the aperture field distribution on the dual frame functions. Thus, they may be interpreted as extracting the local radiation spectrum of \( \mathbf{x}_m \) around the point \( \mathbf{x}_m \) in the direction \( \mathbf{p}_n \).

The representation in equation (A.4) is a priori sparse both in the propagators and amplitudes. Giving a point over the subsurface only the beams which pass near this point have to take into account in the summation. As discussed after equation (A.6), \( \Delta \) extracts the local spectrum of \( \mathbf{x}_m \)(x). Thus, we may filter out beams with low amplitudes, which reduces the number of beams that need to be taken into account in the summation.

A.2. Special case: Iso-diffracting Gaussian windows

As was shown in [32, 51], a favorable choice for the function is the use iso-diffracting Gaussian windows. These are windows where their width is scaled with frequency by the frequency independent parameter \( b \). These windows are given by

\[
\hat{\psi}_{ID}^{0}(x) = e^{-\omega/\omega_0|x|^2/b}, \quad \omega > 0, \tag{A.8}\]

For this case the dual window function is given by

\[
\hat{\tilde{\phi}}(x) \simeq \left(\frac{\nu_{\text{max}}^2}{\pi\nu_0 b\omega_{\text{max}}}\right) \omega^3 \hat{\psi}(x). \tag{A.9}\]

The resulting beam propagators propagation parameters are frequency independent and need to be calculated only once in the frequency band. In addition, by choosing \( b = \sqrt{\frac{\nu_0}{\omega_0}} \) we obtain the snuggest frame for all frequencies. Combining this result with equation (A.3), the phase-space unit cell dimensions are given by

\[
\left(\hat{x}, \hat{p}\right) = \sqrt{\frac{2\pi\nu_{\text{max}}^2}{\omega_{\text{max}}}} \left(\sqrt{\nu_0 b}, 1/\sqrt{\nu_0 b}\right), \tag{A.10}\]

thus one obtains a frequency-independent for all frequencies. These properties simplify the calculations and also very efficient for UWB applications, and enable to transform the method to the time domain directly.

**References**

[1] Claerbout J F 1976 *Fundamentals of Geophysical Data Processing* (New York, NY: McGraw-Hill Book Co.)
[2] Wu C, Nemet T and Schuster G T 1999 Least-squares migration of incomplete reflection data *Geophysics* 64 208–21
[3] Lailly P 1983 The seismic inverse problem as a sequence of before stack migrations *Conference on Inverse Scattering, Theory and Application* (Society for Industrial and Applied Mathematics, Expanded Abstracts) pp 206–20
[4] Tarantola A 1984 Inversion of seismic reflection data in the acoustic approximation *Geophysics* 49 1259–66
[5] Schuster G T 1993 Least-squares cross-well migration. In SEG Technical Program Expanded Abstracts 1993 Society of Exploration Geophysicists 63rd Annual International Meeting 110–113
[6] Wang J and Sacchi M D 2007 High-resolution wave-equation amplitude-variation-with-ray-parameter (avp) imaging with sparseness constraints *Geophysics* 72 S11–S8
[7] Henning K and Sacchi M D 2003 Least-squares wave-equation migration for avp/ava inversion *Geophysics* 68 262–73
[8] Valenciano A, Biondo A and Guitton A 2006 Target-oriented wave-equation inversion *Geophysics* 71 A35–A8
[9] Dong S, Cai J, Guo M, Suh S, Zhang Z, Wang B and Li E Z 2012 Least-squares reverse time migration: Towards true amplitude imaging and improving the resolution. In SEG technical program expanded abstracts 2012 Society of Exploration Geophysicists 83rd Annual International Meeting 1425–9
[10] Luo S and Hale D 2014 Least-squares migration in the presence of velocity errors *Geophysics* 79 S153–61
[11] Zhang Y, Duan L and Xie Y 2015 A stable and practical implementation of least-squares reverse time migration *Geophysics* 80 V23–31
[12] Plessix R E 2006 A review of the adjoint-state method for computing the gradient of a functional with geophysical applications *Geophys. J. Int.* 167 495–503
[13] Jianhua Y, Hu J, Schuster G T and Estill R 2006 Prestack migration deconvolution *Geophysics* 71 S53–S62
[14] Aoki N and Schuster G T 2009 Fast least-squares migration with a deblurring filter *Geophysics* 74 WCA83–WCA93
[15] Ren H, Wu R S and Wang H 2011 Wave equation least square imaging using the local angular hessian for amplitude correction *Geophysics* 76 651–61
[16] Zhang Y, Sun J, Notfors C, Gray S H, Chernis L and Young J 2005 Delayed-shot 3D depth migration *Geophysics* 70 E21–8
[17] Liu F, Hanson D W, Whitmore N D, Day R S and Stolt R H 2006 Toward a unified analysis for source plane-wave migration *Geophysics* 71 S129–39
[18] Gao F, Atle A and Williamson P 2010 Full waveform inversion using deterministic source encoding. In SEG Technical Program Expanded Abstracts 2010 Society of Exploration Geophysicists 29 1013–7
[19] Schuster G T, Wang X, Huang Y, Dai W and Boonyasiriwat C 2011 Theory of multisource crosstalk reduction by phase-encoded statics *Geophys. J. Int.* 184 1289–1303
[20] Dai W, Wang X and Schuster G T 2011 Least-squares migration of multisource data with a deblurring filter *Geophysics* 76 R135–46
[21] Pratt R G 1999 Seismic waveform inversion in the frequency domain; part 1, theory and verification in a physical scale model Geophysics 64 888–901
[22] Plessix R E, Mulder W A and Pratt R G 2001 Frequency-domain finite-difference migration with only few frequencies? In SEG Technical Program Expanded Abstracts 2001 Society of Exploration Geophysicists 21 1103–6
[23] Sigurgeirsson L and Pratt R G 2004 Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies Geophysics 69 231–48
[24] Mulder W A and Plessix R E 2004 How to choose a subset of frequencies in frequency-domain finite-difference migration Geophys. J. Int. 158 801–12
[25] Ren H, Wang H and Chen S 2013 Least-squares reverse time migration in frequency domain using the adjoint-state method J. Geophys. Eng. 10 035002
[26] Zhao Z and Sen M K 2019 Frequency-domain double-plane-wave least-squares reverse time migration Geophys. Prospect. 67 2061–84
[27] Operto S, Hustedt B and Viret J 2004 Mixed-grid and staggered-grid finite-difference methods for frequency-domain acoustic wave modelling Geophys. J. Int. 157 1269–96
[28] Zhao Z, Sen M K and Stofka P L 2016 Double-plane-wave reverse time migration in the frequency domain Geophysics 81 S367–82
[29] Zhao Z, Sen M K and Stofka P L 2017 Double plane-wave reverse-time migration Geophys. Prospect. 65 1541–58
[30] Sen M K and Fraser L N 1991 Multifold phase space path integral synthetic seismograms Geophys. J. Int. 104 679–87
[31] Stofka P L, Sen M K, Seifoulaee R K, Pestana R C and Fokkema J T 2006 Plane-wave depth migration Geophysics 71 S261–72
[32] Heyman E and Felsen L B 2001 Gaussian beam and pulsed beam dynamics: complex-source and complex spectrum formulations within and beyond paraxial asymptotics J. Acoust. Soc. Am. 118 1588–611
[33] Hill N R 1990 Gaussian beam migration Geophysics 55 1416–28
[34] Hill N R 2001 Prestack gaussian-beam depth migration Geophysics 66 1240–50
[35] Nowack R L 2011 Dynamically focused gaussian beams for seismic imaging International Journal of Geophysics 2011 316581
[36] Gray S H and Bleistein N 2009 True-amplitude gaussian-beam migration Geophysics 74 S11–23
[37] Bleistein N and Gray S H 2010 Amplitude calculations for 3D Gaussian beam migration using complex-valued traveltimes Inverse Prob. 26 085017
[38] Popov M M, Semtchenok N M, Popov P M and Verde A R 2010 Depth migration by the gaussian beam summation method Geophysics 75 S81–93
[39] Hu H, Liu Y, Zheng Y, Liu X and Lu H 2016 Least-squares gaussian beam migration Geophysics 81 S587–100
[40] Yuan M, Huang J, Liao W and Jiang F 2017 Least-squares gaussian beam migration J. Geophys. Eng. 14 184–96
[41] Yang J, Zhu H, McMechan G and Yue Y 2018 Time–domain least-squares migration using the gaussian beam summation method Geophys. J. Int. 214 548–72
[42] Zhong R, Huang J P, Zhuang S B and Li Z C 2019 Target–oriented gaussian beam migration using a modified ray tracing scheme Pet. Sci. 16 1301–19
[43] Popov M M 1982 A new method of computation of wave fields using Gaussian beams Wave Motion 4 85–97
[44] Cerveny V 1982 Expansion of a plane wave into gaussian beams Stud. Geophys. Geol. 26 120–31
[45] Cerveny V, Popov M M and Pienčík I 1982 Computation of wave fields in inhomogeneous media Gaussian beam approach Geophys. J. Int. 70 109–28
[46] Tuvi R, Zhao Z and Sen M K 2020 frequency beam-based migration in inhomogeneous media using windowed Fourrier transform frames Geophys. J. Int. 08 1086–99
[47] Bastiaans M J 1984 Signal description by means of a local frequency spectrum. In Transformations in optical signal processing International Society for Optics and Photonics 47 49–62
[48] Steinberg B Z, Heyman E and Felsen L B 2001 Gaussian beam and pulsed beam dynamics: complex-source and complex spectrum formulations Geophys. J. Int. 157 1269–96
[49] Bastiaans M J, Heyman E and Felsen L B 2001 Gaussian beam and pulsed beam dynamics: complex-source and complex spectrum formulations Geophys. J. Int. 157 1269–96
[50] Lugaro D and Letrou C 1998 Alternative to gabor's representation of plane aperture radiation Electron. Lett. 34 2286–7
[51] Shvltinski A, Heyman E, Boag A and Letrou C 2004 phase–beam summation formulation for ultra wideband radiation IEEE Trans. Antennas Propag. 52 2042–56
[52] Melamed T 2009 Exact gaussian beam expansion of time–harmonic electromagnetic waves J. Electromagn. Waves Appl. 23 975–86
[53] Melamed T 2011 T e and T m beam decomposition of time–harmonic electromagnetic waves J. Opt. Soc. Am. A 28 401–9
[54] Leibovich M and Heyman E 2017 Beam summation theory for waves in fluctuating media, part i: The beam frame and the beam-domain scattering matrix IEEE Trans. Antennas Propag. 65 5431–42
[55] Tuvi R, Heyman E and Melamed T 2019 Beam frame representation for ultrawideband radiation from volume source distributions: Frequency–domain and time–domain formulations IEEE Trans. Antennas Propag. 67 1010–24
[56] Pratt R G, Changsuo S and Hick G J 1998 Gauss–newton and full newton methods in frequency–space seismic waveform inversion Geophys. J. Int. 133 541–62
[57] Plessix R E and Mulder W A 2004 Frequency–domain finite–difference amplitude–preserving migration Geophys. J. Int. 157 975–87
[58] Pestana R C and Stofka P L 2010 Time evolution of the wave equation using rapid expansion method Geophysics 75 T121–31
[59] Shvltinski A, Heyman E and Boag A 2005 A pulsed beam summation formulation for short pulse radiation based on windowed Radon transform (WRT) frames IEEE Transactions on Antennas and Propagation 53 3030–48
[60] Shvltinski A, Heyman E and Boag A 2005 A phase–space beam summation formulation for ultra wideband radiation–part ii: A multi–band scheme IEEE Trans. Antennas Propagation 53 948–57