Robust seismic imaging by full-waveform inversion with model time extension: time-domain and frequency-domain formulations

Biondo Biondi[1,*], Guillaume Barnier,[1], Rustam Akmadhiev[1], Robert G. Clapp[1], Ettore Biondi[1,2]

[1]Geophysics Department, Stanford University, Stanford, USA
[2]Seismological Laboratory, Caltech, Pasadena, USA

*Email: biondo@stanford.edu

Abstract

Seismic full waveform inversion is formulated as non-convex optimization problem that converges to useful solutions only when the starting velocity is close to the true velocity and/or the data contain unrealistic low frequencies. A time-extension of the velocity model leads to waveform inversion algorithms that are more robust than the conventional full waveform inversion algorithms. Waveform inversion with time extension can be formulated in either the time domain or the temporal-frequency domain. The time-domain algorithm has been successfully tested with several datasets. A simple numerical example illustrates the characteristics of the proposed method and provides intuition on the convergence properties of the long wavelengths of the velocity model.

Keywords: seismic, waveform inversion, non-convex optimization

1 Introduction

Since its introduction in the early 80s [12, 20] full waveform inversion (FWI) has been appealing for the simplicity of its formulation. FWI is defined as the search of the velocity model that minimizes the differences between the data recorded in the field and data modeled by using a waveform modeling operator. FWI practical applications have been hampered by three main challenges: 1) computational cost, 2) data quality (e.g. poor spatial sampling), and 3) unreliable convergence to a useful model when the starting model is inaccurate and the data do not contain unrealistic low frequencies. Modern computational and data-acquisition technology have mostly solved the first two of these challenges. Progress in ray-based velocity estimation methods and progress in low-frequency and long-offsets data acquisition have reduced the practical impact of the third and enabled the successful application of FWI to many datasets. However, there are many datasets acquired over complex geology that still defy modern FWI algorithms because of lack of convergence.

Several examples of algorithmic solutions of the convergence problem have been proposed; for some useful algorithmic solutions see [4, 9, 14,21]. This paper presents two methods, which are closely related to each other, to overcome the convergence challenges of conventional FWI by using the concept of velocity-model extension. The idea of using velocity-model extensions to estimate the long wavelengths of the velocity model is rooted in the concept of prestack images. The goal of conventional velocity estimation is to focus prestack images obtained by seismic migration [6]. In a waveform inversion framework, prestack images are estimates of the short-wavelength component of an extended velocity model. Adding to the data-fitting goal of FWI the additional goal of focusing prestack images leads to waveform inversion methods that can robustly update the long wavelength of the velocity model [5, 15–18].

Symes [19] generalized the velocity-model extension idea beyond extended images to a velocity space-extension that includes all scales of the model. Biondi and Almomin [7] showed that extending the model in time yields a practical algorithm for robustly estimating all the wavelengths in the model. However, their optimization algorithm requires the tuning of several hyperparameters. Barnier et al. [6] introduced a more robust and general algorithm based on variable projection [11] that can be applied to both space and time extension of the model.

In this paper we focus on the time extension of the model and present two waveform-inversion methods. The first one is based on the time-domain solution of the acoustic wave equation and the second one on the temporal-frequency domain solution. The time-domain method is the time-extension instance of the

more general FWI with model extension (FWIME) method presented in [1–3].

2 Full waveform inversion with time extension (FWITE)

Conventional full-waveform inversion (FWI) is performed by solving the following optimization problem

$$\min_{\mathbf{s^2}} J\left(\mathbf{s^2}\right),$$

where:

$$J\left(\mathbf{s^2}\right) = \frac{1}{2}\left\|\mathcal{L}\left(\mathbf{s^2},\mathbf{f}\right) - \mathbf{d}\right\|_2^2, \qquad (1)$$

and $\mathbf{s^2}$ is a vector of gridded slowness-squared values, $\mathcal{L}$ is a wave-equation operator whose solutions are nonlinear with respect to slowness perturbations but linear with respect to the source function, f, which is function of time, t. The data vector $\mathbf{d}$ is a subset of the pressure-field vector $\mathbf{p}$ that is defined on the same spatial grid as $\mathbf{s}$ and on a discretized time. The data are extracted from the pressure field at the receivers' locations through a linear sampling operator $\mathbf{K}$; such as $\mathbf{d} = \mathbf{Kp}$.

For the sake of simplicity, we assume that $\mathcal{L}$ is the acoustic and isotropic wave-equation operator. A generalization of the concepts presented in this paper to problems that require more complex wave-equation operators such as elastic and/or anisotropic is possible in principle, although it is not straightforward.

To define full waveform inversion with time extension (FWITE) we introduce a new "extended-model" wave operator $\tilde{\mathcal{L}}$ defined by the extended acoustic wave equation:

$$\nabla^2 p\left(t,x\right) - \tilde{s}^2\left(\tau,x\right) \overset{\tau,t}{*} \partial_t^2 p\left(t,x\right) = f\left(t\right). \quad (2)$$

In equation 2 the symbol $\overset{\tau,t}{*}$ signifies convolution in time between the pressure field and the time-extended slowness-squared model, $\tilde{s}^2$. When $\tilde{s}$ id different from zero only at $\tau = 0$ equation 2 reduces to the conventional wave equation that describes the physical phenomenon of wave propagation in an acoustic medium.

We define FWITE as the minimization of the following objective function,

$$\tilde{J}\left(\tilde{\mathbf{s}}^{\mathbf{2}}\right) = \frac{1}{2}\left\|\tilde{\mathcal{L}}\left(\tilde{\mathbf{s}}^{\mathbf{2}}\right) - \mathbf{d}\right\|_2^2 + \epsilon\left\|\tau\tilde{\mathbf{s}}^{\mathbf{2}}\right\|_2^2. \qquad (3)$$

We introduce the second term in 3 to constrain the solution to be "physical"; that is, to satisfy the conventional wave-equation without extension, and ϵ is a trade-off parameter between the two terms in the objective function.

The formulation of FWITE as the minimization of the objective function 3 presents several theoretical and practical problems. The most fundamental challenge concerns the stability of the solution of the extended-model wave equation in 2. In the next section we present practical solution of these challenges based on a first-order Born linearization of 2 that leads to a robust, though computationally expensive, variable-projection inversion algorithm.

Frequency-domain time extension

To avoid the hurdles of the time-domain formulation of FWITE and to reduce the computational cost of the method, we are developing a frequency-domain based on the one-way approximation of the Helmholtz equation. Claerbout [10] introduced one-way wave propagation in the early 1970s. Since then, the seismic exploration community has developed many efficient numerical schemes for one-way wave propagation in the frequency domain. The use of the one-way wave equation for waveform inversion is more recent [4]. We believe that our frequency-domain method could be generalized to solutions of the the full Helmholtz equation, which is more accurate, but also substantially more expensive to solve, in particular in three dimensions.

FWITE can be defined in the temporal frequency, ω, domain by extending a complex valued $\tilde{s}_\omega^2$ along ω. In that case the convolution in equation 2 becomes a simple multiplication and the weight by τ in the second term of the objective function 3 becomes a derivative with respect to frequency.

FWITE in the frequency domain minimizes the following objective function,

$$\tilde{J}\left(\tilde{\mathbf{s}}_\omega^{\mathbf{2}}\right) = \frac{1}{2}\left\|\tilde{\mathcal{L}}_\omega\left(\tilde{\mathbf{s}}_\omega^{\mathbf{2}}\right) - \mathbf{d}_\omega\right\|_2^2 + \epsilon\left\|i\partial_\omega\tilde{\mathbf{s}}_\omega^{\mathbf{2}}\right\|_2^2, \quad (4)$$

where $\tilde{\mathcal{L}}_\omega$ is a frequency-domain wave operator based on the one-way wave equation and $\mathbf{d}_\omega$ is the recorded data in the frequency domain. The stability of the one-way modeling operator can be ensured by imposing the constraint $\Im\left(\tilde{\mathbf{s}}_\omega^{\mathbf{2}}\right) \leq 0$.

3 FWITE algorithms

In this section we discuss two different optimization algorithms for solving the time-domain and the temporal-frequency domain formulations of FWITE.

Variable-projection optimization for time-domain FWITE

As mentioned in the previous section, an important challenge of minimizing the objective function 3 is computing numerical solutions to equation 2. Because of the large dimension of the computational grid needed in reflection seismology, practical algorithms for solving the wave equation in time domain are based on explicit finite differences; implicit finite-differences algorithms are too computationally intensive in 3D for being practical. Explicit finite-differences solutions of equation 2 would be unstable.

The solution to this problem is to approximate the wave operator $\tilde{\mathcal{L}}$ with its first-order Born linearization, $\tilde{\mathbf{L}}$, around a background physical (i.e. not extended) slowness model $\mathbf{s_b^2}$; that is, using the following approximation

$$\tilde{\mathcal{L}}\left(\mathbf{s_b^2}+\mathbf{\Delta\tilde{s}^2}\right)\approx\mathcal{L}\left(\mathbf{s_b^2}\right)+\tilde{\mathbf{L}}\left(\mathbf{s_b^2}\right)\mathbf{\Delta\tilde{s}^2}. \quad (5)$$

The objective function 3 then becomes

$$\begin{aligned}\tilde{J}_l\left(\mathbf{s_b^2},\mathbf{\Delta\tilde{s}^2}\right) &= \frac{1}{2}\left\|\mathcal{L}\left(\mathbf{s_b^2}\right)+\tilde{\mathbf{L}}\left(\mathbf{s_b^2}\right)\mathbf{\Delta\tilde{s}^2}-\mathbf{d}\right\|_2^2 \\ &+ \epsilon\left\|\left(|\tau|+\alpha\right)\mathbf{\Delta\tilde{s}^2}\right\|_2^2. \quad (6)\end{aligned}$$

This new objective function is then minimized by applying a variable projection scheme [11]. Notice the addition of the small positive scalar α in the objective function 6. It ensures that the Hessian of the variable projection step is a positive-definite matrix.

At each iteration of the optimization, we first fix $\mathbf{s_b^2}$ and iteratively solve the quadratic problem in $\mathbf{\Delta\tilde{s}^2}$ by applying a conjugate gradient method to estimate an optimal $\mathbf{\Delta\tilde{s}}_o^2$. A conjugate-gradient solution requires applications of $\tilde{\mathbf{L}}$ and of its adjoint $\tilde{\mathbf{L}}^*$ to vectors in the slowness space and data space, respectively. These matrix-vector products are computed using algorithms based on adjoint-state methods that are similar to the ones used for conventional FWI. Given the solution $\mathbf{\Delta\tilde{s}}_o^2$ of the variable-projection step, we then perform a single step of L-BFGS to update the background model $\mathbf{s_b^2}$. We stop the iterative process when the differences between the data modeled using the new background model and the recorded data are sufficiently small.

The analysis of the structure of the gradient for the outer iterations illuminates the nature of the FWITE process and provides intuition on the way that the methods overcome the convergence challenges of conventional FWI [1–3]. The gradient is the sum of two different components and can be written as:

$$\nabla\tilde{J}_l=\left(\mathbf{T}^*+\mathbf{L}^*\right)\left(\mathcal{L}\left(\mathbf{s_b^2}\right)+\tilde{\mathbf{L}}\left(\mathbf{s_b^2}\right)\mathbf{\Delta\tilde{s}}_o^2-\mathbf{d}\right), \quad (7)$$

where $\mathbf{L}^*$ is the adjoint of the first-order Born linearization of $\mathcal{L}$, and $\mathbf{T}^*$ is the adjoin of a data-domain tomographic operator that connects the long-wavelength of the slowness model to data perturbations [7]. When $\mathbf{s_b^2}$ is far from the "true" model and $\mathbf{\Delta\tilde{s}}_o^2$ is far from being focused around $\tau=0$, the first term drives the long-wavelengths of the model towards the correct solution. In contrast, when $\mathbf{s_b^2}$ is close to the "true" model and $\mathbf{\Delta\tilde{s}}_o^2$ is well focused around $\tau=0$, the contributions of the tomographic term to the gradient are small, and the iterations approximate standard FWI iterations.

Constrained optimization for frequency domain FWITE

The frequency-domain objective function 4 can be directly minimized using a gradient-based optimization scheme for complex variables, such as an L-BFGS optimization scheme. To avoid instability in the computation of the forward operator $\tilde{\mathcal{L}}_\omega$, we ensure that the condition $\Im\left(\tilde{\mathbf{s}}_\omega\right)\leq 0$ is always fulfilled by projecting the gradients onto the feasible subspace.

4 Numerical example

The time-domain algorithm outlined in the previous section has successfully been tested on several challenging synthetic examples as well as on a 3D field dataset [3]. In this section we show one of the examples from Barnier [3] that is an archetypal example in reflection seismology [13]. It has the advantage of simplicity and thus it lends itself to illustrating some of the salient characteristic of the method.

Figure 1 shows the 2D velocity model assumed for numerical modeling a reflection dataset solving a constant-density acoustic wave equation. The velocity in the circle in the middle is substantially lower (2.25 km/s) than the back-

ground velocity (2.7 km/s). The source had the unrealistic frequency range between 20 and 50 Hz to ensure that conventional multi-scale FWI [9] fails to retrieve the true model. Indeed, when the starting model is homogeneous and set to the background velocity of 2.7 km/s, conventional FWI employing the well-known frequency bootstrapping procedure that starts with the low frequency to improve convergence [9]. produces the model shown in Figure 2. In contrast, a multi-scale FWITE algorithm [2] yields the accurate model shown in Figure 3, after two scale refining steps. The graphs shown in Figures 4 and 5 compare the result of FWI (magenta lines) and FWITE (blue lines) at the constant fixed horizontal location of 2 km and at the fixed depth of 0.6 km, respectively. The FWITE result is close to the true model (black lines) notwithstanding the starting model was far from accurate and the missing low frequencies in the data.

Examining the tomographic component of the first outer-loop iteration gradient (equation 7) provides some intuition on the reason why FWITE converges to an excellent model even when the starting model is grossly inaccurate and the data miss the low frequencies. Figure 6 shows the tomographic component (first term in equation 7) of the search direction (opposite sign of the gradient). The tomographic component is already moving the long-wavelengths of the model in the right direction of decreasing the velocity in the low-velocity anomaly in the middle. The gradient tomographic component does not cycle skip as it would the conventional FWI gradient because the data residuals that are back-projected by equation 7 have been corrected by the addition of the term $\tilde{\mathbf{L}}\mathbf{\Delta}\tilde{\mathbf{s}}_o^2$.

The nature of the long-wavelength contribution of the tomographic component is further illustrated by the complex modulus of the Fourier transform of the search direction shown in Figure 7. Comparing this wavenumber-domain spectrum with the similar spectrum of the difference between true and starting models shown in Figure 8, we can see how the tomographic component starts to fill in the low wavenumbers, in particular around the horizontal direction, of the velocity anomaly missing from the starting model. These results are consistent with the generalization of the classical wavenumber-domain analysis of first-order scattering [22] by Biondi et al. [8] and summarized by the diagram shown in Figure 9. The area shaded in orange in the Figure corresponds to the model wavenumber components that are illuminated by second order scattering when the data are recorded with source and receivers located at the surface.

5 Conclusions

Full waveform inversion with model time extension can be a powerful algorithmic solution to the convergence problems of conventional full waveform inversion. The time-domain version of the method is solved with a variable projection algorithm after a modification of the objective function based on a Born linearization of the modeling operator with respect to the extended model. The frequency-domain formulation holds the promise of enabling a direct solution of the optimization problem formulated with non-linear extended modeling operator.

6 Acknowledgements

We would like to thank the Stanford Exploration Project affiliate companies for financial support and the Stanford Center for Computational Earth and Environmental Sciences (CEES) for providing computational resources.

References

[1] G. Barnier, E. Biondi and B. Biondi, Full waveform inversion by model extension, *Proceedings of SEG Annual Meeting, Anaheim, USA*, 2018, pp. 1183-1187.

[2] G. Barnier and E. Biondi, Full waveform inversion by model extension using a model-space multi-scale approach, *Proceedings of SEG Annual Meeting, Houston, USA*, 2020, pp. 646-650

[3] G. Barnier, *Full waveform inversion by model extension: a robust method to estimate the seismic propagation velocity in the subsurface from recorded seismograms*, Ph.D. Thesis, Stanford University, Stanford, 2022.

[4] A.J. Berkhout, Review Paper: An outlook on the future of seismic imaging, Part III: Joint Migration Inversion, *Geophysical Prospecting* **62** (2014), 950-971.

[5] B. Biondi and P. Sava, Wave-equation migration velocity analysis, *Proceedings of SEG Annual Meeting, Houston, USA*, 1999, pp. 1723-1726.

[6] B. Biondi, *3-D Seismic Imaging*, 2nd edition, Society of Exploration Geophysicists, Tulsa, 2012.

[7] B. Biondi, and A. Almomin, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion, *Geophysics* **79** (2014), pp. WA129-WA140.

[8] B. Biondi, G. Barnier and E. Biondi, Can full waveform inversion image all scales of the velocity model? *Stanford Exploration Project report* **170** (2017), pp. 111-122.

[9] C. Bunks, F.M. Saleck, S. Zaleski, and G. Chavent, Multiscale seismic waveform inversion, *Geophysics* **60** (1995), pp. 1457-1473.

[10] J.F. Claerbout, Coarse grid calculations in inhomogeneous media with applications to delineation of complicated seismic structures, *Geophysics* **35** (1970), pp. 407-418.

[11] G.H. Golub and V. Pereyra, The differentiation of pseudo-inverses and nonlinear least-squares problems whose variables separate, *SIAM Journal on numerical analysis* **10** (1973), pp.413-432.

[12] P. Lailly, The seismic inverse problem as a sequence of before stack migration, in *Proceedings of SIAM Conference on Inverse Scattering, Theory and Applications, Philadelphia, USA*, 1983.

[13] P. Mora, Inversion = migration + tomography, *Geophysics* **54** (1989), pp. 1575-1586.

[14] L. Métivier, R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux, An optimal transport approach for seismic tomography: Application to 3D full waveform inversion, *Inverse Problems* **32** (2016), 115008.

[15] P. Sava and B. Biondi, Wave-equation migration velocity analysis – I: Theory, *Geophysical Prospecting* **52** (2004), 593-606.

[16] P. Shen and W.W. Symes, Automatic velocity analysis via shot profile migration, *Geophysics* **73** (2009), pp. VE49-VE59.

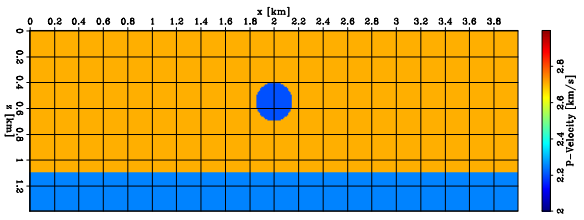

Figure 1: Velocity model assumed to compute synthetic seismograms for test dataset.

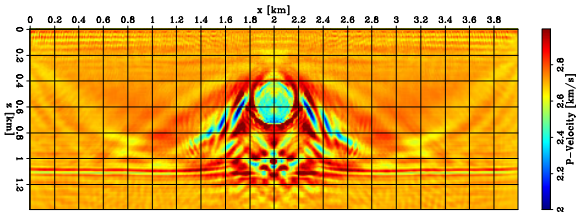

Figure 2: Velocity model produced by conventional FWI employing a frequency bootstrap algorithm to improve convergence [9].

[17] W. Symes and J.J. Carazzone, Velocity inversion by differential semblance optimization, *Geophysics* **56** (1991), pp. 654-663.

[18] W.W. Symes and M. Kern, Inversion of refection seismograms by differential semblance analysis: Algorithm structure and synthetic examples, *Geophysical Prospecting* **42** (1994), pp. 565-614.

[19] W.W. Symes, Migration velocity analysis and waveform inversion: *Geophysical Prospecting* **56** (2008), pp. 765-790.

[20] A. Tarantola, Inversion of seismic reflection data in the acoustic approximation, *Geophysics* **49** (1983), pp. 1259-1266.

[21] M. Warner and L. Guasch, Adaptive waveform inversion: Theory, *Geophysics* **81** (2016), pp. R429-R445.

[22] R. Wu and M.N. Toksoz, 1987, Diffraction tomography and multisource holography applied to seismic imaging, *Geophysics* **52** (1987), pp. 11-25.

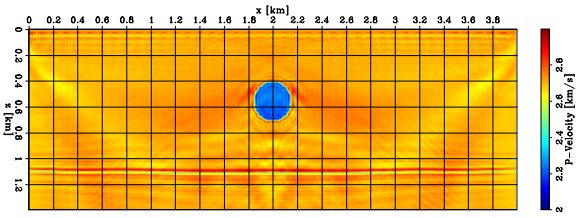

Figure 3: Velocity model produced by the time-domain FWITE algorithm presented in this paper.

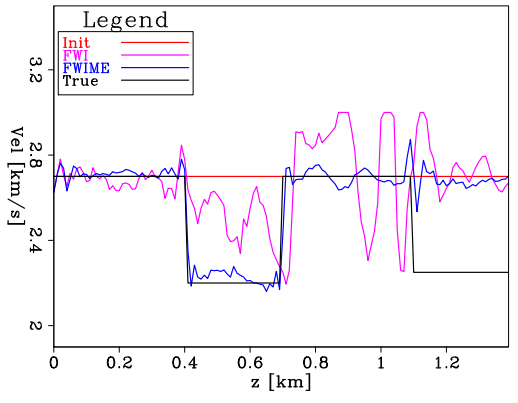

Figure 4: Vertical slice taken at the lateral location of 2 km through the initial model (red), FWI model (magenta), FWITE model (blue), and true model(black).

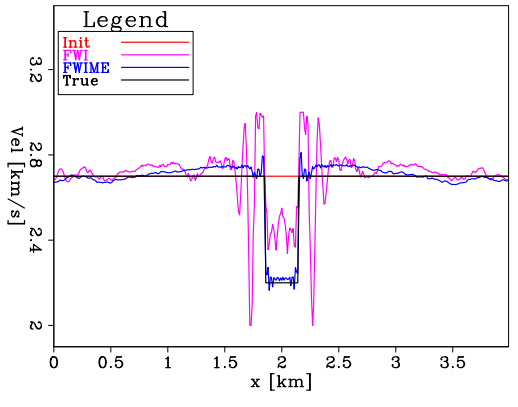

Figure 5: Horizontal slice taken at depth of 0.6 km through the initial model (red), FWI model (magenta), FWITE model (blue), and true model(black).

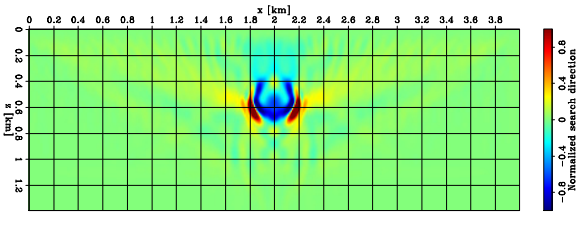

Figure 6: Tomographic component of the first search direction of the FWITE optimization process.

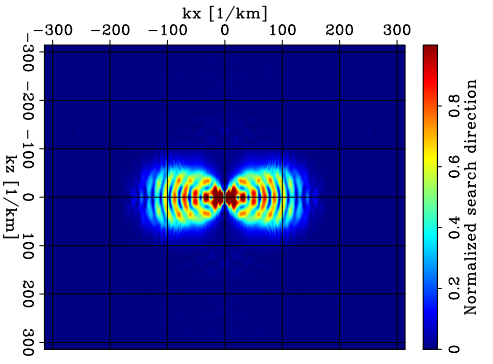

Figure 7: Complex modulus of the Fourier transform of the search direction shown in Figure 6

.

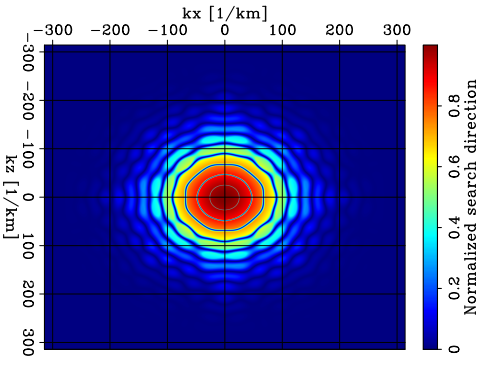

Figure 8: Complex modulus of the Fourier transform of the difference between the true and the starting models.

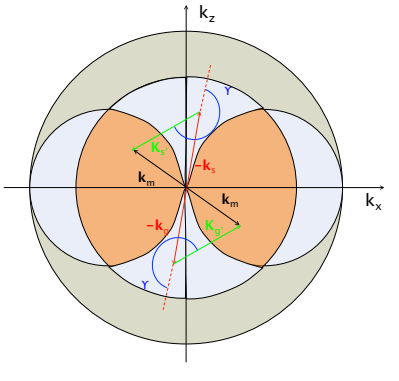

Figure 9: Diagram showing the areas (orange) in the wavenumber plane that are illuminated by second order scattering (tomographic operator) when the data are recorded with source and receivers located at the surface.