Journal of Seismology manuscript No.
(will be inserted by the editor)

Direct and indirect inversions

Jean Virieux · Romain Brossier · Ludovic Métivier · Stéphane Operto · Alessandra Ribodetti

Received: date / Accepted: date

Abstract A bridge is highlighted between the direct inversion and the indirect inversion. They are based on fundamental different approaches: one is looking after a projection from the data space to the model space while the other one is reducing a misfit between observed data and synthetic data obtained from a given model. However, it is possible to obtain similar structures for model perturbation and we shall focus on P-wave velocity reconstruction. This bridge is built up through the Born approximation linearizing the forward problem with respect to model perturbation and through asymptotic approximations of the Green functions of the wave propagation equation. We first describe the direct inversion and its ingredients and then we focus on a specific misfit function design leading to a indirect inversion.

R. Brossier
ISTerre, University Grenoble Alpes, BP 53, 38041 Grenoble CEDEX 09
Tel.: +33-4-76-63-51-74
Fax: +33-4-76-63-52-52
E-mail: Romain.Brossier@ujf-grenoble.fr

L. Métivier
LJK, University Grenoble Alpes, BP 53, 38041 Grenoble CEDEX 09
Tel.: +33-4-76-51-45-83
Fax: +33-4-76-63-12-63
E-mail: Ludovic.Metivier@ujf-grenoble.fr

S. Operto
GeoAzur, IRD, CNRS, UNS, 250 Rue Albert Einstein, 06560 Valbonne, France
Tel.: +33-4-83-61-87-52
Fax: +33-4-83-61-86-10
E-mail: operto@geoazur.unice.fr

A. Ribodetti
GeoAzur, IRD, CNRS, UNS, 250 Rue Albert Einstein, 06560 Valbonne, France
Tel.: +33-4-83-61-87-69
Fax: +33-4-83-61-86-10
E-mail: ribodetti@geoazur.unice.fr

J. Virieux
ISTerre, University Grenoble Alpes, BP 53, 38041 Grenoble CEDEX 09
Tel.: +33-4-76-63-52-54
Fax: +33-4-76-63-52-52
E-mail: Jean.Virieux@ujf-grenoble.fr

Finally, we shall compare this indirect inversion with more standard least-squares inversion as the FWI, enabling the focus on small weak velocity perturbations on one side and the speed-up of the velocity perturbation reconstruction on the othe side. This bridge has been proposed by the group led by Raul Madariaga in the early nineties, emphasizing his leading role in efficient imaging workflows for seismic velocity reconstruction, a drastic requirement at that time.

Keywords Seismology · seismics · seismic inversion · seismic migration

PACS PACS 07.05Pj · PACS 02.30.Zz · 91.30.Jk

1 Introduction

Recovering properties of the Earth through seismic records is a challenging problem which has been tackled with different methods as computers become more and more accessible with improvements in data manipulation: this is the so-called seismic migration with early systematic graphical attempts back to the middle of the XX century. At the end of the XX century, migration based on the wave equation has been promoted starting from the seminal article of Claerbout (1971). Gazdag (1978) has proposed an efficient technique when one considers the computer facilities at that time. Reverse propagation has been recognized at that time as the ingredient for imaging velocity perturbation with imaging condition (Berkhout, 1984; Stolt and Weglein, 1985). As computers become more and more powerful, two trends have emerged with two different ways of thinking. As underlined recently in a quite interesting and fruitful article, although written in a controversial manner, by Weglein (2013), estimating parameters from seismic data may be handled by two fundamentally different strategies: direct inversion methods and indirect inversion methods. Many definitions exist for these terms and, therefore, we rely on the following definition. The direct inversion is seeking for a transformation from the data space into the model space while the indirect inversion is related to the definition of a misfit function quantifying precisely the disagreement between observed data and synthetic data computed by the model. In other words, the direct inversion does not look at a misfit function to be reduced.

The first group devoted to direct inversion has been particularly fruitful with available computer facilities at that time because they minimize the computer resources need. Many functions have direct inversion as the Fourier transform, for example. Different researchers have been working out velocity reconstruction based on integral expressions of the available data projecting them in the model space. The Herglotz-Wiechert-Bateman (HWB) method (Aki and Richards, 1980, section 12.1) based on the Abel transform and its inverse transform is the typical illustration of direct inversion in seismology. Often the inverse transform cannot be built and should be approached iteratively. Bleistein et al (1985) has proposed an integral through asymptotic wave solutions while Beylkin (1985) has introduced an inverse operator of the generalized Radon transform. Inspired by the work of Beylkin, Bleistein (1987) has built up the inverse operator for band-limited data. Since the brilliant paper of Beylkin (1985), intensive investigations have been performed: Bleistein et al (1987) have constructed the 2.5D inversion formula while Wu and Toksöz (1987) have described the link with diffraction tomography (Devaney, 1984). Cohen et al (1986); Miller et al (1987) have proposed very inter-

esting geometrical interpretations for specific acquisition configuration. Since the approach of generalized Radon transform is based on asymptotic formulation, Beylkin and Burridge (1990) have extended it to elasticity while considering any acquisition configuration. This approach has produced many interesting tools such as those for mitigating free surface multiples and even internal multiples based on inverse scattering series. It has been applied at various scales from seismics (Weglein et al, 2009; Zhang and Weglein, 2009a,b) to seismology (Bostock et al, 2001; Rondenay et al, 2001; Bostock, 2003)

The second group developed indirect inversion under the leadership of Albert Tarantola. These indirect methods were mainly based on least-squares inversion (Tarantola, 1984; Gauthier et al, 1986; Ikelle et al, 1986; Pica et al, 1990). Mora (1987) has moved to elastic parameter reconstruction using the same strategy while Mora (1989) have presented the relation with the diffraction tomography. Woodward (1992) has performed a nice review of these different approaches and highlight differences with travel-time tomography. Unfortunately, reflection seismic data were the most often acquired data and the intermediate wavenumber content of the model space was missing in these data (Jannane et al, 1989), illustrating the well-known difficulty of retrieving intermediate wavenumbers from reflection data (Claerbout, 1985). Fortunately, Pratt et al (1998) have revived this least-squares approach, especially as seismic data have been acquired increasingly with long offsets data and as computers have become more and more powerful. This approach, so-called full waveform inversion (FWI), has encountered a significant success moving from synthetic data illustration to real applications (Sirgue and Pratt, 2004; Vigh and Starr, 2008; Brossier et al, 2009). A review by Virieux and Operto (2009) is recommended for the reader interested in a more detailled description of this indirect inversion.

At the end of the 80's, Raul Madariaga was convinced that there should be a bridge between these two ways of thinking at least when considering asymptotic solutions under the Born approximation. In two distinct papers, one more oriented to the methodological point of view (Jin et al, 1992) and the other one oriented to applications (Lambaré et al, 1992), a bridge has been built up between the Beylkin inverse integral and least-squares inversion and it has been applied to a real dataset. The content of this work is a reminder of this link between direct inversion and indirect inversion which have their own advantages and drawbacks. Moreover, we shall benefit from a recent investigation mixing least-squares inversion and Beylkin asymptotic preconditioning (Métivier et al, 2015).

In this paper, we shall start with a simple example for illustrating the differences between direct and indirect inversion. We shall remind the HWB approach before presenting the Beylkin operator for seismic imaging and tackling the relation with the full waveform inversion.

2 Simple example: finding roots of a polynomial

Although seismic imaging is a complex mixed-determined inversion problem, it might be useful to consider a very simple problem where differences between methods can be clearly seen. Let us start with a naive example as proposed by Weglein (2013). A straightforward problem is seeking roots of a quadratic equation $ax^2 + bx + c = 0$. We consider the model space of the variable x as a real number.

The data space is also a 1D real space. For any model variable x, we associate the value $f(x) = ax^2 + bx + c$ in the data domain. The quantities a, b, c are those specifying the transformation: the physics when considering wave equation. The observed data we have collected is only one value equal to zero.
The direct inversion will provide model solutions which could be zero value, one value or two values depending on the sign of the discriminent $b^2 - 4ac$ through the formulation

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}, \tag{1}$$

when the discriminent is positive or null. One can see that any bias in the transformation (values of coefficients a, b and c) we consider may change dramatically expected solutions. We may underline that the mapping could be subjective and that this mapping assumes that the data itself (value zero here) has been well measured.
The indirect inversion will start with an initial value x_1 to reduce the distance between $f(x_1)$ and the data value which is zero: we consider a misfit function to be minimized. We may consider a least-squares inversion which can face better measurement errors than direct inversion. An iterative procedure could be used from this initial guess x_1 with a predicted perturbation δx_1 based on a steepest descent defined by

$$\delta x_1 = -\frac{ax_1^2 + bx_1 + c}{2ax_1 + b}, \tag{2}$$

which can be repeated with the new $x_2 = x_1 + \delta x_1$ with values of the function f going to zero (Press et al, 2007). When considering the full Newton algorithm including the second derivative, we may reach the correct value in one iteration in this particular case of a quadratic equation. For a polynomial of degree n higher than 2, we need to iterate. One can see that we do not find two solutions but only one if successful. The solution, if there is any, will depend on the starting value. The convergence when achieved will be faster as the initial value will be nearby the expected solution.

In this very simple example, we illustrate that both methods have their own pros and cons. The choice will depend on the problem to be solved and on the means at your disposal for doing so. We shall consider now a well-established example of the direct inversion in seismology.

3 Absolute travel time tomography: the HWB method as a direct inversion

A more realistic example comes with the seismic travel-time inversion for 1D medium where one wants to recover the 1D velocity from travel times. The Herglotz-Wiechert-Bateman (HWB) method (Aki and Richards, 1980, section 12.1) is based on travel times and not on travel time delays. We shall introduce it by considering first the projection of a 2D function along a line through the Radon transformation (RT). The RT of a 2D function marked by the symbol ˜ is defined as

$$\tilde{f}(q, \phi) = \int_L f(x(s), y(s)) ds, \tag{3}$$

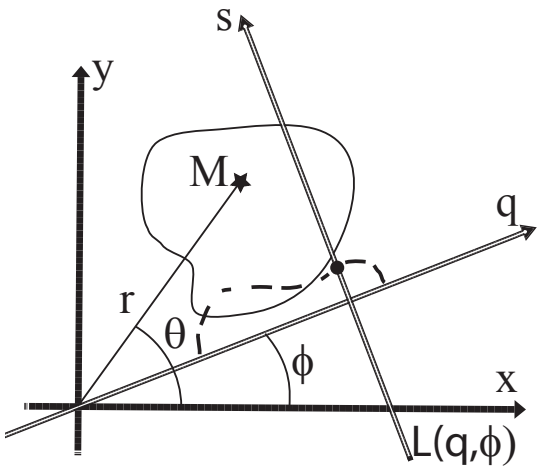

Fig. 1 Radon transform: geometrical projection of an object expressed as values marked by a small black circle corresponding to an integral of values at the point $M(x, z)$ along a line $L(p, \phi)$ described by the sampling s. This projection is the dotted profile with respect to the variable p running on the line defined by an angle ϕ with the x axis.

where the straight line $L(q, \phi)$ is defined by $q = xcos\phi + ysin\phi = rcos(\theta - \phi)$. We may also consider polar coordinates (r, θ) (figure 1). The sampling s is defined in the figure 1 and allows the estimation of the value of the Radon transform at the position q for an integration direction $\phi+\pi/2$. We may introduce the unitary vector $\mathbf{q} = (cos\phi, sin\phi) = (q_x, q_y)$ perpendicular to the line L. The origin is at a intrinsic distance q of the line L. If the function $f(x, y)$ is the slowness, then the function $\tilde{f}(q, \phi)$ is the travel time along the line L in the Fraunhofer configuration (plane-to-plane geometry). By moving into the Fourier domain with following definitions

$$F(k_x, k_y) = \iint_{-\infty}^{+\infty} f(x, y) e^{-2i\pi(k_x.x+k_y.y)} dxdy \tag{4}$$

$$f(x, y) = \iint_{-\infty}^{+\infty} F(k_x, k_y) e^{2i\pi(k_x.x+k_y.y)} dk_x dk_y, \tag{5}$$

one can show that the 2D Fourier transform F of the model is equal to the 1D Fourier transform $\tilde{F}$ of the projection data (for a fixed angle ϕ), in other words the Fourier transform of the Radon transform through the relation

$$F(k_q q_x, k_q q_z) = \tilde{F}(k_q, \phi). \tag{6}$$

By considering the inverse Fourier transform, one can deduce the model

$$f(x, y) = \int_0^{\pi} d\phi \int_{-\infty}^{+\infty} F(k_q, \phi) e^{2i\pi k_q(q_x.x+q_y.y)} |k_q| dk_q = \int_0^{\pi} d\phi \tilde{F}(k_q, \phi) \tag{7}$$

which is just the inverse Radon transform (IRT). The filtered data $\tilde{F}(k_q, \phi)$ can be obtained by the inverse Fourier transform of $|k_q|F(k_q, \phi)$. As underlined by Chapman (1987), the projection data produced by the RT integration is smoother than the model and, therefore, the inverse Radon transform should roughen the

projection data through the multiplication by k_q before smoothing by the radial integration. The numerical evaluation of the inverse Radon transform (7) is difficult and leads to different strategies such as the filtered back projection (FBP) which can be a crude approximation by omitting the roughening step (ignoring the multiplication by k_q), leading to a smoothed version of the model. Although this simple approximation of the IRT is not consistent with the original data (it means that the computed data with the reconstructed model is not the original data, regardless the discretization we use), it is widely applied as, for example, in medical tomography.
Another alternative expression comes with the circular harmonic decomposition (CHD) where the function $f(x, y)$ and its Radon transform are assumed to be well described by Fourier series

$$f(x,y) = \sum_{\infty}^{+\infty} f_n(r) e^{in\theta}$$
$$\tilde{f}(q,\phi) = \sum_{\infty}^{+\infty} \tilde{f}_n(q) e^{in\phi}.$$

Chapman (1987) has discussed precisely the three inverse forms with very different numerical performances. For $n = 0$, the CHD expression of the IRT is reduced to the integral for the HWB method for absolute travel time tomography. One can find

$$\tilde{f}_0(q) = 2 \int_q^{+\infty} \frac{r f_0(r)}{\sqrt{r^2 - q^2}} dr. \tag{8}$$

If the function $f_0(r)$ is the slowness $u(r)$ depending only on the distance r from the center of the Earth, the expression $\tilde{f}_0(q)$ is the travel time

$$T(p) = 2 \int_{r_p}^{r_\oplus} \frac{(ru(r))^2}{\sqrt{(ru(r))^2 - p^2}} \frac{dr}{r} \tag{9}$$

by introducing the ray parameter $p = qu$ and the radius $r_p = q = p/u$ of the turning point. We limit the integration to the radius $r_\oplus$ of the Earth. We often prefer to consider the exit distance $\Delta(p)$ for a spherical Earth, which can be written as

$$\Delta(p) = 2p \int_{r_p}^{r_\oplus} \frac{1}{\sqrt{(ru(r))^2 - p^2}} \frac{dr}{r} \tag{10}$$

The inverse transform of the CHD expression (8) is

$$f_0(r) = -\frac{1}{\pi} \frac{d}{dr} \int_r^{+\infty} \frac{r \tilde{f}_0(q)}{q\sqrt{q^2 - r^2}} dq, \tag{11}$$

which can be written for travel time data as

$$ln \left[\frac{r_\oplus}{r(u)} \right] = \frac{1}{\pi} \int_0^{\Delta(ru)} cosh^{-1}(p/ru) d\Delta, \tag{12}$$

which is known as the Herglotz-Wiechert (HW) formula (Aki and Richards, 1980, section 12.1). One can see that we obtain a profile $r(u)$ which can be inverted to

get the slowness profile $u(r)$ and consequently a velocity profile with respect to the distance at the center of the Earth.

The Radon transform and its inverse expression perform the sampling of the medium through projections along one direction onto a line for 2D case and onto a plane for 3D case, regardless of the velocity structure of the medium: it is a pure geometrical operation. The IRT through the FBP or CHD expressions produces artifacts difficult to mitigate: one must be very careful on numerical error propagation when applying numerically these expressions. If the velocity of the medium depends on a single variable (often the vertical component), the inverse radon transform turns out to be the HWB method connected to the Abel's problem (see Aki and Richards, 1980, p. 646-648). The HWB method is, therefore, a direct inversion providing the velocity profile directly from travel-time measurements. It is worthy underlining that rays connecting the source and the receiver are embedded into the inverse transform: they will be true rays in the wanted unknown medium. There is no link between the pure geometrical projection along straight lines independent of the wanted velocity field. In other words, there is no linearization of the forward problem. Nice geometrical interpretation has been proposed by Garmany et al (1979) leading to linear algebra and providing extremal solutions in a much simpler way than graphical methods of Bessonova et al (1976) also illustrated in the book (Aki and Richards, 1980). Chapman (1987) called this approach the long wavelength strategy. Unfortunately, there is no extension of this direct inversion method to velocity fields depending on more than one variable. As illustrated by different authors (Bessonova et al, 1976; Orcutt, 1979), care should be taken when managing the dataset as this one has to honor the physics we consider.

Chapman (1987) has proposed possible extensions to higher dimensions by new definitions of the model and data spaces based on perturbations. This allows iterative approximations as the one we consider later on. These extensions have never been used in seismology and, therefore, we shall not focus our attention on them. Anyway, it is worth mentioning that there is a possible road for direct inversion when considering higher dimensions. Again, let us underline that the new data should honor the transform we consider and pre-processing will be even more complex in these cases compared to 1D configuration.

This is the exemplary case in seismology where the direct inversion based on strict data-to-model mapping transforms has been used very efficiently during the XX^{th} for the reconstruction of the radial structure of the Earth. One must underline that nowadays velocity reconstruction is more tackled by indirect inversion based on misfit reduction than by using the HW formula (12) even for 1D cases: outliers could be handled in an easier way by indirect inversion while we need to remove them when considering direct inversion. The pre-processing of the data is a crucial step for direct inversion.

4 Generalized Radon transform

Through different publications the references of which can be found in the fundamental paper (Beylkin, 1985), Beylkin has proposed to consider local projections at each point of the medium. These local projections are based on rays linked to the current velocity model: they do not sample geometrically the model as they are now based on a particular velocity structure. Therefore, we can use these rays for

a wavefield estimation. We are now focused on the diffracted wavefield, considered as wavefield perturbations.

For specific geometries (zero-offset configuration, common-source gather, ...), he has developped an asymptotic inverse of the integral describing diffraction wavefield with respect to velocity anomalies. Beylkin and Burridge (1990) have shown how to extend to any arbitrary acquisition geometry including folding. The asymptotic inverse is related to the sampling of the sphere around a particular point of the medium, regardless the acquisition geometry. This paper intends to underline that this asymptotic inverse could be elaborated also through indirect inversion by considering a specific misfit function.

We shall first review the linearized Born and its asymptotic expression which will define the generalized Radon transform which will be at the core of the two types of inversion we want to compare. Then, we shall build up the leading term of the inverse transform giving the velocity perturbation from the diffracted wavefield, which will be called the Beylkin inverse transform: this direct mapping from data space to model space essentially captures singularities of perturbations. We shall omit mathematical details related to Fourier Integral operators and we closely follow presentations by Cohen et al (1986); Miller et al (1987) more oriented to physical understanding of these transformations for seismic applications.

4.1 Born approximation and asymptotic expression

The first-order Born approximation is commonly used in seismic inverse problems for the linearization of the forward problem (see for instance Symes (1995)). Because we are essentially interested in pointing relations between direct and indirect inversions, we shall consider here the acoustic case with constant density for simplicity although extensions to acoustic case with variable density and to elastic case has been performed by Beylkin (1985). We have the scalar wave equation

$$\frac{1}{c^2(\mathbf{x})}\frac{\partial^2 p(\mathbf{x},t;\mathbf{x}_s,m)}{\partial t^2} - \nabla^2 p(\mathbf{x},t;\mathbf{x}_s,m) = S(\mathbf{x}_s,t) \tag{13}$$

where the variable velocity is denoted by $c(\mathbf{x})$ describing the 3D model $m(\mathbf{x})$ while we assume the density constant in the medium. The source term $S(\mathbf{x}_s,t)$ is a temporal function applied at the position of the source $\mathbf{x}_s$. If we apply a Fourier transform to the pressure field $p(\mathbf{x},\omega;\mathbf{x}_s,m)$, we obtain the Helmholtz equation

$$-\frac{\omega^2}{c^2(\mathbf{x})}p(\mathbf{x},\omega;\mathbf{x}_s,m) - \nabla^2 p(\mathbf{x},\omega;\mathbf{x}_s,m) = S(\mathbf{x}_s,\omega). \tag{14}$$

Then, the pressure wavefield $p(\mathbf{x},\omega;\mathbf{x}_s,m)$ is decomposed into a reference pressure wavefield $p_0(\mathbf{x},\omega;\mathbf{x}_s,m)$ solution of (14) in a reference model m_0 defined through the relation $m_0 = 1/c_0^2(\mathbf{x})$ where the velocity is denoted by c_0 and a perturbation $dp(\mathbf{x},\omega;\mathbf{x}_s,m_0,dm)$, such that

$$p(\mathbf{x},\omega;\mathbf{x}_s,m) = p_0(\mathbf{x},\omega;\mathbf{x}_s,m) + dp(\mathbf{x},\omega;\mathbf{x}_s,m_0,dm). \tag{15}$$

The perturbed wavefield $dp(\mathbf{x},\omega;\mathbf{x}_s,m_0,dm)$ is solution of equation (14) with a new source term depending on the model perturbation dm and the reference wavefield

$p_0(\mathbf{x},\omega;\mathbf{x}_s,m_0)$ through the equation

$$-\frac{\omega^2}{c^2(\mathbf{x})}dp(\mathbf{x},\omega;\mathbf{x}_s,m_0,dm)-\nabla^2 dp(\mathbf{x},\omega;\mathbf{x}_s,m_0,dm)=\omega^2 dm(\mathbf{x})p_0(\mathbf{x},\omega;\mathbf{x}_s,m_0) \tag{16}$$

where the model perturbation $dm(\mathbf{x})$ can be written as $-2dc(\mathbf{x})/c_0^3(\mathbf{x})$ where $dc(\mathbf{x})$ is the velocity perturbation. For a point source $\mathbf{x}_s$, we may consider the source excitation $S(\omega)$ and the Green function $G(\mathbf{x},\omega;\mathbf{x}_s,m)$, solution of the equation (14) for the medium m and an impulsive point source. The pressure perturbation could be expressed as an integral over the diffraction domain Ω where model perturbations exist through the expression

$$dp_B(\mathbf{x},\omega;\mathbf{x}_s,m_0,dm)=\omega^2\int_\Omega d^3y\, G(\mathbf{y},\omega;\mathbf{x}_s,m)dm(\mathbf{x})G(\mathbf{y},\omega;\mathbf{x}_s,m_0)S(\omega), \tag{17}$$

for one frequency. The first-order Born approximation, sometimes called Born approximation, will replace the Green function in the model m by the Green function in the model m_0, leading to the linearized expression of the pressure perturbation dp_B with respect to model perturbation.

We now assume that the Green function could be approximated by the WKBJ approximation assuming a single phase through the expression

$$G(\mathbf{y},\omega;\mathbf{x},m)=a(\mathbf{y},\mathbf{x},m)e^{i\omega\tau(\mathbf{y},\mathbf{x},m)}, \tag{18}$$

where we have introduced the travel-time τ satisfying the eikonal equation

$$(\nabla\tau(\mathbf{y},\mathbf{x},m))^2=1/c^2(\mathbf{x})$$

and the amplitude a satisfying the transport equation

$$2\nabla\tau(\mathbf{y},\mathbf{x},m)\cdot\nabla a(\mathbf{y},\mathbf{x},m)+a(\mathbf{y},\mathbf{x},m)\nabla^2\tau(\mathbf{y},\mathbf{x},m)=0.$$

The imaginary complex i is such that $i^2=-1$. At a given circular frequency ω, the expression (17) could be rewritten with asymptotic expressions as

$$dp_B^a(\mathbf{x}_r,\mathbf{x}_s,\omega,m_0,dm)=\omega^2 S(\omega)\int_\Omega d^3y\, dm(\mathbf{y})A(\mathbf{x}_r,\mathbf{y},\mathbf{x}_s,m_0))e^{i\omega T(\mathbf{x}_r,\mathbf{y},\mathbf{x}_s,m_0)}. \tag{19}$$

The asymptotic expression dp_B^a has a Fourier integral form with the total travel time (by extension we call it phase) from the source $\mathbf{x}_s$ to the receiver $\mathbf{x}_r$ through the diffracting point $\mathbf{y}$ as the following sum

$$T(\mathbf{x}_r,\mathbf{y},\mathbf{x}_s,m_0)=\tau(\mathbf{x}_r,\mathbf{y},m_0)+\tau(\mathbf{y},\mathbf{x}_s,m_0) \tag{20}$$

and the total amplitude as the following product

$$A(\mathbf{x}_r,\mathbf{y},\mathbf{x}_s,m_0)=a(\mathbf{x}_r,\mathbf{y},m_0)a(\mathbf{y},\mathbf{x}_s,m_0). \tag{21}$$

The expression (19) is the asymptotic Born relation (linearized relation) between the model perturbation *dm* and the pressure perturbation *dp* for one source and one receiver. We shall use these general expressions for the two approaches we consider: the direct inversion through a mapping and the indirect inversion through a misfit function. In this particular case, the indirect inversion will be based on a least-squares inversion.

4.2 Beylkin inverse transform

Thanks to the Fourier integral structure of the expression (19), one could argue that the inversion operator for this (source,receiver) couple should have the negative of the phase of this Fourier expression. We may consider the measured pressure perturbation $dp_{obs}(\mathbf{x}_r, \mathbf{x}_s, \omega)$ as the asymptotic quantity dp^a_B, making the mapping between the model space and the data space explicit. Consequently we may write the inverse operator as

$$dm(\mathbf{y}, m_0) = \int_{\Omega_\xi} d^2\boldsymbol{\xi}\; b(\mathbf{y}, \boldsymbol{\xi}, m_0) \int_\omega d\omega\; S(\omega) dp_{obs}(\mathbf{x}_r, \mathbf{x}_s, \omega) e^{-i\omega T(\mathbf{x}_r, \mathbf{x}_s, m_0)}, \quad (22)$$

where the factor $b(\mathbf{y}, \boldsymbol{\xi}, m_0)$ has to be estimated. We need also to estimate dimensions of the vector $\boldsymbol{\xi}$ which will be shown to span a space Ω_ξ of two dimensions in relation with the data domain as noted in the above expression. The phase $T(\mathbf{x}_r, \mathbf{x}_s, m_0)$ should be also computed and will correspond to a phase matching contribution. Injecting the asymptotic expression of the pressure perturbation leads to a mapping between a model perturbation $dm(\mathbf{y})$ and another model perturbation $dm(\mathbf{y}')$ for this couple

$$\begin{aligned} dm(\mathbf{y}, m_0) = &\int_{\Omega_\xi} d^2\boldsymbol{\xi}\; b(\mathbf{y}, \boldsymbol{\xi}, m_0) \int_\omega d\omega\; S(\omega)\omega^2 \\ &\int_\Omega d^3y' dm(\mathbf{y}', m_0) A(\mathbf{x}_r, \mathbf{y}', \mathbf{x}_s, m_0) e^{i\omega\left(T(\mathbf{x}_r, \mathbf{y}', \mathbf{x}_s, m_0) - T(\mathbf{x}_r, \mathbf{x}_s, m_0)\right)}, \end{aligned} \quad (23)$$

which should correspond to the following limit

$$dm(\mathbf{y}, m_0) = \int d^3y'\; dm(\mathbf{y}', m_0)\delta(\mathbf{y}' - \mathbf{y}), \quad (24)$$

at high frequencies with a full-band spectrum. Consequently, we may device the function $b(\mathbf{y}, \boldsymbol{\xi}, m_0)$ such that

$$\begin{aligned} \int_{\Omega_\xi} d^2\boldsymbol{\xi} \quad &b(\mathbf{y}, \boldsymbol{\xi}, m_0) \int_\omega d\omega\; S(\omega)\omega^2 \\ &A(\mathbf{x}_r, \mathbf{y}', \mathbf{x}_s, m_0) e^{i\omega\left(T(\mathbf{x}_r, \mathbf{y}', \mathbf{x}_s, m_0) - T(\mathbf{x}_r, \mathbf{x}_s, m_0)\right)} \to \delta(\mathbf{y} - \mathbf{y}'). \end{aligned} \quad (25)$$

The key proposition of Beylkin (1985) is a local Taylor expansion of the phase around the position $\mathbf{y}$ which could be connected to discontinuity structures of the model perturbation dm. Therefore, we may consider the Taylor approximation of the phase matching contribution

$$T(\mathbf{x}_r, \mathbf{x}_s, m_0) \sim T(\mathbf{x}_r, \mathbf{y}, \mathbf{x}_s, m_0) + \nabla_{\mathbf{y}} T(\mathbf{x}_r, \mathbf{y}, \mathbf{x}_s, m_0) \cdot (\mathbf{y}' - \mathbf{y}), \quad (26)$$

corresponding to an approximation of the phase T around the position $\mathbf{y}$ as well as a rather simple approximation of the amplitude term

$$A(\mathbf{x}_r, \mathbf{y}', \mathbf{x}_s, m_0) \sim A(\mathbf{x}_r, \mathbf{y}, \mathbf{x}_s, m_0) \quad (27)$$

leading to the expression

$$\begin{aligned} \int d^2\boldsymbol{\xi} \quad &b(\mathbf{y}, \boldsymbol{\xi}, m_0) \int_\omega d\omega S(\omega)\omega^2 \\ &A(\mathbf{x}_r, \mathbf{y}, \mathbf{x}_s, m_0) e^{i\omega\left(\nabla_{\mathbf{y}} T(\mathbf{x}_r, \mathbf{y}, \mathbf{x}_s, m_0) \cdot (\mathbf{y}' - \mathbf{y})\right)} \to \delta(\mathbf{y}' - \mathbf{y}). \end{aligned} \quad (28)$$

Following Beylkin (1985), we introduce a new vector $\mathbf{k}$ at the position of interest $\mathbf{y}$ through the change of variables

$$\mathbf{k} = \omega \nabla_{\mathbf{y}} T(\mathbf{x}_r, \mathbf{y}, \mathbf{x}_s, m_0), \tag{29}$$

from (ω, ξ_1, ξ_2) to (k_1, k_2, k_3) which justifies the dimension two of the sampling vector $\boldsymbol{\xi}$. Beylkin (1985) introduces the Jacobian defined as

$$h(\mathbf{y}, \boldsymbol{\xi}) = \frac{1}{\omega^2} \frac{\partial \mathbf{k}}{\partial(\omega, \xi_1, \xi_2)} \tag{30}$$

which allows to write the expression

$$\int d^3\mathbf{k}\ \frac{A(\mathbf{x}_r, \mathbf{y}, \mathbf{x}_s, m_0) b(\mathbf{y}, \boldsymbol{\xi}(\mathbf{k}), m_0)}{h(\mathbf{y}, \boldsymbol{\xi}(\mathbf{k}))} S(\omega(\mathbf{k})) e^{i\mathbf{k}\cdot(\mathbf{y}'-\mathbf{y})} \rightarrow \delta(\mathbf{y}' - \mathbf{y}). \tag{31}$$

We may choose now the following expression

$$b(\mathbf{y}, \boldsymbol{\xi}(\mathbf{k}), m_0) = \frac{1}{8\pi^3} \frac{h(\mathbf{y}, \boldsymbol{\xi}(\mathbf{k}))}{A(\mathbf{x}_r, \mathbf{y}, \mathbf{x}_s, m_0)}, \tag{32}$$

inspired from the spectrum of the dirac function. The expression (31) becomes

$$\frac{1}{8\pi^3} \int d^3\mathbf{k}\ S(\omega(\mathbf{k})) e^{i\mathbf{k}\cdot(\mathbf{y}'-\mathbf{y})}. \tag{33}$$

The source spectrum $S(\omega)$ induces a band-limited expression of the dirac function where one can see that the vector $\mathbf{k}$ is the wavenumber vector. Often, we introduce the illumination vector $\mathbf{q} = \mathbf{k}/\omega$ which has a simple geometrical interpretation with rays connecting in the forward direction the source to the diffracting point and in the backward direction the receiver to the diffracting point (figure 2). By inserting the expression (32) into the model perturbation estimation (22), we find

$$\begin{aligned} dm(\mathbf{y}, m_0) = \frac{1}{8\pi^3} \int_{\Omega_\xi} d^2\boldsymbol{\xi}\ \frac{h(\mathbf{y}, \boldsymbol{\xi})}{A(\mathbf{x}_r, \mathbf{y}, \mathbf{x}_s, m_0)} \\ \int_\omega d\omega\ S(\omega)\, dp_{obs}(\mathbf{x}_r, \mathbf{x}_s, \omega) e^{-i\omega T(\mathbf{x}_r, \mathbf{y}, \mathbf{x}_s, m_0)}, \end{aligned} \tag{34}$$

which is the relation mapping the data space through the Ω_ξ sampling into the model space: a direct inversion result. This operator is sometimes called the Beylkin operator and the Jacobian h the Beylkin determinant. This sampling is directly related to the double integral over the sphere surrounding the scattering point (Burridge and Beylkin, 1988), allowing to sample the sphere around the diffracting point with general distribution of sources and receivers.

Cohen et al (1986); Miller et al (1987) have nicely expressed the mapping through the Jacobian for specific geometries as zero-offset (resp. constant-offset) configuration or common-source (resp. common-receiver) configuration with interesting geometrical interpretations. These mappings are linked with the single scattering interpretation of the figure 2, although we may consider the entire seismic cube (Beylkin and Burridge, 1990). A geometrical interpretation is provided in the figure 3 where one can see that two different (source,receiver) couples at different frequencies sample exactly the same wavenumber for the reconstruction of the velocity perturbation at the diffracting point.

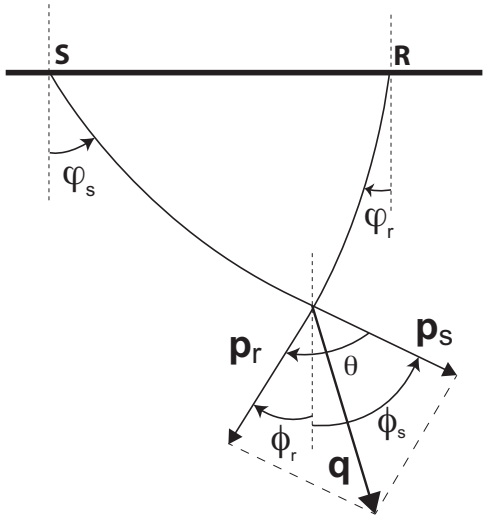

Fig. 2 Single scattering geometrical configuration. At the point of diffraction where model properties have to be reconstructed, the incident ray arrives with an angle Φ_s providing a wavenumber vector $\mathbf{p}_s$. Corresponding receiver data or residuals are back-propagated in the medium and will arrive with an angle Φ_r at the diffracting point providing a wavenumber vector $\mathbf{p}_r$. The angle θ between these two vectors is the illumination angle and the vectorial composition of these vectors provides the illumination vector $\mathbf{q}$, a key ingredient for parameter imaging at the diffracting point. Waves leave the source with an angle ϕ_s and arrive with an angle ϕ_r at the receiver.

Under an asymptotic approximation, we have built a direct inversion as an integral operator combining the data (pressure perturbation) in order to reconstruct the 3D model (velocity perturbation). This direct inversion is linked to the background model m_0: its performance depends on the quality of the background model we are free to choose. In order to obtain this expression (34), we have linearized the forward problem through the Born approximation. We have created a mapping from the data space to the model space which is the velocity perturbation without considering a misfit function which characterizes indirect inversion.

In order to emphasize the relation between direct and indirect inversions which is the aim of this paper, we would like to recover the relation (34) through an indirect inversion while considering the wave equation (Tarantola, 1984). Using standard least-squares misfit function will not provide this relation. We need to design a specific misfit function such that the expression giving the velocity perturbation should be the same: therefore, using the same asymptotic forward problem is likely to be needed to recover this relation while the Born approximation is already embedded into the approach proposed by Tarantola (1984). This is the purpose of the next section.

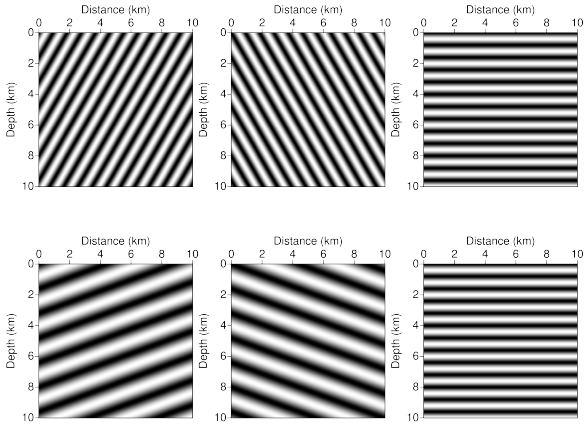

Fig. 3 Two different planar decomposition are considered at a given diffracting point. They provide the same gradient (two right vertical panels) although wavefields have different frequency contentsand that they have different illumination at the diffracting point (four left panels). The top panel represents a planar wavefield contribution with a high frequency content coming from a source (left map) with a rather flat angle and leaving to a receiver (middle map) again with a rather flat angle. The right map represents the product of these two wavefields which is just the gradient contribution of these two plane waves at the diffracting point. The bottom panel represents another planar contribution with more dipping illumination at a lower frequency. When constructing the gradient contribution on the right map by making the product, one can see that we obtain the same contribution as for the previous case of the top panel. Consequently, these two contributions are identical and will sum up coherently for the medium property reconstruction through the gradient. One can see the strong relation between illumination and frequency when waves interact locally with the medium.

5 Ray + Born approach

We shall proceed with the same approximation for the forward problem as the one considered in the GRT but we shall not try to construct the inverse transform from the data space toward the model space. We shall consider instead an indirect inversion by fitting observed data with synthetic data as promoted by Tarantola (1984). Let us define a misfit function through the standard least-squares expression

$$\mathcal{C}(dm(\mathbf{y})) = \frac{1}{2}\int_\omega d\omega \int_S d^3\mathbf{x}_s \int_R d^3\mathbf{x}_r \mathcal{Q}\left[dp_{obs}(\mathbf{x}_r,\omega;\mathbf{x}_s) - dp_B^a(\mathbf{x}_r,\omega;\mathbf{x}_s,m_0,dm(\mathbf{y}))\right]^2, \tag{35}$$

where the weight $\mathcal{Q}$ has to be designed. A very specific weight has been introduced by Jin et al (1992); Lambaré et al (1992) through the following expression

$$\mathcal{Q}(\mathbf{x}_r,\mathbf{x}_s,\mathbf{y}_0) = \frac{1}{8\pi^3}\frac{(\nabla_{\mathbf{y}_0} T(\mathbf{x}_r,\mathbf{y}_0,\mathbf{x}_s,m_0))^2}{A^2(\mathbf{x}_r,\mathbf{x}_s,\mathbf{y}_0)}, \tag{36}$$

which varies from point to point where the velocity perturbation should be imaged. The misfit function in the data space depends on the point of the model space we consider for the reconstruction: we may underline this dependence through the notation $\mathcal{C}_{\mathbf{y}_0}(dm(\mathbf{y}))$. Making the misfit function only depends on data values will

require to consider the total misfit function

$$\mathbb{C}(dm(\mathbf{y})) = \sum_{\mathbf{y}_0} \mathcal{C}_{\mathbf{y}_0}(dm(\mathbf{y})), \tag{37}$$

which needs to be minimized under the constraint of the linearized asymptotic forward problem given by the equation (19) and could be compactly noted by

$$\mathbf{dp} = \mathcal{G}\mathbf{dm}, \tag{38}$$

where the matrix $\mathcal{G}$ represents the Born modelling operator or the Jacobian operator. The minimization of the misfit function (37) will end up to solve the Gauss-Newton equation

$$\mathcal{G}^\dagger \mathcal{G}\ \mathbf{dm} = \mathcal{G}^\dagger \mathbf{dp}, \tag{39}$$

where the symbol † denotes the adjoint matrix which is just the transpose for real matrices. We need the gradient $-\mathcal{G}^\dagger \mathbf{dp}$ and, depending on the approximation for the inversion of the Hessian $\mathcal{G}^\dagger\mathcal{G}$, we may perform one or more iterations in order to find the solution through conjugate-gradient methods. The gradient of this misfit function has been shown to be

$$\gamma(\mathbf{y}_0) = \frac{1}{8\pi^3} \int_S d^3\mathbf{x}_s \int_R d^3\mathbf{x}_r \frac{1}{A(\mathbf{x}_r, \mathbf{x}_s, \mathbf{y}_0)} (\nabla_{\mathbf{y}_0} T(\mathbf{x}_r, \mathbf{y}_0, \mathbf{x}_s, m_0))^2 \int_\omega d\omega\ S(\omega)\ (dp_{obs}(\mathbf{x}_r, \mathbf{x}_s, \omega) - dp_B^a(\mathbf{x}_r, \mathbf{x}_s, \omega, m_0, 0))\, e^{-i\omega T(\mathbf{x}_r, \mathbf{y}_0, \mathbf{x}_s, m_0)} \tag{40}$$

by Jin et al (1992); Lambaré et al (1992). Initially the solution dp_B^a is zero as there is no model perturbation dm. We are back to the Beylkin expression (34). We have now identified what is the misfit function which has the Beylkin expression as the gradient. This gradient is a mapping between the data space to the model space and acts identically through the direct inversion or this specific indirect inversion. The Hessian of this specifig misfit function is shown to converge to the identity operator when acquisition is fully illuminating the target with full spectral content (Jin et al, 1992). The sampling is naturally over sources and receivers but we have introduced the illumination of the scattering point through the very specific weight we have defined intuitively (Forgues, 1996). A bridge between the direct inversion proposed by Beylkin (1985) and the indirect (least-squares) inversion proposed by Lailly (1983); Tarantola (1984) has been built and highlights similarities and differences between these two complementary approaches.

The background model m_0 is often a smooth model while the model perturbation dm has a high wavenumber content. They contain different information of the medium coming from scale separation we are familiar in seismics. We improve the reconstruction of m_0 to more and more detailed structures because of the dramatic increase of data coverage and quality. We may succeed in filling of the wavenumber gap and, therefore, the model perturbation dm could be used for updating the model m_0. Unfortunately, the ray tracing become a quite challenging task when the medium perturbation introduces small-scale features in the background model. A better strategy should come from the wave equation we shall consider in the next paragraph.

To our knowledge, it was a first attempt to bring a bridge between apparent disconnected methods by designing a very specific misfit function where the

asymptotic inverse transform proposed by Beylkin (1985) is identical to the gradient obtained by minimising a specific least-squares data weighted misfit. It has been a contribution of the group led by Raul Madariaga in the early nineties.

6 Mixing asymptotic solutions and numerical solutions for full waveform inversion

The Ray+Born approach faces the challenge of tracing rays in updated models when more and more small-scale heterogeneities are introduced. Up to now, stable ray tracing as proposed by Vinje et al (1993); Lambaré et al (1996); Lucio et al (1996); Forgues and Lambaré (1997) requires smooth media and faces often difficulties as the wavefront shape becomes more and more complex in this asymptotic assumption. This difficulty has prevented any iteration on the background model m_0 which is often taken smoother than required by any initial velocity analysis or high-resolution travel-time tomography because of ray tracing limitations.

An alternative to the Ray+Born approach could be to consider the wave equation for the forward problem while keeping the same elaborated misfit function through ray mapping performed in a background model. This workflow will belong to the so-called full waveform inversion (FWI) family as we minimize a seismic waveform misfit.

The ray mapping allows to sum the contribution of the wavefields in a very specific way for building the gradient. We may not trace rays at each iteration and rely on rays from a background velocity model computed before hand and, therefore, without no computational impact on the inversion procedure. By this way, we perform a preconditioning of the Full Waveform Inversion (FWI) as shown by Métivier et al (2015), speeding the convergence. Hence, it could be considered as a target-oriented imaging, thanks to the ray mapping. This particular workflow does not intend to overcome inherent difficulties of the FWI for reconstructing the low-wavenumber content of the medium: it will be sensitive to the high-wavenumber content, especially when the phase difference between the observed data and the synthetic data in the background model m_0 goes to zero. This phase matching is the kind of prior information we foster in the gradient reconstruction at the expense of the low-wavenumber content. This phase sensitivity may induce amplification of incorrect phases or noises when artificially matched.

Let us remind the asymptotic model perturbation could be expressed as

$$\mathcal{B}(m_B)\left(p_B^a(m_0) - p_{obs}\right) \tag{41}$$

from the Beylkin operator given by the equation (34) defined in a background model m_B. Métivier et al (2015) suggest to minimize this model perturbation and, therefore, we need to consider the following misfit function

$$\mathcal{D}(m) = \frac{1}{2}||\mathcal{B}(m_B)\left(p_{cal}(m) - p_{obs}\right)||^2_{\mathcal{M}}, \tag{42}$$

where the model space is denoted by the symbol $\mathcal{M}$ and seismic data $p_{cal}(m)$ are computed now by the wave equation in the model m. The expression of the gradient of the misfit function $\mathcal{D}$ is

$$\nabla\mathcal{D}(m) = J^\dagger(m)\mathcal{B}(m_B)^\dagger\mathcal{B}(m_B)\left(p_{cal}(m) - p_{obs}\right) \tag{43}$$

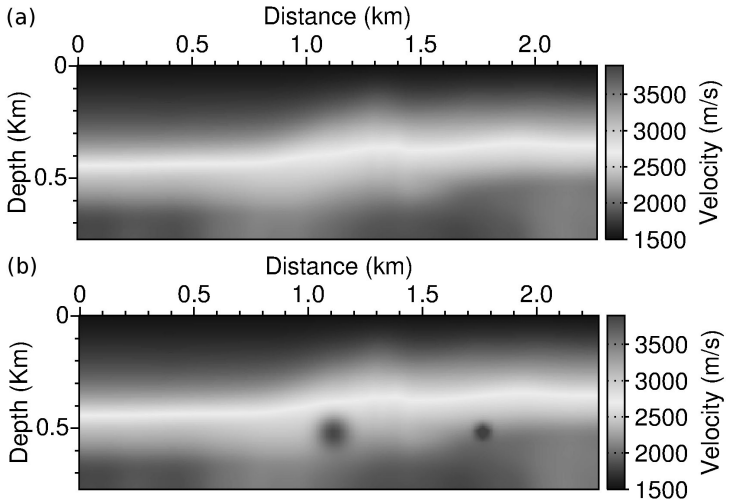

Fig. 4 Smooth P-wave model used as background and initial models (top). Smooth P-wave model with two gaussian perturbation added in depth to be recovered (bottom)

where the operator J is the Fréchet derivative of the standard FWI misfit function

$$C(m) = \frac{1}{2}||p_{cal}(m) - p_{obs}||^2 \tag{44}$$

in the data domain. When considering this new misfit function, the Gauss-Newton approximation of the related Hessian tends towards the identity improving the behaviour of the inversion (Métivier et al, 2015). Of course, the background model could be different from the initial model but, in the following illustration, we consider that this background model is the initial model m_0 where first-arrival rays are traced only once.

We consider a rather simple canonical test where two inclusions of different sizes should be reconstructed at a depth of $500\,m$ in a medium of a length of $2275\,m$ and a depth of $775\,m$. The large perturbation is at an horizontal distance of $1100\,m$ and the small one is at a distance of $1750\,m$ as shown in the figure 4. An fixed-spread acquisition of 84 sources and receivers is deployed at a depth of $50\,m$. We consider 10 frequencies from $3\,Hz$ to $12\,Hz$ with a spacing of $1\,Hz$. Forward modeling is performed by a finite difference technique.

The gradient of the standard least-squares misfit function is shown on the top of the figure 5 while the one of the modified misfit function is shown at the bottom of the figure 5. Least-squares inversion attribute greater importance to high residues than to small residues. Therefore, the least-squares gradient mainly focuses on the largest inclusion which a low frequency imprint coming from the acquisition design: the second inclusion is difficult to identify. In contrast, the modified gradient though the Beylkin preconditioning provides a more balanced update for the large and small inclusions. It is also less sensitive to the acquisition geometry. Associated to this more balanced focusing, positive/negative oscillations around inclusions are observed, coming from the wavenumber filtering we perform. This filtering could be connected to the asymptotic linearized inverse weighting introducing bandwidth effects.

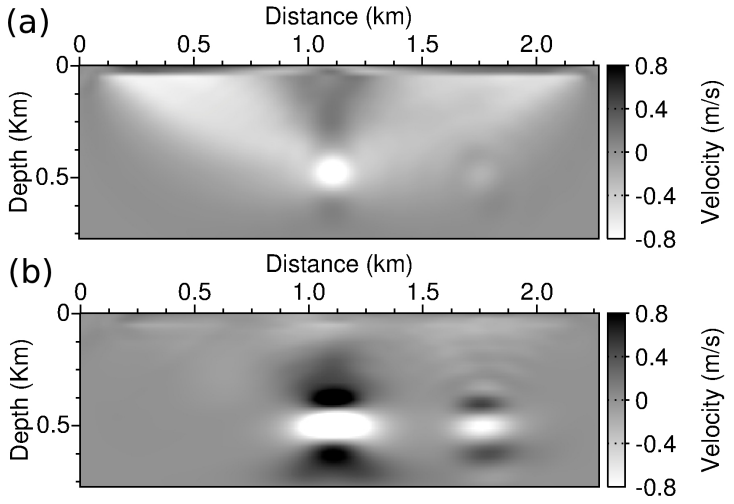

Fig. 5 Gradients computed for the two inclusions case study. Standard L^2 norm gradient (top). Asymptotic preconditioned gradient (bottom).

The fast build-up of the velocity perturbation is related to the ray structure which corresponds to first-arrival rays in this example. By selecting the appropriate background model, we may accelerate the reconstruction of detailed structures in a specific velocity model, especially weak and small structures. When considering more complex ray pattern as triplication or shadow zones, we may extract from weak phases the interesting signal for velocity perturbation. This is a way to mitigate effects of least-squares optimization which have a tendency for fitting large amplitudes.

7 Conclusion

In this paper, we have shown relations between direct and indirect inversions in the framework of seismic migration approaches. By considering the so-called Born approximation, the forward problem is linearized, making the inversion easier to be performed. Moreover, we may consider asymptotic solutions for wave propagation, enhancing the key importance of phases (travel times) for locating velocity perturbations. Amplitudes are carefully considered for the estimation of velocity anomalies.

We have shown that the integral inverse operator, called the Beylkin operator, could be obtained also through an indirect inversion with the definition of a specific misfit function. This bridge between these two different strategies of interpretating the data highlights differences and similarities of these approaches when considering their standard formulation.

The direct inversion may reconstruct unwanted features when unexpected waveforms are recorded while not included in the forward modeling engine. The indirect inversion mitigates effects of outliers, i.e. observed waveforms which are not predicted by the forward modeling engine, but it does not eliminate their influence, especially the least-squares method. Moreover, it requires an accurate

model in order to focus high-frequency waveforms nearby the places they come from. This is particularly true for the least-squares methods. If such accurate initial model is designed, we can combine the two approaches in order to speed up the reconstruction of details in the velocity model once a low-wavenumber model has been reconstructed. We expect this combined strategy to improve reconstruction in areas which provide only weak signals inside the acquisition recording.

We may conclude that both techniques are useful and extract complementary information from seismic data by emphasizing specific features in the high wavenumber part of the model spectrum. They have different tolerance behaviours to noise in the data: the direct inversion may require a rather important preprocessing for extracting waveforms we expect to be modeled by the forward engine while the indirect inversion is less sensitive, reducing the importance of the preprocessing. We should not oppose them.

Acknowledgements This study was granted access to the HPC resources of the Froggy platform of the CIMENT infrastructure (https://ciment.ujf-grenoble.fr), which is supported by the Rhône-Alpes region (GRANT CPER07_13 CIRA), the OSUG@2020 labex (reference ANR10 LABX56) and the Equip@Meso project (reference ANR-10-EQPX-29-01) of the programme Investissements d'Avenir supervised by the Agence Nationale pour la Recherche, and the HPC resources of CINES/IDRIS under the allocation 046091 made by GENCI. This study was partially funded by the SEISCOPE consortium, sponsored by BP, CGG, CHEVRON, EXXONMOBIL, JGI, PETROBRAS, SAUDI ARAMCO, SCHLUMBERGER, SHELL, SINOPEC, STATOIL, TOTAL and WOODSIDE (*http://seiscope2.osug.fr*).

References

Aki K, Richards P (1980) Quantitative Seismology: Theory and Methods. W. H. Freeman & Co, San Francisco

Berkhout AJ (1984) Seismic Migration-Imaging of acoustic energy by wave field extrapolation, Volume 14B. Elsevier Science Publ. Co., Inc

Bessonova E, Fishman V, Shnirman M, Sitnikova G, Johnson L (1976) The tau method for the inversion of travel times, II, earthquake data. Geophysical Journal of the Royal Astronomical Society 46:87–108

Beylkin G (1985) Imaging of discontinuities in the inverse scaterring problem by inversion of a causal generalized Radon transform. Journal of Mathematical Physics 26:99–108

Beylkin G, Burridge R (1990) Linearized inverse scattering problems in acoustics and elasticity. Wave motion 12:15–52

Bleistein N (1987) On the imaging of reflectors in the Earth. Geophysics 52(7):931–942

Bleistein N, Cohen JK, Hagin FG (1985) Computational and asymptotic aspects of velocity inversion. Geophysics 50(8):1253–1265

Bleistein N, Cohen JK, Hagin FG (1987) Two and one-half dimensional Born inversion with an arbitrary reference. Geophysics 52:26–36

Bostock M (2003) Linearized inverse scattering of teleseismic waves for anisotropic crust and mantle structure: I. theory. Journal of Geophysical Research 108-B5:2258

Bostock MG, Rondenay S, Shragge J (2001) Multiparameter two-dimensional inversion of scattered teleseismic body waves 1. theory for oblique incidence. Journal of Geophysical Research 106(12):30,771–30,782

Brossier R, Operto S, Virieux J (2009) Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. Geophysics 74(6):WCC105–WCC118, DOI 10.1190/1.3215771, URL http://link.aip.org/link/?GPY/74/WCC63/1

Burridge R, Beylkin G (1988) On double integrals over spheres. Inverse problems 4:1–10

Chapman C (1987) The Radon transform and seismic tomography, D. Reidel Publishing Company, chap 2, pp 25–47

Claerbout J (1971) Towards a unified theory of reflector mapping. Geophysics 36:467–481

Claerbout J (1985) Imaging the Earth's interior. Blackwell Scientific Publication

Cohen JK, Hagin FG, Bleistein N (1986) Three-dimensional Born inversion with an arbitrary reference. Geophysics 51:1552–1558

Devaney A (1984) Geophysical diffraction tomography. Geoscience and Remote Sensing, IEEE Transactions on GE-22(1):3–13

Forgues E (1996) Inversion linearisée multi-paramètres via la théorie des rais. PhD thesis, Institut Français du Pétrole - University Paris VII

Forgues E, Lambaré G (1997) Parameterization study for acoustic and elastic ray+born inversion. Journal of Seismic Exploration 6:253–278

Garmany J, Orcutt J, Parker R (1979) Travel time inversion: A geometrical approach. Journal of Geophysical Research 84:3615–3622

Gauthier O, Virieux J, Tarantola A (1986) Two-dimensional nonlinear inversion of seismic waveform : numerical results. Geophysics 51:1387–1403

Gazdag J (1978) Wave equation migration with the phase-shift method. Geophysics 43:1342–1351

Ikelle L, Diet JP, Tarantola A (1986) Linearized inversion of multioffset seismic reflection data in the $\omega - k$ domain. Geophysics 51:1266–1276

Jannane M, Beydoun W, Crase E, Cao D, Koren Z, Landa E, Mendes M, Pica A, Noble M, Roeth G, Singh S, Snieder R, Tarantola A, Trezeguet D (1989) Wavelengths of Earth structures that can be resolved from seismic reflection data. Geophysics 54(7):906–910

Jin S, Madariaga R, Virieux J, Lambaré G (1992) Two-dimensional asymptotic iterative elastic inversion. Geophysical Journal International 108:575–588

Lailly P (1983) The seismic problem as a sequence of before-stack migrations. In: Bednar J (ed) Conference on Inverse Scattering: Theory and Applications, SIAM, Philadelphia

Lambaré G, Virieux J, Madariaga R, Jin S (1992) Iterative asymptotic inversion in the acoustic approximation. Geophysics 57:1138–1154

Lambaré G, Lucio PS, Hanyga A (1996) Two-dimensional multivalued traveltime and amplitude maps by uniform sampling of ray field. Geophysical Journal International 125:584–598

Lucio P, Lambaré G, Hanyga A (1996) 3D multivalued travel time and amplitude maps. Pure and Applied Geophysics 148:113–136

Métivier L, Brossier R, Virieux J (2015) Combining asymptotic linearized inversion and full waveform inversion. Geophysical Journal International 201(3):1682–1703, DOI 10.1093/gji/ggv106

Miller D, Oristaglio M, Beylkin G (1987) A new slant on seismic imaging: Migration and integral geometry. Geophysics 52(7):943–964
Mora PR (1987) Nonlinear two-dimensional elastic inversion of multi-offset seismic data. Geophysics 52:1211–1228
Mora PR (1989) Inversion = migration + tomography. Geophysics 54(12):1575–1586
Orcutt J (1979) Joint linear, extremal inversion of seismic kinematic data. Journal of Geophysical Research 85:2649–2660
Pica A, Diet JP, Tarantola A (1990) Nonlinear inversion of seismic reflection data in laterally invariant medium. Geophysics 55(3):284–292
Pratt RG, Shin C, Hicks GJ (1998) Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. Geophysical Journal International 133:341–362
Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) Numerical Recipes 3rd Edition: The Art of Scientific Computing, 3rd edn. Cambridge University Press
Rondenay S, Bostock MG, Shragge J (2001) Multiparameter two-dimensional inversion of scattered teleseismic body waves. 3. Application to the Cascadia 1993 data set. Journal of Geophysical Research 106(12):30,795–30,807
Sirgue L, Pratt RG (2004) Efficient waveform inversion and imaging : a strategy for selecting temporal frequencies. Geophysics 69(1):231–248
Stolt RH, Weglein AB (1985) Migration and inversion of seismic data. Geophysics 50(12):2458–2472
Symes WW (1995) Mathematics of reflection seismology. Rice Inversion Project
Tarantola A (1984) Inversion of seismic reflection data in the acoustic approximation. Geophysics 49(8):1259–1266
Vigh D, Starr EW (2008) 3D prestack plane-wave, full waveform inversion. Geophysics 73:VE135–VE144
Vinje V, Iversen E, Gjøystdal H (1993) Traveltime and amplitude estimation using wavefront construction. Geophysics pp 1157–1166
Virieux J, Operto S (2009) An overview of full waveform inversion in exploration geophysics. Geophysics 74(6):WCC1–WCC26
Weglein A, Zhang H, Ramrez A, Liu F, Lira J (2009) Clarifying the underlying and fundamental meaning of the approximate linear inversion of seismic data. Geophysics 74(6):WCD1–WCD13, DOI 10.1190/1.3256286, URL http://library.seg.org/doi/abs/10.1190/1.3256286, http://library.seg.org/doi/pdf/10.1190/1.3256286
Weglein AB (2013) A timely and necessary antidote to indirect methods and so-called p-wave fwi. The Leading Edge October:1192–1204
Woodward MJ (1992) Wave-equation tomography. Geophysics 57:15–26
Wu RS, Toksöz MN (1987) Diffraction tomography and multisource holography applied to seismic imaging. Geophysics 52:11–25
Zhang H, Weglein A (2009a) Direct nonlinear inversion of 1d acoustic media using inverse scattering subseries. GEOPHYSICS 74(6):WCD29–WCD39, DOI 10.1190/1.3256283, URL http://library.seg.org/doi/abs/10.1190/1.3256283, http://library.seg.org/doi/pdf/10.1190/1.3256283
Zhang H, Weglein A (2009b) Direct nonlinear inversion of multiparameter 1d elastic media using the inverse scattering series. Geophysics 74(6):WCD15–WCD27, DOI

10.1190/1.3251271, URL http://library.seg.org/doi/abs/10.1190/1.3251271, http://library.seg.org/doi/pdf/10.1190/1.3251271