Nonlinear sensitivity operator and inverse thin-slab propagator for tomographic waveform inversion

Ru-Shan Wu*, Chunhua Hu1 and Benfeng Wang2, Modeling and Imaging Laboratory, Earth & Planetary Sciences, Univ. of California, Santa Cruz, CA, USA; 1visiting from Tsinghua University, Beijing, China; 2visiting from China Petroleum University, Beijing, China.

Summary
We derived the nonlinear sensitivity operator and the related inverse thin-slab propagator (ITSP) for nonlinear tomographic waveform inversion based on the theory of nonlinear partial derivative operator. The inverse propagator is based on a renormalization procedure to the forward and inverse T-matrix series. The inverse thin-slab propagator solves the divergence problem of the inverse series for strong perturbations by stepwise partial summation (renormalization). Numerical tests showed that the inverse Born T-series starts to diverge at 20% perturbation (for the given model), while the inverse thin-slab propagator has no convergence problem for up to 50% perturbation. This convergence improvement has potential applications to the iterative procedure of waveform inversion.

Introduction
The gradient method in full waveform inversion (FWI) is based on a linearization of the full nonlinear functional partial derivative (NLPD) operator (See Tarantola, 1984, 2005), and can be considered as a quasi-linear inversion. NLPD can be expanded into a Taylor series which corresponds to a full scattering series, the Born series. The convergence problem of the iterative procedure of quasi-linear inversion, problem of local minima, and the starting model dependence, are all deeply rooted in the well-known convergence problem of the Born series and inverse Born series (see e.g., Morse and Feshback, 1953; Moses, 1956; Prosser, 1969; Aki and Richards, 1980; Weglein et al., 1997, 2003). For the real Earth, the wave equation is strongly nonlinear with respect to the medium parameter changes. Wu and Zheng (2012, 2014) introduced the higher order Fréchet derivatives and the theory of nonlinear partial derivative (NLPD) operator for the acoustic wave equation. Our precious work (Wu and Zheng, 2012, 2014; Wu et al., 2013) have reported the renormalization procedure using De Wolf approximation to improve the convergence of forward scattering series. In this paper we will report the progress in removing the divergence of inverse Born T-series by renormalization procedure and the derivation of the inverse thin-slab propagator. Numerical tests proved that the inverse thin-slab propagator (ITSP) has no divergence and is very efficient in solving inverse problem in tomographic inversion.

Nonlinear sensitivity operator
Assume an initial model \(m_0\), we want to quantify the sensitivity of the data change \(\delta d\) (also called “data residual”) to the model perturbation \(\delta m\),

\[
\delta d = d - d_0 = A(m_0 + \delta m) - A(m_0) = A^{(NLPD)}(m_0, \delta m) \delta m
\]

(1)

where \(A^{(NLPD)}\) is the nonlinear partial derivative (NLPD) operator which is \(\delta m\) dependent,

\[
A^{(NLPD)} = A'(m_0) + \frac{1}{2!} A''(m_0)(\delta m)^2 + \ldots
\]

+ \[
\frac{1}{n!} A^{(n)}(m_0)(\delta m)^n + \ldots
\]

(2)

where \(A', A'', \ldots\) and \(A^{(n)}\) are the first, second, and the \(n\)th order Fréchet derivatives. If we split the scattering operator into forward scattering and backscattering parts, \(S = S' + S''\) and substitute it into the Fréchet series, we can have all combinations of higher order forward and backward derivatives. The De Wolf approximation corresponds to neglecting multiple backscattering (reverberations), i.e. dropping all the terms containing two or more backscattering operators but keeping all the forward scattering terms untouched (De Wolf, 1985; Wu, 2003; Wu et al., 2007; Wu and Zheng, 2012, 2014). In this paper we deal with the transmission tomography, so we concern only the forward-scattering.

Forward and inverse scattering series in T-matrix approach
IBS (inverse Born series) originally is formulated directly in the data space. An alternative way, which is more convenient and computational efficient, is to formulate the inverse scattering in the image space (model space). This is the approach of contrast-source approach and T-matrix approach (e.g., Prosser, 1969; Weglein et al., 2003; Jakobsen, 2012; Jakobsen and Ursin, 2012). Here, we will apply the T-matrix approach to the NLSO (nonlinear sensitivity operator).

\(\mathbf{T}\)-matrix (transition matrix) \(\mathbf{T}\) is defined through the following equation

\[
\mathbf{V}(\mathbf{x}, \mathbf{x}') g(\mathbf{x'}, \mathbf{x}) = \mathbf{T}(\mathbf{x}, \mathbf{x}') g_0(\mathbf{x'}, \mathbf{x})
\]

(3)

where \(g_0\) is the background Green’s function, \(g\) is Green’s function in the inhomogeneous media and \(\mathbf{V}(\mathbf{x}, \mathbf{x}')\)
Nonlinear sensitivity operator and inverse thin-slab propagator

is the scattering potential defined as (in the scalar wave case)
\[ V(x, x') = S_v(x, x') e_v(x') = k^2 \int_{\omega} d^3 x' e_v(x') \delta(x - x'). \] (4)
with \( e_v = (v_0/v)^2 - 1 \) as the velocity perturbation function (dimensionless). In the following we formulate the forward scattering and inverse scattering in terms of T-matrix. First we discuss the forward problem. From the relation
\[ T_{g_0} = V_g = V_{g_0} + V_G T_{g_0} \]
we have
\[ T = [I - V G]^{-1} V = PV \] (5)
where \( P \) is the propagator. This is in a form of integral equation similar to the Lippmann-Schwinger equation. From the above equations, we can write
\[ T = V + V_G T \] (6)
In the case of weak scattering, in which \( \|V G_0\| < 1 \), an iterative procedure can be used to get a Born series of T-matrix from (6) for the forward scattering solution. For strong perturbations, we can apply the forward scattering renormalization, or the De Wolf approximation to eliminate the divergence problem.

Now we consider the inversion process. Assume we know the data residual \( \delta d \), and try to map it back to a model perturbation \( \delta m \). To predict the model perturbation from a measured data residual is also a mapping operator, a nonlinear inverse sensitivity operator. The kernel of the operator is the nonlinear sensitivity kernel (NLSK), which plays a similar role as the linear sensitivity kernel in the linear inversion.

First, we can say that the ISO (inverse sensitivity operator) must be nonlinear one to recover the model perturbations from the data. After obtaining the T-matrix based on data of experiments, e.g. by a linear inversion or imaging, we can invert the T-matrix for the scattering potential and perturbation function \( e_v \). The scattering data are kept in the T-matrix and stay in the model space (image space), and the acquisition process is peeled off. Of course, the knowledge of acquisition process is needed to estimate the T-matrix by inversion. Knowing \( T \), the scattering potential \( V \) can be derived as
\[ V = T(1 + G_s T)^{-1} = T \sum_{n=0}^{\infty} (-1)^n (G_s T)^n \] (7)
where \( P \) is defined as the inverse propagator. In case of weak scattering, the above equation can be solved by a series solution, which is the inverse Born T-series (IBS). However, for strong scattering (strong-contrast or large-scale perturbations), the IBS is highly oscillating and may diverge, which we will shown later in the examples. Also we know that IBS is closely related to the iterative inversion using the linear Fréchet derivative (e.g., Markel et al., 2003), so the divergence problem of inversion may be rooted to that of the IBS.

Renormalization of ISS (inverse scattering series) and the inverse thin-slab propagator (ITSP)

As shown in Wu and Zheng (2012, 2014), we can split the scattering operator into a forward part and a backward part so that the Born series can be reformed into a De Wolf series. After applying the De Wolf approximation to the forward T-matrix formulation (5) by the split of the V-operator, we can form the split the T-matrix into a part due to forward scattering and a part due to single backscattering (neglecting multiples)
\[ T = T' + T'' \] (8)
where \( T' \) is the transition matrix due to forward-scattering, and \( T'' \) is the transition matrix due to backscattering. In this work, we treat only the forward scattering problem such as in the case of smooth media, so only \( T' \) is involved. For T-matrix due to forward scattering, the T-matrix for any point \( \mathbf{x} \) in the medium can be decomposed into one derived from the interaction with the upper half-space velocity potential (up-scattering) and one from the lower half-space velocity potential (down-scattering), plus a part from the same level,
\[ T'(x, x') = T_f(x, x') + T_g(x, x') + T_b(x, x'), \]
\[ \begin{align*}
T_f(x, x') &\triangleq T_f(x, x'), \quad x' \in \{ z' < z \} \\
T_g(x, x') &\triangleq T_g(x, x'), \quad x' \in \{ z' = z \} \\
T_b(x, x') &\triangleq T_b(x, x'), \quad x' \in \{ z' > z \}
\end{align*} \] (9)
Since only forward scattering is involved, we can recover the velocity potential from the corresponding \( T' \), \( T' \) and \( T' \). Substitute the decomposition (9) into the inverse T-series (7), and apply the forward-scattering renormalization, resulting in a De Wolf approximation for the inverse T-matrix series. We mainly concern the diagonal terms of the recovered V-matrix (all off-diagonal elements should be zeros due to mutual cancellations by multiple inverse-scattering), therefore,
\[ \tilde{V}(x, x) = k^2 \tilde{e}(x) = \text{Diag} \{ TP' \} \]
\[ \approx \text{Diag} \{ T_f \} + \text{Diag} \{ T_g P_g' \} + \text{Diag} \{ T_b P_b' \} + \text{Diag} \{ T_f P_f' \} \] (10)
where \( P' \) is the inverse propagator,
\[ P' = \sum_{n=0}^{\infty} (-1)^n (G_s T)^n \] (11)
and \( P_g', P_b' \) and \( P_f' \) are the corresponding inverse propagators for the decomposed T-matrices. In deriving these propagators we drop the cross-coupling terms between \( T' \) and \( T'' \), since the reverberations of up- and
Nonlinear sensitivity operator and inverse thin-slab propagator

down-going waves will be cancelled out eventually. This is consistent with the physics that no multiples exist within a smooth medium. Also we neglect the contribution of $T'$ altogether. This is due to our forward-scattering approximation. The T-matrix correction in the same level is from lateral waves (side-walking waves) which are often absent in the real data. This approximation is also tolerable in the viewpoint of inversion accuracy, since the error of this approximation can be reduced by decreasing the thickness of the thin-slab. It will become clear in the numerical examples that in the case of missing data due to the unavailability of lateral waves, the ITSP method will have better approximation than the IBS even the series converges. In this way we have the approximation:

\[
P_a = \sum_{n=1}^{\infty} (-1)^n (G^T x_n)^+ \quad \quad P_d = \sum_{n=1}^{\infty} (-1)^n (G x_n)^+ \tag{12}
\]

where $G^T$ and $G$ are the corresponding one-way Green's operators, which are triangular matrices in this case.

Now we apply the renormalization procedure to the inverse propagator $P_a$ and $P_d$. The general procedure is similar to the forward thin-slab propagator and can be written in a form as (Wu, 2003; Wu et al., 2012)

\[
P' (x_n, x_0) = \frac{1}{1 + G a T} = \prod_{i=1}^{N} T_i \quad T_i = 1 - G_i T_i \tag{13}
\]

where $T_i$ is the thin-slab operator, $T_i$ is the column T-matrix of the thin-slab, and $G_i$ is the free propagator for the thin-slab. Although with similarity, the inverse propagator is more complicated in structure and operations, since T-matrix is full, while V-matrix is a diagonal one.

Convergence tests of the inverse Born T-matrix series and the inverse thin-slab propagator (the renormalized series)

Figure 1 The model of Gaussian ball ($a = 5\lambda$) with different perturbation strengths (15%, 20% and 50%).

Figure 2 T-matrix in the kernel representation. Each small ball is a kernel $T(x, x')$ for a fixed $x$ (a point spreading function). On the right are the two kernels we used for the V($x$) recovery tests.

In the following we show some simple examples to demonstrate the convergence property of the renormalized inverse T-series: the inverse thin-slab propagator. The test model is a Gaussian ball ($a = 5\lambda$) with different perturbation strengths from the homogeneous background (Figure 1). The matrix $T$ is a complex-value frequency-dependent matrix of $N$ by $N$. The whole model space is of 200 by 200 in grid size. The perturbation area is about 50 by 50, and $N$ is about 1900. We produce the T-matrix by both finite difference method and the matrix inverse method. They agree well and we adopt the matrix inverse method for the tests. In Figure 2 we plot the sparsely sampled column vectors of the full T-matrix (kernel representation) ($f_0 = 20hz$). In the Figure, each small ball is a representation of a kernel, corresponding to a point spreading function (only real part is shown). On the right is the two kernels we used in the tests to recover the V($x$) at the two corresponding points. This exact T-matrix corresponds to a full-aperture measurement and contains all the information in full-aperture acquired data. Normally, T-matrix is derived by a linear inversion from the data with limited aperture. In order to test the convergence of the inversion scattering series, we use the exact T-matrix here. The influence of data aperture will be studied in the future work.

Figure 3, 4 and 5 give the results of convergence tests. Here we only plot the convergence of velocity perturbation value at some fixed points of the model. In Figure 3(1) and (2) we plot the convergence curves from the inverse Born series using the full-T and the T-matrix with missing data (lateral waves). The vertical axis is the perturbation strength $\epsilon$, and the horizontal axis is the series summation orders. We see high oscillation nature of the IBS in relatively weak scattering (15% perturbation with $a = 5\lambda$). Note that using the full T-matrix series converges to a correct value after many terms. However, with the missing
data, it converges to a wrong value. In Figure 3(3) is the result from the ITSP (inverse thin-slab propagator). For the ITSP calculation, the corrections by $P_{\nu}$ and $P_{\varphi}$ are computed separately from the top (the curve on the left) and the bottom (the curve on the right), respectively (the horizontal axis is labeled with the slab ordering numbers). Due to the stepwise renormalization of ITSP, the convergence is almost monotonically. Although there is a minor error compared with the result using the full T-matrix, but the error is smaller than the result of IBS using the incomplete T-matrix. For 20% perturbation (Figure 4), we test the recovery for the 2nd point. We see the IBS starts to diverge (the left of Figure 4), but ITSP (the right) has a similar convergence as the case of weak scattering. In order to show the detailed behavior, we plot the zoomed figure in Figure 4(4) by amplifying the vertical scale. In the same manner, we plot the comparison for the case of strong perturbation (50%) in Figure 5. We see the fast divergence of IBS, but a similar convergence curve for the ITSP. In principle, the renormalization procedure of ITSP can remove the divergence for any strong perturbations.

Figure 3 Convergence comparison for weak perturbation (15%) between inverse Born series (IBS) using full T-matrix (1); IBS without $T$ (2) and the inverse thin-slab propagator (3). The pink dashed lines are the correct value of velocity perturbation $\varepsilon_v$.

Figure 4 Convergence comparison for 20% perturbation (at point 2) between inverse Born series (IBS) using full T-matrix (1), IBS without $T$ (2), and the inverse thin-slab propagator (3). To see the details we plot the zoomed figure in (4) by amplifying the vertical scale. The pink dashed lines are the correct value of velocity perturbation $\varepsilon_v$.

Conclusion

Based on the nonlinear sensitivity operator in nonlinear tomographic waveform inversion, we apply the renormalization procedure to the forward and inverse T-matrix series and derive the corresponding thin-slab propagators. In this paper we report the results of inverse thin-slab propagator. Numerical tests proved that the renormalized inverse scattering series has much better convergence property than the inverse Born series and the inverse thin-slab propagator is an efficient method and has no divergence problem. This convergence improvement has great potential in applying to the iterative procedure of waveform inversion.

Figure 5 Convergence tests (at point 1) for strong perturbations (50%). Others are the same as in Figure 4.

Acknowledgment: We thank Yingcai Zheng, Lingling Ye, Morten Jakobsen for discussions and collaborations. This research is supported by the WTOPI Research Consortium at the University of California, Santa Cruz.
EDITED REFERENCES
Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2014 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES
Morse, P. M., and H. Feshback, 1953, Methods of theoretical physics, I and II: McGraw-Hill.

Wu, R. S., L. Ye, and Y. Zheng, 2013, Nonlinear partial functional derivative and nonlinear LS seismic inversion: Presented at the 75th Annual International Conference and Exhibition, EAGE.

Wu, R. S. and Y. Zheng, 2012, Nonlinear Fréchet derivative and its de Wolf approximation: Presented at the 82nd Annual International Meeting, SEG.