**Introduction**

Most optimization approaches for reconstruction of velocity models used in Full Waveform Inversion (FWI) are based on smooth techniques such as the Tikhonov regularization. However, realistic velocity profiles often present various discontinuities, sharp interfaces and high contrasts, as in the important case of the presence of salt bodies. Such discontinuities arise from the fact that sound waves travel with greater velocity inside salt bodies compared to the neighboring sediments.

In this work we propose to use nonsmooth optimization techniques for the reconstruction of sharp interfaces. We consider a simplified setting where the velocity is piecewise constant, with known constant but distinct values in the salt body and in the sediment region. In this way, the optimization problem is recast as a shape optimization problem (see Sokołowski and Zolésio (1992)), where the interface of the salt region becomes the unknown. The problem is formulated as the minimization of a tracking-type cost functional with respect to the geometry of the interface, and the evolution of the interface is performed using a level set method, see Osher and Sethian (1988). On one hand, shape optimization and level set-based approaches have been developed recently for FWI in the frequency domain, see Lewis et al. (2012); Guo and de Hoop (2013); Lewis and Vigh (2016); Shi et al. (2017); Kadu et al. (2017); on the other hand we propose here a level set-based algorithm for time-domain FWI.

One difficult challenge arising in the shape optimization framework is the low regularity of the solution to the wave equation due to the discontinuity of the coefficients, which affects both the state and adjoint state. To address this issue, we use the concept of distributed shape derivative which allows working with nonsmooth domains and functions with low regularity. The calculation of the distributed shape derivative is achieved using the averaged adjoint method introduced in Sturm (2015), which is a Lagrangian approach for shape optimization problems; see also Laurain and Sturm (2016). The shape derivative depends on the time and space derivatives of the state and adjoint state, where the adjoint is the solution of a wave equation with terminal conditions, with the residual on the right-hand side. We show the efficiency of the method through two examples of reconstruction with synthetic data.

**Acoustic wave equation**

Let $\mathcal{D} \subset \mathbb{R}^n$, $n = 2, 3$, be an open, bounded and piecewise $C^2$-domain. Homogeneous Dirichlet and Neumann boundary conditions are imposed on $\Gamma_d \subset \partial \mathcal{D}$ and $\Gamma_n = \partial \mathcal{D} \setminus \Gamma_d$, respectively, where $\partial \mathcal{D}$ denotes the boundary of $\mathcal{D}$. The set $\Gamma_n$ corresponds to the surface of the Earth. For $T > 0$, we consider the acoustic wave equation with damping

$$
\rho \frac{\partial u}{\partial t} - \Delta u + \eta u = f \quad \text{in } \mathcal{D} \times [0,T],
$$

$$
u(0) = 0 \quad \text{in } \mathcal{D},
$$

$$
u_t(0) = 0 \quad \text{in } \mathcal{D},
$$

$$
u = 0 \quad \text{on } \Gamma_d \times [0,T],
$$

$$
\partial_n u = 0 \quad \text{on } \Gamma_n \times [0,T],
$$

where the index $t$ denotes partial differentiation with respect to time. Here, $\rho$ denotes the square slowness defined as $\rho = 1/c^2$, where $c$ is the speed of sound in the given physical media. We optimize in the class of piecewise constant $\rho \in L^\infty(\mathcal{D})$, to be more precise we assume $\rho = \rho_0 \chi_\Omega + \rho_1 \chi_{\mathcal{D}\setminus\Omega}$, where $\Omega \subset \mathcal{D}$, which represents the salt body, is not touching $\partial \mathcal{D}$. The function $\eta \in C^1(\mathcal{D})$ is a damping term equal to 0 inside the physical domain and positive inside a small layer in the neighborhood of $\partial \mathcal{D}$; this damping term is used to model an unbounded domain.
Shape functional and adjoint state for FWI

Let \( d_r \in L^2([0, T], \mathbb{R}) \), \( r = 1, \ldots, N_R \), denote the measured data at the receptors, where \( N_R \) is the number of receptors, located on the surface \( \Gamma_n \). We define the following shape functional

\[
J(\Omega) := \frac{1}{2} \sum_{r=1}^{N_R} \int_0^T (u(x_r,t) - d_r(t))^2.
\]

where \( \Omega \in \mathcal{O} \) and \( \mathcal{O} \) is a suitable set of admissible shapes. Note that \( J \) depends on \( \Omega \) through \( u \), since \( \rho = \rho_{\Omega} = \rho_0 \chi_{\Omega} + \rho_1 \chi_{\overline{\mathcal{O}} \setminus \Omega} \). Introduce the spaces

\[
X := L^2(0, T; H^1_{\Gamma_n}(\mathcal{O})) \cap H^1(0, T; L^2(\mathcal{O})), \quad X_0 := \{ \psi \in X \mid \psi(0) = 0 \}, \quad X_T := \{ \psi \in X \mid \psi(T) = 0 \}.
\]

Given \( \rho \in L^\infty(\mathcal{O}) \) and \( f \in L^2(0, T; L^2(\mathcal{O})) \), the variational formulation for the wave equation with damping corresponding to the strong formulation (1) is: find \( u \in X_0 \) such that

\[
\int_0^T \int_\mathcal{O} \nabla u \cdot \nabla \psi - \rho u_t \psi_t + \eta u_t \psi \, dx \, dt = \int_0^T \int_\mathcal{O} f \psi \, dx \, dt, \quad \forall \psi \in X_T.
\]

Combining the cost functional and the variational formulation (2) of the PDE constraint, we introduce the Lagrangian \( \mathcal{L} : \mathcal{O} \times X_0 \times X_T \to \mathbb{R} \) as

\[
\mathcal{L}(\Omega, \varphi, \psi) := \frac{1}{2} \sum_{r=1}^{N_R} \int_0^T (\varphi(x_r,t) - d_r(t))^2 \, dt + \int_0^T \int_\mathcal{O} \nabla \varphi \cdot \nabla \psi - \rho \varphi_t \psi_t + \eta \varphi_t \psi - f \psi \, dx \, dt.
\]

Following the averaged adjoint method (see Sturm (2015)), the equation for the adjoint is given by

\[
\partial_\varphi \mathcal{L}(\Omega, u, p)(\hat{\varphi}) = 0 \text{ for all } \hat{\varphi} \in X_0.
\]

This yields the following adjoint equation in variational form: find \( p \in X_T \) such that

\[
\int_0^T \int_\mathcal{O} \nabla \hat{\varphi} \cdot \nabla p - \rho \hat{\varphi}_t p_t + \eta \hat{\varphi}_t p \, dx \, dt = -\sum_{r=1}^{N_R} \int_0^T (u(x_r,t) - d_r(t)) \hat{\varphi}(x_r,t) \text{ for all } \hat{\varphi} \in X_0.
\]

Note that the adjoint \( p \) is the solution of a wave equation with terminal conditions \( p(T) = 0 \) and \( p_t(T) = 0 \), instead of initial conditions.

Calculation of the distributed shape derivative

Consider a vector field \( \theta \in C^0(\mathcal{O}, \mathbb{R}^2) := \{ V \in C^0(\mathcal{O}, \mathbb{R}^2) \mid V \text{ has compact support in } \mathcal{O} \} \) and the associated flow \( \Phi_s : \mathcal{O} \to \mathcal{O}, s \in [0, s_0] \) defined for each \( x_0 \in \mathcal{O} \) as \( \Phi_s(x_0) := x(s) \), where \( x : [0, s_0] \to \mathbb{R}^2 \) solves \( \dot{x}(s) = \theta(x(s)) \) for \( s \in [0, s_0] \) and \( x(0) = x_0 \). Then for \( \Omega \) open and compactly contained in \( \mathcal{O} \subset \mathbb{R}^2 \), we consider a family \( \Omega_s := \Phi_s(\Omega) \) of perturbed domains. For Lagrangian approaches in shape optimization, one needs to introduce the so-called shape-Lagrangian \( G \) defined as (see Laurain and Sturm (2016))

\[
G(s, \varphi, \psi) := \mathcal{L}(\Omega_s, \varphi \circ \Phi_s^{-1}, \psi \circ \Phi_s^{-1}).
\]

It is shown in Sturm (2015) that, under appropriate regularity assumptions on the data, the shape derivative is given by

\[
dJ(\Omega)(\theta) = \partial_\varphi G(0, u, p, q).
\]

This yields the following expression of the distributed shape derivative in tensor form:

\[
dJ(\Omega)(\theta) = \int_\mathcal{O} S_1 : D\theta + S_0 \cdot \theta \, dx,
\]
the data acquisition, with disconnected blocks representing the salt body. We use order to guarantee the CFL condition. In the first experiment (Figure 1), the ground truth consists of two receivers. In the second experiment (Figure 2) the ground truth consists of one block representing the transport equation:

\[ \frac{\partial \theta}{\partial t} + \nabla \phi \cdot \nabla \theta = 0 \quad \text{in } \mathcal{D} \times \mathbb{R}^+ . \]  

where \( \mathcal{D} \) is the identity matrix. Note that \( S_1 \) is matrix-valued while \( S_0 \) is vector-valued. Since the support of \( \eta \) is concentrated near \( \partial \mathcal{D} \), it is often the case that \( \text{supp}(\theta) \subseteq \{ \eta = 0 \} \). Also, \( f \) is concentrated on the surface \( \Gamma_n \) in FWI, so it is reasonable to assume that \( \theta = 0 \) on the support of \( f \) if \( \Omega \) does not touch the boundary of \( \mathcal{D} \). When these additional assumptions are satisfied, we get \( S_0 = 0 \) and

\[
S_1 = \left[ \int_0^T -\rho u_t p_t + \nabla u \cdot \nabla p + \eta u_t p - fp dt \right] \mathbb{I}_d - \int_0^T \nabla u \otimes \nabla p + \nabla p \otimes \nabla u dt ,
\]

\[
S_0 = \int_0^T pu_t \nabla \eta - p \nabla f dt ,
\]

where \( \mathbb{I}_d \) is the identity matrix. Note that \( S_1 \) is matrix-valued while \( S_0 \) is vector-valued. Since the support of \( \eta \) is concentrated near \( \partial \mathcal{D} \), it is often the case that \( \text{supp}(\theta) \subseteq \{ \eta = 0 \} \). Also, \( f \) is concentrated on the surface \( \Gamma_n \) in FWI, so it is reasonable to assume that \( \theta = 0 \) on the support of \( f \) if \( \Omega \) does not touch the boundary of \( \mathcal{D} \). When these additional assumptions are satisfied, we get \( S_0 = 0 \) and

\[
S_1 = \left[ \int_0^T -\rho u_t p_t + \nabla u \cdot \nabla p dt \right] \mathbb{I}_d - \int_0^T \nabla u \otimes \nabla p + \nabla p \otimes \nabla u dt .
\]

Numerical algorithm

In order to obtain a smooth descent direction \( \theta \), i.e. a vector field satisfying \( dJ(\Omega)(\theta) < 0 \), we solve the elliptic equation: find \( \theta \in H^1(\mathcal{D})^2 \) such that

\[
\int_\mathcal{D} \alpha_1 g D \theta : D \xi + \alpha_2 g \theta \cdot \xi \ dx = -dJ(\Omega)(\theta), \quad \forall \xi \in H^1(\mathcal{D})^2 ,
\]

(3)

where \( \alpha_1 = 0.01, \alpha_2 = 0.9 \) and \( g : \mathbb{R}^2 \to \mathbb{R} \) is defined as \( g(x,z) := 2 - x^2 \). The role of \( g \) is to compensate the loss of sensitivity as one moves away from the surface. We use a level set method, introduced by Osher and Sethian (1988), to model the evolution of the domain. The core idea of this method is to represent the boundary of the moving domain \( \Omega \subseteq \mathcal{D} \subseteq \mathbb{R}^N \) as the zero level set of a Lipschitz continuous function \( \phi : \mathcal{D} \times \mathbb{R}^+ \to \mathbb{R} \). A family of domains \( \Omega_t \subseteq \mathcal{D} \) is defined as

\[ \Omega_t := \{ x \in \mathcal{D} \mid \phi(x,t) < 0 \}, \quad \text{so that} \quad \partial \Omega_t = \{ x \in \mathcal{D} \mid \phi(x,t) = 0 \}, \]

where we assume \( |\nabla \phi(\cdot,t)| \neq 0 \) on \( \partial \Omega_t \). It can be shown that \( \phi \) is related to \( \theta \) through the following transport equation:

\[
\partial_t \phi(x,t) + \theta(x) \cdot \nabla \phi(x,t) = 0 \quad \text{in } \mathcal{D} \times \mathbb{R}^+. \]

(4)

The algorithm consists in first calculating \( \theta \) using (3), and then to solve (4) to update the domain \( \Omega_t \).

Results

We present two numerical experiments with \( c_0 = 4.12 \) km/s and \( c_1 = 1.95 \) km/s, where \( \rho_0 = 1/c_0^2 \) and \( \rho_1 = 1/c_1^2 \). The domain \( \mathcal{D} \) is a rectangle of length 1 km in \( x \)-axis and depth 0.65 km in \( z \)-axis, which is meshed using a regular grid. We use the time step \( \Delta t = 0.001 \) s and the grid step \( \Delta x = 0.01 \) km in order to guarantee the CFL condition. In the first experiment (Figure 1), the ground truth consists of two disconnected blocks representing the salt body. We use 40 shots with Ricker wavelet of 5 Hz to simulate the data acquisition, with 80 receivers placed on the top surface \( \Gamma_n \) and a spacing of 0.01 km between the receivers. In the second experiment (Figure 2) the ground truth consists of one block representing the salt body. We use 25 shots for the data acquisition, with 81 receivers and a spacing of 0.01 km between the receivers. In both simulations, seismograms of \( T = 0.799 \) seconds are recorded. The salt body \( \Omega \) is updated at each iteration, starting from an initial guess. We observe that the interface reconstruction is very accurate on the upper part of the salt body, while the reconstruction is reasonable but less accurate on the lower part. This result was expected since the data is measured only at the surface.

References

Figure 1: Reconstruction using 40 shots with Ricker wavelet of $5Hz$. Initialization (red curves - left), reconstruction (red curves - right). The black block is the ground truth. Initial value of cost functional: 1461.522, final value: 6.03

Figure 2: Reconstruction using 25 shots with Ricker wavelet of $5Hz$. Initialization (red curves - left), reconstruction (red curves - right). The black block is the ground truth. Initial value of cost functional: 520.069, final value: 3.89


