Incorporating multiple a priori information for full waveform inversion

Da Li, Michael P. Lamoureux and Wenyuan Liao

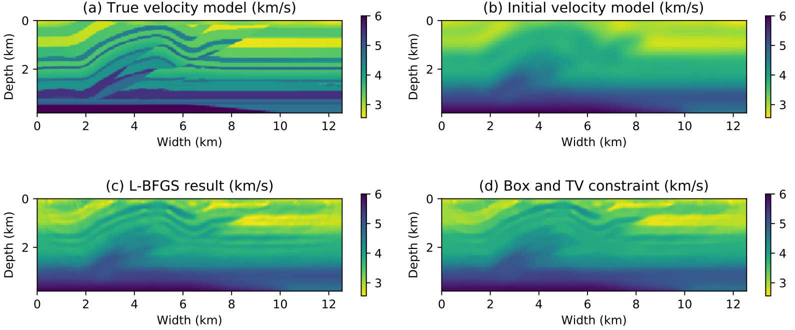
December 2020

NSERC CRSNG
UNIVERSITY OF CALGARY
FACULTY OF SCIENCE
Department of Geoscience

Table of contents

Table of Contents

Motivation



Initial idea:

- Total variation regularization and box constraint are very useful in the seismic inverse problem.
- Find a flexible way to incorporating multiple a priori information.

Formulation of the full waveform inversion

The full waveform inversion (FWI) problem is a PDE constrained optimization problem [Virieux and Operto, 2009]:

$$\min_{(y,u)\in Y\times U_{\text{ad}}} J(y,u) = \frac{1}{2}\|Qy - y_d\|_Y^2, \qquad (1)$$
$$\text{such that } e(y,u) = \mathcal{L}(u)y - s = 0,$$

Since the PDE $e(y,u) = 0$ is well-posed, the parameter-to-state map can be defined as $F(u) = y$. The problem (1) has a reduced form:

$$\min_{u\in U_{\text{ad}}} f(u) = J(F(u), u). \qquad (2)$$

The gradient of $f(u)$ can be achieved through the adjoint state method:

$$\nabla f(u) = \iint v(x,t)\partial_{tt} y(x,t)\, \mathrm{d}x\mathrm{d}t,$$

where v is the adjoint wavefield which is achieved by solving the adjoint equation.

In this work, we transform the a priori information of the model into convex constraint sets and then construct the feasible set as the intersection of the constraint sets. The seismic inverse problem is solved as a constraint optimization problem.

Table of Contents

Scaled gradient projection (SGP) method

Write the constrained optimization problem:

$$\min_{u\in\mathbb{R}^n} f(u), \quad \text{such that } u \in U_{\text{ad}}.$$

The scaled gradient projection (SGP) method is given by:

$$\bar{u}^k = \arg\min_{u\in U_{\text{ad}}} \left\langle \nabla f(u^k), u - u^k \right\rangle + \frac{1}{2\beta^k}\left\langle B_k(u-u^k), u-u^k \right\rangle, \tag{3}$$

$$u^{k+1} = u^k + \alpha^k(\bar{u}^k - u^k). \tag{4}$$

Let $\tilde{u}^k = u^k - B_k^{-1}\nabla f(u^k)$, the equation (3) is equivalent to

$$\bar{u}^k = \arg\min_{u\in U_{\text{ad}}} \frac{1}{2}\|u-\tilde{u}^k\|_{B_k}^2 - \frac{1}{2}\left\langle \nabla f(u^k), B_k^{-1}\nabla f(u^k)\right\rangle.$$

The SGP method is equivalent to

$$\begin{aligned}
\tilde{u}^k &= u^k - B_k^{-1}\nabla f(u^k),\\
\bar{u}^k &= P_{B_k,U_{\text{ad}}}(\tilde{u}^k),\\
u^{k+1} &= u^k + \alpha^k(\bar{u}^k - u^k).
\end{aligned}$$

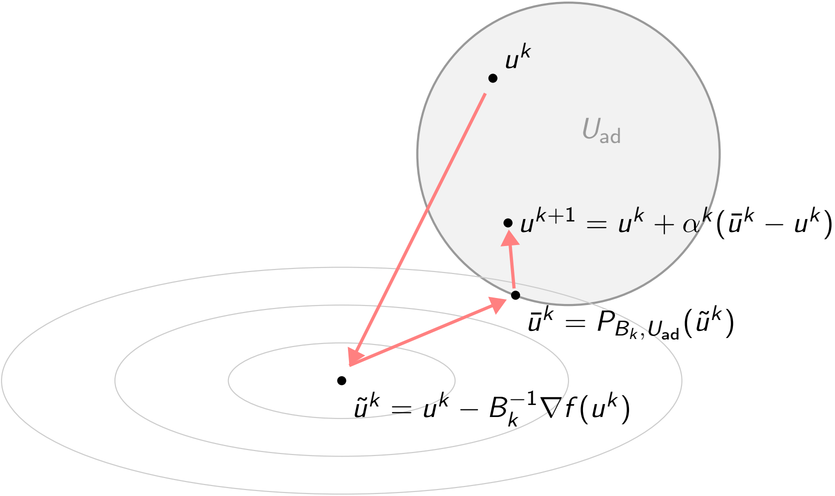


Figure: SGP method at the k-th iteration.

Scaled gradient projection (SGP) method

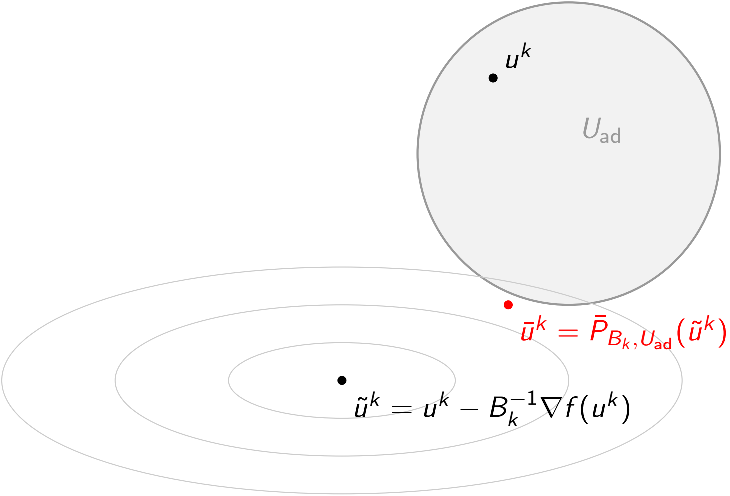


Figure: When an inexact projection result is generated by the projection algorithm, $\bar{u}^k$ may not in the feasible set U_{ad}, then u^{k+1} is not guaranteed in the feasible set U_{ad}.

Scaled gradient projection (SGP) method

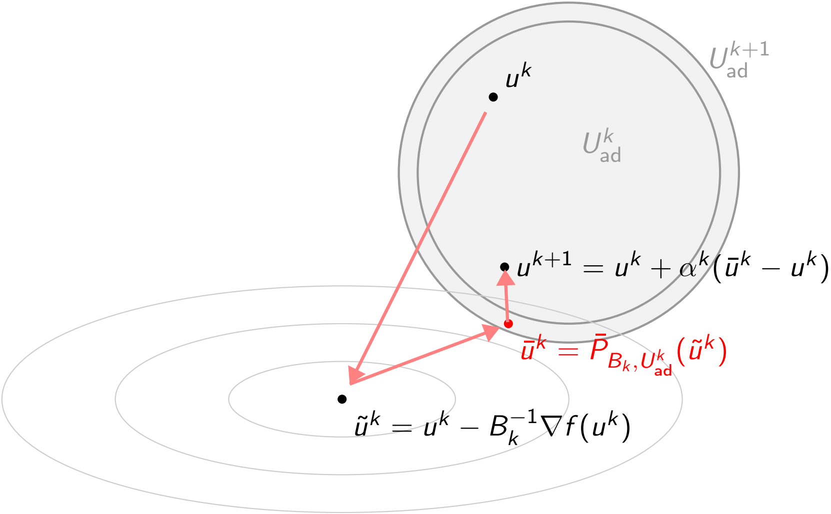


Figure: To overcome the inexact projection issue, we expand the feasible set at each iteration. Find the inexact projection $\bar{u}^k \in U_{\text{ad}}^{k+1}$ first, then update $u^{k+1} \in U_{\text{ad}}^{k+1}$.

Scaled gradient projection (SGP) method

The SGP method on the increasing sequence of feasible sets is: at k-th iteration, given symmetric positive definite matrix B_k (L-BFGS method),

1. Compute $\tilde{u}^k = u^k - B_k^{-1}\nabla f(u^k)$.
2. Evaluate the inexact projection operator $\bar{u}^k = \bar{P}_{B_k, U_{\mathrm{ad}}^k}(\tilde{u}^k)$, until the following equations are satisfied

$$\bar{u}^k \in U_{\mathrm{ad}}^{k+1}, \tag{5}$$

$$\left\langle \tilde{u}^k - \bar{u}^k, u^k - \bar{u}^k \right\rangle_{B_k} \leq 0. \tag{6}$$

   The equation (6) is a condition used in the convergence analysis and guarantee that the $\bar{u}^k - u^k$ is a decreasing direction.

3. Update $u^{k+1} = u^k + \alpha^k(\bar{u}^k - u^k)$, here α^k is determined by the linesearch algorithm.
4. Set $k = k + 1$, the feasible set at $k+1$-th iteration is U_{ad}^{k+1}.

Table of Contents

Constraint sets with projection function in closed form

We need the sets expanding as the iteration goes on, but not expand to infinity large. Given $\varepsilon > 0$ and $\xi \in (0,1)$, define a threshold function as

$$\theta(h) = \begin{cases} 0, & \text{if } h = 0, \\ \sum_{i=1}^{h} \xi^i \varepsilon, & \text{if } h \geq 1, \\ \frac{\xi}{1-\xi}\varepsilon, & \text{if } h \to \infty. \end{cases}$$

Given the box constraint set:

$$U_{\text{box}} = \{u \in \mathbb{R}^n \mid a \leq u_i \leq b, \; i = 1, \cdots, n\}.$$

The projection function is given by:

$$P_{\text{box}}(u)_i = \max(a, \min(u_i, b)).$$

Construct the sequence of box constraint sets as:

$$U_{\text{box}}^h = \{u \in \mathbb{R}^n \mid a - \theta(h) \leq u_i \leq b + \theta(h), \; i = 1, \cdots, n\}, \quad h \in \mathbb{N}.$$

Constraint sets with projection function in closed form

Given $p \in \mathbb{R}^n$ and $\kappa \in \mathbb{R}$, the affine hyperplane is

$$U_{\text{plane}} = \{u \in \mathbb{R}^n \mid \langle u, p \rangle = \kappa\}.$$

The projection function is given by:

$$P_{\text{plane}}(u) = u + \frac{\kappa - \langle u, p \rangle}{\|p\|^2} p.$$

Construct the sequence of hyperplane constraint sets as:

$$U^h_{\text{plane}} = \left\{u \in \mathbb{R}^n \mid \left\|u - P_{\text{plane}}(u)\right\| \leq \theta(h)\right\}, \quad h \in \mathbb{N}.$$

Constraint sets with subgradient projection

Given a continuous convex function g and a height η, a convex and closed lower level set can be constructed

$$C = \mathrm{lev}_{\eta} g = \{u \in \mathbb{R}^n \mid g(u) \leq \eta\}.$$

Approximate the set C with a half-space:

$$H_u = \{v \in \mathbb{R}^n \mid g(u) + \langle u^*, v - u\rangle \leq \eta\},$$

where $u^* \in \partial g(u)$, i.e. a subgradient of function g at the point u.

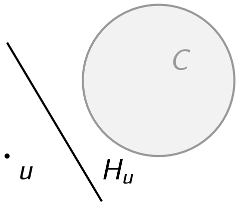


Figure: Approximation of C with half-space H_u.

Constraint sets with subgradient projection

Given a continuous convex function g and a height h, a convex and closed lower level set can be constructed

$$C = \mathrm{lev}_h g = \{u \in \mathbb{R}^n \mid g(u) \leq h\}.$$

Approximate the set C with a half-space:

$$H_u = \{v \in \mathbb{R}^n \mid g(u) + \langle u^*, v - u \rangle \leq h\},$$

where $u^* \in \partial g(u)$, i.e. a subgradient of function g at the point u.

The subgradient projection is given by

$$P_C(u) = \begin{cases} u + \frac{\eta - g(u)}{\|u^*\|^2} u^*, & \text{if } g(u) > \eta, \\ u, & \text{if } g(u) \leq \eta. \end{cases}$$

Constraint sets with subgradient projection: TV function

Consider the discrete total variation (TV) function for $u \in \mathbb{R}^{N_x \times N_y}$:

$$g_{\text{tv}}(u) = \|u\|_{\text{tv}} = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \left| (Du)_{i,j} \right|,$$

where the differential operator D is given by

$$(Du)_{i,j,1} = \begin{cases} u_{i+1,j} - u_{i,j}, & \text{if } 0 \leq i < N_x, \\ 0, & \text{if } i = N_x, \end{cases}, \quad (Du)_{i,j,2} = \begin{cases} u_{i,j+1} - u_{i,j}, & \text{if } 0 \leq j < N_y, \\ 0, & \text{if } i = N_y. \end{cases}$$

Given the radius τ_{tv}, we can construct a sequence of TV constraint sets:

$$U_{\text{tv}}^h = \{u \in \mathbb{R}^n \mid g_{\text{tv}}(u) \leq \theta(h) + \tau_{\text{tv}}\}.$$

The subgradient projection function of g_{tv} at point u is:

$$P_{U_{\text{tv}}^h}(u) = \begin{cases} u + \frac{\theta(k) + \tau_{\text{tv}} - g_{\text{tv}}(u)}{\|u^*\|^2} u^*, & \text{if } g_{\text{tv}}(u) > \theta(k) + \tau_{\text{tv}}, \\ u, & \text{if } g_{\text{tv}}(u) \leq \theta(k) + \tau_{\text{tv}}. \end{cases}$$

The subgradient of g_{tv} at point u is given by

$$\begin{aligned} u^* = & \sum_{(i,j)\in I_1} \sqrt{(u_{i+1,j}-u_{i,j})^2 + (u_{i,j+1}-u_{i,j})^2} \\ & \times \left[(u_{i+1,j}-u_{i,j})\, e_{i+1,j} - (u_{i+1,j} - 2u_{i,j} + u_{i,j+1})\, e_{i,j} + (u_{i,j}-u_{i,j+1})\, e_{i,j+1} \right] \\ & + \sum_{(i,j)\in I_2} \operatorname{sgn}(u_{N_x,j+1}-u_{N_x,j})\,(e_{N_x,j+1}-e_{N_x,j}) \\ & + \sum_{(i,j)\in I_3} \operatorname{sgn}\left(u_{i+1,N_y}-u_{i,N_y}\right)\left(e_{i+1,N_y}-e_{i,N_y}\right), \end{aligned}$$

where the index sets are given by

$$\begin{aligned} I_1 &= \{(i,j) \mid u_{i,j} \neq u_{i+1,j} \text{ and } u_{i,j} \neq u_{i,j+1},\ 1 \le i < N_x,\ 1 \le j < N_y\}, \\ I_2 &= \{(i,j) \mid u_{N_x,j} \neq u_{N_x,j+1},\ 1 \le j < N_y\}, \\ I_3 &= \left\{(i,j) \,\middle|\, u_{i,N_y} \neq u_{i+1,N_y},\ 1 \le i < N_x\right\}. \end{aligned}$$

Constraint sets with subgradient projection: l^1 function

Given a matrix $\Phi \in \mathbb{R}^{m \times n}$, with $\Phi = [\phi_1, \cdots, \phi_m]$, here each ϕ_i, $i = 1, \cdots, m$ is a n-dimensional row vector represents some basis of $\mathbb{R}^n$. Define the l^1 function:

$$g_{l^1}(u) = \|\Phi u\|_1 = \sum_{j=1}^{m} |\langle \phi_j, u \rangle| = \sum_{j=1}^{m} \left| \sum_{i=1}^{n} \Phi_{i,j} u_i \right| .$$

Given initial radius τ_{l^1}, construct the sequence of l^1 constraint sets:

$$U_{l^1}^h = \{ u \in \mathbb{R}^n \mid g_{l^1}(u) \leq \theta(h) + \tau_{l^1} \} ,$$

The subgradient projection function for the l^1 constraint set $U_{l^1}^h$ is

$$P_{U_{l^1}^h}(u) = \begin{cases} u + \frac{\theta(h) + \tau_{l^1} - g_{l^1}(u)}{\|u^*\|^2} u^*, & \text{if } g_{l^1}(u) > \theta(h) + \tau_{l^1}, \\ u, & \text{if } g_{l^1}(u) \leq \theta(h) + \tau_{l^1}. \end{cases}$$

A subgradient of g_{l^1} at point u is given by

$$u^* = \sum_{i=1}^{n} \sum_{j=1}^{m} \operatorname{sgn}\left(\langle \phi_j, u \rangle \right) \Phi_{i,j} e_i ,$$

Projection onto the intersection of convex sets

Given the convex sets U_i and $\tilde{u}$ and the symmetric positive definite matrix B, the projection problem is:

$$\text{find } \bar{u} = \arg\min_{u \in \mathbb{R}^n} \|u - \tilde{u}\|_B^2, \quad \text{such that } u \in U = \cap_{i \in I} U_i.$$

Algorithm 1: The projection algorithm provided in [Combettes, 2003]

Initialization: nonempty convex closed sets U_i, $x^0 = \tilde{u}$, $\{\omega_i\}$ with $\sum_{i \in I} \omega_i = 1$.

while *Not converge* **do**

Step 1: Compute the projection of x^k onto each of U_i with $p_i = P_{U_i}(x^k)$, where P_{U_i} is the projection function or subgradient projection function.

Step 2: Set $a_i = p_i - x^k$, $v = \sum_{i \in I} \omega_i a_i$, $\lambda = \sum_{i \in I} \omega_i \|a_i\|^2$. If $\lambda = 0$, $x^{k+1} = x^k$, break; otherwise $b = x^0 - x^k, c = Bb, d = B^{-1}v, \lambda = \lambda / \langle d, v \rangle$.

Step 3: Set $d = \lambda d$, $\pi = -\langle c, d \rangle, \mu = \langle b, c \rangle, \nu = \lambda \langle d, v \rangle, \rho = \mu\nu - \pi^2$, update

$$x^{k+1} = \begin{cases} x^k + d, & \text{if } \rho = 0 \text{ and } \pi \geq 0, \\ x^0 + \left(1 + \frac{\pi}{\nu}\right) d, & \text{if } \rho > 0 \text{ and } \pi\nu \geq \rho, \\ x^k + \frac{\nu}{\rho}\left(\pi b + \mu d\right), & \text{if } \rho > 0 \text{ and } \pi\nu < \rho. \end{cases}$$

end

Result: $\bar{u} = x^k$

Algorithm 2: Scaled gradient projection on sequence of multiple constraint sets

Given: the objective function f and initial value u^0; a family of nonempty, closed, convex constraint sets U_i.

For each U_i construct the set sequence $\{U_i^j\}_{j\in\mathbb{N}}$; set $U_{\mathrm{ad}}^k = \cap_{i\in I} U_i^k$ for all k.

while *Not converge* **do**

Step 1: Compute $\nabla f(u^k)$.

Step 2: Update the L-BFGS coefficients s_k, y_k, S_k, Y_k, R_k, D_k, U_k.

Step 3: Compute $\tilde{u}^k = u^k - H_k \nabla f(u^k)$.

Step 4: Compute $\bar{u}^k = \bar{P}_{B_k, U_{\mathrm{ad}}^k}(\tilde{u}^k)$, i.e., project $\tilde{u}^k$ to U_{ad}^k in $\mathcal{H}_{B_k}$ with Algorithm 1, until the following equations are satisfied:

$$\bar{u}^k \in U_{\mathrm{ad}}^{k+1},$$
$$\left\langle \tilde{u}^k - \bar{u}^k, u^k - \bar{u}^k \right\rangle_{B_k} \leq 0.$$

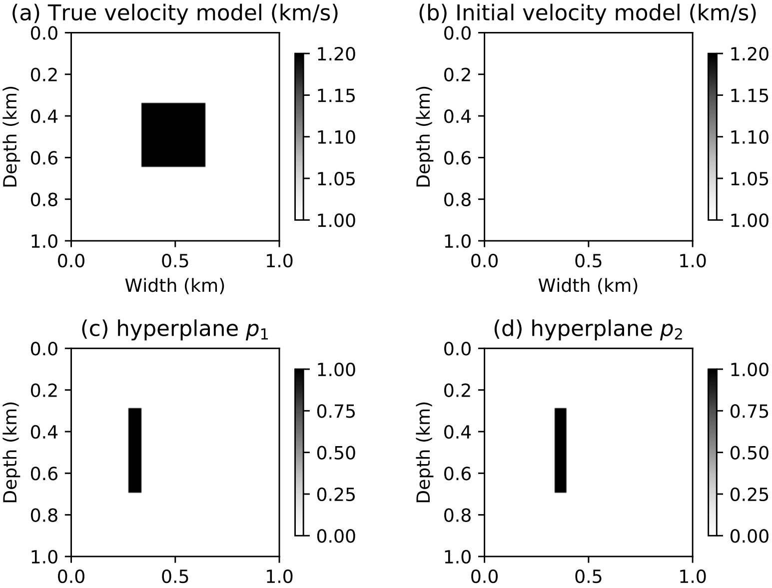
Step 5: Update $u^{k+1} = u^k + \alpha^k(\bar{u}^k - u^k)$, here α^k is the linesearch parameter.

Step 6: Set $k = k + 1$.

end

Table of Contents

Numerical example 1

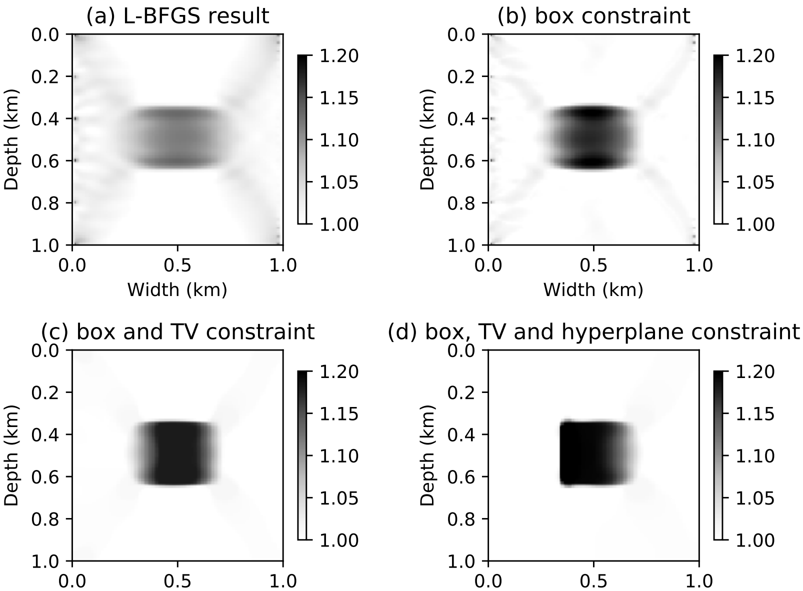


Numerical example 1

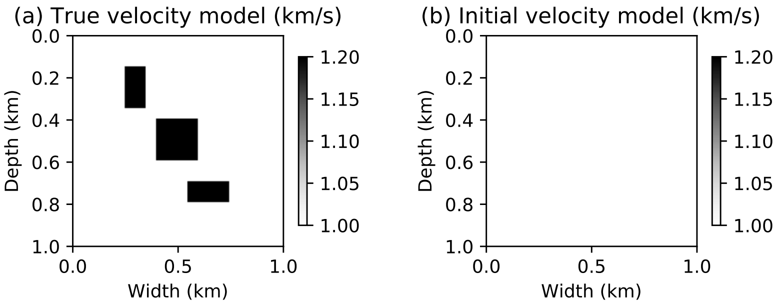
Settings of the sequences of constraint sets:

- $U_1^h = \{u \in \mathbb{R}^n \mid 1 - \theta_1(h) \leq u_i \leq 1.2 + \theta_1(h),\ i = 1, \cdots, n\}$, where $\theta_1(h) = \sum_{i=1}^h 0.001 \times 0.9^i$.
- $U_2^h = \{u \in \mathbb{R}^n \mid g_{\text{tv}}(u) \leq 24 + \theta_2(h)\}$, where $\theta_2(h) = \sum_{i=1}^h 0.24 \times 0.9^i$.
- $U_3^h = \{u \in \mathbb{R}^n \mid \|u - P_1(u)\| \leq \theta_3(h) + 0.01\}$, where
$$P_1(u) = u + \frac{\langle u_{\text{true}}, p_1\rangle - \langle u, p_1\rangle}{\|p_1\|^2} p_1, \quad \theta_3(h) = \sum_{i=1}^h 0.01 \times 0.9^i.$$
- $U_4^h = \{u \in \mathbb{R}^n \mid \|u - P_2(u)\| \leq \theta_4(h) + 0.01\}$, where
$$P_2(u) = u + \frac{\langle u_{\text{true}}, p_2\rangle - \langle u, p_2\rangle}{\|p_2\|^2} p_2, \quad \theta_4(h) = \sum_{i=1}^h 0.01 \times 0.9^i.$$

Numerical example 1



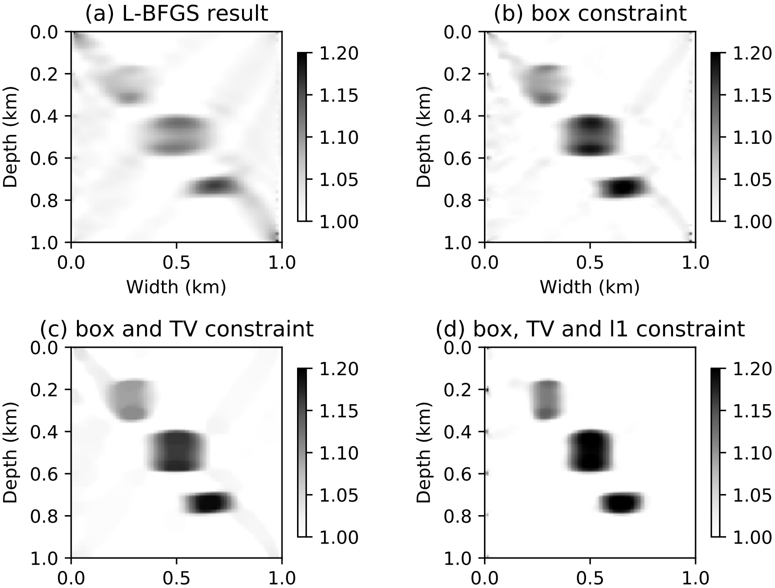
Numerical example 2

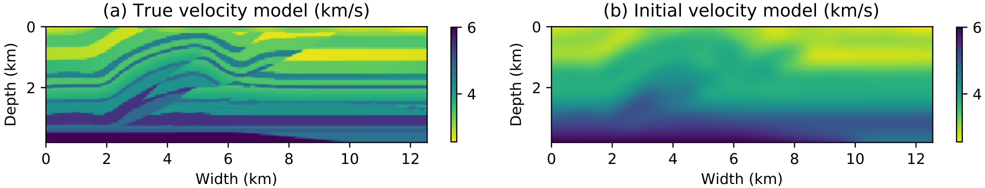


Settings of the sequences of constraint sets:

- $U_1^h = \{u \in \mathbb{R}^n \mid 1 - \theta_1(h) \leq u_i \leq 1.2 + \theta_1(h),\ \ i = 1, \cdots, n\}$, where $\theta_1(h) = \sum_{i=1}^h 0.001 \times 0.9^i$.
- $U_2^h = \{u \in \mathbb{R}^n \mid g_{\mathrm{tv}}(u) \leq 39.5 + \theta_2(h)\}$, where $\theta_2(h) = \sum_{i=1}^h 0.395 \times 0.9^i$.
- $U_3^h = \left\{u \in \mathbb{R}^n \,\middle|\, \|u - u^0\|_1 \leq 128 + \theta_3(h)\right\}$, where $\theta_3(h) = \sum_{i=1}^h 1.28 \times 0.9^i$.

Numerical example 2

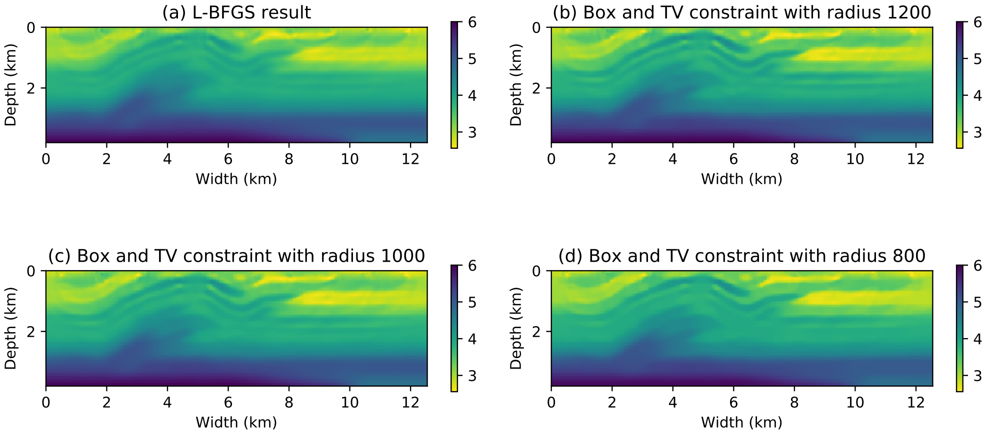




Settings of the sequences of constraint sets:

- $U_1^h = \{u \in \mathbb{R}^n \mid 2.5588 - \theta_1(h) \leq u_i \leq 6 + \theta_1(h),\ i = 1, \cdots, n\}$, where $\theta_1(h) = \sum_{i=1}^{h} 0.02 \times 0.9^i$.
- $U_2^h = \{u \in \mathbb{R}^n \mid g_{\mathrm{tv}}(u) \leq 800 + \theta_2(h)\}$, where $\theta_2(h) = \sum_{i=1}^{h} 40 \times 0.9^i$.
- $U_3^h = \{u \in \mathbb{R}^n \mid g_{\mathrm{tv}}(u) \leq 1000 + \theta_3(h)\}$, where $\theta_3(h) = \sum_{i=1}^{h} 50 \times 0.9^i$.
- $U_4^h = \{u \in \mathbb{R}^n \mid g_{\mathrm{tv}}(u) \leq 1200 + \theta_4(h)\}$, where $\theta_4(h) = \sum_{i=1}^{h} 60 \times 0.9^i$.

Numerical example 3



Conclusions

- The proposed method is flexible for incorporating multiple a priori information for the seismic inverse problem.
- The sequences of constraint sets behave like soft constraints, which encourage the inverse results to move towards the right direction instead of the fastest decreasing direction.
- More realistic reflective seismic inverse examples are needed.

Acknowledgement

- CREWES sponsors
- CREWES researchers
- China Scholarship Council and NSERC

Acknowledgement

Thank You!

References

Combettes, P. L. (2003).
A block-iterative surrogate constraint splitting method for quadratic signal recovery.
IEEE Transactions on Signal Processing, 51(7):1771–1782.

Virieux, J. and Operto, S. (2009).
An overview of full-waveform inversion in exploration geophysics.
Geophysics, 74(6):WCC1–WCC26.