Convergence Analysis for Multi-step One-shot Inversion Methods

Marcella Bonazzoli, Houssem Haddar, Tuan Anh Vu
INRIA, IDEFIX Project Team, Institut Polytechnique de Paris, France
tuan-anh.vu@polytechnique.edu

Abtract

We analyze the convergence of the gradient descent (GD) method to solve large-scale inverse problems, where the corresponding forward and adjoint problems are solved iteratively by fixed-point iteration methods.

1. Introduction

We study the *linear forward problem:*

$$u = Bu + M\sigma + F$$

where $u \in \mathbb{R}^{n_u}$ the state variable; $\sigma \in \mathbb{R}^{n_\sigma}$ the design variable; B, M, H are real matrices.

Inverse problem: Find σ from $f = Hu(\sigma) \in \mathbb{R}^{n_f}$.

Assumption: $\rho(B) < 1$ and $H(I-B)^{-1}M$ is injective.

Method of least squares with the cost function $J(\sigma) = \frac{1}{2}\|Hu(\sigma) - f\|^2$.

Lagrangian technique to define the adjoint state $p = p(\sigma)$: $p = B^*p + H^*[Hu(\sigma) - f]$.

Usual GD with fixed step $\tau > 0$:

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n, \\ u^n = Bu^n + M\sigma^n + F, \\ p^n = B^* p^n + H^*(Hu^n - f). \end{cases}$$

Shifted GD with fixed step $\tau > 0$:

$$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n, \\ u^{n+1} = Bu^{n+1} + M\sigma^n + F, \\ p^{n+1} = B^* p^{n+1} + H^*(Hu^{n+1} - f). \end{cases}$$

One-shot inversion methods: solve the forward and adjoint problems iteratively and *iterate at the same time* on u, p and σ.

Multi-step one-shot inversion methods: do k fixed-point interations on u and p. We study two variants of them (see section 2).

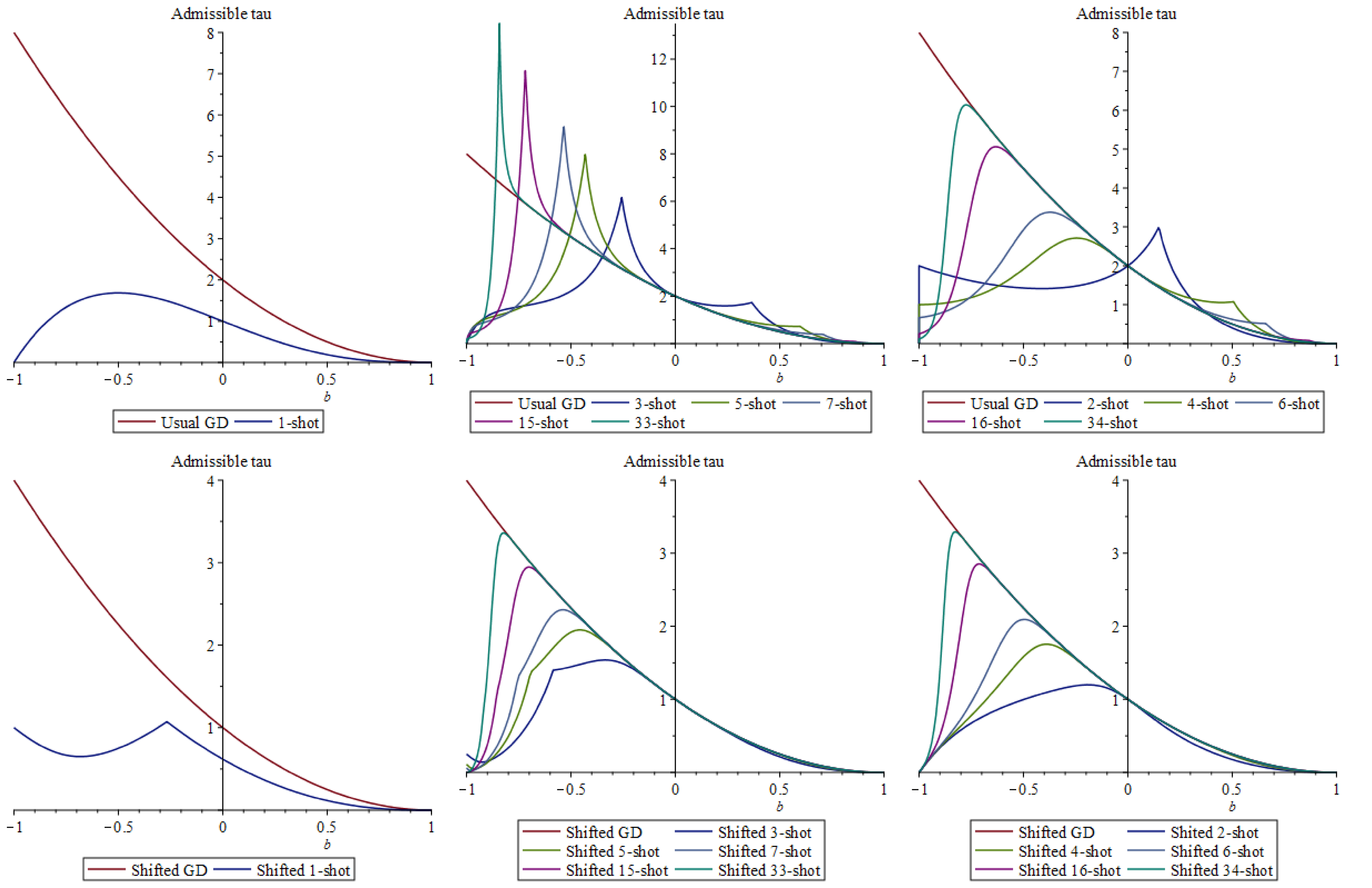
2. Multi-step one-shot algorithms

k-step one-shot	Shifted k-step one-shot
$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n, \\ u_0^{n+1} = u^n, p_0^{n+1} = p^n, \\ \quad u_{\ell+1}^{n+1} = Bu_\ell^{n+1} + M\sigma^{n+1} + F, \\ \quad p_{\ell+1}^{n+1} = B^* p_\ell^{n+1} + H^*(Hu_\ell^{n+1} - f), \\ u^{n+1} = u_k^{n+1}, p^{n+1} = p_k^{n+1} \end{cases}$	$\begin{cases} \sigma^{n+1} = \sigma^n - \tau M^* p^n, \\ u_0^{n+1} = u^n, p_0^{n+1} = p^n, \\ \quad u_{\ell+1}^{n+1} = Bu_\ell^{n+1} + M\sigma^{n} + F, \\ \quad p_{\ell+1}^{n+1} = B^* p_\ell^{n+1} + H^*(Hu_\ell^{n+1} - f), \\ u^{n+1} = u_k^{n+1}, p^{n+1} = p_k^{n+1} \end{cases}$
Converge to the usual GD as $k \to \infty$	Converge to the shifted GD as $k \to \infty$
Wait for σ before updating u, p	Update σ, u, p at the same time

3. Convergence analysis in 1D

Necessary and sufficient condition for the convergence			
Usual GD	Shifted GD	k-step one-shot	Shifted k-step one-shot
$\tau < \frac{2(1-b)^2}{h^2m^2}$	$\tau < \frac{(1-b)^2}{h^2m^2}$	$\tau < \frac{\eta(k,b)}{h^2m^2}$	$\tau < \frac{\kappa(k,b)}{h^2m^2}$

where κ and η are plotted below ($m = h = 1$):



*Note: B, M, H become $b, m, h \in \mathbb{R}$.

4. Main results

Theorem 1. *$\exists \tau > 0$ such that k-step one-shot converges. If $0 \le \|B\| < 1$, take*

$$\tau < \frac{\psi(k, \|B\|)}{\|H\|^2\|M\|^2}, \quad \psi \text{ is an explicit function.}$$

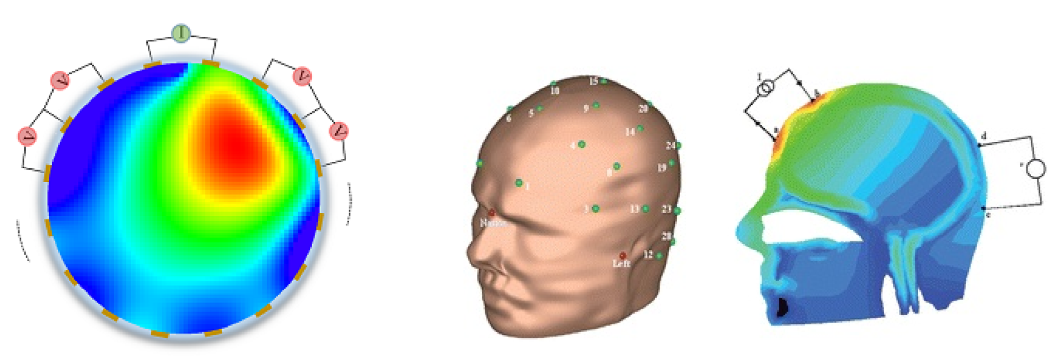
Theorem 2. *$\exists \tau > 0$ such that shifted k-step one-shot converges. If $0 \le \|B\| < 1$, take*

$$\tau < \frac{\chi(k, \|B\|)}{\|H\|^2\|M\|^2}, \quad \chi \text{ is an explicit function.}$$

5. Numerical results for a toy problem

Linearized conductivity inverse problem

Medical application to EIT (Electrical Impedance Tomography).



Forward problem ($\delta > 0$):

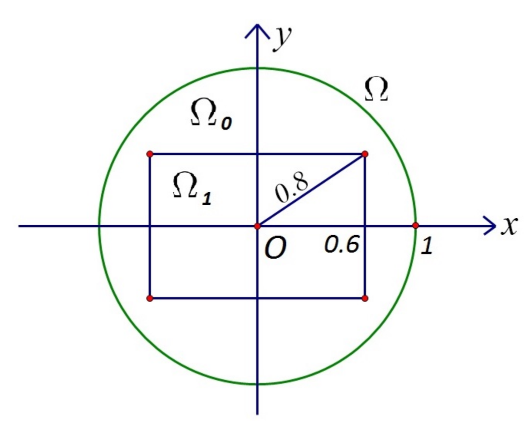
$$\begin{cases} -(1+\delta)\,\mathbf{div}(\sigma_0 \nabla u) + u = -\mathbf{div}(\sigma \nabla u_0) \text{ in } \Omega, \\ \sigma_0 \dfrac{\partial u}{\partial \nu} = 0, \sigma = 0 \text{ on } \partial\Omega \end{cases}$$

where u_0 satisfies

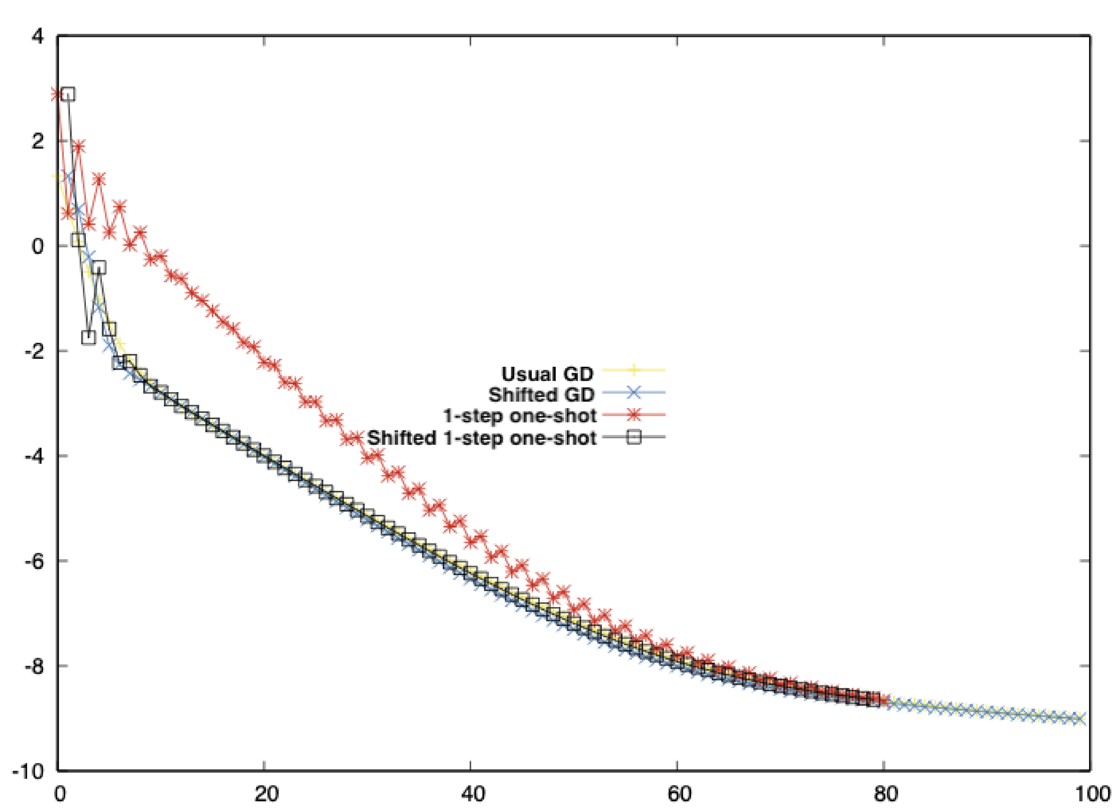
$$\begin{cases} -(1+\delta)\,\mathbf{div}(\sigma_0 \nabla u_0) + u_0 = 0 \text{ in } \Omega, \\ \sigma_0 \dfrac{\partial u_0}{\partial \nu} = g \text{ on } \partial\Omega. \end{cases}$$

Inverse problem: Recover σ from the measurement $f = Hu(\sigma) := u(\sigma)|_{\partial\Omega}$.

- $\delta = 0.5$
- Exact $\sigma^\star = 10 \cdot \mathbb{1}_{\Omega_1}$
- Multiple data $g_m = \cos(m\phi), 1 \le m \le 5$
- Initial guess $\sigma^0 = 15 \cdot \mathbb{1}_{\Omega_1}$
- $\tau = 1$



Log-plot of the cost function for each methods at different iterations:



6. Perspectives

1) Study the convergence when the iterative solver comes from DDM (domain decomposition methods)
2) Extend the analysis to non-linear inverse problems

7. References

[1] A. Griewank, *Projected Hessians for Preconditioning in One-Step One-shot Design Optimization* in Large-Scale Nonlinear Optimization, Springer (2006), 151-171
[2] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, Frontiers in Applied Mathematics (1997)
[3] M. Marden, *The geometry of the zeros of a polynomial in a complex variable*. Amer. Math. Soc., N.Y. (1949), 152