Preconditioning for the Hessian-free Gauss-Newton Full-waveform Inversion

Wenyong Pan*, Kris Innanen and Wenyuan Liao
wpan@ucalgary.ca

Summary

The gradient-based methods for FWI promise to converge globally but suffer from slow convergence rate. The Newton-type methods provide a quadratic convergence, but the computation, storage and inversion of the Hessian are beyond the current computation ability for large-scale inverse problem. The Hessian-free (HF) optimization method represents an attractive alternative to these above-mentioned optimization methods. At each iteration, it obtains the search direction by approximately solving the Newton linear system using a conjugate-gradient (CG) algorithm with a matrix-free fashion. One problem of the HF optimization method is that the CG algorithm requires many iterations. The main goal of this paper is to accelerate the HF FWI by preconditioning the CG algorithm. In this research, different preconditioning schemes for the HF Gauss-Newton optimization method are developed. The preconditioners are designed as Hessian approximations (e.g., diagonal pseudo-Hessian and diagonal Gauss-Newton Hessian) or its inverse approximations. We also developed a new pseudo diagonal Gauss-Newton Hessian approximation for preconditioning based on the reciprocal property of the Green's function. Furthermore, a quasi-Newton l-BFGS inverse Hessian approximation preconditioner with the diagonal Hessian approximation as initial guess is proposed and developed.

Hessian-free Gauss-Newton FWI

The Hessian-free optimization method (truncated-Newton or inexact-Newton method) represents an attractive alternative to the traditional optimization methods. At each iteration, the search direction is computed by approximately solving the Newton equations through a matrix-free fashion of the conjugate-gradient (CG) algorithm:

$$\mathbf{H}_k \Delta\mathbf{m}_k = -\mathbf{g}_k$$

This linear iterative solver only requires computation of the Hessian-vector products instead of forming the Hessian operator explicitly. In this paper, the full Hessian is replaced with the Gauss-Newton Hessian, which is always symmetric positive definite. One problem of the HF optimization method is that obtaining the search direction approximately requires a large number of CG iterations. Our main goal in this paper is to precondition the CG algorithm for reducing the CG iterations and accelerating the HF Gauss-Newton full-waveform inversion.

Preconditioning

One problem of the CG iterative algorithm is that it requires many iterations when obtaining the approximate solution of a linear system **Wx=b**. The convergence rate of the CG method depends on the spectral properties (e.g., its eigenvalues) of the coefficient matrix **W**. It is often convenient to transform the equation system into a system that has the same solution but has more favorable spectral properties. This can be achieved by applying a suitable preconditioner **M** on the linear system. Thus, the preconditioned Newton system for the HF Gauss-Newton FWI is given by:

$$\mathbf{M}_k^{-1}\left(\tilde{\mathbf{H}}_k + \epsilon\mathcal{A}\right)\Delta\mathbf{m}_k = -\mathbf{M}_k^{-1}\mathbf{g}_k$$

The preconditioner for the CG method is always devised to approximate the Hessian or the inverse Hessian. We first consider the traditional Hessian approximations (e.g., diagonal pseudo-Hessian and diagonal Gauss-Newton Hessian) as the preconditioners for the CG inner iteration.

We also develop an L-BFGS preconditioning scheme for the HF optimization method, namely the L-BFGS-GN method. Furthermore, the L-BFGS preconditioner is constructed with the diagonal Hessian approximations as initial guess.

Stopping Criteria

Newton's method is based on the Taylor series approximation. If this approximation is inaccurate then it may not be suitable to solve the Newton equations accurately and "over-solving" the Newton equation will not produce a better search direction. The CG algorithm should be terminated with an appropriate stopping criteria:

$$\gamma_{\tilde{k}} = \frac{\|\tilde{\mathbf{H}}_{\tilde{k}}\Delta\mathbf{m}_{\tilde{k}} + \mathbf{g}_{\tilde{k}}\|}{\|\mathbf{g}_{\tilde{k}}\|}$$

where $\tilde{k}$ indicates the CG inner iteration index. The inner iteration is stopped when $\gamma_{\tilde{k}} < \gamma_{min}$, where γ_{min} indicates the relative residual tolerance.

Numerical Results

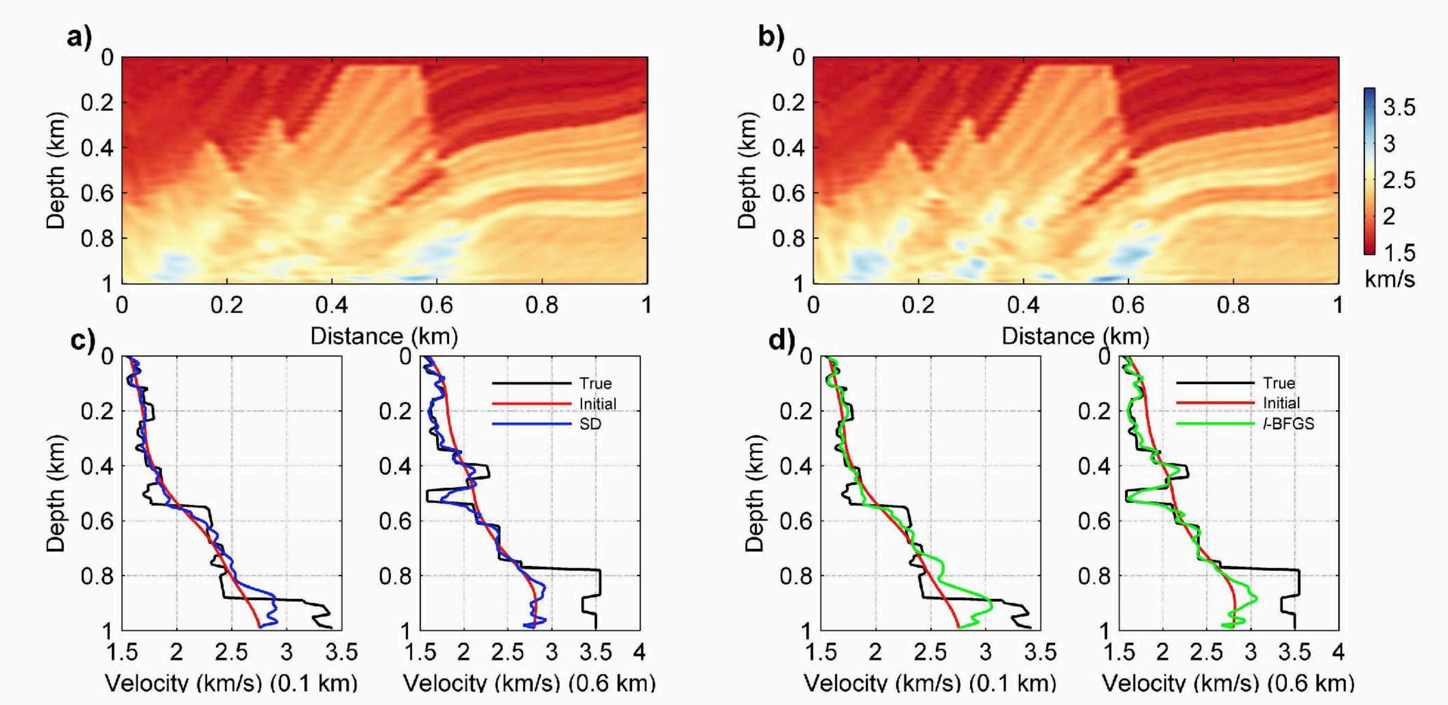

Figure 1. (a) SD method; (b) L-BFGS method; (c) and (d) are the well log data comparison.

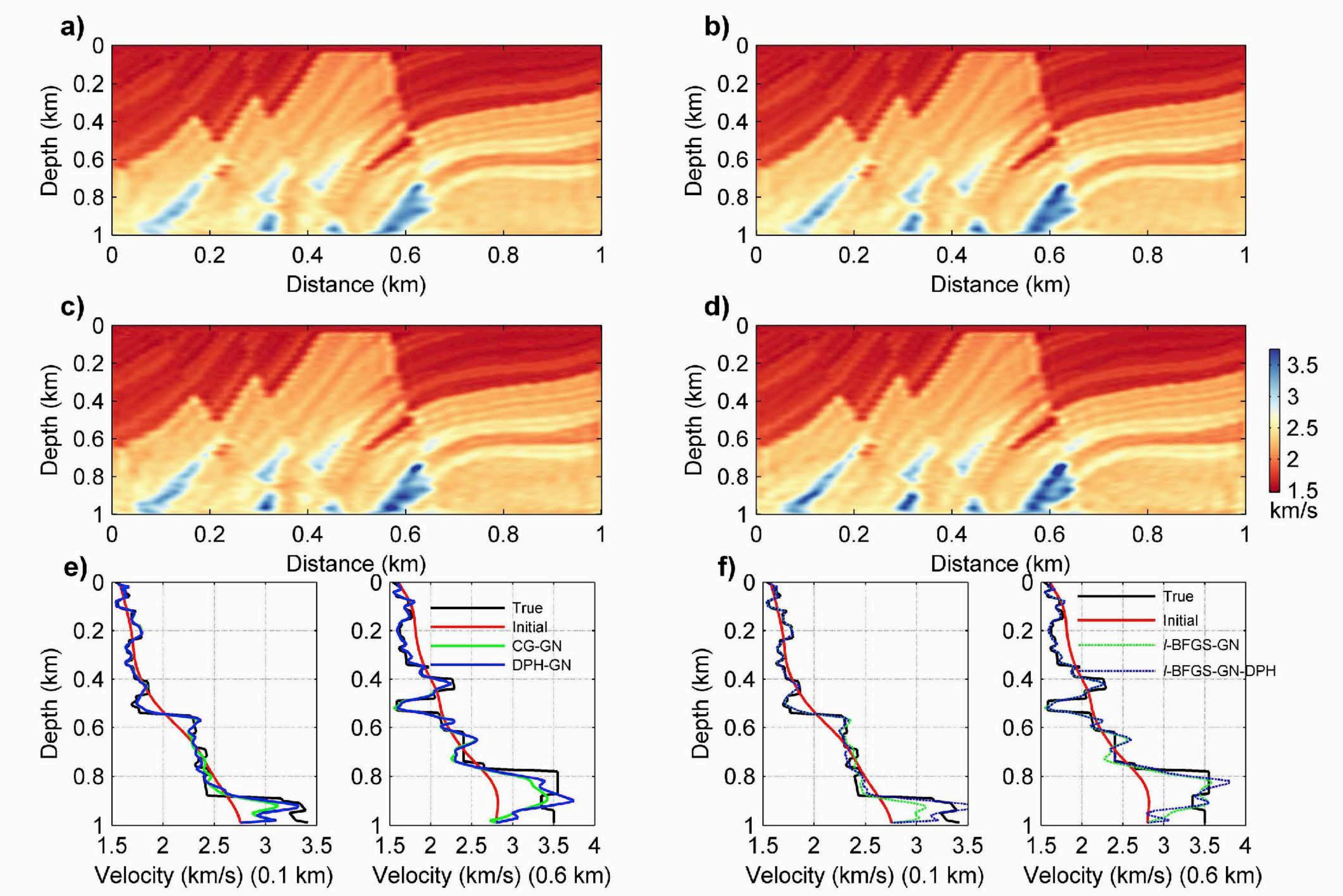

Figure 2. (a) GN-CG method; (b) L-BFGS-GN method; (c) DPH-GN method; (d) L-BFGS-GN-DPH method.

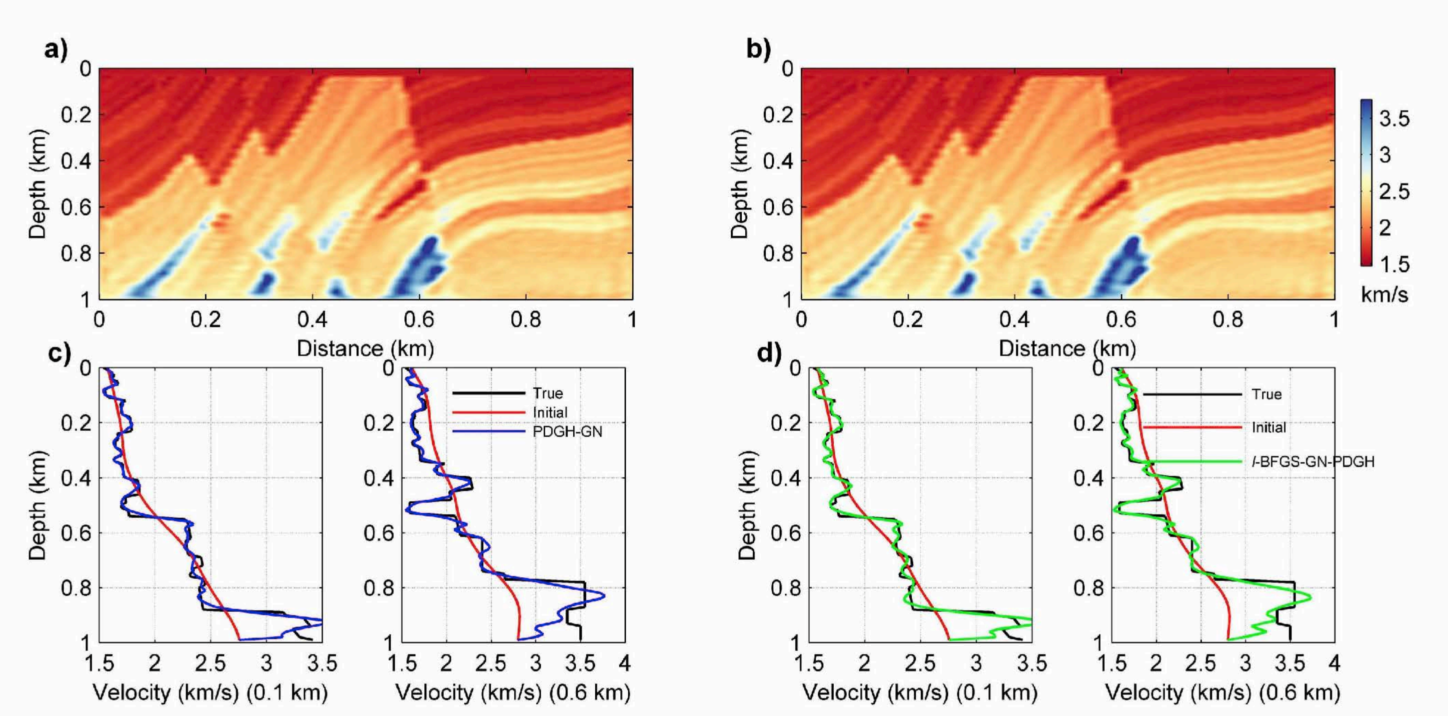

Figure 3. (a) PDGH-GN method; (d) L-BFGS-GN-PDGH method; (c) and (d) show the well log data comparison.

Acknowledgements

This research was supported by the Consortium for Research in Elastic Wave Exploration Seismology (CREWES).