The failure mode of correlation focusing for model velocity estimation

Hyoungsu Baek[1], Henri Calandra[2], Laurent Demanet[1]
[1] MIT Mathematics department, [2] TOTAL S.A.

January 15 2013

Abstract

We analyze the correlation focusing objective functional introduced by van Leeuwen and Mulder to avoid the cycle-skipping problem in full waveform inversion. While some encouraging numerical experiments were reported in the transmission setting, we explain why the method cannot be expected to work for general reflection data. We characterize the form that the adjoint source needs to take for model velocity updates to generate a time delay or a time advance. We show that the adjoint source of correlation focusing takes this desired form in the case of a single primary reflection, but not otherwise.

The correlation-focusing objective functional

The model velocity problem consists in inverting the low-wavenumber components of a wave speed profile c from waveform data $d_s(x_r, t)$ indexed by source s, receiver r, and time t. Call the corresponding prediction $u_s(x_r, t)$. The correlation-focusing objective functional of van Leeuwen and Mulder is

$$J[c] = \sum_{s,r} \frac{\displaystyle\int W(t) C_{s,r}^2(t)dt}{\displaystyle\int C_{s,r}^2(t)dt}, \tag{1}$$

where $C_{s,r}(t) = \int u_s(x_r, \tau) d_s(x_r, t + \tau)d\tau$. Note that $J[c]$ depends on the wave speed profile c through u_s. The weight function is chosen as $W(t) = t^2$ in this note, but our conclusions do not depend on this particular form. Provided u_s shows a delay/advance with respect to d_s as a function of t, but is an otherwise comparable waveform, their cross-correlation $C_{s,r}$ will peak at a time offset from zero by this delay [2]. Consequently, minimizing (1) is heuristically expected to resolve traveltime discrepancies.

In all our tests the reflectors are supposed known. The velocity model is updated with the steepest descent method, then smoothed by projection onto a space of B-splines. Gradients are computed using the adjoint state method. Define the adjoint source $f_{adj}(t)$ as the input in the right-hand-side of the adjoint-state wave equation, whose solution is then used in the imaging condition in a standard fashion. The adjoint source is simply the residual $d - u$ in least-squares minimization, but for the correlation-focusing objective it is (in prestack form)

$$f_{adj}(t) = 2 \sum_r E_{s,r}^{-1} \int (W(\tau) - J) C_{s,r}(\tau) d(x_r, t + \tau)d\tau = \sum_r k(t, t')u_s(x_r, t')dt'. \tag{2}$$

The kernel $k(t, t') = 2E_{s,r}^{-1} \displaystyle\int (W(\tau) - J)d_s(x_r, t + \tau)d_s(x_\tau, t' + \tau)d\tau$ is symmetric.

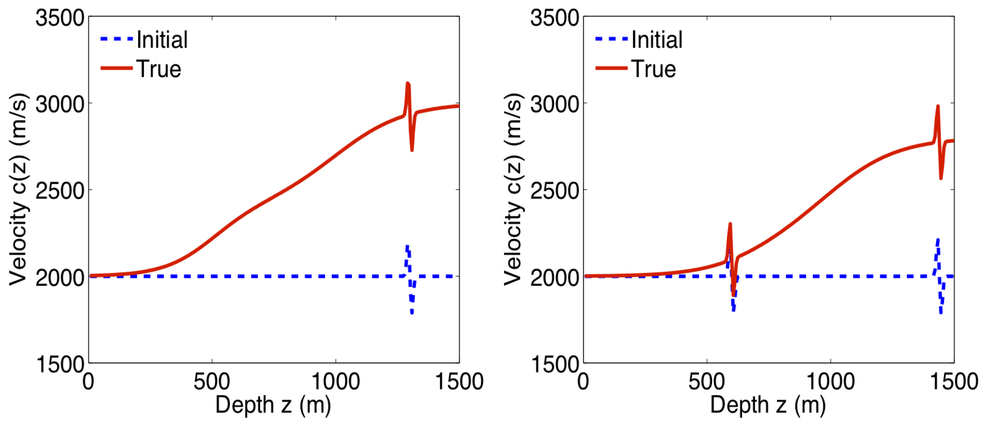

Figure 1: Layered velocity models used in the numerical examples. (Left) velocity models for example 1. A flat reflector is located at $z = 1300$ m. (Right) velocity models for example 2. Two flat reflectors are at $z = 600$ and 1440 m.

Traces are (ideally) composed of separate waves that correspond to individual reflection events. Let us first consider the case of a single wave arriving at time t_d for d_s, and time t_u for u_s. If in addition the waves are assumed impulsive in the sense that $d_s(x_r, t) \sim \delta(t - t_d)$, $u_s^2(x_r, t) \sim \delta(t - t_u)$, up to multiplicative scalars, then the adjoint source becomes $f_{adj}(t) \sim (t - t_u)(t - 2t_d + t_u)u_s(t)$. For t close to t_u, we further simplify $f_{adj}(t) \sim (t - t_u)(t_u - t_d)u_s(t)$. The combination $f_{adj}(t)u_s(t)$ is particularly informative: its support coincides with that of $u_s(t)$, and its sign goes from negative to positive at $t = t_u$ if $t_u > t_d$, or from positive to negative if $t_u < t_d$.

We claim that this sign property of $f_{adj}(t)u_s(t)$ is precisely what guarantees a good model velocity update. In the next section, we demonstrate that the reasoning above is corroborated by the numerical experiments when each trace contains a single reflected wave. In that case, the "good" model update consists of a sensitivity kernel concentrated along the broken rays linking sources to receivers, and is entirely positive or entirely negative (after smoothing.) However, when each trace contains *two* reflected waves, the sign property of $f_{adj}(t)u_s(t)$ no longer holds. In that case, the "bad" model update is still located near the broken ray, but has oscillations in the transverse direction that make it act as a waveguide. Rather than slowing down or speeding up the waves, a bad model update adjusts the amplitude and the shape of each wave in an unintended manner.

Numerical examples

The acoustic wave equation in an isotropic heterogeneous medium is solved for both observed data and predicted data. The velocity models used in this study are layered with a magnitude gradually increasing in depth, as plotted in Figure 1. Data are generated in the media labelled "True", while the inversion is initialized in the media labelled "Initial". Reflectors are assumed to be known a priori in this study. The forward problem and adjoint state equations are solved with a finite difference solver of second order in time and fourth order accuracy in space; the step size for time marching is 2.5×10^{-4} sec. The size of the computational domain is 500×250 grid points with grid spacing 6 m in both directions. The center frequency of the source Ricker wavelet is 20 Hz. Perfectly matched layer (PML) boundaries are used to avoid spurious reflections.

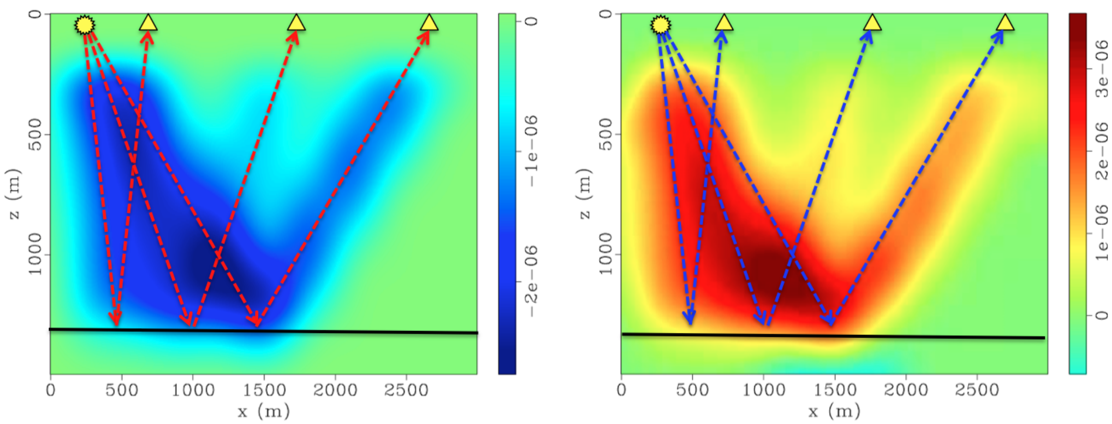

Figure 2: This is a color figure. Velocity updates (gradients) projected onto B-splines spaces. (Left) velocity update when the velocity models are given as in Figure 1. (Right) velocity update when the initial/true velocity models in Figure 1 are swapped. Black lines correspond to reflectors. Dotted red and blue lines schematically show the wave paths.

Example 1: the single reflector case

A source at $x = 200$ m and three receivers at $x = 700, 1700, 2700$ m near the surface are marked in Figure 2. A velocity update in Figure 2 (left) is obtained using the reference and initial velocity models in Figure 1 (left). Figure 2 (right) shows a velocity update with *swapped* initial and true models. Such swapping of models generates an update of opposite sign and similar magnitude. As a result, the velocity is correctly updated along the broken wave paths in both situations. For a convergence study, we refer the reader to a numerical example in [1].

We confirm the properties of the adjoint source $f_{adj}(t)$ in a case when $t_d > t_u$. Figure 3 (top) compares one trace of predicted data with the corresponding adjoint source. Notice that $f_{adj}(t)$ becomes zero near the peak of $u(t)$, mostly has the same sign as $u(t)$ to the left of the root, and mostly has the opposite sign of $u(t)$ to the right of the root. This observation agrees with our analytical result that $f_{adj}(t) \sim -(t - t_u)u(t)$ for t close to t_u, where t_u is the arrival time of the wave in the predicted trace $u(t)$. As suggested earlier, it is also instructive to form the combination $f_{adj}(t)u(t)$; we see from Figure 3 (bottom) that it is in very good agreement with the theorized $\sim -(t-t_u)u^2(t)$. As long as the data have a single wave in the trace, the adjoint source has such a pattern, which we have seen results in a “good” velocity update. (The actual zero crossing of the adjoint source is slightly offset from the arrival time t_u due to the limited accuracy of the numerical simulation.)

Example 2: the multiple reflector case

In the following example, the velocity model is shown in Figure 1 (right) with two flat reflectors at $z = 600, 1300$ m. Hence the traces have two waves and their cross-correlation has three peaks; the one marked with ‘Second’ in Figure 4 (top) is spurious (cross-talk). The adjoint source does not show any signs of multiplication by $(t - t_u)(t_d - t_u)$ as seen in the previous example; the first (third) piece are in and out of phase with respect to the predicted data $u(t)$.

In order to see the effect of each correlation peaks on the gradient, the adjoint source is split into three pieces. The pieces of gradient shown in Figure 5 (left, center) are the results of feeding the “First” and “Third” pieces in the adjoint state equation, respectively. The gradient from the

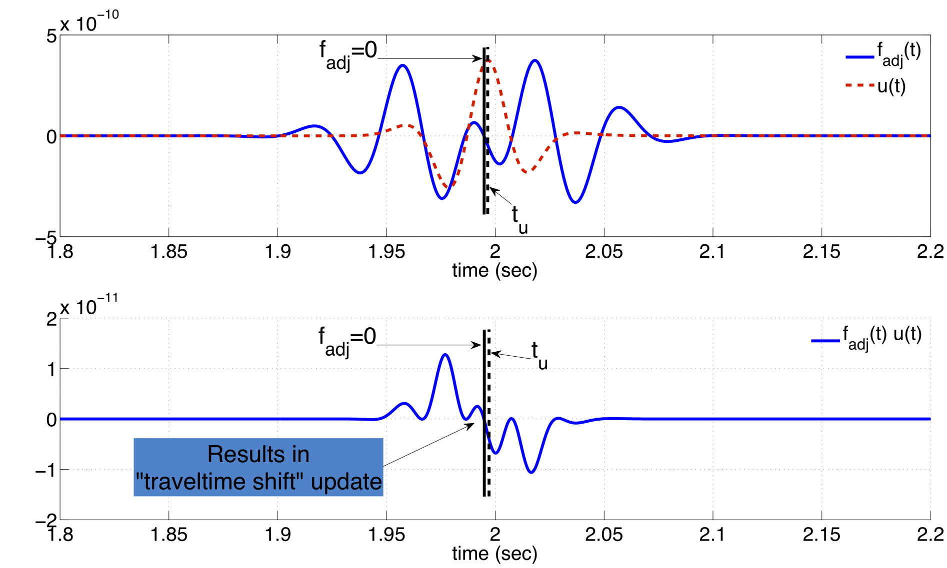

Figure 3: Top: comparison of the computed adjoint source $f_{adj}(t)$ (blue solid line) and the predicted data $u(t)$ (red dotted line). Bottom: plot of $f_{adj}(t)u(t)$. The sign changes from positive to negative near the peak of predicted data. The black solid (dashed) vertical line marks the instant when the adjoint source becomes zero (when the peak of predicted data $u(t)$ arrives), respectively.

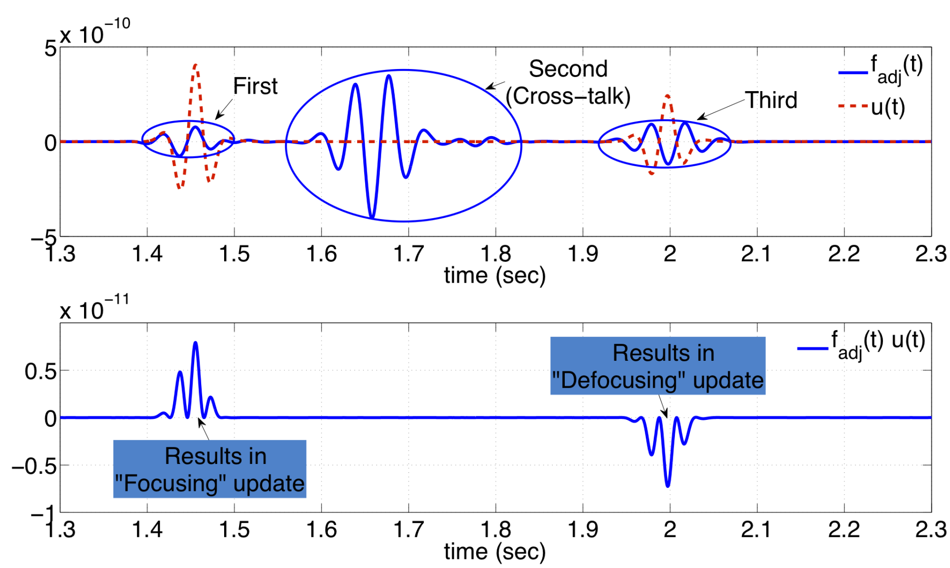

Figure 4: Top: comparison of the computed adjoint source $f_{adj}(t)$ (blue solid line) and the predicted data $u(t)$ (red dotted line). 'First' and 'Third' indicate the pieces of the source which are used to form the partial gradients in Figure 5 (left) and (center). Bottom: multiplication of source and predicted data. There is no sign change near the peaks of the predicted data.

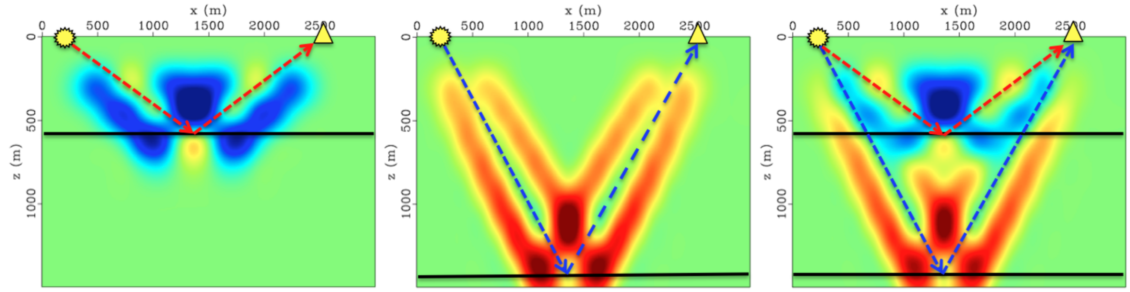

Figure 5: Gradients: (left) gradient from the first piece of the adjoint source, (center) gradient from the third pieceof the adjoint source, and (right) gradient from the entire source.

cross-talk signal "Second" in Figure 4 (top) as an adjoint source is negligibly small compared to others shown in Figure 5 (left and center) and is not shown here.

These updates help minimize the objective function in an unintended way. The update direction in Figure 5 does not shift the peak of correlation toward zero, but instead *increases the denominators* $\int C_(s,r)^2(t)dt$. Physically, this effect is achieved because the model update is the superposition of two waveguide-like sensitivity kernels: one that focuses the rays to strengthen the first wave (Figure 5, left), and one that defocuses the rays to weaken the second wave (Figure 5, center). The explanation for this phenomenon can in turn be traced back to the loss of the $(t - t_u)$ multiplication pattern in the adjoint source. Note that the cross-talk in the correlation is not responsible for this behavior, since it does contribute meaningfully to the update in this case. Note also that iterating wrong model updates does not in general salvage their deficiencies.

One possible solution to avoid this failure mode would be to forbid the minimization from selectively shifting energy between the different waves making up the traces. This could be achieved by an objective of the form

$$J = \sum_{s,r,k} \frac{\int w(t) C^2_{s,r,k}(t)dt}{\int C^2_{s,r,k}(t)dt}, \tag{3}$$

where k indexes each wave in the trace. However, this idea would involve manually picking and matching events between the predicted and observed data.

Conclusions

Correlation-focusing waveform inversion can update low-wavenumber velocity models in the reflection setting, but only in the case of a single reflected wave. The explanation of success vs. failure lies in the sign structure of the adjoint source, not in the presence of correlation cross-talk.

Acknowledgements

The authors are grateful to the authors of Madagascar, and to TOTAL S.A. for support.

References

[1] H. Baek and L. Demanet. Conditioning bounds for traveltime tomography in layered media. *Inverse Problems*, 28(5):055008, 2012.

[2] Y. Luo and G. T. Schuster. Wave-equation traveltime inversion. *Geophysics*, 56(5):645–653, 1991.

[3] T. van Leeuwen and W. A. Mulder. Velocity analysis based on data correlation. *Geophysical Prospecting*, 56(6):791–803, 2008.

[4] T. van Leeuwen and W. A. Mulder. A correlation-based misfit criterion for wave-equation traveltime tomography. *Geophysical Journal International*, 182(3):1383–1394, 2010.