The failure mode of correlation focusing for model velocity estimation

Hyoungsu Baek[1(*)], Henri Calandra[2], and Laurent Demanet[1]
[1]Dept. of Mathematics and Earth Resources Lab, Massachusetts Institute of Technology, [2]TOTAL S.A.

SUMMARY

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. Ultimately, failure owes to the specific form of the normalization present in the correlation focusing objective.

INTRODUCTION

Traveltime inversion has been traditionally used to recover background velocity models, the indispensable starting guesses for full waveform inversion (Virieux and Operto, 2009; Prieux et al., 2012; Bregman et al., 1989). It exploits traveltime shifts in the arrivals of waves in observed data and predicted data. More recently, waveform data have also been incorporated into traveltime inversion in order to relax the high frequency approximation. For example, traveltime inversion from peaks of crosscorrelations in the data domain has been proposed and studied in Luo and Schuster (1991). Crosscorrelations in the image domain, between forward wavefield from a source and back-propagated wavefield from corresponding receivers, have also been used to extract traveltime shifts with the aid of time windowing (Zhang et al., 2011).

In the same spirit of hybridizing traveltime inversion and waveform inversion (Pratt and Goulty, 1991), a different approach was proposed by van Leeuwen and Mulder (2008, 2010). Instead of minimizing the traveltime discrepancies, defocusing of the crosscorrelation is minimized. Some very encouraging numerical examples were reported in the transmission setting, yet the method may occasionally run into trouble with the inversion of reflection data (van Leeuwen and Mulder, 2010). To better understand the issue, we analyze the form of the adjoint source using the high frequency approximation, and explain which form yields good vs. bad velocity updates, from data with multiple waves in the reflection setting.

THE CORRELATION-FOCUSING OBJECTIVE FUNCTIONAL

The model velocity problem consists in inverting the low wavenumber components of a wave speed profile c, or slowness squared $m = 1/c^2$ from waveform data, $d_s(x_r, t)$, indexed by source s, receiver r, and time t. We call the corresponding prediction $u_s(x_r, t)$. The correlation-focusing objective functional of van Leeuwen and Mulder is

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

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 (Luo and Schuster, 1991). 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)

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

The kernel

$$k(t, t') = 2 E_{s,r}^{-1} \int (W(\tau) - J) d_s(x_r, t + \tau) d_s(x_\tau, t' + \tau) d\tau \tag{3}$$

is symmetric.

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

$$d_s(x_r, t) \sim \delta(t - t_d), \quad u_s^2(x_r, t) \sim \delta(t - t_u), \tag{4}$$

up to multiplicative scalars, then the adjoint source becomes

$$f_{adj}(t) \sim (t - t_u)(t - 2t_d + t_u) u_s(t). \tag{5}$$

For t close to t_u, we further simplify

$$f_{adj}(t) \sim (t - t_u)(t_u - t_d) u_s(t). \tag{6}$$

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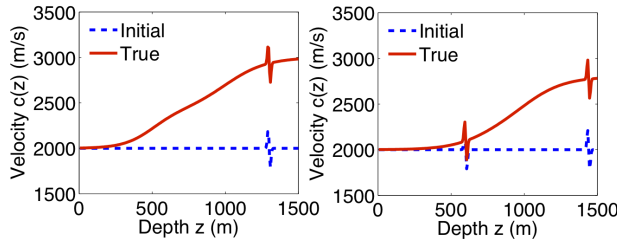
Figure 1: Layered velocity models used in the numerical examples. (left) velocity models for example 1. A flat reflector is located at $z = 1300$ m. (right) velocity models for example 2. Two flat reflectors are at $z = 600$ and 1440 m.

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(a) is obtained using the reference and initial velocity models in Figure 1 (left).Figure 2(b) 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 Baek and Demanet (2012).
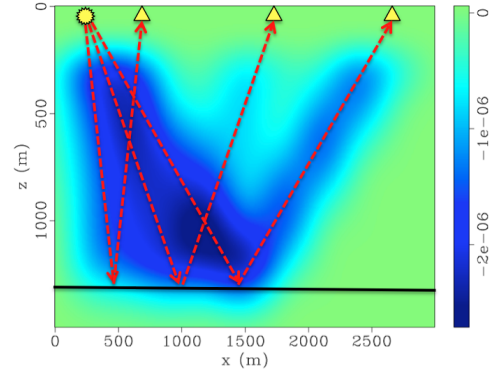
(a)
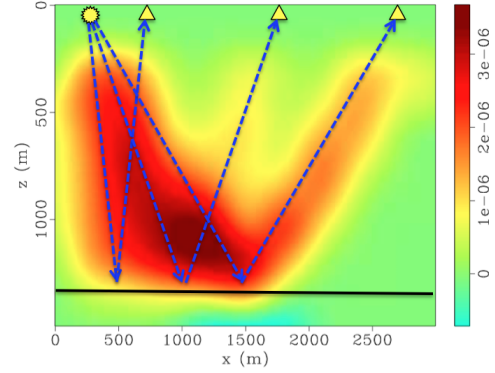
(b)

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

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 Fig-

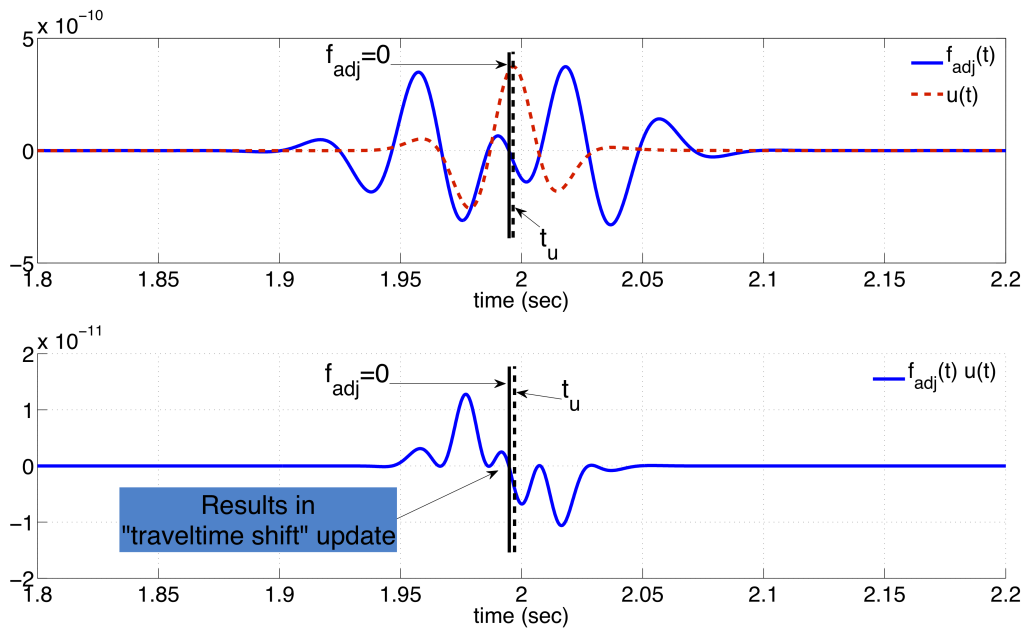**Correlation focusing for model velocity estimation**

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.
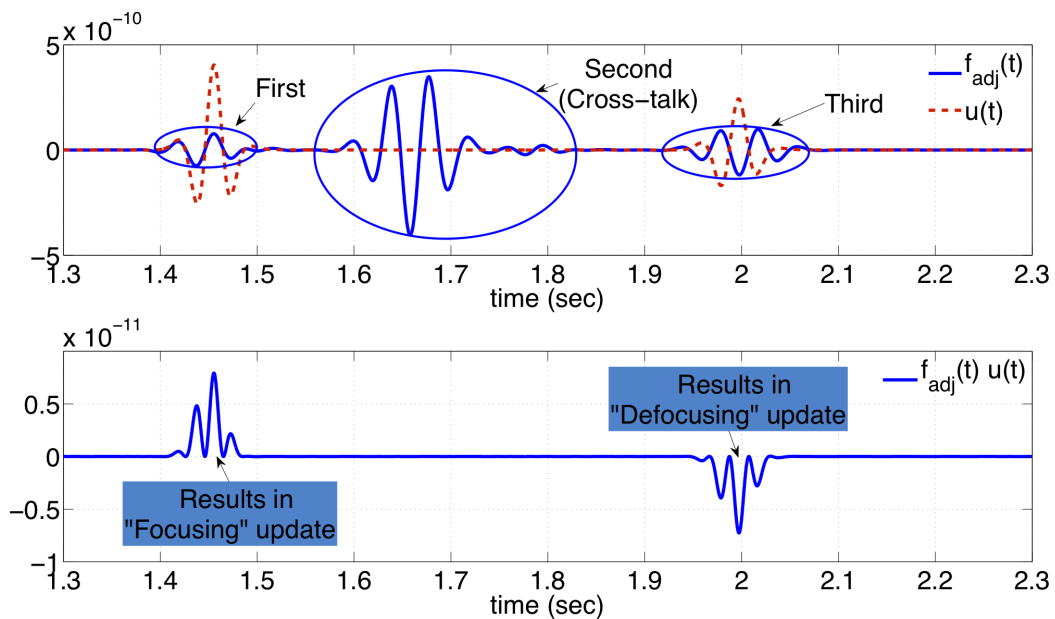

Figure 4: Top: comparison of the computed adjoint source $f_{adj}(t)$ (blue solid line) and the predicted data $u(t)$ (red dotted line). "First" and "Third" indicate the pieces of the source which are used to form the partial gradients in Figures 5(a) and 5(b). Bottom: $f_{adj}(t)u(t)$, multiplication of the adjoint source and the predicted trace. There is no sign change near the peaks of the predicted data.

Correlation focusing for model velocity estimation

ure 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 is in (out of) phase with respect to the predicted data $u(t)$.
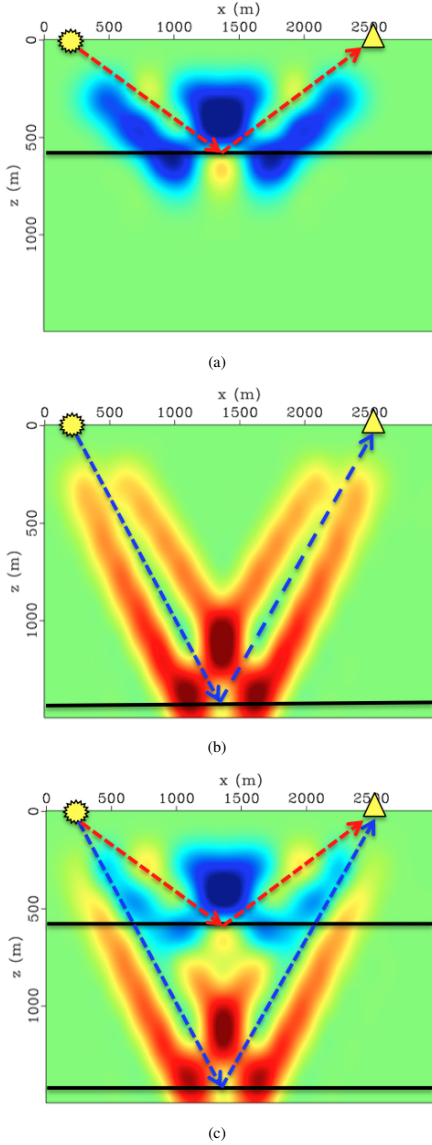

(a)

(b)

(c)

Figure 5: Gradients: (a) gradient from the first piece of the adjoint source, (b) gradient from the third piece of the adjoint source, and (c) gradient from the entire source.

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 Figures 5(a) and 5(b) are the results of feeding the "First" and "Third" pieces in the adjoint state equation, respectively. The gradient from the cross-talk signal "Second" in Figure 4 (top) as an adjoint source is negligibly small compared to others shown in Figures 5(a) and 5(b) and is not shown

here.

These updates decrease the objective function in an unintended way. The update direction in Figure 5(c) does not shift the peak of correlation toward zero, but instead *increases the denominators* $\int C_{s,r}^2(t)dt$. Physically, this behavior results from the model update being the superposition of two waveguide-like sensitivity kernels: one that focuses the rays to strengthen the first wave (see Figure 5(a)), and one that defocuses the rays to weaken the second wave (see Figure 5(b)). 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 not contribute meaningfully to the update in this case. Note also that iterating wrong model updates does not in general salvage their deficiencies.

We anticipate that moment-matching methods with a similar normalizing denominator as in (7) would suffer from the same failure mode **in the reflection setting**. Such would be the case of the normalized moment method presented in Liu et al. (2011), consisting in minimizing

$$J = \sum_{s,r} \int (m_u^p(t) - m_d^p(t))^2 dt, \tag{7}$$

where $m_u^p(t) = \dfrac{\int_0^t |u_s(x_r, \tau)|^p d\tau}{\int_0^T |u_s(x_r, \tau)|^p d\tau}$, and $m_d^p(t)$ is defined in the same way using $d_s(x_r, \tau)$. We hasten to add that a successful application of moment matching in the transmission setting is given in Liu et al. (2011).

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 considering an objective of the form

$$J = \sum_{s,r,k} \frac{\int w(t) C_{s,r,k}^2(t) dt}{\int C_{s,r,k}^2(t) dt}, \tag{8}$$

where k indexes each wave in the trace. However, this idea would involve manually picking and matching events between the predicted and observed data. Time-windowing the cross-correlations could have the same effect as the splitting into separate waves in the objective functional.

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.

ACKNOWLEDGMENTS

REFERENCES

Baek, H., and L. Demanet, 2012, Conditioning bounds for traveltime tomography in layered media: Inverse Problems, **28**, 055008.

Bregman, N. D., R. C. Bailey, and C. H. Chapman, 1989, Crosshole seismic tomography: Geophysics, **54**, 200–215.

Liu, J., H. Chauris, and H. Calandra, 2011, The normalized integration method - an alternative to full waveform inversion?: Near surface 2011 - the 1t7th European meeting of environmental and engineering geophysics.

Luo, Y., and G. T. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.

Pratt, R. G., and N. R. Goulty, 1991, Combining wave-equation imaging with traveltime tomography to form high-resolution images from crosshole data: Geophysics, **56**, 208–224.

Prieux, V., G. Lambaré, S. Operto, and J. Virieux, 2012, Building starting models for full waveform inversion from wide-aperture data by stereotomography: Geophysical Prospecting, no–no.

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

———, 2010, A correlation-based misfit criterion for wave-equation traveltime tomography: Geophysical Journal International, **182**, 1383–1394.

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

Zhang, S., G. Schuster, and Y. Luo, 2011, Wave-equation reflection traveltime inversion: 81th Annual International meeting, SEG, Expanded Abstracts, 2705–2708.