Geophysical Journal International

Geophys. J. Int. (2017) **209,** 695–712
Advance Access publication 2017 February 3
GJI Seismology

doi: 10.1093/gji/ggx042

Parsimonious refraction interferometry and tomography

Sherif Hanafy[1,2] and Gerard T. Schuster[1]

[1] *Department of Earth Science and Engineering, King Abdullah University of Science and Technology, Thuwal, Jeddah* 23955-6900, *Saudi Arabia. E-mail: sherif.geo@gmail.com*
[2] *Geophysics Department, Faculty of Science, Cairo University, Cairo, Egypt*

Accepted 2017 February 2. Received 2017 January 25; in original form 2016 September 24

SUMMARY

We present the theory of parsimonious refraction interferometry and tomography where a densely populated refraction data set can be obtained from two reciprocal and several infill shot gathers. The assumptions are that the refraction arrivals are head waves, and a pair of reciprocal shot gathers and several infill shot gathers are recorded over the line of interest. Refraction traveltimes from these shot gathers are picked and spawned into $O(N^2)$ virtual refraction traveltimes generated by N virtual sources, where N is the number of geophones in the 2-D survey. The virtual traveltimes can be inverted to give the velocity tomogram. This enormous increase in the number of traveltime picks and associated rays, compared to the many fewer traveltimes from the reciprocal and infill shot gathers, allows for increased model resolution and a better condition number with the system of normal equations. A significant benefit is that the parsimonious survey and the associated traveltime picking can be an order-of-magnitude less time consuming than that for a standard refraction survey with a dense distribution of sources.

Key words: Tomography; Interferometry; Controlled source seismology; Body waves.

1 INTRODUCTION

Refraction tomography is an important imaging tool in earthquake studies (Stein & Wyession 2003), crustal-mantle imaging (Prodehl & Mooney 2012), exploration geophysics (Bishop *et al.* 1985; Yilmaz 2001) and engineering seismology (Yilmaz 2015). For 2-D engineering seismology, receivers are deployed along a line and common shot gathers (CSGs) are recorded for sources at selected positions along the line. To save costs, a parsimonious survey is carried out where sources are located only at each end of the receiver line as shown in Fig. 1. The first breaks are picked from each shot gather and the traveltimes are inverted by a simple formula that assumes a layered medium with an unknown dip angle for each interface. These two shot gathers are denoted as a pair of reciprocal shot gathers. We often require several short-offset infill shot gathers that can be used to find the near-offset and direct traveltimes. Therefore, parsimonious surveys can significantly reduce acquisition costs and survey efforts, but at the cost of less slowness resolution and certainty in the estimate of the subsurface model.

For a 2-D survey, we now show that two reciprocal shot gathers and several short-offset infill shot gathers can be decomposed into many virtual shot gathers that have many more traveltime picks available for inversion. In some cases, the virtual shot gathers give as much refraction information as a full survey with N shots, where a shot is located at each of the N geophone locations. The result is a tomogram with a much denser ray coverage and better slowness resolution compared to that from the original data. We call this procedure parsimonious refraction interferometry because it uses a stationary-phase principle to decompose the reciprocal traveltimes into many virtual traveltimes associated with shorter ray paths. Unlike the original pair of reciprocal shots, the virtual shot locations are at all of the interior geophone locations.

The next section describes the theory of parsimonious refraction interferometry. It is an application of Fermat's interferometric traveltime principle (Schuster 2005) and closure phase (Schuster *et al.* 2014) that allows for the decomposition of long ray paths and traveltimes into, respectively, shorter ray paths and traveltimes. Instead of body-wave traveltimes we now apply it to head-wave traveltimes. Section 3 presents the results of applying parsimonious refraction interferometry to both synthetic data and field data. The field data are from 2-D surveys conducted next to the Gulf of Aqaba and over a basin in East Africa. The final section presents a summary and conclusions.

2 THEORY

Assume two reciprocal sources (**A** and **D**) and the irregularly layered medium in Fig. 1, where head waves propagate along the interface between the upper and lower layers. There can be lateral velocity variations in the upper medium and there are N evenly spaced geophones on the recording surface between the two sources. The head-wave traveltime from the source at **A** to the geophone at **C** is defined as

$$\tau_{AC} = \tau_{Ax'} + \tau_{x'x} + \tau_{xC}, \tag{1}$$

and the reciprocal traveltime from **D** to **B** is

$$\tau_{DB} = \tau_{Dx} + \tau_{xx'} + \tau_{x'B}, \tag{2}$$

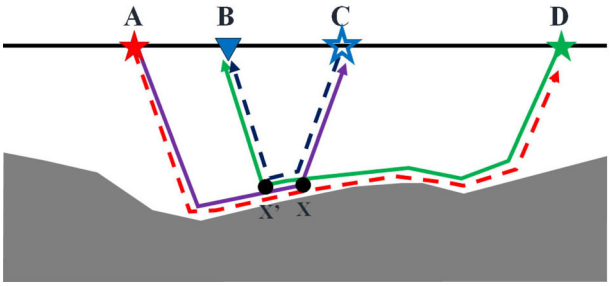

Figure 1. Two-layer model where the grey medium is faster than the top layer; the reciprocal sources are at **A** and **D** and are associated with the dashed red ray. The dashed blue ray path is associated with the virtual refraction ray that is excited by the virtual source (blue star) at **C** and terminates at **B**. The refraction traveltimes associated with the reciprocal shots (green and red stars) can be decomposed into the virtual refraction traveltime generated by the blue star. Here, x and x' are two points along the interface between the upper and lower layers.

where $\tau_{xx'}$ is the traveltime from $\mathbf{x}$ to $\mathbf{x}'$ along the refraction ray, and reciprocity says $\tau_{xx'} = \tau_{x'x}$. The traveltimes τ_{AC} and τ_{DB} are denoted as a pair of reciprocal traveltimes.

2.1 Virtual-traveltime equation

To create virtual sources and receivers within the array in Fig. 1, we assume that there is a post-critical distance between the geophone positions **B** and **C**, and both traveltimes τ_{AC} and τ_{DB} are associated with the same refractor. Subtracting the reciprocal traveltime $\tau_{AD} = \tau_{Ax'} + \tau_{x'x} + \tau_{xD}$ from the sum $\tau_{AC} + \tau_{DB}$ gives the virtual traveltime $\delta\tau_{CB}$:

$$\begin{aligned}\delta\tau_{CB} &= \tau_{AC} + \tau_{DB} - \tau_{AD},\\ &= [\tau_{Ax'} + \tau_{x'x} + \tau_{xC}] + [\tau_{Dx} + \tau_{xx'} + \tau_{x'B}]\\ &\quad - [\tau_{Ax'} + \tau_{x'x} + \tau_{xD}],\\ &= \tau_{Cx} + \tau_{xx'} + \tau_{x'B}. \end{aligned} \tag{3}$$

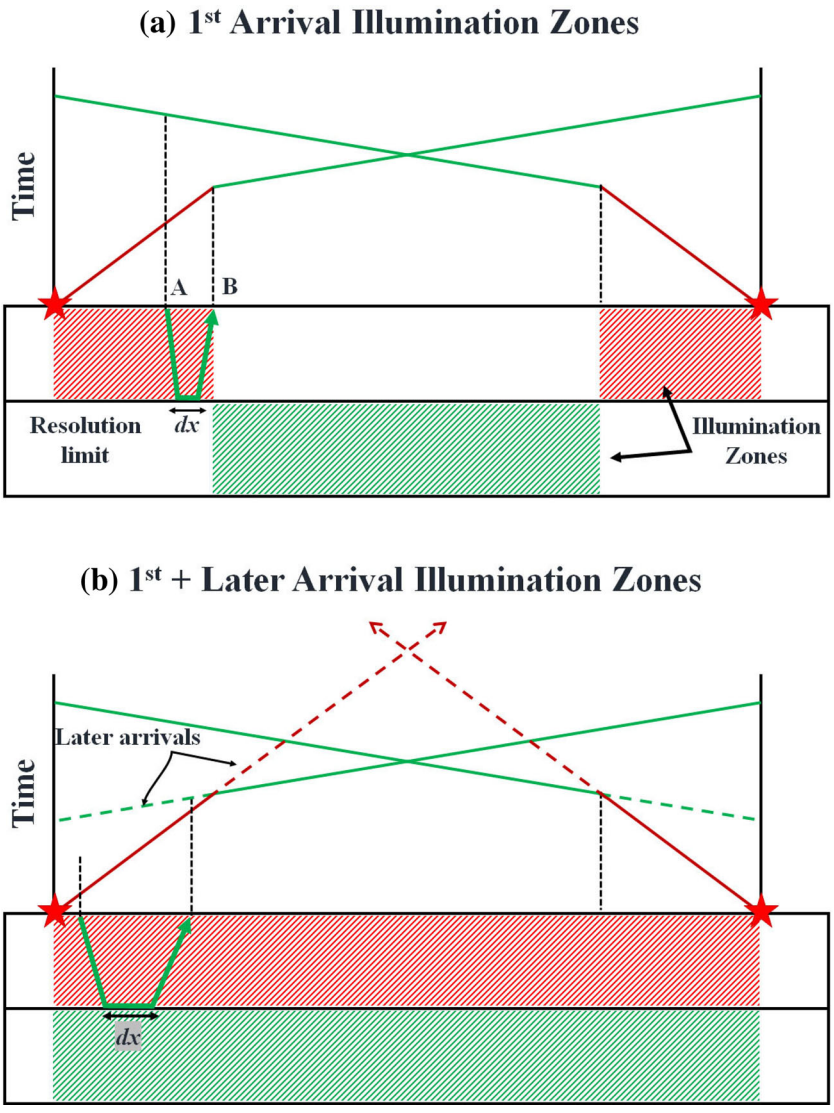

Figure 2. The red and green zones approximate the areas of illumination associated with traveltimes picked from (a) first arrivals and (b) both first and the dashed later arrivals. The red (green) lines indicate direct-wave (refraction) arrivals. The horizontal resolution limit dx is denoted by two-sided arrows for different portions of the subsurface. The resolution limit dx cannot be less than the distance between adjacent geophones.

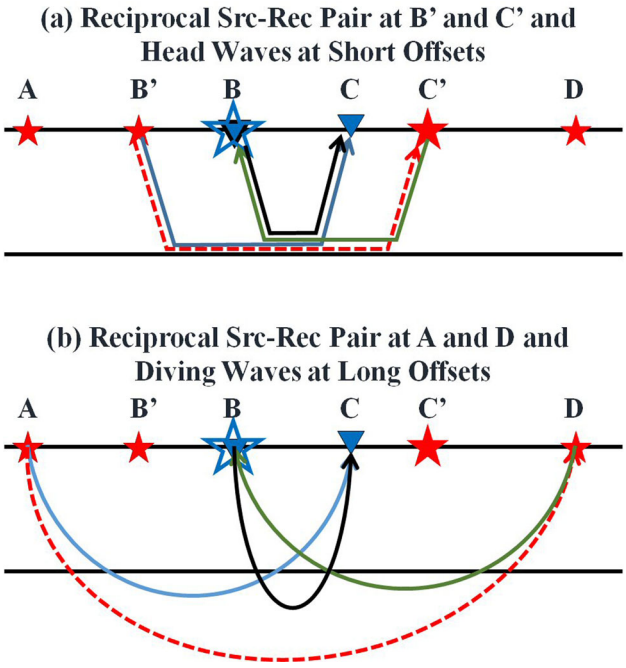

Figure 3. Virtual traveltimes created with (a) short-offset $T_{BC} = T_{B'C} + T_{BC'} - T_{B'C'}$ and (b) long-offset $\bar{T}_{BC} = T_{AC} + T_{DB} - T_{AC}$ virtual traveltimes. The long-offset (short-offset) sources generate diving waves (head waves) as first arrivals so that $T_{BC} \neq \bar{T}_{BC}$. The black solid ray corresponds to the virtual ray path associated with either T_{BC} or $\bar{T}_{BC}$.

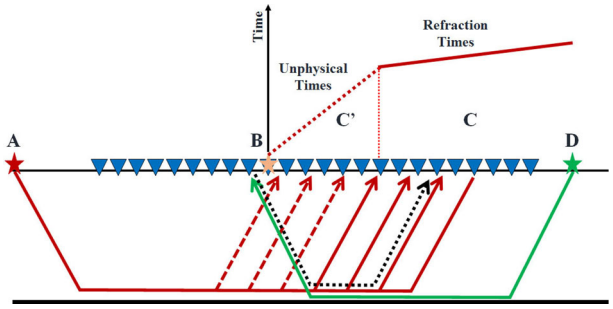

Figure 5. Ray diagram showing the dashed non-physical rays such that $\tau_{BC'} = \tau_{DB} + \tau_{AC'} - \tau_{DA}$ is a non-physical virtual traveltime which does not correspond to a physical head wave. In comparison, $\tau_{BC} = \tau_{DB} + \tau_{AC} - \tau_{DA}$ corresponds to the virtual traveltime of a physical head wave along the dotted black ray. The zone of non-physical virtual traveltimes is between the origin at **B** and the sudden change in the traveltime slope.

Here, $\delta\tau_{CB}$ is denoted as an interferometric stationary traveltime because the reciprocal ray path $\overline{Ax'xD}$, marked by the dashed red ray in Fig. 1, cancels the phase associated with the common ray paths of the purple $\overline{Ax'xC}$ and green $\overline{Dxx'B}$ rays. The result is the virtual traveltime $\delta\tau_{CB}$ associated with the much shorter ray path $\overline{Cxx'B}$ denoted by the dashed blue ray. Equivalently, $\delta\tau_{CB}$ is the traveltime of a virtual refraction excited by a virtual source at **C** and recorded at **B**. This natural redatuming operation is the key principle underlying seismic interferometry (Snieder 2004; Wapenaar 2004).

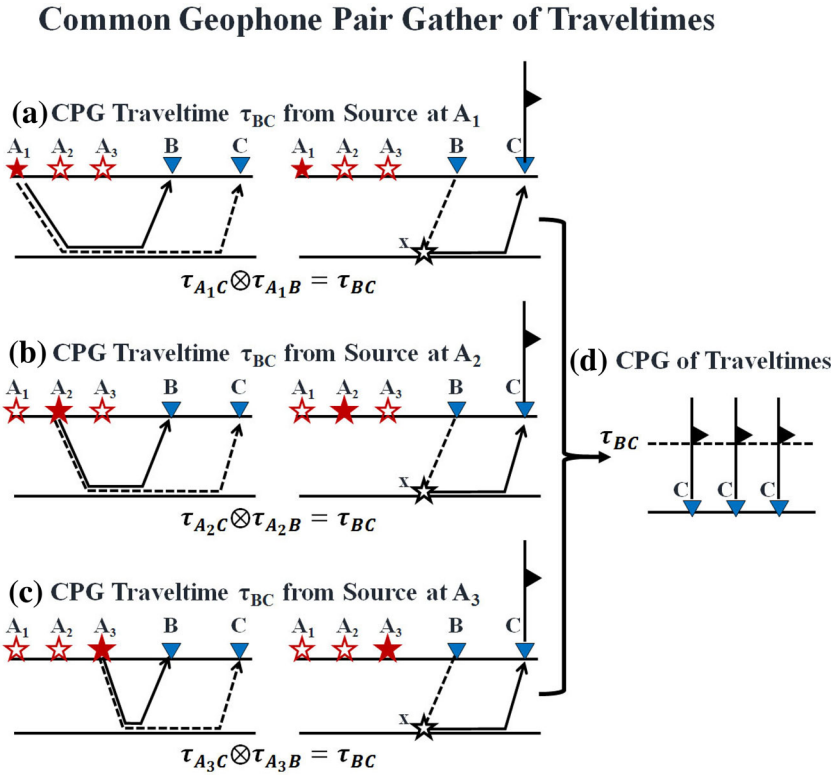

Figure 4. Virtual traveltimes τ_{BC} created by computing (a) $\tau_{A_1C} - \tau_{A_1B}$, (b) $\tau_{A_2C} - \tau_{A_2B}$ and (c) $\tau_{A_3C} - \tau_{A_3B}$. If the input traveltimes are those of head waves, then the output (d) traveltime τ_{BC} will be that of a virtual head wave recorded at *C*. The common geophone pair gather (CPG) of virtual traces in a–c share the same pair of geophones at **B** and **C**, but have different source locations.

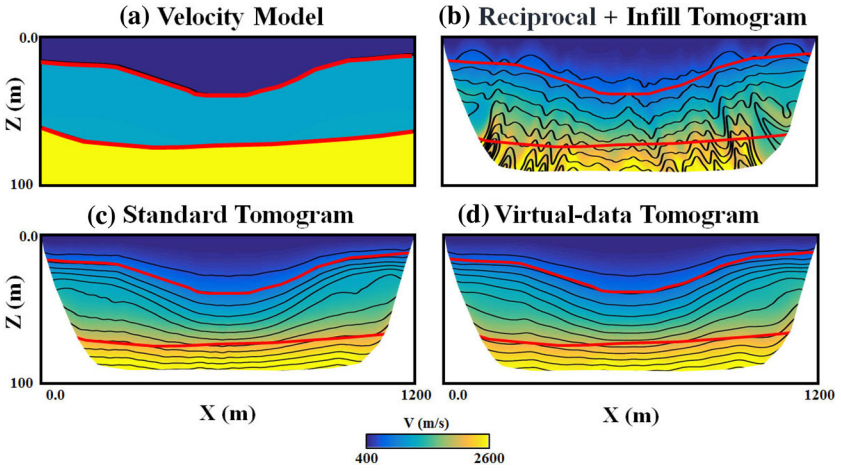

Figure 6. (a) Three-layer model, (b) tomogram inverted from the two reciprocal and five infill shot gathers, (c) standard tomogram inverted from the 57 600 actual traveltimes in 240 shot gathers, with a shot at each geophone location. (d) Tomogram inverted from $O(57\ 000)$ virtual traveltimes created from the first-arrival traveltimes picked from two reciprocal and five infill shot gathers.

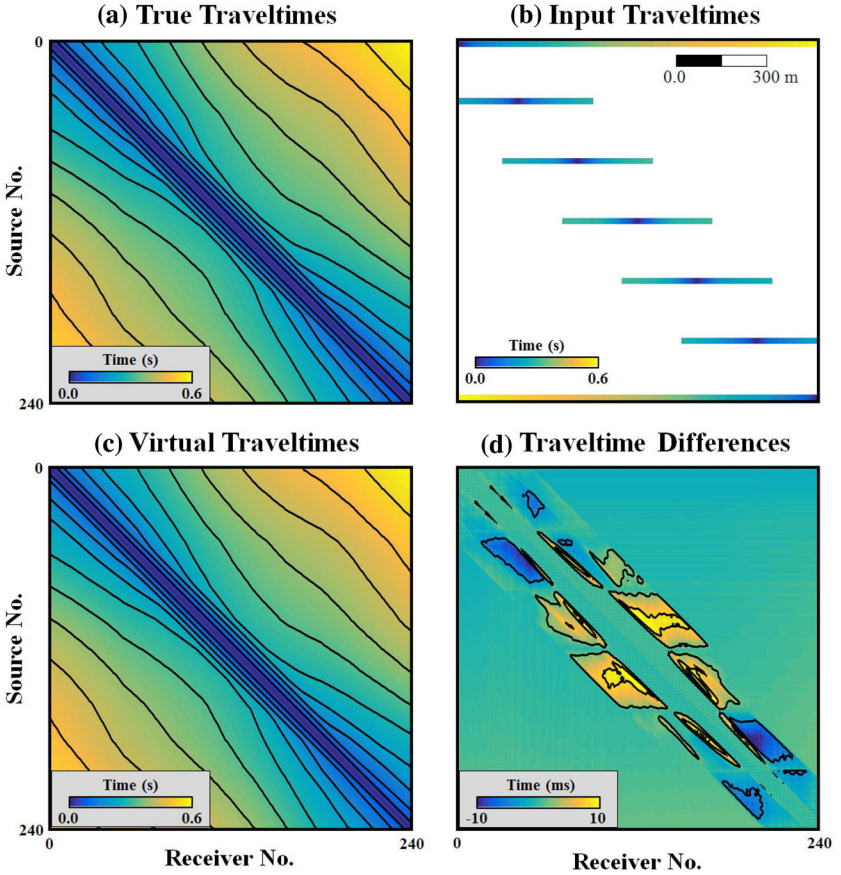

Figure 7. (a) The actual traveltime picks, (b) traveltimes for two reciprocal and five infill shot gathers used as input to the proposed method, (c) the virtual traveltimes calculated from eq. (3) and (d) the difference between the actual and virtual traveltime picks.

Eq. (3) can be used to generate $O(N)$ virtual shot gathers of head-wave traveltimes, where the number of geophone pairs with post-critical separation is nearly equal to the number N of geophones in the survey, assuming that a few geophones are recording direct waves as first arrivals. Each virtual shot gather will, on average, contain $O(N)$ virtual traveltimes to create $O(N^2)$ virtual refraction traveltimes from the $2N$ traveltimes in two reciprocal shot gathers. This abundance of new traveltimes can be used to invert for the subsurface velocity model with much greater ray density and better model resolution compared to that from the original data set.

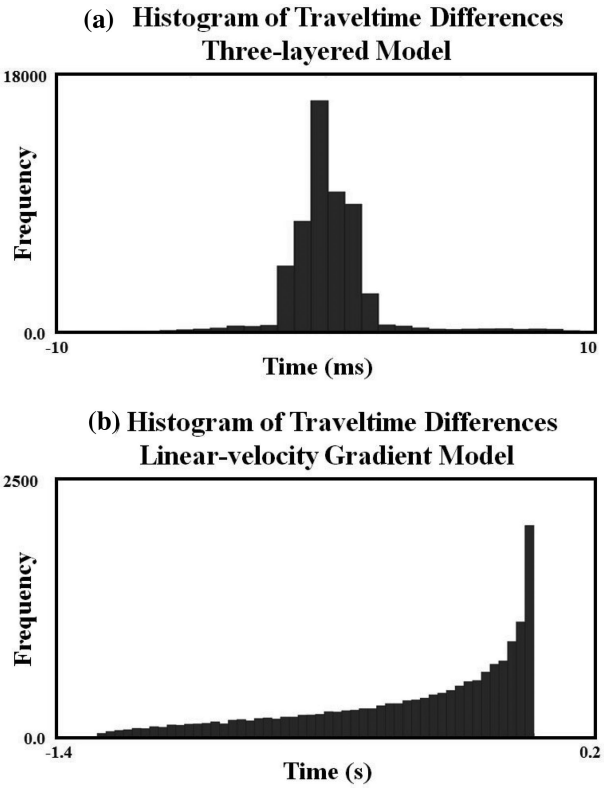

Figure 8. The histogram of the difference between the true and virtual traveltimes of the (a) three-layer and (b) gradient synthetic models. The average differences in (a) are less than 5 ms, while in (b) it is as high as 1.4 s.

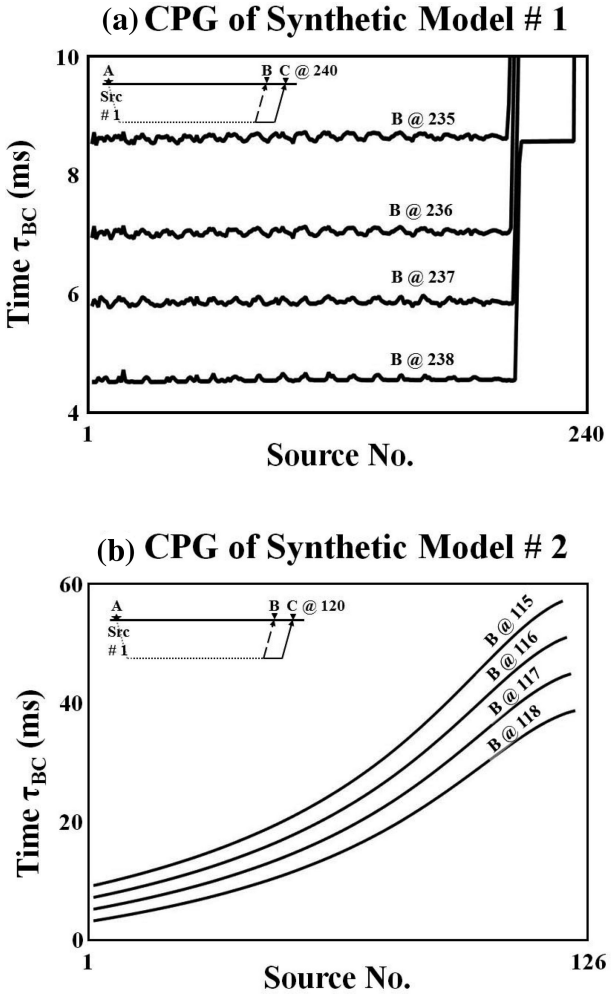

Figure 9. Consistency test results for the (a) three-layer and (b) gradient synthetic models.

Eq. (3) satisfies Fermat's interferometric principle (Schuster 2005) because the subtraction of τ_{AD} (red dashed ray) from $\tau_{AC} + \tau_{DB}$ (solid green and purple rays) gives the same value of $\delta\tau_{CB}$ for all post-critical, i.e. stationary, locations of the reciprocal sources. Eq. (3) is also a special case of the closure phase condition discussed in Schuster *et al.* (2014).

The reciprocal shots are typically too far apart from one another, so the first-arrivals will not be direct waves for far-offset receivers. These direct waves and their traveltimes are needed to estimate the velocities at the near-surface. Therefore, several infill shots are required to get the near-surface velocities from the direct-wave traveltimes. The direct-arrival traveltimes are picked from these infill shot gathers, and then interpolated to other source locations to fill in the missing direct-wave traveltimes in the other CSGs. The number of infill shot gathers depends on the complexity of the subsurface. We suggest that the offset between the shots of the infill shot gathers should be less than the cross-over distance offset of the direct-wave/refraction arrivals in the recorded data. However, in the case of a simple near-surface velocity distribution, this rule can be relaxed and the offset between infill shot gathers can be larger than the cross-over distance.

2.2 Condition number, illumination area and slowness uncertainty

The resolution properties associated with the virtual-traveltime equations are now compared to those from the reciprocal traveltime equations. These properties are (1) the condition number of the normal equations, (2) the enhancement of illumination with virtual-traveltime equations and (3) the slowness uncertainty associated with errors in traveltime picking.

2.2.1 Condition number

As Appendix A shows, the extra traveltime equations generated by eq. (3) act as a preconditioner that can reduce the condition number of the normal matrix $[\mathbf{L}^T\mathbf{L}]$ by a factor of 3 in some cases, where $\mathbf{L}$ is the ray path matrix associated with the traveltime equations. The dense set of virtual traveltimes does not create new information about the subsurface because it is created by adding and subtracting traveltime equations from the reciprocal data. However, the synthetic examples in the next section suggest that the enhanced subsurface illumination and smaller condition number in the virtual $[\mathbf{L}^T\mathbf{L}]$ matrix can significantly reduce the number of velocity models that can approximately satisfy the same data.

2.2.2 Illumination zones and spatial resolution

Fig. 2(a) shows that a reciprocal pair of shot gathers can give rise to gaps in the illumination zones of the direct arrivals and refractions.

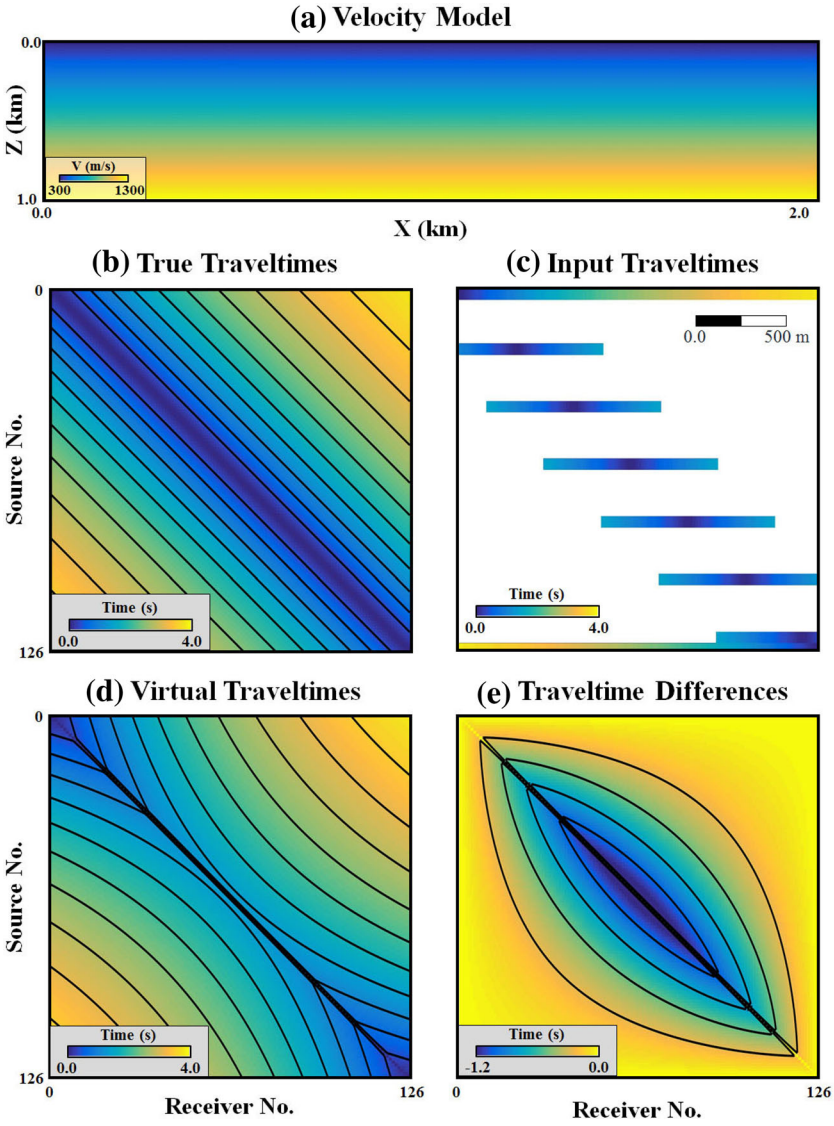

Figure 10. (a) Model with velocity linearly increasing in depth. (b) True, (c) reciprocal and (d) virtual traveltimes. (e) The difference between the true and virtual traveltimes.

For refractions, these gaps can be caused by the inability to pick far-offset traveltimes as indicated by the truncated green lines in Fig. 2(a). This problem can be partly remedied if the later-arriving direct waves and refractions can be picked, as illustrated by the dashed lines in Fig. 2(b). In practice, however, such picking is both tedious and time consuming.

A large illumination zone does not necessarily mean a good spatial resolution limit dx. Here, dx is the minimum width of a slowness box whose velocity can be resolved from the data. For virtual data, dx depends on the midpoint position and minimum offset between a reciprocal pair of geophones. As an example, the pair of dashed vertical lines at **A** and **B** in Fig. 2(a) indicates the reciprocal refraction traveltimes with minimum offset $dx = |x_A - x_B|$ that are pickable from the leftgoing and rightgoing refractions. This means that picking later refraction arrivals in Fig. 2(b) will lead to wider subsurface illumination, but the estimated velocity model can still suffer from poor resolution as demonstrated by the wide double-sided black arrow. If both the first-arrival and later-arrival refraction times can be picked almost everywhere between the two reciprocal shots, then the spatial resolution dx is equal to the spacing between adjacent geophones. This assumes that the typical wavelength of the arrivals is much shorter than a geophone spacing.

2.2.3 *Slowness variance*

Traveltime picking is prone to random errors, and so the variation of these data errors give rise to an uncertainty in the estimated slowness model, denoted by the slowness variance σ_i^2 in the ith cell. Appendix B shows that the slowness variance associated with the reciprocal data is proportional to $\frac{1}{l^2}$ (l is the width of a slowness cell in Fig. A1, see Appendix A for more details), while that for the virtual data is proportional to $\frac{2}{Nl^2}$. Thus, the variance of the virtual slowness can be much less than that for the reciprocal data if N is large and there is only weak correlation in the data errors.

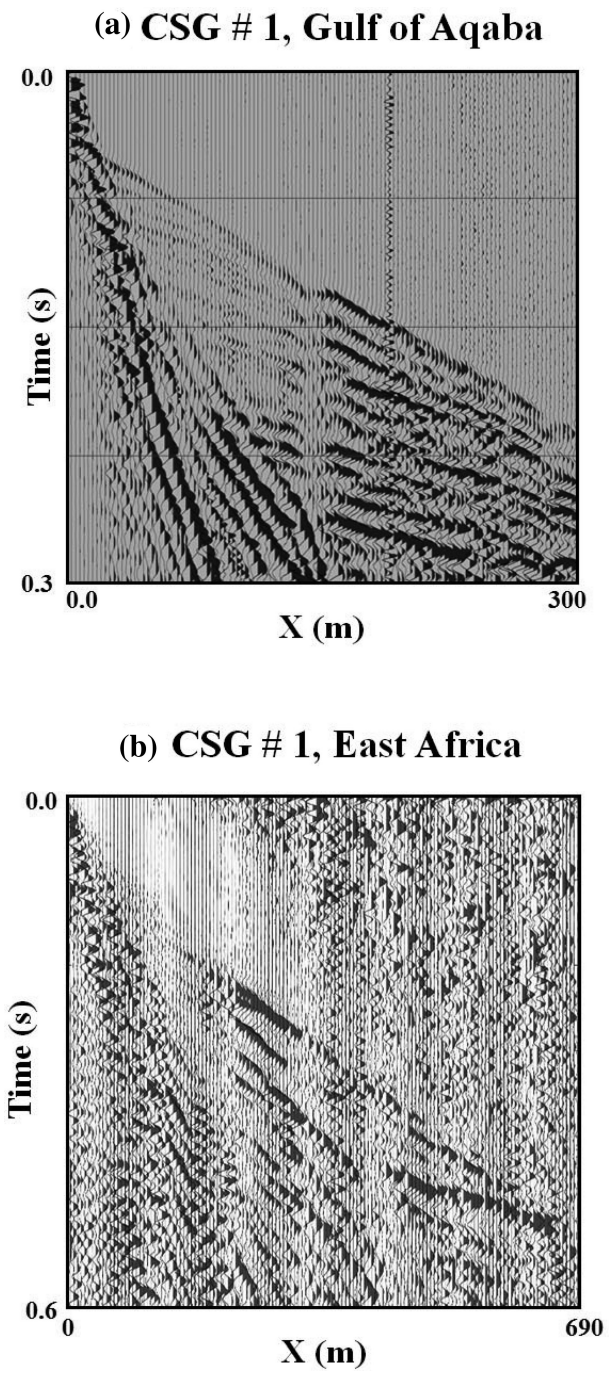

Figure 11. Typical common shot gathers recorded at the (a) Aqaba and (b) East Africa field sites.

2.3 Consistency test

A consistency test is now described which assesses the validity of the assumption that the recorded refraction arrivals are those of head waves.

Assume head wave arrivals in Fig. 3(a) so that

$$\bar{\tau}_{BC} = \tau_{B'C} + \tau_{C'B} - \tau_{B'C'}, \tag{4}$$

is the traveltime of a virtual head wave excited at **B** and recorded as a 1st-arrival at **C**. However, this might not be a head-wave traveltime if the source at **A** is far from the receivers and there is a velocity gradient linearly increasing in depth. As an example, Fig. 3(b) shows a diving wave recorded as a 1st-arrival at **B** and **C** when there is a strong velocity gradient that linearly increases in depth. Diving waves violate the head-wave assumption underlying eq. (4) so the resulting traveltime $\bar{\tau}_{BC}$ is that of a non-physical diving wave between **B** and **C** (see the solid black ray in Fig. 3b).

Therefore, traveltimes from recorded shot gathers must be tested for consistency with the head-wave assumption. We propose a consistency test identical to that given in Dong *et al.* (2006) where the traveltimes in a common geophone pair gather (CPG) must equal one another if they are for head waves. As an example, the head-wave traveltimes in Figs 4(a)–(c) are transformed by

$$\bar{\tau}_{BC} = \tau_{AC} - \tau_{AB}, \tag{5}$$

into the traveltimes $\bar{\tau}_{BC}$ for the traces in Fig. 4(d). If the difference between these traveltimes is greater than the picking uncertainty for different positions of **A** on the left side of **B** then they should be rejected as head-wave traveltimes associated with the same refractor.

Some of the head-wave traveltimes associated with the same refractor are not physically related to refraction times because the receiver–receiver offset is not at a post-critical offset (Fig. 5). Such non-physical traveltimes are identified by a sudden jump in the CPG. These non-physical traveltimes are easily identified[1] and eliminated, but they leave a gap in the arrival times of the direct waves. The interpolated direct-wave traveltimes from the infill CSGs will automatically fill this gap.

3 NUMERICAL RESULTS

Parsimonious refraction interferometry is now tested for traveltimes generated for two synthetic models and two seismic field surveys. One of the synthetic models is a three-layered velocity model with a velocity gradient and the second one has a strong velocity gradient that increases with depth to demonstrate the diving wave problem. In each synthetic example the traveltimes were generated using a finite-difference solution to the eikonal equation (Qin *et al.* 1992) and inverted by the multiscale gradient method described in Nemeth *et al.* (1997). The field examples are for refraction data recorded near the Gulf of Aqaba and over a deep basin in East Africa. Two more data examples are presented in the Supporting Information.

3.1 Three-layer synthetic model

The first example is a three-layer model (Fig. 6a), where the first layer has a constant velocity of 400 m s^{-1}, the second layer has a horizontal velocity gradient of 0.2 m s^{-1} m^{-1}, while the third layer has a vertical velocity gradient of 4 m s^{-1} m^{-1} with a minimum velocity value of 2600 m s^{-1}. A finite-difference solution (Qin *et al.* 1992) to the 2-D eikonal equation is used to compute 240 shot gathers of first-arrival traveltimes (Fig. 7a), where the source and receiver points are located every 5 m on the surface.

We first inverted the first-arrival traveltime values of two shot gathers, with shots located at $X = 0$ m and $X = 1200$ m, as well as the near-offset traveltimes picked from five infill shot gathers (Fig. 7b) to get the tomogram shown in Fig. 6(b). In this case there is a poor correspondence between the actual velocity model and the tomogram constructed from the reciprocal and infill data. For

[1] The non-physical traveltimes can be visually identified by identifying sudden changes in their slopes. The non-physical traveltimes can automatically be eliminated by the consistency test.

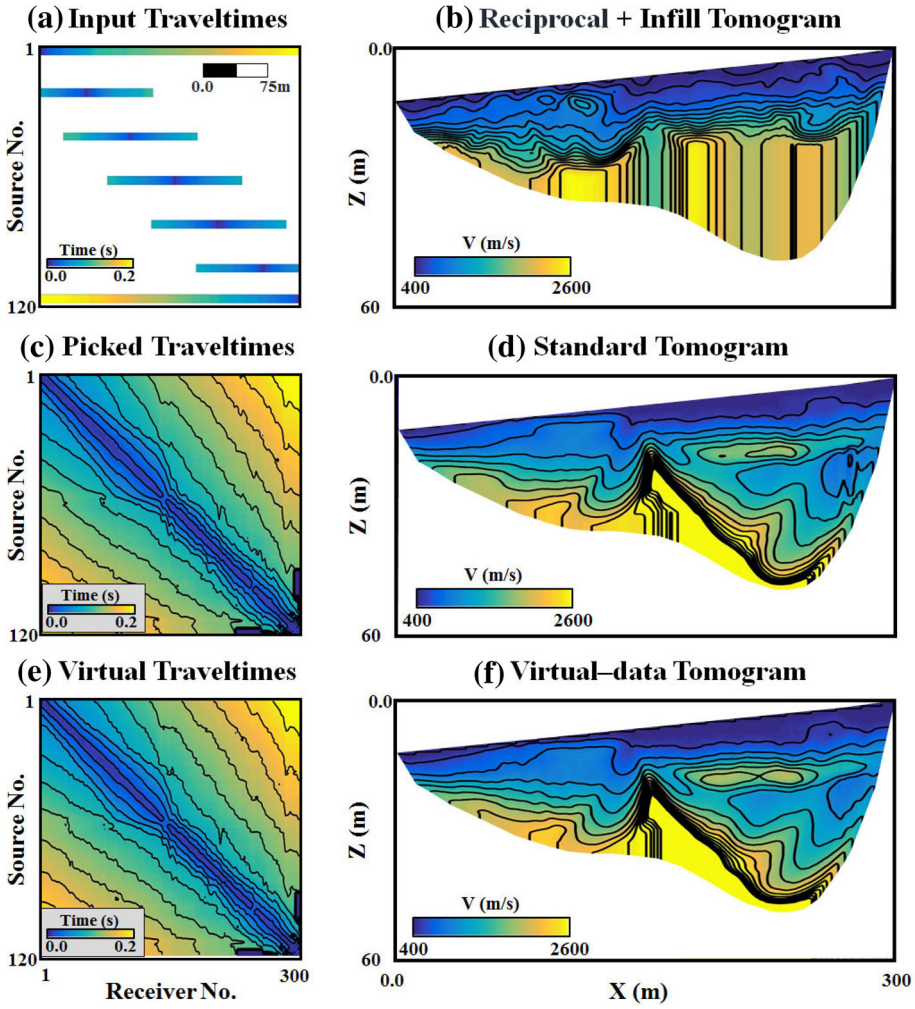

Figure 12. Gulf of Aqaba traveltime picks for the (a) reciprocal, (c) standard and (e) virtual field data. The corresponding tomograms are on the right.

comparison, Fig. 6(c) shows the standard tomogram inverted from the *actual* traveltimes generated by exciting shots at each of the 240 geophones. As expected, the standard tomogram mostly agrees with the actual velocity model.

Eq. (3) is then used to compute the virtual traveltimes from the 480 traveltimes associated with the two reciprocal shot gathers. To account for the fact that the direct-wave traveltimes could not be generated at long offsets, the direct-wave traveltimes were picked from five infill CSGs with the shots evenly located at 200 m intervals between the endpoints of the receiver line. These direct-wave traveltimes were then interpolated to other source locations to fill in the missing direct-wave traveltimes in the other CSGs. The result is the creation of $O(57\ 000)$ virtual traveltimes computed for virtual shots at each of the geophones (Fig. 7c). These virtual traveltimes are inverted to give the tomogram shown in Fig. 6(d). This tomogram agrees with the standard tomogram, which is not surprising because (1) the virtual traveltimes are nearly the same as the true traveltimes shown in Figs 7(d) and 8(a), (2) there are no strong velocity gradients in the model and (3) there is a dense illumination in the subsurface from the virtual rays.

The validity of the head-wave assumption could has been predicted by determining whether or not the traveltimes in the first-arrival traveltimes in the CPGs are flat. Fig. 9(a) shows that the CPG traveltimes mostly agree with one another for the same CPG. For specified locations of **B** and **C** (see ray diagram in upper left of Fig. 9a), the thin black line represents the traveltimes τ_{BC} for all actual shot locations to the left of **B**. Here, the location **C** is always at location #240 and **B** is progressively moved to its left to give rise to larger traveltimes τ_{BC}.

3.2 Linear velocity gradient model

Eq. (3) is not valid if the recorded traveltimes are mainly associated with diving waves and not with head waves. To demonstrate this point, we used the linear-velocity-gradient model of Fig. 10(a) to calculate the first-arrival traveltimes for 126 shot gathers with a shot interval of 8 m. The calculated traveltimes are shown in Fig. 10(b). Only the first-arrival traveltimes from the two reciprocal shots at $x = 0$ and $x = 2$ km (Fig. 10c) are converted by eq. (3) to virtual traveltimes for virtual shots at every grid point along the surface (Fig. 10d). The differences between the actual and virtual traveltimes range between 0 and -1.4 s as shown in Figs 10(e) and 8(b). This is consistent with the disagreement between the traveltime matrices in Figs 10(b) and (d). There are two reasons for this disagreement, (1) the diving waves violate the head-wave assumption and (2) in the case of diving waves, sources are not at

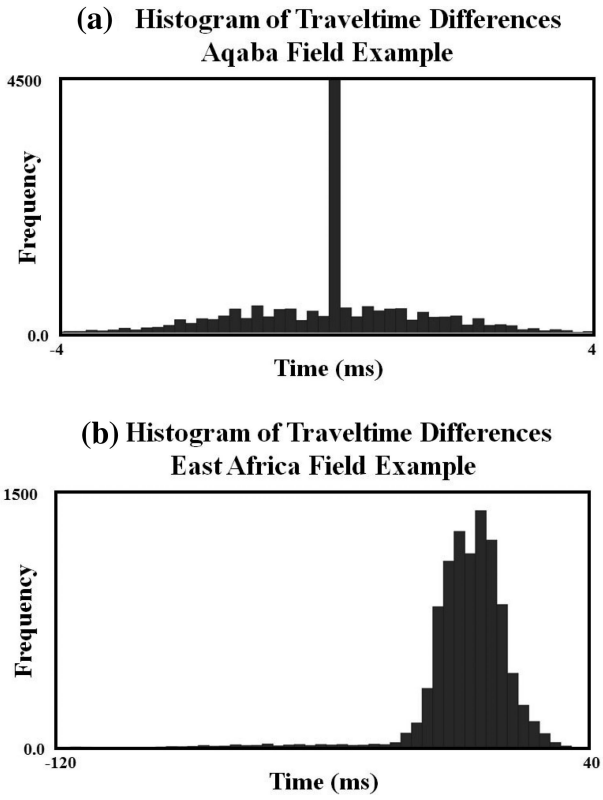

Figure 13. The histogram of the differences between the true and virtual traveltimes obtained from the (a) Aqaba and (b) East Africa data. The average differences in (a) are less than 5 ms, while in (b) it is as high as 120 ms.

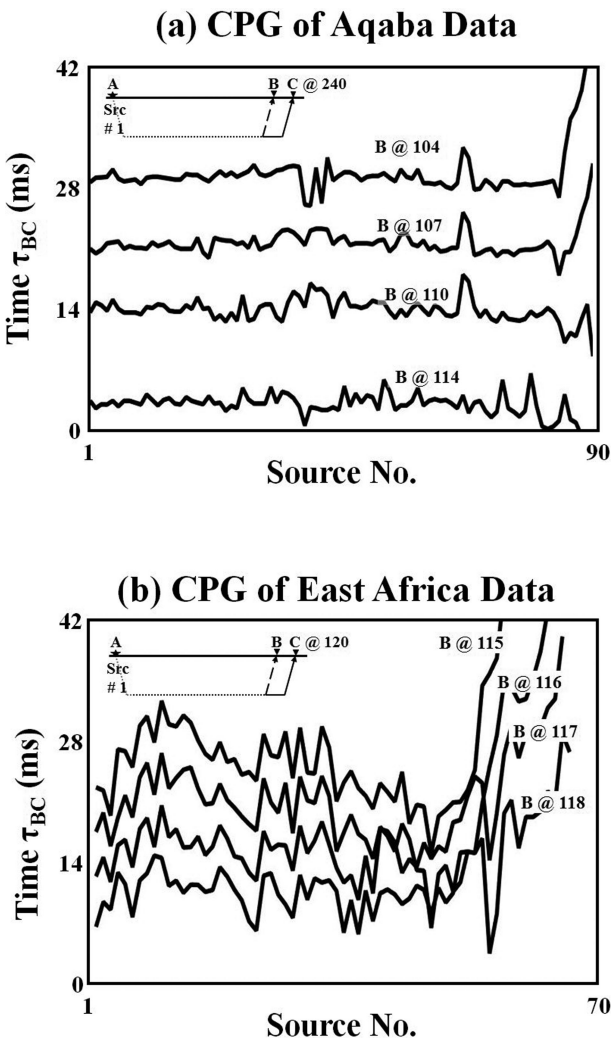

Figure 14. Consistency test results for the (a) Aqaba and (b) East Africa data. The East Africa data in (b) fail the consistency test while it is passed for the Aqaba data. In all of the figures, the input data were traveltimes picked from the standard survey where there is a dense distribution of shots.

stationary positions. In contrast, if the first arrivals are those of head waves then all post-critical sources are located at stationary locations. Fig. 9(b) shows that the CPG traveltimes disagree with one another, as expected for diving waves.

3.3 Aqaba field data

A seismic survey was carried out near the Gulf of Aqaba (Hanafy *et al.* 2014), where 120 vertical-component geophones were deployed at 2.5 m intervals along a line. A 90-kg accelerated-weight drop was used for a source at every geophone position to record 120 common shot gathers. A typical shot gather is shown in Fig. 11(a), and more than 14 000 first-arrival traveltimes were picked from the 14 400 traces in the CSGs (Fig. 12c). The signal-to-noise ratio (SNR) is excellent so traveltimes could be picked at almost every trace, as verified by the densely populated traveltime matrix in Fig. 12(c).

A source at each end of the line was used to generate two reciprocal shot gathers. The first-arrival traveltimes were picked and used to compute more than 14 000 virtual traveltimes from eq. (3). As in the synthetic example, some of these traveltimes are not physically related to refraction times because the receiver-receiver offset is not at a post-critical offset. These non-physical traveltimes are easily identified and eliminated, but they leave a gap in the arrival times of the direct waves. To fill in this gap, we included the direct-wave traveltimes from five infill CSGs with a shot interval of 50 m (Fig. 12a).

Inverting the traveltimes by a multiscale tomography method (Nemeth *et al.* 1997) gives the tomograms shown in Figs 12(b), (d) and (f). It is clear that there is a mostly excellent agreement between the virtual and standard tomograms shown in Figs 12(d) and (f), respectively. In the standard tomogram, the shooting effort required 1.5 d, and more than a day of effort was needed for picking (14 400) traveltimes. For the virtual tomogram, it is estimated that we would need less than 1 hr of shooting time and less than 30 min of picking time for the reciprocal and infill traveltimes. Moreover, the five shot gathers recorded for the direct-wave traveltimes would not require much more effort to record, because only one strike of the weight-drop source is needed to record short-offset traces with adequate SNR. The subsequent generation of the $O(14\,000)$ virtual traveltimes is automatically computed by the computer in less than a second. Fig. 13(a) shows that the average differences between the virtual and actual traveltimes are well below the estimated picking error of 4 ms in the data.

The consistency test results for the recorded Aqaba data are depicted in Fig. 14(a). For specified locations of **B** and **C** (see ray

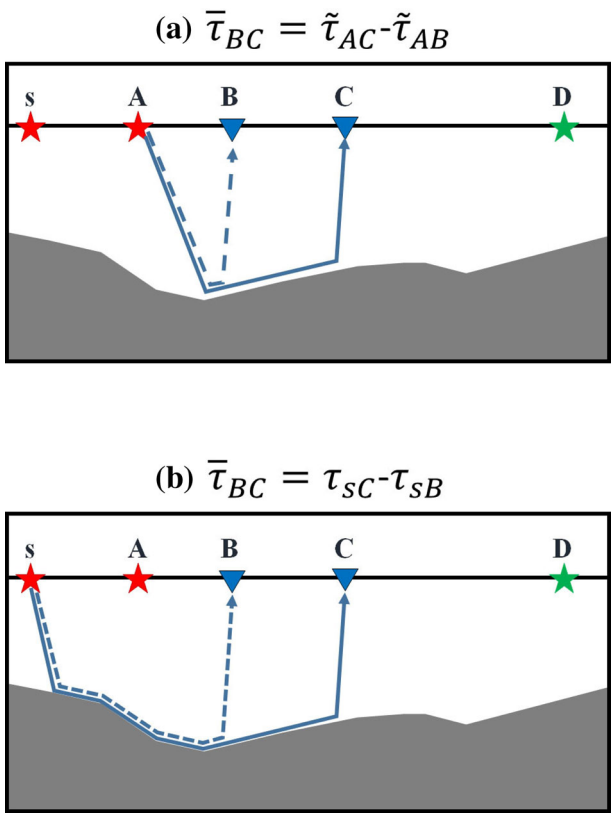

Figure 15. A sketch shows the calculation of common pair gather $\bar{\tau}_{BC}$ from (a) virtual ($\tilde{\tau}_{AC}$ and $\tilde{\tau}_{AB}$) and (b) recorded data (τ_{sC} and τ_{sB}).

diagram in upper right), the thin black line represents the traveltimes τ_{BC} for all actual shot locations to the left of **B**. Here, **C** is always at location #120 and **B** is progressively moved to its left to give rise to larger traveltimes τ_{BC}.

Most of the recorded traveltimes τ_{BC} to the left are flat to within a quarter of the dominant period ($\frac{1}{4}$T, where T = 32 ms), and therefore indicate passage of the consistency test. However, the traveltimes associated with the bumps are rejected. This plot is for one location of **C**, and so the traveltimes for other locations of **C** should also be tested for consistency with the head-wave assumption.

It is tempting to apply the consistency test to the virtual traveltimes $\tilde{\tau}_{ij}$, not just the recorded traveltimes τ_{ij}. However, this will not be useful as shown in the following exercise. Assume a source at **s** to the left of **A** (Fig. 15). The receiver locations are at **B** and **C**. Recorded traveltimes, from two reciprocal shots at **s** and **D**, are used to compute the virtual traveltimes τ_{ij} for all interior locations associated with the i and j indices. These virtual traveltimes are used to generate the CPG traveltime $\bar{\tau}_{BC}$ in eq. (5) (see Fig. 15a). Plugging eq. (4) into eq. (5) we get

$$\begin{aligned}\bar{\tau}_{BC} &= \tilde{\tau}_{AC} - \tilde{\tau}_{AB},\\ &= \tau_{sC} + \tau_{DA} - \tau_{sD} - [\tau_{sB} + \tau_{DA} - \tau_{sD}],\\ &= \tau_{sC} - \tau_{sB},\end{aligned} \tag{6}$$

which is true for any post-critical location **s** to the left of **A** (Fig. 15b). Even if the recorded reciprocal traveltimes are those for diving waves, the virtual CPG time $\bar{\tau}_{BC}$ will *appear* to be a head-wave time because it is the same for any virtual source location **s** to the left of **A**. Thus, the consistency test cannot be applied to *virtual* traveltimes to distinguish head-wave arrivals from diving waves.

3.4 East Africa field data

Parsimonious refraction interferometry is now applied to a data set consisting of strong diving waves. In this case, the survey is conducted over a deep basin where the two reciprocal shots at the end of the line are separated by 690 m, and there are 139 vertical-component geophones evenly distributed along the line at 5 m intervals. The standard survey is conducted with the accelerated weight drop source located at 10 m intervals along the line. A typical shot gather is shown in Fig. 11(b).

The reciprocal traveltimes are picked and used to generate the virtual traveltimes. The comparison between the reciprocal, virtual and standard traveltimes is depicted along the left column of Fig. 16. Unlike the Aqaba data, there are notable discrepancies between the virtual and standard traveltimes (Fig. 13b). These discrepancies are manifested in the notable dissimilarities between the standard and virtual tomograms in the right column. The reason for these discrepancies is revealed by the consistency test results illustrated in Fig. 14(b), which shows strong violations of the head-wave assumption.

4 CONCLUSIONS

The theory of parsimonious refraction interferometry and tomography is presented where a dense set of virtual refraction traveltimes is computed from refraction traveltimes picked from a pair of reciprocal and several infill shot gathers. In theory, a virtual shot gather of refraction traveltimes can be computed for a virtual shot at each of the N geophones in the 2-D reciprocal survey. This means that $O(N^2)$ virtual traveltimes can be created from virtual shots placed at each of the N geophones. An advantage of virtual data is that they can reduce the condition number of $[\mathbf{L}^T\mathbf{L}]$ by a factor of three compared to that of the reciprocal data. The accompanying increase in ray coverage reduces the number of slowness models that satisfy the traveltime equations. In addition the slowness variance associated with virtual data can be less than that obtained from reciprocal data.

Tests with synthetic data and field data verify that virtual refraction traveltimes can be inverted to give tomograms that closely resemble those computed from traveltimes recorded in a dense survey. If the first arrivals are mostly head-wave arrivals, not strong diving waves, then dense 2-D refraction surveys might be replaced by inexpensive reciprocal surveys with as few as two shots placed at each end of the line and several infill shot gathers.

Some opportunities presented by parsimonious refraction interferometry include the following.

(i) The virtual traveltimes can be used to design narrow picking windows for an autopicker to pick first-arrival traveltimes from the dense survey data. For example, the Aqaba data only used 240 traveltime picks to automatically generate O(14 000) virtual traveltimes.

(ii) More efficient surveys can be carried out because a reciprocal survey that includes several infill shot gathers is much less time consuming than a standard survey.

(iii) Traveltimes from refractions can be used to estimate statics (Yilmaz 2001). The problem is that land surveys have coarse spacing between recording stations and so prevent an estimate of the shallow

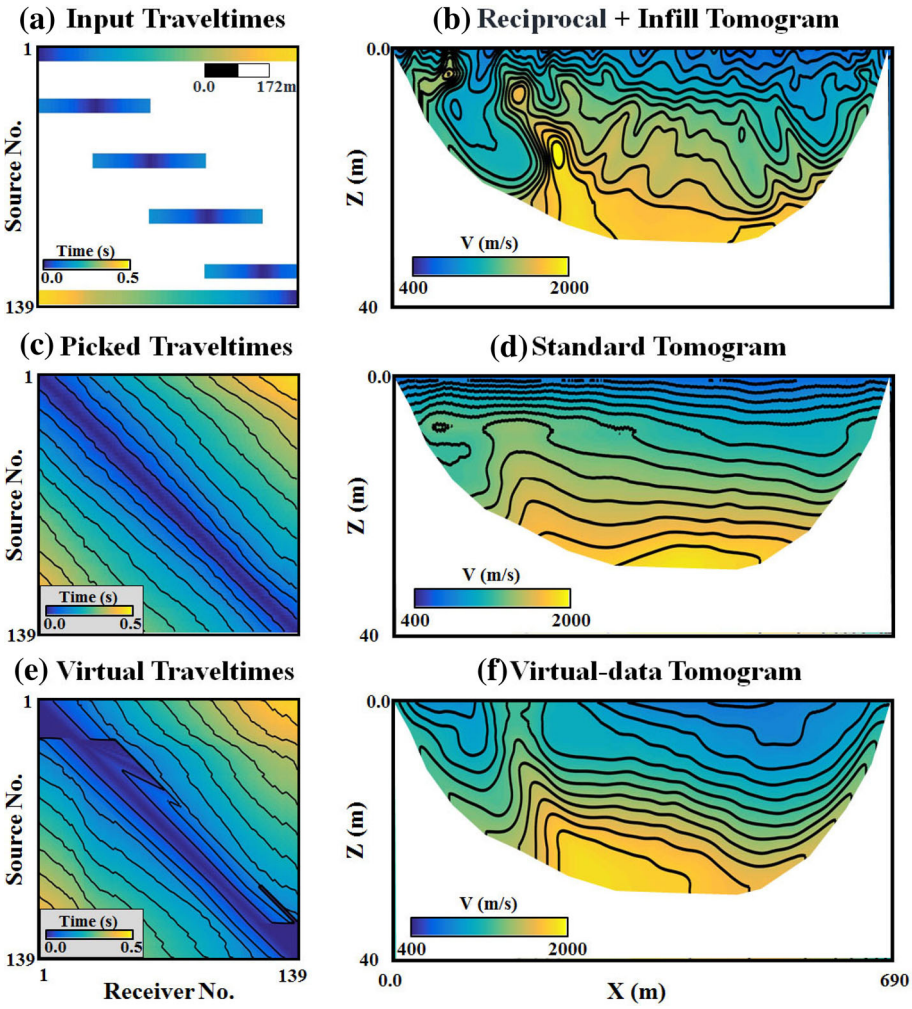

Figure 16. East Africa traveltime picks for the (a) reciprocal, (c) standard and (e) virtual data. The corresponding tomograms are on the right.

velocity model. Can an inexpensive source truck towing a short-offset land streamer be used to quickly obtain short-offset reciprocal gathers? If so, the virtual traveltimes can be inverted to give a detailed estimate of the shallow velocity distribution. A source is required at both the front and back of the land streamer.

(iv) Refraction surveys in environmentally sensitive sites, such as archaeological surveys, might use a parsimonious survey to minimize the number of sources.

Some limitations of parsimonious refraction interferometry are the following.

(i) The virtual traveltimes can have up to three times the variance of the recorded traveltimes for random uncorrelated picking errors.

(ii) The first arrivals are assumed to be mostly head waves, which is strictly not true for velocities that strongly increase with depth. Surprisingly, our tests suggest that parsimonious tomography has some tolerance to traveltimes associated with diving waves. We do not always expect this type of tolerance for wide-offset refractions from the deep part of a basin or crust.

(iii) First-arrivals with low SNRs will contain large picking errors at far-offset traces. These large picking errors are expected to propagate to all source-receiver offsets in the virtual data.

(iv) If the reciprocal pair of shots is too far apart then this will lead to poor lateral resolution and gaps in the illumination of the subsurface. This problem can be partly mitigated by picking later direct and refraction arrivals, or by conducting additional reciprocal surveys with shorter offsets across the survey area.

(v) Out-of-the-plane refractions in the reciprocal survey will degrade the accuracy of the virtual refractions.

(vi) The recorded first-arrival traveltimes must satisfy the consistency assumption, which is difficult if only two reciprocal shot gathers are recorded. One possibility is to examine these two reciprocal gathers for evidence of head waves, where first arrivals have the same slope for a large range of source-receiver offsets. Diving waves might be recognized by the continuous decrease of the slope dt/dx with increasing offset. Alternatively, one can record some extra shots in the interior over a portion of the line and test the traveltimes from these shot gathers for head-wave consistency.

REFERENCES

Bishop, T.N. *et al.*, 1985. Tomographic determination of velocity and depth in laterally varying media, *Geophysics,* **50,** 903–923.

Dong, S., Sheng, J. & Schuster, G.T., 2006. Theory and practice of refraction interferometry, *SEG Expanded Abstracts,* pp. 3021–3025.

Hanafy, S.M., Jonsson, S. & Klinger, Y., 2014. Imaging normal faults in alluvial fans using geophysical techniques: field example from the coast of gulf of aqaba, saudi arabia, *SEG Technical Program Expanded Abstracts,* pp. 4670–4674.

Menke, W., 1984. *Geophysical Data Analysis: Discrete Inverse Theory,* Academic Press Inc.

Nemeth, T., Normark, E. & Qin, F., 1997. Dynamic smoothing in crosswell traveltime tomography, *Geophysics,* **62**(1), 168–176.

Prodehl, C. & Mooney, W.D., 2012. *Exploring the Earth's Crust: History and Results of Controlled-source Seismology,* Geological Soc. Am. Mem.

Qin, F., Luo, Y., Olsen, K.B., Cai, W. & Schuster, G.T., 1992. Finite-difference solution of the eikonal equation along expanding wavefronts, *Geophysics,* **57,** 478–487.

Schuster, G.T., 1988. An analytic generalized inverse for common depth point and vertical seismic profile traveltime equations, *Geophysics,* **53,** 314–325.

Schuster, G.T., 2005. Fermat's interferometric principle for target-oriented traveltime tomography, *Geophysics,* **70,** U47–U50.

Schuster, G.T., Huang, Y., Zhou, M., Yu, J., Hanafy, S.M., Alhagan, O. & Dai, W., 2014. Review on improved seismic imaging with closure phase, *Geophysics,* **79,** W11–W25.

Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Phy. Rev. E,* **69,** 046610-1–046610-8.

Stein, S. & Wyession, M., 2003. *An Introduction to Seismology, Earthquakes, and Earth Structure,* Blackwell Publishing Company.

Tao, D. & Yasuda, M., 2002. A spectral characterization of generalized real symmetric centrosymmetric and generalized real symmetric skew-centrosymmetric matrices, *SIAM J. Matrix Anal. Appl.,* **23,** 885–895.

Wapenaar, K., 2004. Retrieving the electrodynamic Green's function of an arbitrary inhomogeneous medium by cross-correlation, *Phys. Rev. Lett.,* **93,** 254301, doi:10.1103/PhysRevLett.93.254301.

Yilmaz, O., 2001. *Seismic Data Analysis,* SEG Publishing.

Yilmaz, O., 2015. *Engineering Seismology with Applications to Geotechnical Engineering,* SEG Publishing.

SUPPORTING INFORMATION

Supplementary data are available at *GJIRAS* online.

Figure 1. (a) Aqaba velocity model, (b) tomogram inverted from the two reciprocal and some additional direct-wave traveltimes, (c) standard tomogram inverted from the 14 400 actual traveltimes in 240 shot gathers, with a shot at each geophone location, and (d) tomogram inverted from virtual and interpolated direct-wave traveltimes. The Aqaba model is based on results from a seismic survey near the Gulf of Aqaba.
Figure 2. (a) The actual traveltimes, (b) two reciprocal and five infill shot gathers used as input to the proposed method, (c) the virtual traveltimes calculated from eq. (3) and (d) the difference between the actual and the virtual traveltimes.
Figure 3. (a) The traveltime differences and (b) the corresponding histogram of the synthetic model.
Figure 4. The CPG of the (a) synthetic model and (b) Wadi Qadid field example.
Figure 5. A sample CSG from Wadi Qadid field data.
Figure 6. Wadi Qadid traveltime picks for the (a) reciprocal, (c) standard and (e) virtual traveltimes. The corresponding tomograms are on the right-hand column of figures.
Figure 7. (a) The traveltime differences and (b) the corresponding histogram of Wadi Qadid example.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

APPENDIX A: CONDITION NUMBER OF $[\mathbf{L}^T\mathbf{L}]$

We will derive the formulas for the condition number κ of the $[\mathbf{L}^T\mathbf{L}]$ associated with the reciprocal and virtual-traveltime equations. The traveltimes are for the refractions in the Fig. A1 slowness model. By definition, the 1-norm condition number is

$$\kappa(\mathbf{L}^T\mathbf{L}) = ||[\mathbf{L}^T\mathbf{L}]||\,||[\mathbf{L}^T\mathbf{L}]^{-1}||,$$

$$= \left\{\max_j \sum_{i=1}^{N} |[\mathbf{L}^T\mathbf{L}]_{ij}|\right\}\left\{\max_j \sum_{i=1}^{N} |[\mathbf{L}^T\mathbf{L}]^{-1}_{ij}|\right\}. \tag{A1}$$

A1 κ^{recip} for Reciprocal Data

Assume the two-layer model in Fig. A1, where the second layer has lateral heterogeneities with an unknown slowness s_i in each of the square boxes. The slowness of the top layer is homogeneous and is assumed to be known, the thickness of this top layer is also known, and each box has a width of l. The goal is to estimate the condition number of $[\mathbf{L}^T\mathbf{L}]$ associated with the reciprocal refraction traveltimes where a source is located at each end of the model.

The traveltime equations for a shot at position 0 in Fig. A1(a) and one at $X = 5$ can be represented by $\mathbf{Ls} = \mathbf{t}$, where

$$\mathbf{Ls} = \begin{bmatrix} l & 0 & 0 & 0 & 0 \\ l & l & 0 & 0 & 0 \\ l & l & l & 0 & 0 \\ l & l & l & l & 0 \\ l & l & l & l & l \\ 0 & 0 & 0 & 0 & l \\ 0 & 0 & 0 & l & l \\ 0 & 0 & l & l & l \\ 0 & l & l & l & l \\ l & l & l & l & l \end{bmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{pmatrix} = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ \tilde{t}_1 \\ \tilde{t}_2 \\ \tilde{t}_3 \\ \tilde{t}_4 \\ \tilde{t}_5 \end{pmatrix} = \mathbf{t}, \tag{A2}$$

where t_i ($\tilde{t}_i$) is the traveltime for a rightgoing (leftgoing) refraction along a horizontal ray path. The traveltimes t_i and $\tilde{t}_i$ are considered to be the recorded traveltimes after subtracting the upgoing and downgoing traveltimes in the top layer, which are assumed to be known.

The $\mathbf{L}$ matrix can be decomposed into

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_0 \\ \tilde{\mathbf{L}}_0 \end{bmatrix}, \tag{A3}$$

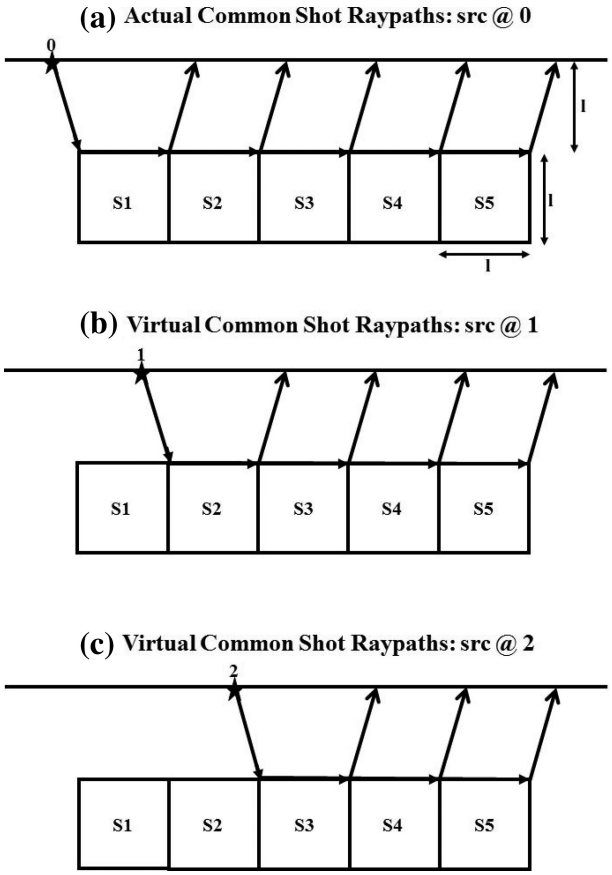

Figure A1. Ray paths for a common shot gather with the source at positions (a) 0, (b) 1 and (c) 2. Here, it is assumed that the source at position 0 is an actual source and the sources at 1 and 2 are virtual sources generated from a reciprocal pair of shot gathers.

where

$$\mathbf{L}_0 = l\mathbf{Q} = l\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}; \quad \bar{\mathbf{L}}_0 = l\mathbf{Q}\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{A4}$$

where the inverse to $\mathbf{Q}$ is the backward difference matrix (Schuster 1988)

$$\mathbf{Q}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}. \tag{A5}$$

Therefore the normal matrix $[\mathbf{L}^T\mathbf{L}] = \mathbf{L}_0^T\mathbf{L}_0 + \bar{\mathbf{L}}_0^T\bar{\mathbf{L}}_0$ is

$$[\mathbf{L}^T\mathbf{L}] = l^2\begin{bmatrix} 6 & 5 & 4 & 3 & 2 \\ 5 & 6 & 5 & 4 & 3 \\ 4 & 5 & 6 & 5 & 4 \\ 3 & 4 & 5 & 6 & 5 \\ 2 & 3 & 4 & 5 & 6 \end{bmatrix}. \tag{A6}$$

and its inverse is

$$[\mathbf{L}^T\mathbf{L}]^{-1} = \frac{1}{l^2}\begin{bmatrix} 0.5625 & -0.5 & 0 & 0 & 0.0625 \\ -0.5 & 1.0 & -0.5 & 0 & 0 \\ 0 & -0.5 & 1.0 & -0.5 & 0 \\ 0 & 0 & -0.5 & 1.0 & -0.5 \\ 0.0625 & 0 & 0 & -0.5 & 0.5625 \end{bmatrix}. \tag{A7}$$

In general, the normal matrix $[\mathbf{L}^T\mathbf{L}]$ for $2N$ reciprocal equations is the Toeplitz matrix

$$[\mathbf{L}^T\mathbf{L}] = l^2\begin{bmatrix} N+1 & N & N-1 & . & . & . & . & 3 & 2 \\ N & N+1 & N & N-1 & . & . & . & . & 3 \\ N-1 & N & N+1 & N & . & . & . & . & . \\ N-2 & N-1 & N & N+1 & . & . & . & . & . \\ . & . & . & . & \ddots & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & N+1 & N & N-1 \\ 3 & . & . & . & . & . & N & N+1 & N \\ 2 & 3 & . & . & . & . & N-1 & N & N+1 \end{bmatrix}, \tag{A8}$$

which is defined as

$$[\mathbf{L}^T\mathbf{L}]_{ij} = l^2\begin{cases} N+1-(j-i) & i \le j \\ N+1-(i-j) & i > j. \end{cases} \tag{A9}$$

Therefore, for even values of N, the 1-norm of the normal matrix for reciprocal data is

$$\begin{aligned}\max_j \sum_{i=1}^{N} |[\mathbf{L}^T\mathbf{L}]_{ij}| &= \sum_{i=1}^{N} |[\mathbf{L}^T\mathbf{L}]_{i\,\frac{N}{2}}|,\\ &= l^2\left\{\sum_{i=1}^{N/2}[N/2+1+i]+\sum_{i=N/2+1}^{N}[3N/2+1-i]\right\},\\ &= l^2\left[\frac{3N^2}{4}+N\right].\end{aligned} \tag{A10}$$

The inverse matrix for eq. (A8) is

$$[\mathbf{L}^T\mathbf{L}]^{-1} = \frac{1}{l^2}\begin{bmatrix} \alpha & -0.5 & 0 & 0 & \cdots & 0 & \beta \\ -0.5 & 1 & -0.5 & 0 & \cdots & 0 & 0 \\ 0 & -0.5 & 1 & -0.5 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ . & . & . & . & . & . & . \\ . & . & . & -0.5 & 1 & -0.5 & 0 \\ 0 & 0 & \cdots & 0 & -0.5 & 1 & -0.5 \\ \beta & 0 & \cdots & 0 & 0 & -0.5 & \alpha \end{bmatrix}, \tag{A11}$$

where $\alpha = \frac{0.5N+2}{N+3}$ and $\beta = -\frac{\alpha N}{3} + \frac{0.5(N+1)}{3}$ for $N \geq 3$. The 1-norm of this inverse is

$$\max_j \sum_{i=1}^{N} |[\mathbf{L}^T\mathbf{L}]^{-1}_{ij}| = \frac{2}{l^2}. \tag{A12}$$

Substituting eqs (A10) and (A12) into eq. (A1) gives the condition number for the reciprocal traveltimes equations:

$$\kappa(\mathbf{L}^T\mathbf{L})^{\text{recip}} = \frac{3N^2}{2} + 2N, \tag{A13}$$

so that $\kappa(\mathbf{L}^T\mathbf{L})^{\text{recip}}$ grows quadratically with the number of unknowns N.

A2 κ^{virt} for virtual Data

We will now derive the condition number $\kappa^{\text{virt}} = N^2/2 + N$ for the $\mathbf{L}^T\mathbf{L}$ from virtual data, which is about three times smaller than that for reciprocal data. This suggests that there will be many fewer slowness models that satisfy the noisy virtual data compared to the reciprocal data.

As discussed previously, virtual refraction traveltimes can be created from the reciprocal traveltimes. Thus, the virtual-traveltime equations for the ray paths depicted in Fig. A1(b) for a virtual source at position 1 are given by

$$\mathbf{L}_1\mathbf{s} = l\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix}\begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{pmatrix} = \begin{pmatrix} t'_1 \\ t'_2 \\ t'_3 \\ t'_4 \\ t'_5 \end{pmatrix} = \mathbf{t}_1, \tag{A14}$$

where $\mathbf{t}_1$ is the vector of virtual traveltimes. In this case $\mathbf{L}_1$ can be decomposed into the product of the summation $\mathbf{Q}$ matrix in eq. (A4) and the column-mask matrix $\mathbf{M}_1$:

$$\mathbf{L}_1 = l\mathbf{Q}\mathbf{M}_1 = l\mathbf{Q}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{A15}$$

The role of the column-mask matrix $\mathbf{M}_i$ is to mask columns $(1, 2, \ldots\ i)$ of $\mathbf{Q}$.

In general, for N unknown slowness cells, N virtual shot gathers can be created similar to those shown in Fig. A1. The picked traveltimes give rise to the overdetermined system of equations

$$l\begin{bmatrix} \mathbf{QM_0} \\ \mathbf{QM}_1 \\ . \\ . \\ . \\ \mathbf{QM}_{N-1} \end{bmatrix}\mathbf{s} = \begin{pmatrix} \mathbf{t_0} \\ \mathbf{t}_1 \\ . \\ . \\ . \\ \mathbf{t}_{N-1} \end{pmatrix}. \tag{A16}$$

The above system of equations lead to the $N \times N$ normal equations

$$\overbrace{l^2\left[\sum_{i=0}^{N-1}\mathbf{M}_i^T\mathbf{Q}^T\mathbf{Q}\mathbf{M}_i\right]}^{\mathbf{L}^T\mathbf{L}}\mathbf{s} = \overbrace{l\sum_{i=0}^{N-1}\mathbf{M}_i^T\mathbf{Q}^T}^{\mathbf{L}^T}\mathbf{t}_i, \tag{A17}$$

where $[\mathbf{L}^T\mathbf{L}]$ is the symmetric matrix

$$[\mathbf{L}^T\mathbf{L}] = l^2\begin{bmatrix} N & N-1 & N-2 & . & . & . & . & . & . \\ N-1 & 2(N-1) & 2(N-2) & . & . & . & . & . & . \\ N-2 & 2(N-2) & 3(N-2) & . & . & . & . & . & . \\ N-3 & 2(N-3) & 3(N-3) & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & 2(N-2) & N-2 & \\ 2 & . & . & . & . & . & 2(N-1) & N-1 & \\ 1 & . & . & . & . & . & N-1 & N & \end{bmatrix}, \tag{A18}$$

which is also classified as a bisymmetric matrix (Tao & Yasuda 2002), where one symmetry axis is along the matrix diagonal and the other is along the axis that intersects the bottom left and top right of the matrix. In this case

$$[\mathbf{L}^T\mathbf{L}]_{ij} = i(N-j+1) \text{ for } i \le j;\ [\mathbf{L}^T\mathbf{L}]_{ij} = j(N-i+1) \text{ for } i > j. \tag{A19}$$

Therefore, the formula for the 1-norm of $[\mathbf{L}^T\mathbf{L}]$ is

$$\begin{aligned} ||\mathbf{L}^T\mathbf{L}|| = \max_j \sum_{i=1}^{N}[\mathbf{L}^T\mathbf{L}]_{ij} &= \sum_{i=1}^{N}[\mathbf{L}^T\mathbf{L}]_{i\frac{N}{2}}, \\ &= l^2\sum_{i=1}^{\frac{N}{2}} i\left(\frac{N}{2}+1\right) + l^2\sum_{i=\frac{N}{2}+1}^{N}\frac{N}{2}(N-i+1), \\ &= l^2\left(\frac{N}{2}+1\right)\sum_{i=1}^{\frac{N}{2}} i + l^2\frac{N}{2}\sum_{i=1}^{\frac{N}{2}} i, \\ &= l^2\frac{N}{4}(N+1)\left(1+\frac{N}{2}\right), \end{aligned} \tag{A20}$$

where the maximum column sum is at $j = N/2$ for N an even number.

To determine $||[\mathbf{L}^T\mathbf{L}]^{-1}||$, we need the analytic inverse of eq. (A18):

$$[\mathbf{L}^T\mathbf{L}]^{-1} = l^{-2}\left[\sum_{i=0}^{N-1}\mathbf{M}_i^T\mathbf{Q}^T\mathbf{Q}\mathbf{M}_i\right]^{-1}, \tag{A21}$$

where

$$[\mathbf{L}^T\mathbf{L}]^{-1} = \frac{1}{l^2}\begin{bmatrix} \frac{2}{N+1} & -\frac{1}{N+1} & 0 & 0 & 0 & \cdots & . & . & . \\ -\frac{1}{N+1} & \frac{2}{N+1} & -\frac{1}{N+1} & 0 & 0 & \cdots & . & . & . \\ 0 & -\frac{1}{N+1} & \frac{2}{N+1} & -\frac{1}{N+1} & 0 & \cdots & . & . & . \\ 0 & 0 & -\frac{1}{N+1} & \frac{2}{N+1} & -\frac{1}{N+1} & \cdots & . & . & . \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & 0 & 0 & -\frac{1}{N+1} & \frac{2}{N+1} & -\frac{1}{N+1} \\ . & . & . & . & 0 & 0 & 0 & -\frac{1}{N+1} & \frac{2}{N+1} \end{bmatrix}. \tag{A22}$$

Thus, the 1-norm of $[\mathbf{L}^T\mathbf{L}]^{-1}$ is

$$||[\mathbf{L}^T\mathbf{L}]^{-1}|| = \max_j \sum_{i=1}^{N}[\mathbf{L}^T\mathbf{L}]_{ij}^{-1} = \frac{4l^{-2}}{N+1}. \tag{A23}$$

Inserting eqs (A20) and (A23) into the definition of the 1-norm condition number gives

$$\begin{aligned}\kappa([\mathbf{L}^T\mathbf{L}])^{\text{virt}} &= ||[\mathbf{L}^T\mathbf{L}]||\,||[\mathbf{L}^T\mathbf{L}]^{-1}||, \\ &= N\left(1+\frac{N}{2}\right).\end{aligned} \tag{A24}$$

Therefore, the ratio of the condition numbers in eqs (A24) and (A13) is

$$r = \frac{\kappa^{\text{recip}}}{\kappa^{\text{virt}}} = \frac{3N+4}{N+2}, \tag{A25}$$

where $r \to 3$ for large N. This says that for the large N the condition number of the reciprocal data is three times worse than that for the virtual data.

Fig. A2 illustrates this fact by plotting the condition numbers associated with the reciprocal (red stars) and virtual (green stars) shot gathers. Thus, the traveltime misfit function for reciprocal traveltimes will be more characterized by contours associated with long narrow valleys. Unlike virtual-traveltime tomography, reciprocal tomography will suffer from a greater range of velocity models that can explain nearly the same data.

APPENDIX B: SLOWNESS VARIANCE

Menke (1984) shows that the variance of the slowness σ_i^2 in the ith cell is equal to $\Sigma^2[\mathbf{L}^T\mathbf{L}]^{-1}$ if the data errors are uncorrelated zero-mean random variables with variance Σ^2. Therefore, the diagonal component of eq. (A22) is the virtual slowness $\bar{\sigma}_i^2$:

$$\bar{\sigma}_i^2 = \frac{\bar{\Sigma}^2}{l^2}\frac{2}{N+1}, \tag{B1}$$

where $\bar{\Sigma}^2 = 3\Sigma^2$ is the variance of the virtual data.[2] For large N, $\bar{\sigma}_i^2$ can be significantly less than slowness variance σ_i^2 associated with the reciprocal data:

$$\sigma_i^2 = \frac{\sigma^2}{l^2}, \tag{B2}$$

which follows from the diagonal component of the reciprocal inverse matrix in eq. (A11).

However, the formula in eq. (B1) that relates data variance to slowness variance assumes that the data are uncorrelated, which is not true for the virtual traveltimes computed with eq. (3). This is because the random errors in the virtual traveltimes are correlated with one another even though the traveltime errors in the reciprocal traveltimes are uncorrelated. For example, define the reciprocal traveltime error as $\bar{\tau}_{ij}$

[2] Each virtual traveltime is computed by adding two traveltimes and subtracting one traveltime. If the traveltime errors are uncorrelated and zero-mean random variables this means that the variance $\bar{\Sigma}^2$ of the virtual traveltime is three times the variance Σ^2 of the original data.

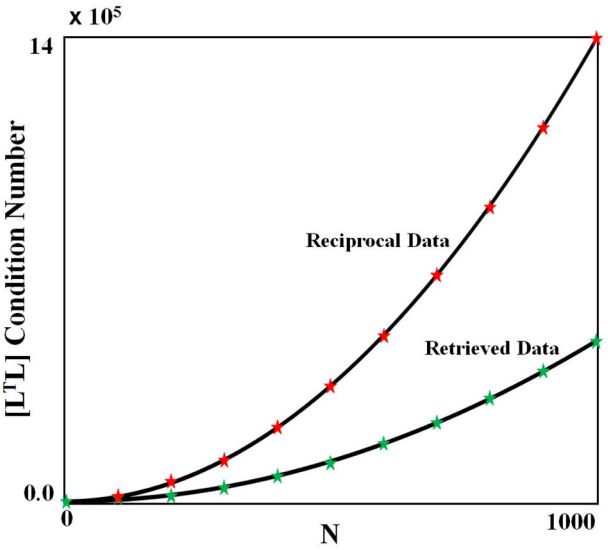

Figure A2. Condition number versus N computed from the 1-norm product $||[\mathbf{L}^T\mathbf{L}]|| \cdot ||[\mathbf{L}^T\mathbf{L}]^{-1}||$, where $[\mathbf{L}^T\mathbf{L}]$ is the $N \times N$ normal matrix numerically computed by a MATLAB code. The red (green) stars denote the condition numbers numerically computed for the reciprocal virtual data associated with the checkerboard model in Fig. A1. The black lines, computed by eqs (A24) and (A13), are in perfect agreement with the numerically computed condition numbers.

associated with the *i*th source and *j*th receiver, where the 'bar' indicates a traveltime picking error. In this case the reciprocal-traveltime errors are uncorrelated, i.e., $\langle \bar{\tau}_{ij}\bar{\tau}_{i'j'}\rangle \propto \delta_{ii'}\delta_{jj'}$, but the virtual-traveltime errors $\langle \delta\bar{\tau}_{CB}\delta\bar{\tau}_{CB'}\rangle = \langle \bar{\tau}^2_{AC}\rangle \neq 0$ are correlated, where B' is to the left of B for the virtual-traveltime errors $\delta\bar{\tau}_{ij}$. Here, $\langle\ \rangle$ represents the averaging operation over the random variables.