

# Compound earthquakes on a bimaterial interface and implications for rupture mechanics

Enning Wang,<sup>1,\*</sup> Allan M. Rubin<sup>1</sup> and Jean-Paul Ampuero<sup>2</sup>

<sup>1</sup>*Department of Geosciences, Princeton University, Princeton, NJ 08544, USA. E-mail: wen8411@gmail.com*

<sup>2</sup>*Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA*

Accepted 2014 February 4. Received 2014 January 31; in original form 2013 August 15

## SUMMARY

Earthquake ruptures on the San Andreas Fault are affected by the material contrast across the fault. Previous observations of microearthquakes at the northern end of the creeping section have found strong signals of asymmetry in both rupture directivity (preferential propagation to the SE), and aftershock asymmetry (many more to the NW, on timescales from 10 s to 9 hr). To understand the aftershock asymmetry, Rubin & Ampuero simulated slip-weakening ruptures on a bimaterial interface and observed differences in the timescales for the two edges to experience their peak stress after being slowed by barriers. This is suggestive of the possibility of asymmetry of subevents in compound earthquakes. A second possible source of subevent asymmetry is that when slowed by barriers, a significant tensile stress pulse is predicted to propagate in the SE but not the NW direction. To study the possible asymmetry of subevent distribution, we search for compound events using an empirical Green's function method. Three sections on the northern San Andreas and part of the Calaveras faults were selected where the events have high spatial density and similar focal mechanisms. About 677 candidate compound events were identified in a 28 869-event catalogue from 1984 to 2009. Most delays between the two subevents cluster around the shear wave transit time over the subevent separation, although with considerable scatter. For subevents on the San Andreas Fault separated by 0.7–2 times the estimated radius of the first subevent (the same spatial separation found to exhibit strong asymmetry of longer term aftershocks), nearly twice as many second subevents occurred to the SE of the first than to the NW. This asymmetry of second subevent distribution is not present on the Calaveras Fault, which does not have a significant across-fault contrast in wave speed in this region. One interpretation is that the extra SE subevents on the San Andreas Fault are representative of the events 'missing' from the 'longer term' (10 s–9 hr) aftershock population because they became part of the main shock.

**Key words:** Earthquake dynamics; Earthquake source observations; Earthquake interaction, forecasting and prediction.

## 1 INTRODUCTION

The right-lateral, strike-slip San Andreas Fault defines the plate boundary between the Pacific Plate to the SW and the North American Plate to the NE. As a result of the large displacement that has accumulated, in many regions the fault juxtaposes rocks with significantly different elastic properties. This material contrast influences the earthquake rupture process and may be important for earthquake hazard evaluation. Several studies have shown that at the northern end of the creeping section of the San Andreas, the Pacific Plate has higher seismic wave speeds at seismogenic depths than the North American Plate (Rubin 2002; McGuire & Ben-Zion 2005;

Thurber *et al.* 2007). Wang & Rubin (2011) took advantage of the tight spatial clustering of microearthquakes (magnitude  $M \lesssim 3.0$ ) in this region to search for rupture directivity by using spectral ratios between nearby events. They found a modest but obvious preference for rupture propagation to the SE, consistent with theoretical predictions and some numerical simulations. The same preferred direction was found by Lengline & Got (2011) and Kane *et al.* (2013) using somewhat different methods and earthquakes at the southern end of the creeping section near Parkfield. In the following, we refer to the preferred rupture propagation direction [SE, along the San Andreas Fault (SAF) in our study area] as the 'positive' direction, and the opposite (NW) as the 'negative' direction.

A second, independent observation of a bimaterial effect on earthquake rupture is the asymmetry of aftershock distribution on both short and long timescales. Rubin & Gillard (2000) and Rubin (2002)

\*Now at: Chevron ETC.

used a cross-correlation technique to relocate the microearthquakes in the same region as Wang & Rubin (2011) and found that on timescales from 10 s to 9 hr and distance ranges from 1 to 2 estimated main shock radii, two to three times more aftershocks occurred on the NW side of the main shock. However, in a portion of this region Zaliapin & Ben-Zion (2011) found that later and more distant aftershocks tended to occur to the SE of the main shock. They interpreted this as resulting from more SE-propagating events and the dominance of aftershocks in the propagation direction of the main shock.

Rubin & Ampuero (2007) simulated slip-weakening ruptures on a 1-D (line fault) bimaterial interface and observed differences in the timescales for the two edges to experience their peak stress after being slowed by stress barriers. The barrier on the ‘negative’ side reaches its peak stress when the *P*-wave stopping phase arrives from the opposite end of the rupture. This causes a delay of  $\sim 20$  ms for an event with a diameter of 100 m, about what is expected for an  $M_{2.2}$  earthquake with a 3 MPa stress drop. This may be long enough for a potential secondary rupture to be observed as a distinct subevent. In contrast, the same timescale for a barrier at the ‘positive’ front is nearly instantaneous (really the distance from the stopped rupture edge to the potential second subevent divided by the shear wave speed), possibly making a secondary subevent there indistinguishable from the main rupture.

Rubin & Gillard (2000) used the pulse width of the first *P* arrivals as an indication of rupture duration and found five compound events in a family of 72 similar earthquakes along the San Andreas Fault in northern California. In all cases, the anomalously wide pulses were visible only at stations to the SE, indicating that the second subevents were located to the NW of the first. Based on their simulations, Rubin & Ampuero (2007) interpreted this as being due to the above-mentioned asymmetry in the dynamic stressing-rate history on the two sides of a rupture on a bimaterial interface. However, other mechanisms could also affect the relative locations of the two subevents and may lead to asymmetry of the opposite sign. Similar to the effect of main shock propagation direction on the spatial distribution of distant aftershocks (Zaliapin & Ben-Zion 2011), preferential rupture propagation to the SE could result in more aftershocks to the SE on dynamic timescales. On a bimaterial interface, this tendency could be amplified by the tensile stress pulse that runs along the fault in the positive direction when that rupture front is stopped by a barrier, although in numerical simulations the pulse amplitude seems to decay quickly with distance (Rubin & Ampuero 2007). If instantaneous SE subevents were close enough to the edge of the initial event, they could also help to explain the scarcity of SE aftershocks observed by Rubin & Gillard (2000); that is, sources that could potentially have been aftershocks became subevents of a complicated main shock instead. It remains to be seen if such short-term and proximal subevents can be detected, however.

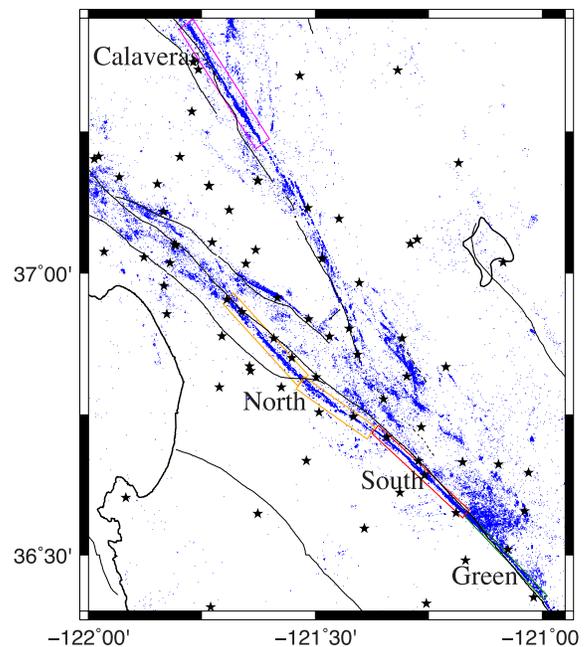
The distinction between continuous but perhaps complex ruptures, compound earthquakes with two separable subevents and main shock/aftershock pairs has not been clearly defined. At the long-timescale end, a plausible working distinction between compound earthquakes and main shock/aftershock pairs is that the secondary subevents in compound earthquakes occur on or near the timescale of the shear wave transit time between the subevents. At the short-timescale end, it is likely that there will be a gradation between ‘complicated’ earthquakes and earthquakes consisting of subevents with a clear separation. A distinction between the two in this case will likely depend upon data quality and the inversion procedure.

In this study, we use an empirical Green’s function (EGF) method to systematically search for compound earthquakes within a

catalogue of many thousands of events. As in Fischer (2005), the basic idea is to treat each earthquake (the ‘target’) as a potential compound earthquake, and each of its near neighbours as a simple EGF. The target waveform at each station was fitted as the sum of two occurrences of each EGF, shifted in time and scaled as needed. If the resulting misfit reduction is sufficiently large at a sufficiently large number of stations, compared to the misfit resulting from a single Gaussian convolution of the EGF, and if the delays between the two arrivals at all stations are azimuthally consistent, the target event is identified as compound and is removed from the pool of potential EGFs. The delays at all the stations were then used to estimate the relative locations and the centroid time difference of the two subevents.

## 2 METHOD

The data used in this study are velocity seismograms with a sampling rate of 100 Hz from the Northern California Seismic Network (NCSN) catalogue. Since the method depends upon finding multiple, closely spaced EGFs, the density of events is critical to the identification of potential compound earthquakes. In addition, the similarity of focal mechanism between a candidate compound event and all its EGFs are important for fitting the waveform of the compound event. With these two considerations in mind, we have selected four boxes that enclose four straight sections of the San Andreas and Calaveras faults with densely distributed events, which could potentially give rise to high-density event clusters with similar focal mechanisms. These boxes and the seismic stations used are shown in Fig. 1. Only data with signal-to-noise ratios higher than 1.5 were processed. The raw seismograms were demeaned and



**Figure 1.** Map of the San Andreas Fault in the study region. Microearthquakes smaller than magnitude 3 (blue dots) are plotted using the double-difference locations of Waldhauser & Schaff (2008); black stars indicate the locations of stations of the Northern California Seismic Network (NCSN). Events used in this study are enclosed in four boxes: on the San Andreas Fault, the two orange boxes together are referred to as the north box and the red one is called the south box. To the SE of the south box is a narrow green box, which is located between  $(-121.17, 36.57)$  and  $(-120.99, 36.43)$ . The magenta box is on the Calaveras Fault.

a Butterworth low-pass filter with a 25-Hz cut-off frequency was applied to remove high frequencies (from previous experience with the seismic data in this region, the spectral coherence between adjacent, as opposed to repeating, events drops sharply above 25 Hz). The filtered waveforms were then interpolated with five times the original sampling rate to increase resolution of the delay measurement between the two arrivals at each station. Since the coherence between the waveforms of the target and the EGF generally decreases with both event separation and the elapsed time since the first  $P$  arrival, 0.66-s signal windows containing the first  $P$  arrival were processed for each target event and EGFs within 400 m at stations where the cross-correlation coefficients are higher than 0.6. At each station, the target waveform  $w_t(t)$  was fitted by the sum of two occurrences of the same EGF  $w_E(t)$ , shifted in time and scaled as needed. To compensate for the differences in size and therefore apparent duration between the target and the EGF, the bigger subevent of the compound earthquake is fitted using the convolution of the EGF with a Gaussian function [ $g(t) = (\sigma\sqrt{2\pi})^{-1} \exp(-t^2/2\sigma^2)$ ] with standard deviation  $\sigma$  to effectively increase the pulse width of the EGF to better match the target, if needed (see Appendix C for an assessment of the magnitude differences between the targets and the EGFs). So ultimately the waveform of a compound event is fitted by the convolution of its EGF with a function consisting of a Gaussian pulse plus a delta function offset in time. We refer to this as a ‘pulse plus spike’ fit below. A grid search method was used to find the best-fitting time offsets between the target and the Gaussian and delta functions ( $t_1$  and  $t_2$ ) as well as the standard deviation of the Gaussian function  $\sigma$ , and the amplitudes of the Gaussian and delta functions were determined using a linear inversion (so five parameters per target-EGF pair). At all stations where the waveforms of both events have high signal-to-noise ratios the two time offsets  $t_1$  and  $t_2$  resulting in the minimum misfits were saved (see Appendix A for details).

Due to various noise sources, only stations with small misfits were used to determine the relative locations of subevents. The original misfit function  $F_0$  is defined as the sum of the absolute values of the target waveform:  $\sum_t |w_t(t)|$ . The misfit functions after subtracting first the larger subevent from the target and then both subevents are given by

$$F_1 = \sum_t |w_t(t) - A_L \cdot w_E(t - t_L) * g(t)|. \quad (1)$$

$$F_2 = \sum_t |w_t(t) - A_L \cdot w_E(t - t_L) * g(t) - A_S \cdot w_E(t - t_S)|. \quad (2)$$

Here  $A_L$  and  $A_S$  are the larger and smaller amplitude scalings for the same EGF, with  $t_L$  and  $t_S$  being the corresponding delays between the EGF and the two subevents. The reduction of misfits by subtracting the larger and both subevents were computed using the logarithmic difference of Tape *et al.* (2010)

$$MR_1 = \ln \left( \frac{F_0}{F_1} \right), \quad (3)$$

$$MR_2 = \ln \left( \frac{F_1}{F_2} \right). \quad (4)$$

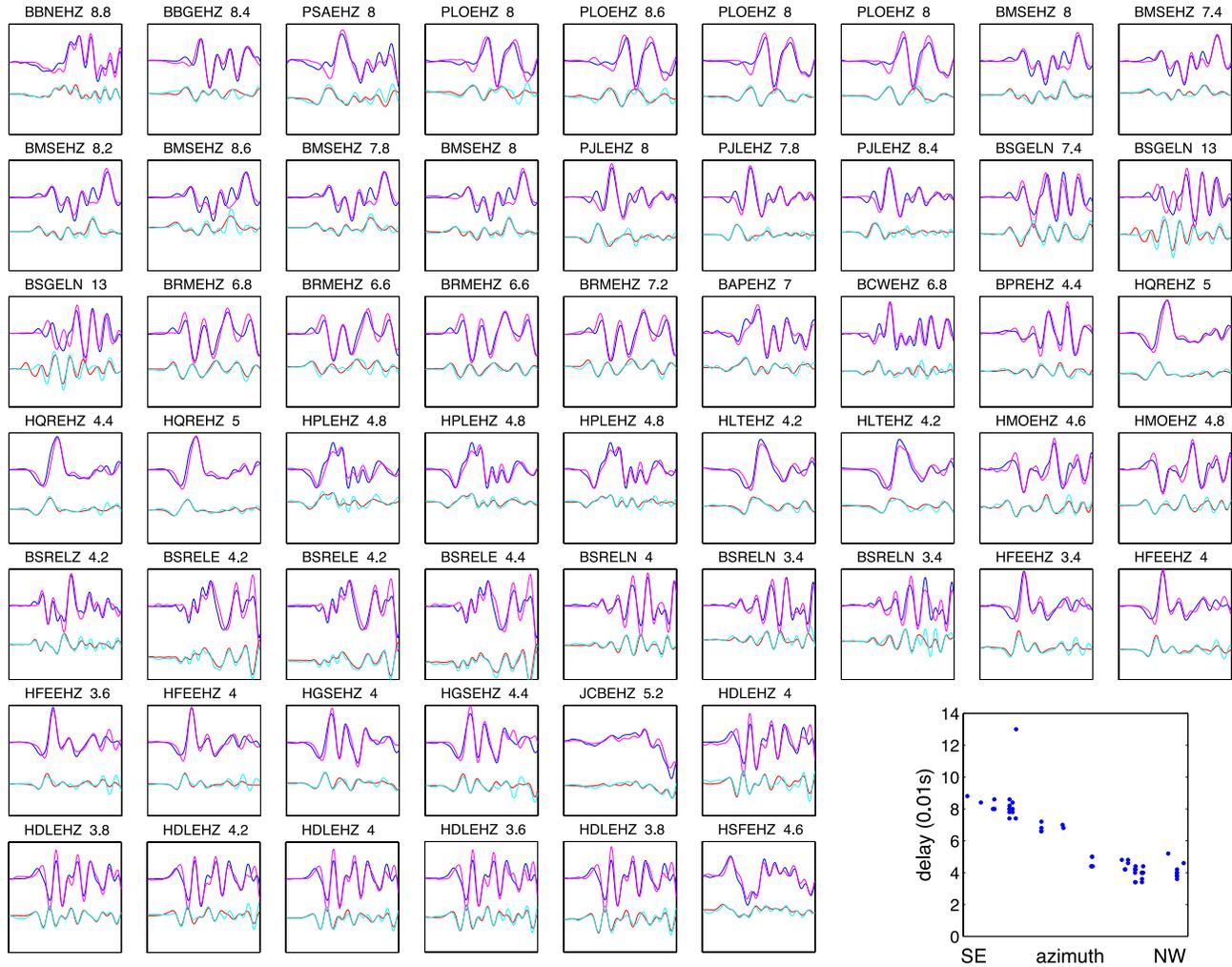
Due to the decreasing coherence between the target and the EGF with time after the first  $P$  arrival, the waveforms closely following the  $P$  arrival are often better fitted than the later part of the 0.66-s window. Since the delays between the two subevents at many stations where only the earlier part can be well fitted are still consistent between stations at nearby azimuths, assessing the quality of the fit

based on the misfit reductions defined on the entire window could potentially throw out many reliable detections. We found that requiring part of the window to have high misfit reductions could more effectively distinguish consistent detections. So the misfit reductions were computed on three equal segments of the 66-sample window. For a target/EGF/station combination to qualify, the highest misfit reduction  $MR_2$  for the three segments has to be higher than 0.85 and the second highest  $MR_2$  must be higher than 0.3. In addition, the ratio of the smaller amplitude to the larger  $A_S/A_L$  must be higher than 0.22 since very small subevents are not expected to be well resolved. Finally, to test for the robustness of the smaller subevent we fit the same target waveform using both a Gaussian pulse plus spike (the five-parameter inversion mentioned above) and a Gaussian pulse alone (only  $t$ ,  $A$  and  $\sigma$ , with a minimum misfit =  $F'_1$ ). The misfit reduction from the single EGF fit to the two-EGF fit,  $\ln(F'_1/F_2)$ , must be higher than 0.1 to justify the introduction of the second EGF.

Based on the criteria mentioned above, in Fig. 2 we plot the waveforms of the target, the two scaled EGFs and the difference between the target and the EGF scaled by the higher amplitude, for one of the identified compound events in the south box at all qualified stations. Each subplot shows the waveforms at one station, with the station name and the delay  $|t_1 - t_2|$  between the shifted EGFs in the title. The delay uses the absolute value since the  $P$  waves of the chronologically first subevent are assumed to arrive prior to that of the second subevent (see Appendix E for the effect of misidentification of the two subevents). Note that a station name can appear once for each different EGF. The blue and magenta curves at the top of each subplot are the target waveform and the smoothed EGF scaled by the higher amplitude that is used to match the larger subevent. At the bottom of each plot, the residual of subtracting the larger subevent from the target is plotted in cyan, while the EGF which was only scaled to match the residual without being smoothed by the Gaussian function is plotted in red. At most stations, the red EGFs used to fit the residuals have lower amplitudes than the magenta EGFs fitting the larger subevent, and the red precede the magenta by 0.04–0.09 s, indicating that in this example the first subevent is smaller. A smaller first subevent often leads to clean double arrivals since the second and bigger subevent is less affected by the coda of the first smaller one.

The delays are summarized as a function of station azimuth in the last plot on the lower right. Two stations symmetric with respect to the fault are defined to have the same azimuth. With the exception of two outliers possibly caused by complex ruptures of two EGFs at station BSGELN, the delays have an obvious azimuthal dependence, suggesting that the two subevents have a horizontal separation in the direction of the fault strike. Shorter delays at stations to the NW mean that the second subevent is on the NW side of the first. Since the delays at each station were determined independently, the azimuthal consistency across multiple stations with a wide range of azimuth gives us confidence in the results. The azimuthal consistency also serves as a guide in determining the criteria used to throw out low-quality stations, with the purpose of keeping as many stations as possible to avoid large azimuthal gaps while lowering the percentage of outliers in delay-azimuth plots.

We applied the grid search method for all the events in the north, south and the green boxes on the San Andreas Fault and the magenta box on the Calaveras Fault (Fig. 1). The total number of events in the four boxes are listed in Table 1. Most of these candidate events have only a small number of qualifying stations or large azimuthal gaps, which is incompatible with the expectation that the detections of double arrivals for a compound earthquake should



**Figure 2.** The upper two waveforms in each subplot are the target waveform (blue) and the convolution between the EGF and a Gaussian function (magenta) scaled to match the larger subevent. Of the two lower waveforms, the cyan is the difference between the target and the scaled, convolved EGF shown in the upper waveforms, and the red is the same EGF scaled by a lower amplitude to fit the cyan. The title of each subplot contains the station name and the difference between the time-shifts applied to the same EGF (magenta and red, unit: 0.01 s). The last plot on the lower right shows the relative EGF delays against the station azimuth. Note that not all data points can be seen because many plot on top of one another.

**Table 1.** Number of candidate events for different criteria in the three boxes on the San Andreas and the Calaveras faults.

	North	South	Green	Calaveras
Total	5877	8385	6732	7875
$\geq 10$ detections & small azimuth gap & $\max(\text{delay}) > 2$ samples	637	648	262	610
$\geq 10$ detections & small azimuth gap & $\max(\text{delay}) > 2$ samples & consistent with azimuth	168	183	67	259

show up at many stations. So we require that for an event to be considered as compound, a minimum of 10 unique stations (one station with multiple EGFs counts only as one station) should have double arrivals and that the largest azimuthal gap should be  $< 90^\circ$ .

Some target events have consistently small subevent time delays ( $< 0.02$  s or two original samples) and misfit reductions for the single pulse relative to the pulse plus spike close to the 0.1 cut-off. In addition, while the two scaling factors  $A_1$  and  $A_2$  are reasonably stable at most stations for the identified compound events with large delays (Fig. E1), those with small delays often have large variations in  $A_1$  and  $A_2$ . This suggests that the inversion algorithm might be having difficulty distinguishing between a clear compound earthquake and a simple rupture with a large pulse width. A visual

inspection of the target waveforms for these events reveals no clear double arrival at any station, unlike the example in Fig. 2 which shows clear double arrivals at several stations. Although these events with small delays might be compound earthquakes, we cannot state this with confidence. To eliminate these suspicious detections, we require that the higher end of the sinusoidal fit must be larger than two original samples. If the second subevent is delayed relative to the first by their spatial separation divided by the  $S$ -wave speed, then this two-sample cut-off corresponds to a spatial separation of  $\sim 40$  m (2/3 coming from the centroid time delay and 1/3 from the additional  $P$ -wave traveltime to stations in the opposite direction—for a discussion of possible differences between centroid times and origin times, see Appendix D). For comparison, 40 m is roughly the

diameter of a magnitude 1 earthquake, assuming circular ruptures with a 3-MPa stress drop.

The constraints on the number of stations, azimuthal gap and higher end of the sinusoidal fit cut the number of candidates from thousands to hundreds (second row in Table 1). The remaining events have significant numbers of stations with double arrivals but no guarantee on the azimuthal consistency of the delays.

### 3 RESULTS

Using the delays between the two subevents at multiple stations for the candidate events with  $\geq 10$  detections and small azimuthal gaps, we invert for the relative along-strike location of the second subevent relative to the first,  $\Delta x$ , and centroid time difference,  $\Delta t$ . Since the relative position vectors between most catalogue main shock–aftershock pairs are parallel to their host fault, either the San Andreas or the Calaveras (Rubin 2002), we force this to be the case when we invert for the relative location of the subevents within individual earthquakes. This is consistent with our observation that the maximum and minimum time delays between the subevents are found at stations lying along the strike of the fault as in Fig. 2. We tried some inversions that included the vertical separation  $\Delta z$  as a third parameter, but found that because there were very few stations with steep take-off angles  $\Delta z$  was poorly constrained. Rubin (2002) found that in the 10 s–9 hr range there were 62 per cent more aftershocks located along strike than along dip for event separations of less than 2 main shock radii, and that this ratio increased for larger event separations, so perhaps the compound earthquakes are dominated by subevents located along strike as well. A Markov chain Monte Carlo method (Tarantola 2004) is used both to search for the best-fitting  $\Delta x$  and  $\Delta t$  and to estimate the uncertainties of the model parameters from the scatter in the data. A norm-I misfit function (sum of absolute values) is used to suppress the effect of outliers, and we also throw out station/EGF pairs with more than six times the average misfit (only once; this is not done iteratively). Different weights were assigned based on data quality and redundancy. Higher weights were given to station/EGF pairs with high misfit reductions  $MR_2$  because of the similarity between the residual and the second application of the EGF. Lower weights were used for stations with many EGFs due to redundancy of information.

Not all the candidate events with more than 10 well-distributed stations have azimuthally consistent delays. Due to small signal-to-noise ratios, clipping of the target waveform, differences in ray path or complex ruptures of either the target or the EGF, the delays of some events have large scatter. So we use empirically chosen criteria to keep only events with obvious sinusoidal trends: The mean misfit at all retained stations needs to be smaller than one sample (0.01 s) and the average misfit normalized by the centroid time difference  $\Delta t$  has to be smaller than 0.2. In addition, the 90 per cent confidence interval of the normalized horizontal separation  $\Delta x/r_1$  has to be smaller than 1.6, where  $r_1$  is the estimated radius of the first subevent (only five events total from all the regions examined were eliminated by this criterion). The size of the bigger subevent is calculated from the catalogue magnitudes using the moment–magnitude relation of Abercrombie (1996), assuming circular ruptures with a 3-MPa stress drop. When the first subevent is the larger, this size is used to estimate  $r_1$ . When the first subevent is the smaller one, the ratio of the smaller subevent to larger subevent size is estimated by the ratio of the smaller scaling factor  $A_S$  and area below the scaled Gaussian pulse from the grid search, assuming that the subevents have the same stress drop.

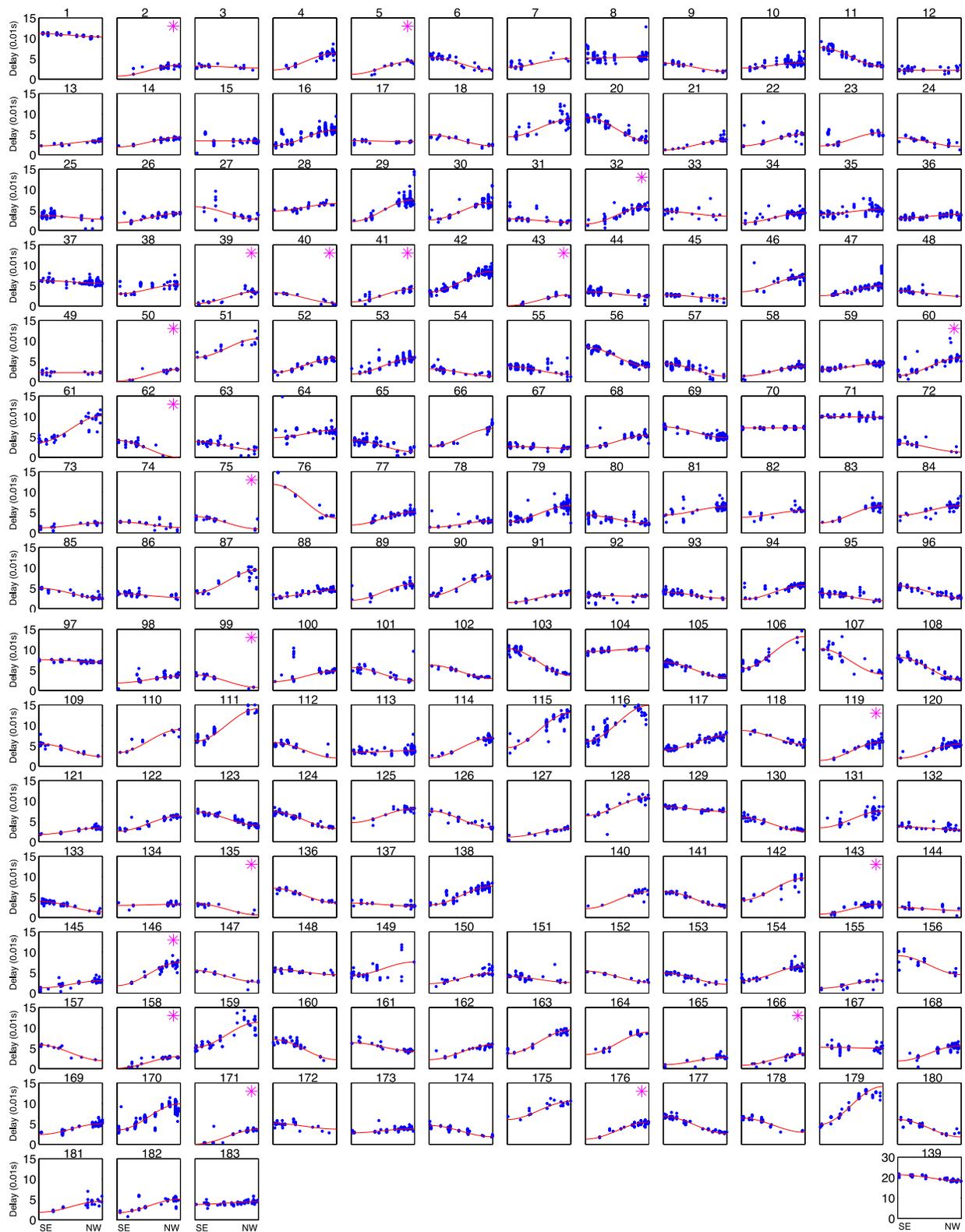
Of the target events with  $\geq 10$  detections and good station coverage (second line in Table 1), 31 per cent also satisfy the azimuthal consistency criteria (last line). Most of the discarded events lie at the low end of the  $\geq 10$  station range, which are possibly false detections caused by noise, or may be complex targets with more than two subevents (events 236 and 256 in Fig. B3, which barely passed the azimuthal consistency threshold, might be examples of earthquakes with three subevents). The delays for surviving candidate events in the south box are plotted against station azimuth as the blue dots in Fig. 3, and the delays for the identified compound events in the other three boxes are shown in Appendix B. The red curve in each plot shows the synthetic delays computed from the best-fitting model parameters:

$$\text{delay}_{\text{syn}} = -\cos(\phi) \cdot \Delta x/c_p + \Delta t, \quad (5)$$

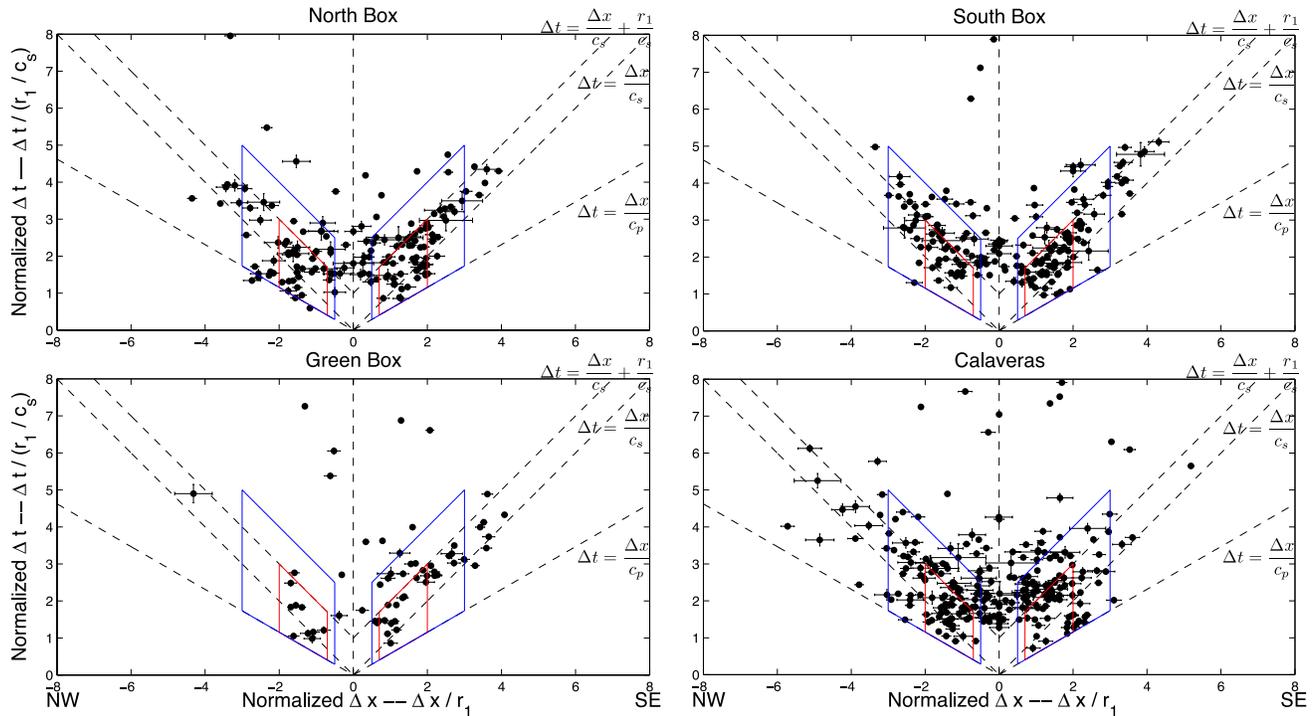
where  $c_p$  is the  $P$ -wave speed and the rays leaving the sources towards all stations are assumed to be horizontal only for plotting purposes (actual take-off angles derived from the 1-D NCSN velocity model were used in the inversion).  $\phi$  is the station azimuth with the SE and NW directions along the fault strike defined as  $0^\circ$  and  $180^\circ$ , respectively. Most of the remaining candidate events can be well fitted using the two-parameter model.

For the San Andreas Fault, 168, 183 and 67 compound events were identified in the north, south and green boxes, respectively. 259 compound events were identified on the Calaveras Fault. The best-fitting  $\Delta x$  and  $\Delta t$  are plotted in Fig. 4 along with 90 per cent confidence intervals from the Monte Carlo sampling.  $\Delta t$  is defined as the delay between the centroid times of the two subevents and is always positive, while positive and negative  $\Delta x$  indicate that the second subevent is on the SE and NW sides of the first, respectively. To better observe the interaction between the two subevents on scales relevant to event size and shear wave speed,  $\Delta x$  is normalized by the estimated radius of the first subevent  $r_1$ , and  $\Delta t$  is normalized by the  $S$ -wave traveltime over the estimated radius of the first subevent  $r_1/c_s$ , where  $c_s$  is the  $S$ -wave speed. Since the two subevents are so close in time and space, it is expected that they interact with each other and are causal. Therefore, we force  $\Delta t > \Delta x/c_p$  so that the  $P$ -wave arrival from the chronologically second subevent always arrives after that from the first. We examine this assumption further in Appendix E. To show this restriction, the  $P$ -wave traveltime as a function of distance is plotted as the bottom dashed lines in Fig. 4. The data points are expected to plot above this, although a few may plot below (if the red curves in Fig. 3 cross the horizontal axis; individual data points in Fig. 3 are constrained to be positive).

There is considerable scatter in Fig. 4. Many points lie close to the two indicated shear wave lines  $\Delta t = \Delta x/c_s$  and  $\Delta t = \Delta x/c_s + r_1/c_s$ , suggesting that these second subevents could have been triggered by the shear waves from the first. For events in the three boxes on the San Andreas Fault, the distribution of the second subevent relative to the first is asymmetric, with more second subevents located on the SE side of the first. Due to uncertainties introduced by directivity and relative size of the subevents, a second subevent triggered immediately by the shear wave of the first may not lie right on the line  $\Delta t = \Delta x/c_s$  in Fig. 4 (see Appendix D for possible effects of relative subevent size and directivity on the inferred normalized  $\Delta x$  and  $\Delta t$ ). In addition, since  $\Delta z$  is fixed at 0 due to lack of resolution, the event separation  $\sqrt{\Delta x^2 + \Delta z^2}$  could be much larger than the estimated  $\Delta x$ . So compound events with non-zero  $\Delta z$  could lie above  $\Delta t = \Delta x/c_s$  even if the second subevent was triggered by the shear wave from the first. Compound earthquakes with mode III second subevents would have small  $\Delta x$  for their  $\Delta t$  and would plot close to the vertical dashed line.



**Figure 3.** Delays between the two subevents as a function of station azimuth for the 183 identified compound earthquakes in the south box (blue dots). The horizontal separation  $\Delta x$  and centroid time difference  $\Delta t$  between the two subevents were inverted from the delays and the best-fitting curves are shown in red. Events with delays larger than 0.15 s are shown at the bottom right using a different scale. Stars are plotted on the upper right corner for events with  $\Delta t < \Delta x/c_s$ , where  $c_s$  is the shear wave speed.



**Figure 4.** Normalized centroid time difference  $\Delta t$  as a function of horizontal separation  $\Delta x$  for events in the three boxes on the San Andreas Fault as well as events on the Calaveras Fault.  $\Delta x$  is normalized using the estimated radius of the first subevent  $r_1$ . The estimated shear wave transit time over the first subevent radius  $r_1/c_s$  is used to normalize the centroid time difference  $\Delta t$ . Error bars indicate the 90 per cent confidence intervals from Monte Carlo sampling. The dashed lines are the traveltimes as a function of event separation for both  $P$  and  $S$ -wave velocities. Red polygons define the events whose second subevents lie close to the expected margin of the first, and which were triggered at close to the expected shear wave arrival from the first. Blue polygons define slightly larger space–time regions to account for uncertainties in  $\Delta x$  and  $\Delta t$  and for possible differences between the subevent centroid time and origin time difference.

**Table 2.** Numbers of mode II second subevents on different sides of the first.

	SE (red)	NW (red)	SE/NW (red)	SE (blue)	NW (blue)	SE/NW (blue)
North	40	22	1.8	67	48	1.4
South	46	22	2.1	75	56	1.3
Green	14	9	1.6	32	10	3.2
Calaveras	33	34	1.0	85	81	1.0

Because of these uncertainties, we use two different definitions of likely mode II second subevents to judge the robustness of our results: (1) events between 0.7 and 2 times the radius of the first subevent and between the  $P$ -wave line and one unit above  $\Delta t = \Delta x/c_s$  (within the red polygon), or (2) events from 0.5 to 3 times the radius of the first subevent and between the  $P$ -wave line and two units above  $\Delta t = \Delta x/c_s$  (blue polygon). 0.7–2 radii was the distance range observed by Rubin (2002) to exhibit pronounced aftershock asymmetry on timescales of 10 s–9 hr. The asymmetry we find is quantified by the number of second subevents in the blue and red polygons in Fig. 4. Table 2 summarizes the numbers of mode II second subevents as defined above. For the red polygon, the three boxes on the San Andreas Fault collectively have nearly twice as many subevents on the SE side as on the NW (100 versus 53). The asymmetry is weaker for the larger blue polygons but the sign is the same. For the north and south boxes, the added events are distributed symmetrically between the NW and the SE, as observed also by Rubin (2002) for catalogue aftershocks farther than 2 radii from the main shock. For the green box, the added events are skewed even more heavily to the SE. For the magenta box on the Calaveras Fault, most of which has only a 2–3 per cent velocity contrast (Zhao

& Peng 2008) and no longer term aftershock asymmetry (Rubin 2002), the asymmetry is negligible in both polygons.

Plots showing the locations of the detected compound earthquakes on the fault plane, and the tendency of the second subevent to occur to the NW or the SE of the first, are included as Supporting Information Fig. S1. Although a few regions appear to show some clustering of NW or SE second subevents, the overall sparsity of compound earthquakes makes it difficult to judge the statistical significance of this.

A substantial number of points in Fig. 4 plot between the  $\Delta t = \Delta x/c_p$  and  $\Delta t = \Delta x/c_s$  lines. This could be interpreted as resulting from supershear rupture of the first subevent, but this need not be the case. As pointed out in Appendix D, if the first subevent propagates unilaterally to the NW or SE, a second subevent located in the opposite direction could plot well below the  $\Delta t = \Delta x/c_s$  line. The events in question are marked by magenta stars in the upper right corners of the delay versus azimuth plots in Fig. 3 and Appendix B (events plotting between the  $\Delta t = \Delta x/c_p$  and  $\Delta t = \Delta x/c_s$  lines in Fig. 4 correspond to those where the delay at an azimuth of  $90^\circ$  ranges from 1 to  $c_p/c_s \approx 1.7$  times the half-amplitude of the red curve). While not all of these are necessarily

well resolved, many (see especially several of those on the Calaveras Fault in Fig. B3) clearly are.

The horizontal separation  $\Delta x$  of most subevents in Fig. 4 is less than 200 m and within three times the radius of the first subevent. The range of the grid search on  $t_1$  and  $t_2$  is  $-0.2$  to  $0.2$  s, where  $t = 0$  is defined to be zero time-shift relative to the picked (if available) or computed (if not) catalogue arrival times for the target and EGF (for details, see Appendix A). This results in a maximum possible delay of 0.4 s, and closer to 0.2 s if the catalogue arrival times for the EGF and first subevent of the target are accurate. The maximum delays occur at stations lying along the fault ( $\phi = 0^\circ$  or  $180^\circ$ ), so for subevents lying along the  $\Delta t = \Delta x/c_s$  lines in Fig. 4, the maximum delay is  $\Delta x(c_p^{-1} + c_s^{-1})$  (eq. 5). For representative values of  $c_s$  ( $\sim 3.5$  km s $^{-1}$ ) and  $c_p$  ( $\sim 6$  km s $^{-1}$ ), this maximum delay for  $\Delta x = 200$  m is 0.09 s, well within our search range of  $\sim 0.2$  s. Thus one interpretation of the lack of distant subevents in Fig. 4 is that short-term triggering at large interevent distances is very inefficient. However, for target subevents and EGFs that are not colocated, waveform similarity degrades with time following the  $P$  arrival, making the sensitivity of our detection method decrease as the delay between subevent arrivals increases. In addition, if the subevents are widely separated in space it may be difficult to find EGFs that are similar to both of them. For both reasons, we cannot rule out the possibility that the lack of widely separated subevents in Fig. 4 is a detection issue rather than an indication of fault behaviour.

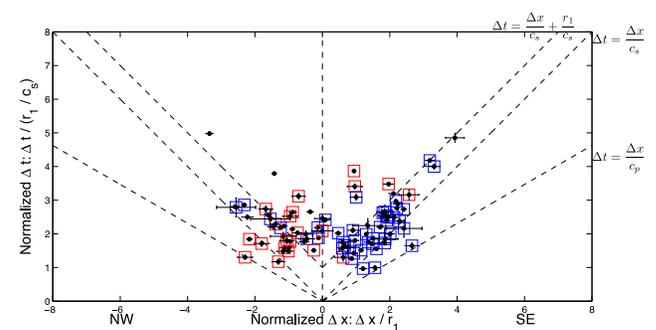
#### 4 DISCUSSION

The asymmetric distribution of secondary subevents on the SAF provides additional observational evidence of a bimaterial effect on earthquake rupture. For the north, south and green boxes on the SAF, the overall asymmetry in the red polygons in Fig. 4 (total #SE/#NW = 100/53 = 1.9) is almost as large but of opposite sign to the longer term aftershock asymmetry in the same distance range observed by Rubin (2002) for the south box (for mode II aftershocks #NW/#SE = 115/44 = 2.6). The excess of secondary subevents to the SE we observe could have been triggered by the tensile pulse expected to propagate in the SE direction after the SE front of the first subevent was slowed by a barrier.

Conceivably, the excess of secondary subevents to the SE found in this study is sufficient to explain the deficit of aftershocks to the SE found by Rubin (2002), once the different sensitivities of the two observational methods are accounted for. The 159 nearby mode II aftershocks found by Rubin (2002) amount to 2.7 per cent of the 5769 potential main shocks examined. Here, we find 153 secondary subevents within the red polygons of 20 990 events within the three SAF boxes in Fig. 1, or 0.73 per cent of the total. This is only one-fourth the detection rate found by Rubin (2002) in the same distance range, but the compound earthquakes identified here are skewed much more heavily towards the largest earthquakes in the catalogue (Appendix C). Overall, we identified 2.3 per cent of the 28 869 earthquakes examined as compound, but the rate was  $\sim 20$  per cent for magnitudes larger than 2.5, very likely because of lower resolving power for very small delay times. The longer term aftershock study does not suffer from this difficulty; the only observational bias is that most aftershocks are smaller than their main shock, and potential main shocks are skewed towards the catalogue cut-off magnitude. The median magnitude of the events found by Rubin (2002) to have nearby mode II aftershocks was 1.8–1.9.

When we began this study, we anticipated that if the tensile stress pulse propagating to the SE triggered a secondary subevent contiguous with the first, we might be unable to distinguish it as a separate subevent. The reason is that peak stress to the SE occurs simultaneously with the arrival of the tensile pulse emanating from the nearest (SE) rupture edge. That we see a significant excess of secondary events to the SE suggests that they might be triggered at some distance from the margin of the first, even if by only a fraction of a subevent radius, and even if some ultimately grow to be contiguous with the first. Based on the slip-weakening simulations of Rubin & Ampuero (2007), we also anticipated that nearby secondary subevents to the NW might be delayed relative to those occurring at the same distance to the SE, because in their simulations peak stress to the NW did not occur until the  $P$ -wave stopping phase arrived from the SE end of the rupture. We see no good evidence of such a delay in Fig. 4. If the NW-moving front of the first subevent propagated as a slip pulse, rather than as the edge of an expanding crack, then the time delay for reaching peak stress to the NW would be much reduced (any ‘stopping phase’ would come from the much closer trailing edge of the pulse). This might do away with any increased delay for subevents to the NW. Pulse-like ruptures on bimaterial faults were studied by Ampuero & Ben-Zion (2008).

Wang & Rubin (2011) used a spectral ratio method to invert for the rupture directivity of  $\sim 3100$  events that occurred from 1984 to 1998 in the south box in Fig. 1. Each event was assumed to be a continuous and bilateral rupture with propagating distances  $L_{SE}$  and  $L_{NW}$  in the SE and NW directions. The black dots in Fig. 5 show the  $\Delta x$  and  $\Delta t$  for the compound events found in this study for which a directivity estimate was also made by Wang & Rubin (2011). The blue boxes enclose events that were found by them to rupture predominantly to the SE ( $L_{NW}/L_{SE} < 0.8$ ) and events in the red boxes propagated farther to the NW ( $L_{SE}/L_{NW} < 0.8$ ). The directivity study found that more events ruptured to the SE, which is the reason for more blue boxes. Most of the blue boxes lie on the right side of the figure, suggesting that from the perspective of spectral ratios, many ‘compound earthquakes’ with their second subevents located to the SE appear as events that propagate preferentially to the SE. From the standpoint of the mechanics underlying that process, this gradation may not be surprising. Both preferential propagation to the SE and secondary subevents located to the SE could be manifestations of the tensile stress pulse associated with the SE-propagating front, with the latter operating on slightly longer timescales.



**Figure 5.** Relationship between rupture directivity from the spectral ratio study of Wang & Rubin (2011) and the relative locations of subevents for compound events identified in the south box in this study. The blue and red boxes enclose events that were found by Wang & Rubin (2011) to have propagated to the SE and NW, respectively.

This leaves unexplained the SE-propagating events with secondary subevents to the NW (blue boxes on the left in Fig. 5) and NW-propagating events with secondary subevents to the SE (red boxes on the right). Mechanically, such behaviour need not be surprising. Nielsen & Madariaga (2003) simulated unilateral ruptures as self-similar, self-healing pulses and noted that as the main pulse propagates, the stress at the point of origin continually increases and could ultimately lead to a second rupture. This would cause the second subevent to occur on the opposite side of the main pulse propagation direction; such behaviour was later observed in the elastodynamic simulations of Gabriel *et al.* (2012). This can also happen by chance on a fault with heterogeneous initial stresses, as illustrated in fig. 13(h) of Ampuero & Ben-Zion (2008). Either possibility could explain the blue boxes on left/red boxes on right if the rupture directivity estimated from spectral ratios captures the directivity of the first rupture (six of the nine events with the second subevents on the ‘wrong’ side have larger first subevents).

The ‘outliers’ in Fig. 5 could also be the result of the complicated spectra of compound events with longer arrival time separations between the two subevents. We have calculated the power spectra of two synthetic Gaussian pulses separated by different time delays. When the two pulses overlap significantly, their spectrum resembles that of a single wider pulse. This leads to generally longer period energy at stations in the direction opposite from the location of the second subevent, which in a spectral ratio inversion could mimic continuous propagation in the direction of the second subevent. However, when the subevent arrivals are more widely separated in time, more high-frequency energy is added to the spectrum. In this case, the propagation direction inferred from the spectral ratio method could be in the direction opposite to the location of the second subevent. Most of the outliers in Fig. 5 have relatively large normalized time delays, possibly consistent with this interpretation.

In principle, it might be possible to determine the directivity of individual subevents from the azimuthal distribution of the source pulse widths at different stations. To compensate for the size difference between the larger subevent and the EGF, in this study we convolve the EGF with a Gaussian function with standard deviation  $\sigma$ . However, because  $\sigma$  only represents the effective difference in pulse width between the larger subevent and the EGF, and because directivity could potentially make the pulse width of the EGF vary with azimuth, interpreting the azimuthal variation of  $\sigma$  in terms of the directivity of the larger subevent alone is problematic. Alternatively, the azimuthal variation of the scaling factors A1 and A2 between the two subevents (Fig. E1) could be another option for estimating the directivity of the larger subevent, but directivity could influence EGF amplitude as well as duration. As pointed out in Appendix A, a few subevents have sinusoidal trends in scaling factor that appear consistent with directivity, but no obvious correlation between subevent directivity and relative locations was observed.

## 5 CONCLUSIONS

We developed an algorithm to automatically detect compound earthquakes that exploits waveform similarity with nearby events. The method appears to work well for microearthquakes in central California, where 418 and 259 compound events were identified on the San Andreas and Calaveras faults, respectively. Most detected second subevents cluster around the time when the shear wave from the first subevent arrives, although with considerable scatter. For compound earthquakes on the San Andreas Fault, there is an asymmetry in the distribution of nearby secondary subevents,

with significantly more occurring on the SE side of the first. We interpret this as being due to the tensile normal stress associated with rupture front moving in the direction of motion of the slower wave speed side of the fault. There is no apparent asymmetry in subevent distribution on the Calaveras Fault, which unlike the San Andreas does not have a significant across-fault velocity contrast in this region. The larger number of SE second subevents on the San Andreas Fault could be representative of those ‘missing’ from the longer term (10 s–9 hr) aftershock population in this region because they became part of the main shock instead. Thus this study provides additional evidence that a large across-fault contrast in seismic velocities exerts a significant influence on earthquake behaviour.

## ACKNOWLEDGEMENTS

The waveform data used in this study were obtained by the U.S. Geological Survey Northern California Seismic Network (NCSN) and made accessible by the Northern California Earthquake Data Center (NCEDC). We thank editor Eiichi Fukuyama and two anonymous reviewers for helpful comments. This research was supported by NSF grant EAR-1113579, and by the U.S. Geological Survey (USGS), Department of the Interior, under USGS award number G11AP20051. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Government.

## REFERENCES

- Abercrombie, R.E., 1996. The magnitude-frequency distribution of earthquakes recorded with deep seismometers at Cajon Pass, southern California, *Tectonophysics*, **261**, 1–7.
- Ampuero, J.-P. & Ben-Zion, Y., 2008. Cracks, pulses and macroscopic asymmetry of dynamic rupture on a bimaterial interface with velocity-weakening friction, *Geophys. J. Int.*, **173**, 674–692.
- Fischer, T., 2005. Modelling of multiple events using empirical Green’s functions: method, application to swarm earthquakes and implications for their rupture propagation, *Geophys. J. Int.*, **163**, 991–1005.
- Gabriel, A.-A., Ampuero, J.-P., Dalguer, L.A. & Mai, P.M., 2012. The transition of dynamic rupture styles in elastic media under velocity-weakening friction, *J. geophys. Res.*, **117**, doi:10.1029/2012JB009468.
- Kane, D.L., Shearer, P.M., Goertz-Allmann, B.P. & Vernon, F.L., 2013. Rupture directivity of small earthquakes at Parkfield, *J. geophys. Res.*, **118**, 212–221.
- Lengline, O. & Got, J.-L., 2011. Rupture directivity of micro-earthquake sequences near Parkfield, California, *Geophys. Res. Lett.*, **38**, L08310, doi:10.1029/2011GL047303.
- McGuire, J.J. & Ben-Zion, Y., 2005. High-resolution imaging of the Bear Valley section of the San Andreas fault at seismogenic depths with fault-zone head waves and relocated seismicity, *Geophys. J. Int.*, **163**, 152–164.
- Nielsen, S. & Madariaga, R., 2003. On the self-healing fracture mode, *Bull. seism. Soc. Am.*, **93**, 2375–2388.
- Rubin, A.M., 2002. Aftershocks of microearthquakes as probes of the mechanics of rupture, *J. geophys. Res.*, **107**, 2142.
- Rubin, A.M. & Ampuero, J.-P., 2007. Aftershock asymmetry on a bimaterial interface, *J. geophys. Res.*, **112**, B05307, doi:10.1029/2006JB004337.
- Rubin, A.M. & Gillard, D., 2000. Aftershock asymmetry/rupture directivity among central San Andreas fault microearthquakes, *J. geophys. Res.*, **105**, 19 095–19 109.
- Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.*, **180**, 433–462.
- Tarantola, A., 2004. *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM.

- Thurber, C.H., Brocher, T.M., Zhang, H. & Langenheim, V.E., 2007. Three-dimensional P wave velocity model for the San Francisco Bay region, California, *J. geophys. Res.*, **112**, B07313, doi:10.1029/2006JB004682.
- Waldhauser, F. & Schaff, D., 2008. Large-scale relocation of two decades of northern California seismicity using cross-correlation and double-difference methods, *J. geophys. Res.*, **113**, B08311, doi:10.1029/2007JB005479.
- Wang, E. & Rubin, A.M., 2011. Rupture directivity of microearthquakes on the San Andreas fault from spectral ratio inversion, *Geophys. J. Int.*, **186**, 852–866.
- Zaliapin, I. & Ben-Zion, Y., 2011. Asymmetric distribution of aftershocks on large faults in California, *Geophys. J. Int.*, **185**, 1288–1304.
- Zhao, P. & Peng, Z., 2008. Velocity contrast along the Calaveras fault from analysis of fault zone head waves generated by repeating earthquakes, *Geophys. Res. Lett.*, **35**, L01303, doi:10.1029/2007GL031810.

## APPENDIX A: INVERSION FOR TIME OFFSETS

To look for the best time offsets between the two subevents and the empirical Green's function (EGF), the waveform of the same EGF  $w_E$  is shifted by  $t_1$  and  $t_2$  and scaled by  $A_1$  and  $A_2$ , respectively. The sum of the two shifted and scaled EGFs is used to fit the target waveform  $w_t$ , which is assumed to be the superposition of the waveforms of two subevents plus noise  $\varepsilon$ :

$$w_t = A_1 \cdot w_E(t - t_1) * g(t) + A_2 \cdot w_E(t - t_2) + \varepsilon. \quad (\text{A1})$$

Here  $A_1 \cdot w_E(t - t_1)$  (the EGF fitting for the larger subevent) is convolved with a Gaussian function  $g(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{t^2}{2\sigma^2})$ , which allows for an increased source duration of the larger subevent as needed. The Gaussian function was truncated at the two tails where the amplitude falls below 10 per cent of the peak value, and was normalized by the sum of the sampled Gaussian function. To ensure the separation of the spike from the pulse, the time offset of the smaller subevent  $t_2$  has to be at least two interpolated samples (0.004 s) away from the truncated tail of the Gaussian pulse.

Fischer (2005) inverted for the relative locations and centroid time differences between the subevents of compound earthquakes by simultaneously fitting the waveforms at all the stations. In this study, we split the process into two separate steps by first solving for the delays at each station independently, and then using those delays to invert for the relative locations and centroid time differences. For each station/EGF pair, our inversion has five parameters:  $t_1$ ,  $t_2$ ,  $A_1$ ,  $A_2$  and  $\sigma$ . In contrast, Fischer (2005) solves for a single global value for each of  $A_1$  and  $A_2$  at all stations, and uses an empirically derived average relation between  $A$  and  $\sigma$  to determine  $\sigma$  for each subevent. While the global inversion is more immune to local minima at individual stations, independent inversions at different stations have their own advantages. First, using different scaling factors  $A_1$  and  $A_2$  at different stations allows for directivity of both the subevents and the EGF (resulting in azimuthal variations of  $A_1$  and  $A_2$ ). Perhaps more significantly, solving for time offsets independently at each station allows us to use the azimuthal consistency of the results as a check on the reliability of a compound earthquake detection. This is particularly important for us as many of the waveforms are clipped. In principle, it might be best to use a global inversion that includes a sinusoidal variation with azimuth of both  $A$  and  $\sigma$  for each subevent, but at present the computational cost of such a scheme is too high.

Because of the periodic nature of the seismic data, the model space contains many local minima, causing difficulties for inversion methods based on local gradients. So a grid search method is used at each station to look for the best-fitting  $t_1$ ,  $t_2$  and  $\sigma$ . For each set of fixed  $t_1$ ,  $t_2$  and  $\sigma$ , the two scaling amplitudes  $A_1$  and  $A_2$  are

determined using a linear inversion. Only the larger subevent is fitted using the smoothed EGF since adding another parameter in the grid search would greatly increase the computational expense.

The seismic data obtained from the Northern California Earthquake Data Center (NCEDC) were then resampled to five times the original sampling rate to increase the resolution in time (new sampling interval is 0.002 s). Unlike Fischer (2005), in which both  $P$  and  $S$  waves are used, our 0.66-s time windows only contain the  $P$  waves due to more frequent and severe clipping for the  $S$  waves. We initially excluded clipped waveforms, but ultimately included them after finding that even clipped stations often yielded delays that were azimuthally consistent with their neighbours.

The time offset search uses a 0.002-s increment in the range of  $-0.2$  to  $0.2$  s. The pulse width of the Gaussian function was searched between 0 and 5.4 original samples using 20 gridpoints. If the source durations, directivity and the path effects of the subevents and the EGF are all similar, the pulse width  $\sigma$  determined by the grid search will be smaller than one interpolated sample and the Gaussian function becomes a spike.

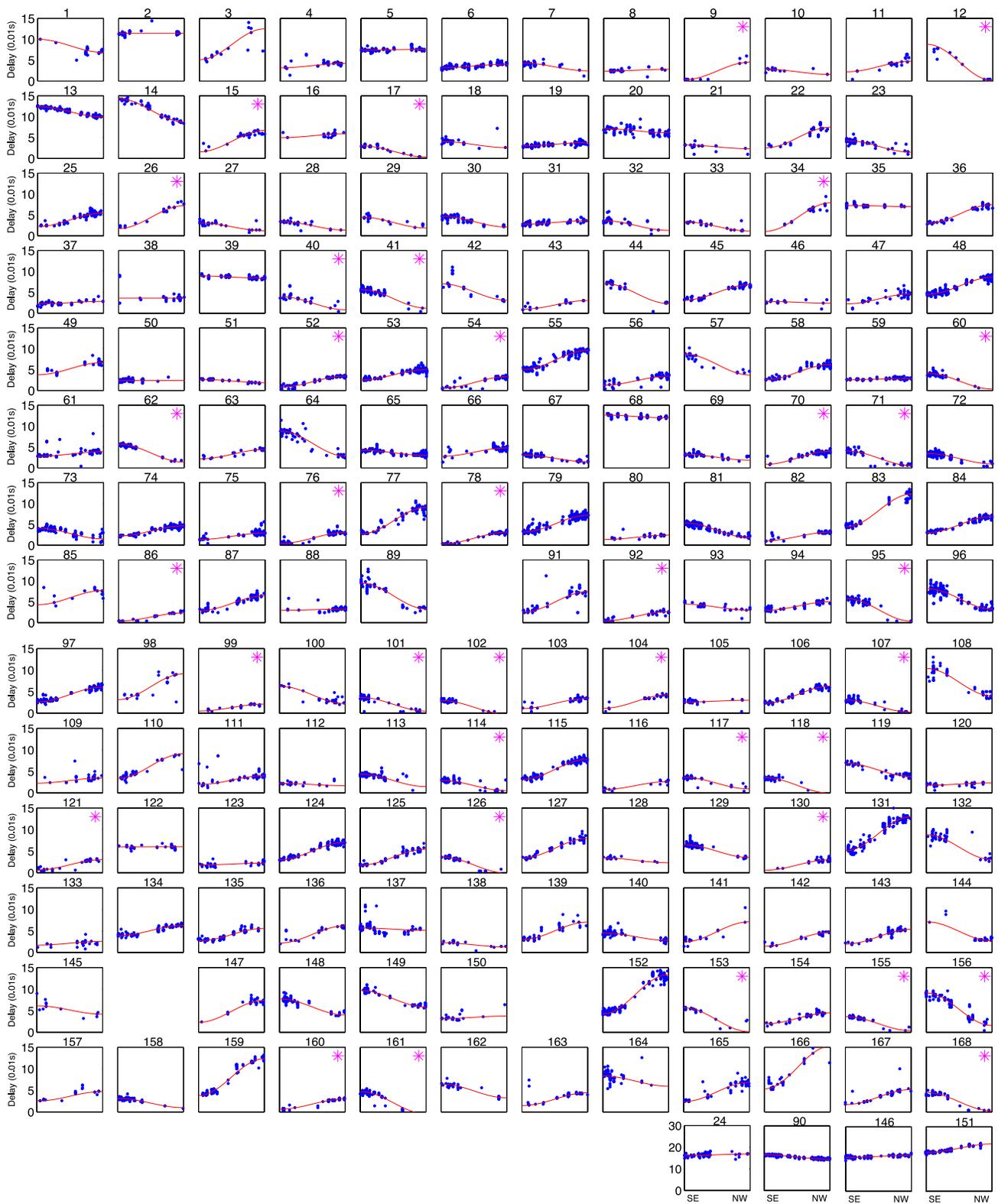
The signal of the chronologically first subevent is assumed to arrive first at all stations. Since there is no prior knowledge of the relative size of the subevents, the possibility of either one being bigger was tested by fitting the chronologically first and second subevents using the smoothed version of the EGF. The misfits produced in both cases were compared at all stations and the subevent which, when convolved with the Gaussian function, more frequently gave rise to the lower misfit was taken as the bigger one.

## APPENDIX B: AZIMUTH-DELAY PLOTS FOR OTHER REGIONS

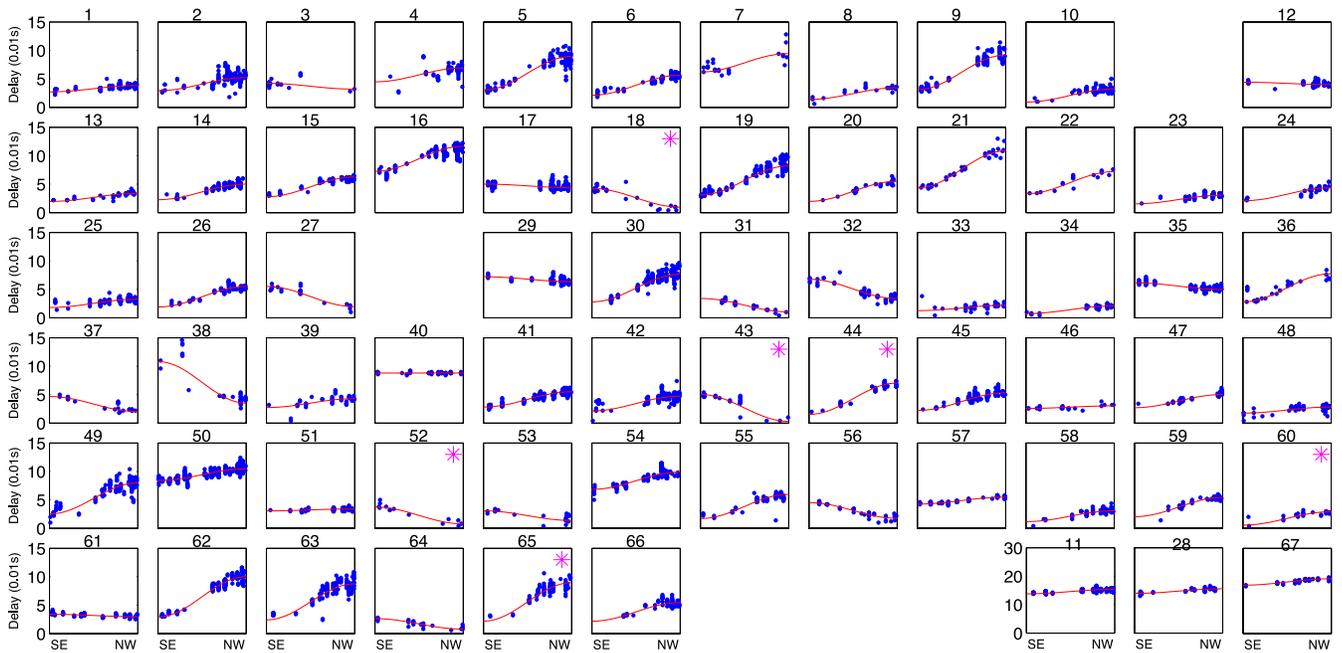
The criteria used to identify compound events in the north and green boxes on the SAF and on the Calaveras Fault are the same as those used for the south box on the SAF. Delay versus azimuth plots for these regions are shown in Figs B1–B3. The maxima and minima of the delays for most events are aligned with the fault strike, however, event 97 on the Calaveras Fault (Fig. B3) may be an example of an off-fault subevent. The delays of most events are smaller than 15 samples; a few exceptions with larger delays are plotted at a different scale in the bottom rows of these figures.

## APPENDIX C: MAGNITUDE OF THE TARGET AND EGF

The earthquakes used in this study are magnitude 0.5–3 events from the NCSN catalogue, but the identified compound events may not span this entire magnitude range. For small events, this may be due to the difficulty in resolving small time delays (the pulse plus spike must fit the target better than a single Gaussian convolution of the EGF); for large events clipping issues may lead to poor misfit reductions at individual stations or azimuthal inconsistency between the stations. Fig. C1(a) plots the magnitude distributions for all 28 869 events examined in Fig. 1 (black), the identified compound events (red) and all the EGFs used in successful inversions (blue). The histogram for all the events has a maximum at  $\sim 0.8$ , and the quick fall off at the smaller end is due to the catalogue detection threshold. The magnitude distribution for the compound events (red) is clearly skewed to larger events. Of the 28 869 events examined, 677, or 2.3 per cent, were identified as compound. This is comparable to the 2.7 per cent (54 of 2000  $M_{1.2}$ –3.2 events) identified as compound by Fischer (2005). About 20 per cent of our events larger



**Figure B1.** Delay versus azimuth plots for the 168 identified compound earthquakes in the north box, as in Fig. 3.



**Figure B2.** Delay versus azimuth plots for the 67 identified compound earthquakes in the green box.

than  $M2.5$  were identified as compound (compare the black and dashed red curves in Fig. C1a). This is a smaller percentage than found by Fischer (2005) for events larger than  $M2.2$  (61 per cent), but clipping is likely more of an issue for the NCSN stations we use.

From the comparison of the blue and red curves, the EGFs generally have smaller amplitudes than the compound events. Fig. C1(b) compares the magnitudes of the identified compound events with all their EGFs used in the inversion for  $\Delta x$  and  $\Delta t$ . Some EGFs are larger than their targets, but most are smaller by about 0.5 magnitude units, and a few are smaller by much more. This is the reason for using the Gaussian convolution of the EGF to fit the larger subevent.

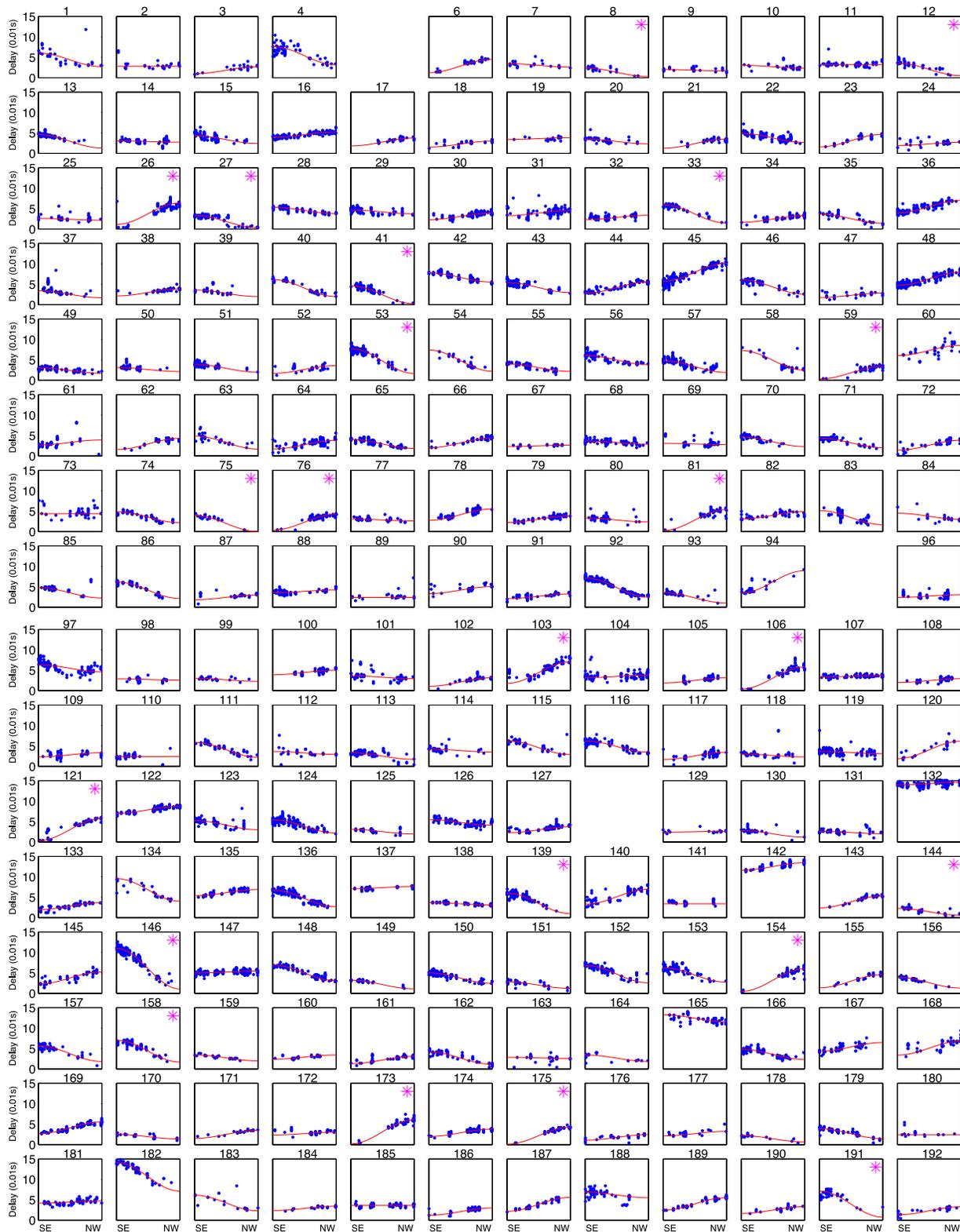
#### APPENDIX D: EFFECT OF SUBEVENT DIRECTIVITY AND SIZE ON THE INFERRED CENTROID TIME DELAY

Since both subevents of the compound earthquake and the EGF have finite source durations, aligning the target waveform and the residual with the EGF is equivalent to aligning the centroids of the source time functions. The calculated separation of the two subevents  $\Delta x$  is interpreted as the separation of their centroids and the delay  $\Delta t$  is the delay between the centroid times. Depending on the directivity and relative sizes of the two subevents, the calculated centroid time difference  $\Delta t$  could be different from the origin time difference  $\Delta t_0$ . To explore the effect of directivity and relative subevent size on  $\Delta t$ , Fig. C2 plots six special cases for subevent sizes and directivity. The rupture areas of the chronologically first and second subevents are represented using the blue and red circles, respectively, the dots in each subplot indicate where the ruptures begin and the arrows point in the propagation directions. We assume that the chronologically second subevent (red) is triggered immediately by the shear wave from the first (blue), so the origin time difference of the two subevents  $\Delta t_0$  is then the transit time of the shear wave over the distance between the nucleation locations of the two subevents. For comparison, the calculated  $\Delta x$  and  $\Delta t$  expected to come out of an

actual inversion are shown in text below each subplot and plotted at the bottom of Fig. C2, and have been normalized by the radius of the first subevent  $r_1$  and the shear wave traveltime across the radius of the first subevent  $r_1/c_s$ , respectively (as in Fig. 4). We assume that ruptures propagate at the shear wave speed.

When two equal-sized subevents have the same or opposite signs of directivity, as those in ‘a’, ‘b’, ‘c’ and ‘d’,  $\Delta t$  and  $\Delta t_0$  are equal to each other. The origin time differences for events ‘a’ and ‘d’ are just the shear wave traveltime over the centroid separation  $2r/c_s$ . Event ‘b’ has a longer  $\Delta t$  than ‘a’, ‘c’ or ‘d’ since the nucleation locations are farther apart; on the other hand,  $\Delta t$  for event ‘c’ is zero. As shown in the bottom plot of ‘ $\Delta x - \Delta t$ ’, the centroid separation  $\Delta x$  of the four compound events ‘a’ to ‘d’ are all  $2r$ . Events ‘a’ and ‘d’ have the same time separation for centroid and nucleation locations and both lie on the shear wave line  $\Delta t = \Delta x/c_s$ . Events ‘b’ and ‘c’ have larger and smaller nucleation time separation than centroid time separation, respectively, so their  $\Delta t$  are 4 and 0 instead of 2 as for ‘a’ and ‘d’.

Different subevent sizes could have two kinds of effects on the inverted  $\Delta x$  and  $\Delta t$ . First, changing  $r_1$  relative to the centroid separation  $\Delta x = r_1 + r_2$  changes the normalization factors  $r_1$  and  $r_1/c_s$  for  $\Delta x$  and  $\Delta t$  and moves the data point in the ‘ $\Delta x - \Delta t$ ’ plot in the radial direction (towards or away from the origin). Secondly, the different source durations makes the inverted centroid delay  $\Delta t$  different from the origin time difference  $\Delta t_0$ . The subevents in the three ‘bilateral’ compound events ‘a’, ‘e’ and ‘f’ in Fig. C2 have different relative sizes, but since  $\Delta x$  and  $\Delta t$  are normalized by  $r_1$  and  $r_1/c_s$ , smaller relative  $r_1$  gives larger normalized  $\Delta x$  and  $\Delta t$ . From smaller to larger relative first subevents, ‘f’, ‘a’ and ‘e’ have larger to smaller  $\Delta x$  and  $\Delta t$ , which place them at three different locations along the lower shear wave line  $\Delta t = \Delta x/c_s$ . In addition to the effect of the scaling factors,  $\Delta t$  for ‘e’ is smaller than the origin time difference  $\Delta t_0 = 1.6r/c_s$  since the peak of the source time function of the first subevent has a longer delay from the first arrival than the second subevent, so event ‘e’ plots a little below the shear wave line. On the other hand,  $\Delta t$  is larger than the origin



**Figure B3.** Delay versus azimuth plots for the 259 identified compound earthquakes on the Calaveras Fault. Event 97 may be an example of a subevent off the fault plane.

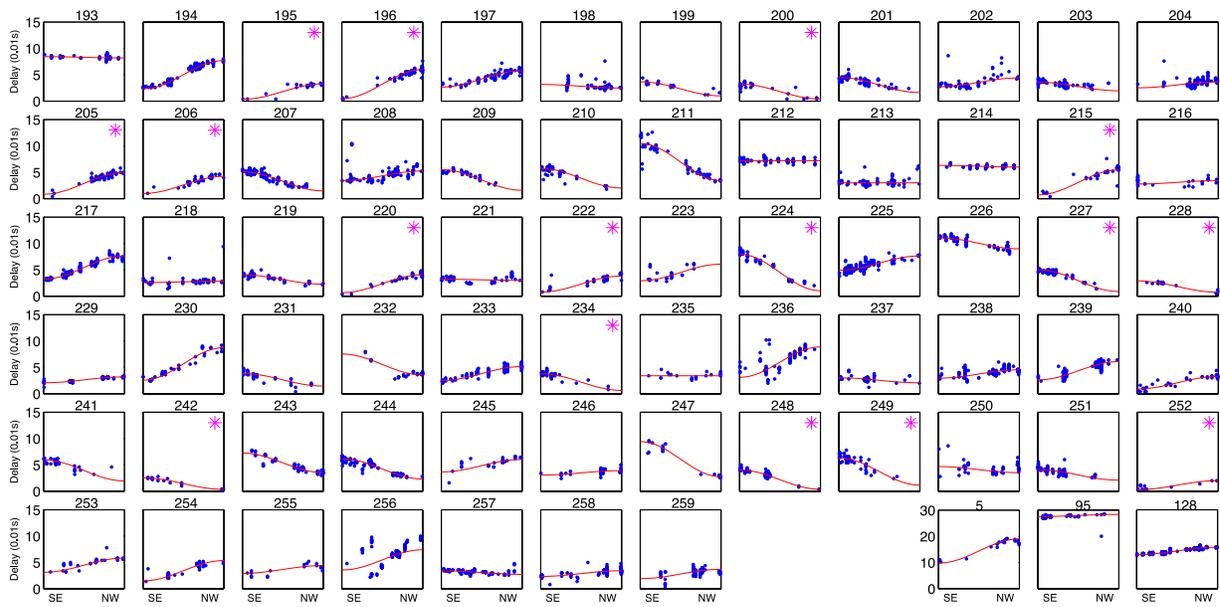
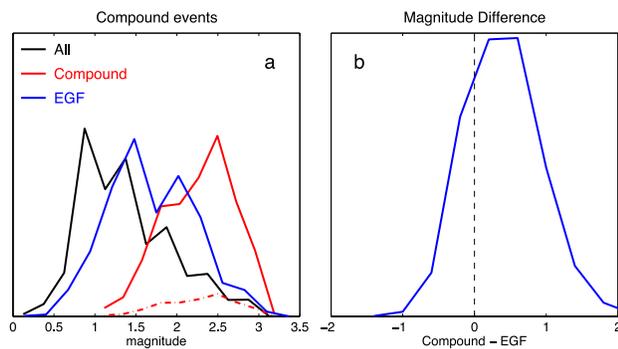


Figure B3. (Continued.)

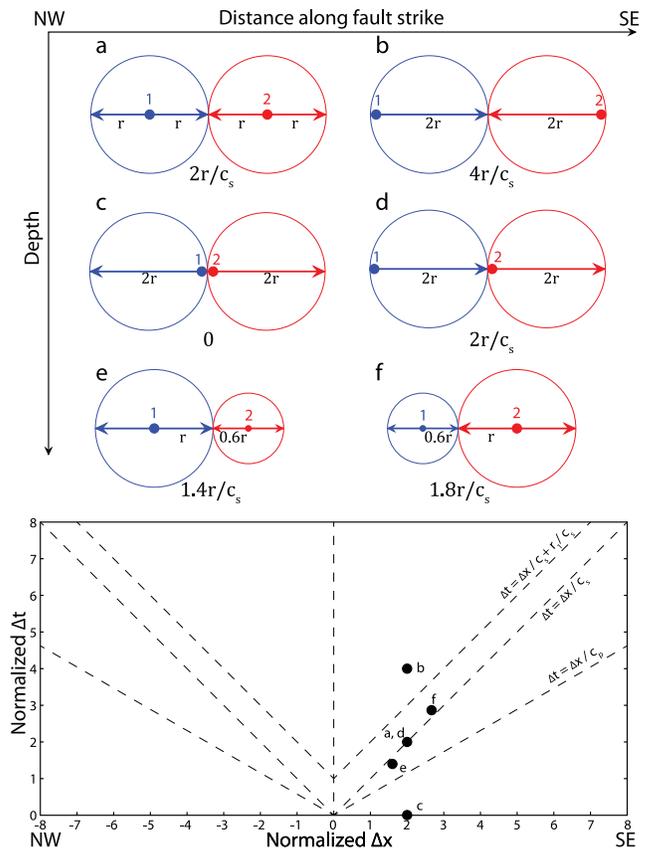


**Figure C1.** Left-hand panel: normalized magnitude distributions for the identified compound events (red), utilized EGFs (blue) and all the events (black) in all the four boxes in Fig. 1. The dashed red curve shows the same distribution of the compound events, but normalized using 20 per cent of all the events (black). Right-hand panel: distribution of the magnitude differences between the compound events and their EGFs.

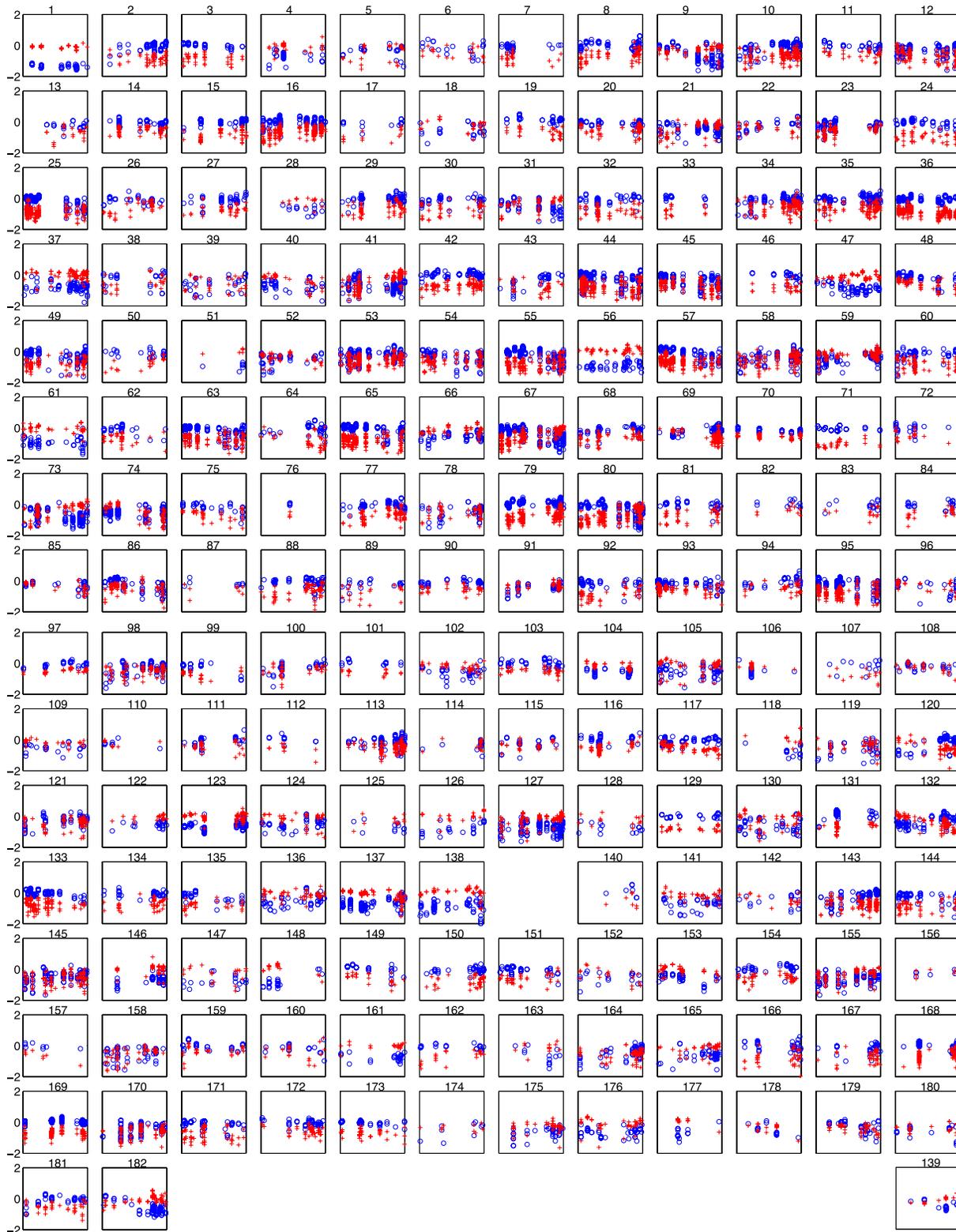
time difference if the first subevent is smaller ('f' in Fig. C2), so 'f' plots above the shear wave line. The two subevents of 'e' and 'f' have the largest permitted relative size difference in the inversion assuming similar stress drops:  $r_S/r_L = 0.6$ , so the upper limit of the difference between  $\Delta t$  and  $\Delta t_0$  caused by this is close to the vertical offset of 'e' and 'f' from the shear wave line in the bottom plot in Fig. C2. This is not as significant as the shift parallel to the shear wave line. Considering all these uncertainties, Fig. 4 plots the dashed line  $\Delta t = \Delta x/c_s$  as well as  $\Delta t = \Delta x/c_s + r_1/c_s$  for reference and uses the red and blue polygons to define second subevents located near the margin of the first.

#### APPENDIX E: AMPLITUDE RATIOS OF THE SUBEVENTS TO THE EGF

To estimate the amplitude ratios of the two subevents, we can fit the target without convolving the EGF by a Gaussian function.



**Figure C2.** Effect of subevent directivity and relative size on the normalized origin time difference. The blue and red circles in each of the six subplots from 'a' to 'f' represent the first and second subevents of a compound earthquake, respectively. The view is towards the fault plane and the horizontal and vertical axis are in the direction of the fault strike and depth. Each subevent starts from the 'dot' and ruptures in the direction shown by the arrows. The figure at the bottom plots the normalized  $\Delta x$  and  $\Delta t$  expected for the six compound events.



**Figure E1.** The logarithmic scaling factors  $\ln(A_1)$  and  $\ln(A_2)$  are plotted versus azimuth for the 183 identified compound events in the south box. Blue circles and red crosses are for  $\ln(A_1)$  and  $\ln(A_2)$ , respectively. Event index is shown on top of each subplot, with left and right side of the figure representing the SE and NW directions along the fault; panels correspond to the same compound earthquakes as in Fig. 3. Individual stations between the two figures may not match exactly due to different criteria.

Such a fit has four parameters for each station/EGF pair: the two time offsets  $t_1$  and  $t_2$  and the two scaling factors  $A_1$  and  $A_2$ . In Fig. E1, these two scaling factors  $A_1$  and  $A_2$  are plotted against azimuth with blue circles and red crosses, respectively. The 183 events are the identified compound events in the south box and are arranged in the same order as in Fig. 3. Note that most of the 183 events have multiple EGFs, so part of the scatter comes from different EGF amplitudes. Most of the  $\ln(A_1)$  and  $\ln(A_2)$  are between 0.5 and  $-1.5$ , and are generally consistent with azimuth, indicating that the bigger subevent is within  $\sim 7$  times the amplitude of the smaller one. A few events seem to have sinusoidal trends that may be caused by directivity of the subevents, but no obvious correlation was found between subevent directivity and the relative locations of the subevents.

In this study, we assume that the  $P$  wave of the chronologically first subevent arrives first at all stations, and distinguish the two subevents based on the order of the first  $P$  arrivals. However, both the order of the two arrivals and the sign of the amplitude ratios  $\ln(A_2/A_1)$  may change with azimuth, making it difficult to judge if the subevents are correctly identified.

For example, for event 'c' in Fig. C2, the order of the arrivals differs in the two directions along the fault strike, even though the second subevent occurs after the  $S$  arrival from the first. Stations to the NW of the earthquakes observe the first subevent (blue) first since it is closer than the second subevent (red) by  $2r$ , while stations to the SE would first see the second subevent. Stations near the nodal plane would only see a single arrival since the distances from the two centroids to the station are approximately the same, giving rise to a gap in detections near  $90^\circ$ . If the arrivals of the two subevents could be correctly identified at all stations, the actual delay-azimuth curve for event 'c' in Fig. C2 would cross zero. However, since we assume

that the chronologically first subevent always arrives first and use the absolute values of delays in the delay-azimuth plots (as in Fig. 3), the negative delays would be flipped to positive leading to a positive and flat azimuth-delay curve. Instead of  $2r$  and 0, the inverted  $\Delta x$  and  $\Delta t$  from this flat curve would be close to 0 and  $2r/c_p$ . Fortunately, many of the blue and red data points in Fig. E1 are well separated, even those that display an azimuthal trend that may result from directivity, indicating that these subevents were identified correctly. Events 67, 120 and 132 in Fig. E1, which generally have small delays ( $< 3$  samples) and few detections orthogonal to the fault plane, might be examples of misidentified subevents. The blue and red data points for these events appear to follow a step function, with blue amplitudes at stations to the SE similar to red amplitudes at stations to the NW (and vice versa), indicating that both  $A_1$  and  $A_2$  could be more uniform at all stations if the arrivals from the chronologically first subevent arrived second at stations to either the NW or SE.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Figure S1.** The relocated catalogue of Waldhauser & Schaff (2008) plotted on cross-sections along the faults (<http://gji.oxfordjournals.org/lookup/suppl/doi:10.1093/gji/ggu047/-/DC1>).

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 article.