**Response to Reviewer 1:**

The one thing that I miss in this paper is clearer reference to data.  In examples from the Calaveras it might be useful to see examples of coda wave trains together with their correlation  (or covariogram) to get a feeling for the relationship between the two.

Point noted – we have added the following text to the manuscript:

“Waveforms, cross correlations and separation estimates for example Calaveras event pairs are given by Robinson et al. (2011)”.

In the interest of brevity we do not illustrate the pairwise examples again in this manuscript.

1. [3.36]  Would be clearer to write "signal-to-noise ratio" as opposed to just "signal-to-noise".

Corrected in paper

2. [3.39]  Suggest insertion of "primarily":  "kilometers because" become "kilometers, primarily because".  In some cases simple measurement uncertainty may have as big an effect on location uncertainty as imperfect velocity information, particularly for S waves.

Corrected in paper

3.  [3.43]  "are" become "is" The verb to be refers to location uncertainty (singular) a few lines above.

Changed “location uncertainty” to “location uncertainties”

4. [4.58-]  Sentence beginning in line 58 "To reduce errors...." mentiones how effects of unknown velocity structure can be reduced by differencing measured arrival times (actually the sentence misleadingly talks about travel times) but refers this to wave forms.  It would be clearer to explain how e.g. Douglas´ JHD applies arrival time differencing in order to reduce effects of unaccountet for velocity structure while if events are close to each other then scattering effects are similar and therefore waveforms are highly correlated leading to the possibility to measure differential arrival time with high precision.  It is this high precision of differential arrival times that renders modern relative location precise to within 15-75 m.

to fix

5. [4.60] The authors refer to Ito (1985), Got et al. (1994), Nadeau and McEvilly (1997) and Waldhauser et al. (1999) as examples of relative relocation studies.  I think it would be in place to refer also to Slunga et al. (1995) [Absolute and relative location of similar events with application to microearthquakes in southern Iceland, Geophys. J. Int. 123, 409-419, 1995].  This paper predates a number of the more popularly cited relative-relocation papers and approaches the problem from an interestingly different angle, but it is not written by members of the Anglosaxon research community.

Thank you for referring us to this interesting paper, which relatively locates microearthquakes in southern Ireland. We have cited it in the manuscript.

6. [5.77]  Again, the authors talk about travel-time difference where the appropriate term is arrival-time difference as relative relocation works with differential times for two nearby events observed at one and the same station.

We accept the reviewer’s suggestion of changing travel time differences to arrival time differences throughout the manuscript.

7. [7.127]  The quantity omega squared with a bar, which occurs in eqn 4, is not defined.

Omega is now defined

8. [13.270]  The mislocations in synthetic example 1 (Figure 1) are stated to be 2 m on average, but appear to be significantly greater in the figure (approximately 3.6 m).

Actually we are commenting on the average coordinate error rather than the average location error. Nevertheless, for the sake of clarity, we now also mention the average error in brackets.

9. [14.279]  The mislocations in synthetic example 2 (Figure 2) are stated to be 2.8 m on average, but appear to be much bigger in the figure (approximately 8.2 m).

Actually we are commenting on the average coordinate error rather than the average location error. Nevertheless, for the sake of clarity, we now also mention the average error in brackets.

10. [15.305]  "principal" should be "principle".

Corrected in paper

11. [15.310]  "phenomena" = plural should be "phenomenon" = singular.

Corrected in paper

12. [15.312]  "Fortunately however," should be "Fortunately, however,".

Corrected in paper

13. [15.314]  "facilitate" refers to "combining" and should be "facilitates".

Corrected in paper

14. [15.315]  "On balance however," should be "On balance, however,".

Corrected in paper

15. [16.331]  "illustrates" refers to Figures 3a and b, i.e. plural and should be "illustrate".

Corrected in paper

16. [16.334 and 335]  black and gray in text refer to blue and red in figure 3.

We now refer to the lines as thick and thin

17. [17.368]  "v = 3 300" change to "v = 3,300" or "v = 3300".

Changed to 3,300

18. [18.378]  black and gray in text refer to blue and red in figure 4.

We now refer to the lines as thick and thin

19. [18.390]  It would be interesting to know what the level of mislocation is in Figure 4.  This is hard to read from the figure.

Agreed. We have computed the level of mislocation and quoted in the figure caption

20. [19.423]  The term "travel time arrivals" is ambiguous.  I suggest "arrival-time data".

Changed throughout the document

21. [20.434]  "introduced in Theory".  I suggest wording be changed to "introduced in the Theory section".

Corrected in paper

22. [20.440]  "travel time differences".  This is inaccurate.  Should be "arrival-time differences".

Changed throughout the document

23. [20.442]  Typo in "uniquesness"

Corrected in paper

24. [20.444]  "for the same event"  suggest "for that event".

Corrected in paper

25. [20.447]  Hyphenation is inconsistently used in this manuscript.  I suggest that it would be clearer to write "travel-time and coda-wave inversion."

We now consistently use "travel-time and coda-wave ” throughout the document

26. [21.454]  "suggests" refers to experioence = singular so should be "suggests".

Corrected in paper

27. [21.454]  "the coda are" suggest rewrite to "the CWI relocations are".

Corrected in paper

28. [21.460]  I do not understand the meaning of the phrase "arrival phases".

Changed to arrival-times

29. [22.478]  Is this a logical use of the word "hypothesis".  Would "claim" be better?

Changed to claim

30. [22.483]  Suggest "less self consistent" instead of "not self-consistent".

Corrected in paper

31. [22.488]  "are" refers to number = singular and should be "is".

Corrected in paper

32. [22.488]  "both ..... are not able" suggest "neither .... is able".

Corrected in paper

33. [23.501]  "exists" refers to differences = plural and should be "exist".

Corrected in paper

34. [23.504]  "located sufficiently" suggest "located sufficiently well".

Corrected in paper

35. [23.512]  The terms "CWI and hypoDD reduced station locations" are odd and misleading. I suggest "statistical comparison of the CWI and hypoDD locations with a reduced number of stations".

Suggestion accepted. Thankyou.

36. [24.535]  "In this paper" and "we" is redundant and the former can be erased.

Corrected in paper

37. [26.562]  "travel time data" should be "arrival-time data".

Changed throughout the document

38. [26.564]  "travel time constraints" should be "arrival-time constraints".

Changed throughout the document

39. [26.575]  Once again the relative relocation is based on differential ARRIVAL times, not travel times.

Changed throughout the document

40. [27.585]  Again, arrival times, not travel times.

Changed throughout the document

41. [27.592]  And again, differential arrival time, not travel time.  Hereafter I will not bother pointing this out!

Thankyou! We are all over this now.

42. [27.598]  "is given by" is a strong statement in this context.  How about "can be modelled with" instead.

Agreed. Corrected in paper.

43. [28.609]  "idea" is redundant and can be removed.

Corrected in paper.

44. [28.617]  "is" refers to data = plural and should be "are".

Corrected in paper.

45. [28.623]  Same, data are plural and the travel-time ghost is there again.

We have it covered now

46.

?

47. [45]  The white star in Figure 5 is barely visible.  There is colour in the figure.  I suggest making the star red.

Changed to a coloured star

48. [46]  The axes in the figure are not defined in the legend.  I assume x is positive eastward and y is positive northward (same for figures 8,9,11,13).

This is now explained in the caption of Fig. 6.

Reviewer #2:

The only one that I'm really concerned about is the apparent inconsistency between your assumptions on page 6 and theory on page 9.

This comment is addressed in detail below

COMMENTS

========

L46: You mean given a local array of receivers? How distant is "local", if so (given that you give absolute numbers for location uncertainties, and your point is that these scale with source-receiver distance)?

How distant?

L120: this assumption is not carried through your theory section. See comments around equations (8) and (9).

This comment is addressed in detail below

L137: delta\_CWIN = delta\_CWI in equation (1)? (what is the extra subscript N?)

The following text has been added to clarify the use of N:

“The use of N in delta\_{CWIN} depicts CWI separations that include noise. The terminology is adopted here to remain consistent with Robinson et al. (2011) who study synthetically generated noise-free delta\_CWI and relate them to noisy estimates delta\_CWIN.”

L142: I don't believe that acronym PDF is defined until the next page.

Thank you, This acronym is now defined at L142.

L179-180: This needs more careful justification. If P(...|e1,e2) and P(...|e1,e3) both depend on e1, how can they be independent? Explain more clearly for the reader.

Wait for feedback from Malcolm

L182: [Here lies my only real concern]: This contradicts the assumption that all events lie on the same fault plane (assumed earlier, L120) for >3 events. [That is, once you have defined 3 events you have defined the fault plane; every other event that is assumed to lie on that plane therefore depends on those three events]. True for 3 events, but not for >3 events.

We do not believe that this is an issue because of the following reasons:

1. The first three events are used to define a plane that in-turn defines the local coordinate system. Events 4 and up are free to move outside of the plane defined by events 1 to 3. Please also note that we do not need to fix the local coordinate system around events 1 to 3 in the examples where arrival-time data are incorporated.
2. Snieder and Vrijlandt (2005) also define a CWI estimate of separation for two double couple sources that are not in the same fault plane. That is, the CWI theory is not restricted to events in the same fault. It is important to note however, that the formula for such a CWI separation estimate is more complicated in that it has two unknowns: the in-fault separation and the out-of-fault separation.

The following changes have been made to the manuscript to avoid other readers becoming confused around the same issue:

1. We now show, in the Theory Section, the CWI separation formula of Snieder and Vrijlandt (2005) for a double couple pair that is not in the same fault plane.
2. We explicitly state that events are free to move outside the plane during the inversion.
3. We clearly articulate that we are approximating the event separation by the in-fault separation of Snieder and Vrijlandt (2005) despite not forcing events to remain in a single plane during the fault.
4. We add more discussion around our approximation at point 3 and invite our readers to explore the impact of either:
   1. Forcing events to lie in the same plane as one may prefer if they believe that the events of interest are aftershocks on the same fault, or
   2. Using the more appropriate formula from Snieder and Vrijlandt (2005) that considers both in-fault and out-of-fault displacement.

L201: I don't see why this is true given my comments for equations (8) and (9).

This comment is addressed in detail above.

L221 - L227: I see why you do this, but what if any of e1,...,e4 are both close together and uncertain? Since you define the axes in this way, don't you risk creating extreme sensitivity of all other locations' coordinates, simply because as e1,...,e4 move within their uncertainties, so does the entire coordinate frame of reference? If they are close together, axes may even reverse in direction! This is likely the cause of your final 'rotational' uncertainties in Figs 1 and 2. Why not then restart the optimisation, fixing the coordinate system using the most distant events in each coordinate direction, as found in the first optimisation? May solve the rotational uncertainty problem...

Double check and comment

L254 - L272: You have not mentioned any receiver distribution here. Is this analysis somehow independent of an exact experimental geometry? Is the geometry hidden inside equation (19), or in sigma^bar\_N? Explain.  [Ignore other comment].

Yes – this experiment is independent of experiment geometry. We have clarified this in the text to avoid confusion by other readers:

1. The first sentence (i.e. L255) is changed to read “ We design a 2D synthetic acoustic experiment (example 1) to test the performance of our CWI based relative location algorithm by randomly …”
2. We add the following description around L259: “ The purpose of these examples is to synthetically test the location algorithm. Therefore, we do not need to generate earthquakes, synthetically model waveforms, compute the waveform cross-correlations or estimate the CWI separations. Rather, we can begin by synthetically generating the CWI estimates directly.”

L286: Really NO difference? If these are indeed due to small eigenvalues, wouldn't you expect them to vary along near-singular directions from experiment to experiment? If they don't vary at all they sound more systematic... Explain! And if they DO vary, then show at least two example solutions.

Agreed. NO difference is not the correct term to use here. This has been changed in the manuscript to negligible difference.

We are not performing different experiments here but rather starting the iterative optimization procedure at different starting locations. The important point is that all starting points lead to the result (i.e. the same minimum) with negligible variation.

Since the results do not vary there is no need to show more than one. The more interesting case is example 3 (L319) where we start to remove the linkage between some event pairs. Here, we observe that different starting locations can lead to different solution. This is already discussed (an illustrated) in example 3 -we have not made changes to example 3.

L329: Reduced by removing the LONGEST existing inter-event linkages? Be explicit.

I am not sure what is meant by LONGEST here. These examples randomly select even pairs and remove the linkage between them. Distance is not relevant. We have attempted to clarify in the manuscript to avoid confusion by other readers. The following text is added: “That is, we randomly select 10% of the event pairs and remove the separation estimates between those pairs to create a data set with 90% linkage. Then, we randomly remove 20% of the links and so on. This is supposed to mimic a realistic recording situation where CWI estimates are not available for all event pairs due to station problems, poor signal-to-noise ratio or any number of other reasons.”

L350: A matter of opinion. Be QUANTITATIVE rather than giving your personal view (at this stage in the paper).

Fair point. Actually, we are more interested in the fact that they appear to be heading towards the true solution but are either getting stuck or making progress so slowly that they never getting there. We have re-worded as follows:

“Despite their failure to converge, the locations at final iteration are close to the actual solution”

Becomes:

“Despite their failure to converge, the locations at the final iteration often resemble the actual solution, even with relatively generous convergence criteria such as ………… Interestingly, increasing the maximum number of iterations does not appear to improve the convergence. A plausible explanation for this is that the gradient of the optimization function L is “flat” around the global minimum in cases of low linkage between event pairs.

L367: Confusing: these are still (I think) synthetic experiments, hence not "actual" events.

We have removed the adjective actual to reduce confusion.

L373: What is the value of epsilon?

Epsilon is the same as that used in examples 2 and 3 earlier in the paper. We have clarified this around L373.

L391 - L400: You haven't mentioned the rotational non-uniqueness in this Summary. Important to do so.

L418: IGNORE MY COMMENTS ABOUT NOT SEEING THE TABLES - I FOUND THEM IN THE END!

Page 20: Note that the "star" in Figure 5 is almost invisible - this needs to be sorted out.

L433: Surely 2.5s < t <= 17.5s  if 5s wide and non-overlapping?

L471: Explained or just listed? [After I found the tables I saw that they were just listed]. I think that information should anyway be displayed in Figure 7 - put anumber 1 to 10 beside each station, "1" being the LAST one to REMAIN.

L497: Still very impressive - congratulations!

L628 - L652: Really nice paper and results. Without doubt will be useful and used. [With respect to your point about only a single station having to record both the main event and the aftershocks...] I think you might also consider discussing the link between this work and [a paper of my own] Curtis et al., (2012) - EPSL, "...Retrospective seismology..." which shows how interferometry can be used to construct the SEISMOGRAMS of the MAIN event at temporary arrays deployed after the main event occurred, again provided that a set of network stations recorded both this event and the aftershocks: you show that as far as the LOCATION of the events is concerned, you don't need the complete seismograms at new stations, nor an array.