Solid Earth Discuss.,
doi:10.5194/se-2016-67-RC2, 2016

Solid Earth Discussions

Interactive comment on "Improved finite-source inversion through joint measurements of rotational and translational ground motions: A theoretical study" *by* Michael Reinwald et al.

W. Debski (Referee)

debski@igf.edu.pl

Received and published: 31 May 2016

During last 30-40 years we are facing in physical sciences a rapid development of a new research methodology namely numerical simulations which become the complementary approach to traditional theoretical and experimental research methods. An importance of this new approach is still not well recognized in geophysics. The reviewed manuscript represents this new approach and for this reason it is quite valuable.

From scientific point of view the manuscript deals with an analysis of possible improvements in kinematic seismic source inversion by including rotational seismic data from modern rotational instruments. This analysis is apparently important, timely and brings interesting results. Another important scientific element of the reviewed manuscript is

use of the probabilistic (Bayesian) inversion framework which has allowed Authors to perform the quantitative analysis of the inversion resolution. This particular feature of the inversion scheme has been efficiently exploit in the presented analysis.

Nevertheless, the manuscript contains some points which should be clarified before publication. There are two important drawbacks of the manuscript. The first one is the an insufficient description of the simulation/inversion performed. After reading the manuscript I really had no idea what data were used for inversion, how they were generated, what was the a priori model, *a priori* pdf, used misfit function, how sampling was performed and so on. I belief that these slightly messing situation is very simply to correct for Authors. The second, more serious in my opinion is connected to errors in presenting the probabilistic inversion methodology. The similar mistakes can be found in many papers dealing with application of the Bayesian inversion so I take a time to discuss this point in more depth.

Authors are using the probabilistic (Bayesian) inversion methodology and already in the introduction state (page 1line 23) " In the last decade . . . because they overcome the drawbacks of regularization techniques like such as local minima...". The similar statement is repeated at the beginning of the Section 2. This statement is unfortunately not true. The probabilistic (Bayesian) approach does not "mysteriously" remove the solution non-uniqueness, existence of secondary minima, null space or other problems like that. They exist no matter what inversion technique is used but the Bayesian approach provides efficient methods to identify them and taking into account when inversion results are interpreted. As an example let us consider an regularization issue which is needed if an inverse problem is ill-posed. Within an algebraic approach (see e.g. Menke) the regularization is needed to assure that the matrix $G^T G$ is invertible - has a non-zero determinant so the inverse matrix can be calculated. In the optimization approach the regularization procedure is also applied typically by modifying an optimized misfit function through adding an additional term with a Lagrange multiplier. Its goal is now to assure convergence of of the optimization algorithm and/or preserving some requested features of the solution which optimization procedure can easily lost. In the probabilistic (Bayesian) approach the *a priori* pdf plays often a similar regularization role. For example (simplifying slightly problem), if some thought parameters are not resolved by data (case very often met in tomography) than the a priori pdf assures that inversion procedure assign a particular (a priori) values to these parameters. In a similar way it can be used to remove the multi-modality of the a posteriori distribution, (which are direct counterparts of of multiple local minima in a classical approach) thus performing some regularization. The real advantage of the probabilistic approach is, however, formulating the solution of the inverse problem in term of the probability distribution over the model space which opens a possibility of the quantitative analysis of problematic (e.g. ill-posed) cases. This particular feature of the Bayesian technique is well illustrated in fig.5 where increasing of width and flatness (so resolution, or inversion errors) of the a posteriori pdf with patches depth is well visible. This effect simply means that for particular patches the slip distributions cannot be uniquely estimated in term of a single "best" value but you can only provide a range of admissible values (flat parts of the distribution).

Describing probabilistic inversion method Authors provide eq. 3 and 4. Unfortunately, equ. 4 contains two mistakes. First, it contains the normalization factor k'. This factor is unnecessary in definition of the likelihood function because the only requirement imposed by the theory is normalization of the a posteriori distribution. Likelihood function does not need to be normalized. Actually, it may not be normalizable at all (then formally $k' = 0$) if the inverse problem at hand exhibits a null space (see, e.g., Debski, 2010 for details). In such case the normalization of the *a posteriori* pdf has to be assured by the *a priori* term in equ.3. The second problem with the equation 4 is the form of the likelihood function (exponent part) which is strictly speaking incorrect or at least badly explained in the main text. First of all, if a form with explicit sum is used you have to explain what the sum is taken over. In a standard notation (for finite dimensional inverse problems) this sum is over all data used for inversion. In such a case, however, the term $\kappa(m)$ is not the misfit function but the norm in data space used to measure

a distance between predicted and observational data and s_i is an estimator of sum of modelling and observational data. On the other hand if you use the notation with $\kappa(m)$ to be the misfit function you should use the formula $L(m) = exp(-\kappa(m))$ with no sum and s_i which are already encompassed in $\kappa(m)$. I guess that an Author's idea was to write the likelihood function in a highly synthetic way what is of course possible. However it has to be very clearly marked and explained in the text. The other issue arise if infinite-dimensional problems are concerned, like, for example a full seismic waveform inversion. The choice of an appropriate norm in data space and so the form of the likelihood function is by no means trivial as discussed, for example, by Kenet 2012.

The another issue mentioned at the beginning of section 2 and also repeated in the conclusion section, namely the non-uniqueness of the kinematic source inversion. In line 21 authors state: *Finite source inversion is non-unique due to noisy data, sparse geographic coverage of the seismic station the non-linearity of the forward problem. . ..* This is only partially true. All the listed factor contribute to a possible inversion non-uniqueness but actually the main source of the aforementioned non-uniqueness in the analysed kinematic source inversion problem is the fact that this is infinite dimensional inverse problem in which the continuous function of the slip distribution over a fault is inferred from the finite number of observations. A discretization of the fault area by patches converts this infinite dimensional problem to the finite size inversion but this does not completely removes the inherent non-uniqueness of the problem. After such dimensionality reduction the aforementioned non-uniqueness manifest itself in a dependence of the inversion results on used parametrization, like in conventional seismic velocity tomography. (for details see, e.g. Debski 2010).

In addition I have some editorial comments.

I would suggest a small change in the title by replacing A theoretical study by A numerical study, A numerical simulation or so. Apparently the manuscript has nothing to do with theoretical (in a classical sense) study.

In the abstract it is unnecessary repeated (line 10) what 6-C component data are. The later part of the abstract (from words The results show. . .) is somehow confusing and not quite clear. For example: which source properties are better resolved by 6C data, what does it mean equally well recovered?. The note on installing logistic is out of the subject of the paper and is not discussed further on. I suggest to remove *We assume...* as misleading because nothing is really assumed about the mentioned effect and leave only This is attributed. . . In the abstract simulations for 2 scenarios (deep slip and strike slip) are mentioned but on page 6 l(lines 4-5) Authors mention the third type of analysis. Some comments on it should be put in abstract as well as more detailed description when describing performed analysis is necessary. The statement that the (third type) experiment was performed with "randomly set of receivers" is not sufficient.

I have found many statements which are imprecise and have to be carefully checked and correct. Some of them are sometimes even surprising like, for example, " . . . *we let probabilistic inversion do random walks*" (probabilistic inversion construct a posteriori pdf only - its use - including sampling method is a completely different issue and apparently performs no random walk). Completely unclear is for me section 4, put before the result section. What means numbers mentioned there. Are they results of simulations (thus why this is not in the results section?) or come from literature (references?).

The sentence (line 29-30) in conclusion section is not justified by the presented analysis simply because it does not cover the non-uniqueness issue. Actually, presented results prove that including rotational data improves an accuracy of a slip distribution estimation, which is of course the important result, but do not discuss the uniqueness issue.

I wish also to make a more general comment on presenting numerical results of the simulations. Talking, for example, about information gains Authors often provides numbers with two decimal digits, something like 17.77

My another general comment refers to the Authors attempt of using bit as a unit of

information gain. In case of the presented analysis it leads to results like 4.31 bit, which numerically is of cause OK, but sounds very strange, because bit as information unit cannot have fractional parts. I suggest to drop using the word bit especially that it brings no profits to the presented analysis at all.

Finally I have a question concerning figs. 6 and 7. This figures shows that the retrieved rupture velocity and rupture rise time significantly different from the true, assumed, values Taking into account that you have performed numerical, fully controlled simulation with a quite low and very convenient non-problematic Gaussian noise, what is a reason of such large discrepancies?. Are you sure that you have properly run the Metropolis sampler with correctly chosen accepting ratio? Did you generate long enough sample series to get behind the burn-in period and avoid influence of the starting values on the final *a posteriori* pdf? Please remember that the Metropolis sampler has relatively poor mixing property and in case of multi-modal distribution the proper sampling *a posteriori* pdf may requires huge number of samples to be generated. What was the *a priori* and starting values for these two parameters. I would be happy to see some comments on this point.

Kenet et all. , 2012, A unified concept for comparison of seismograms using transfer functions, GJI, 10.1111/j.1365-246X.2012.05693.x

Menke, 1989, Geophysical Data Analysis: Discrete Inverse Theory Data Analysis.

Brandt, 1999, Data Analysis, Statistical and Computational Methods for Scientists

Debski 2010, Advances in Geophys., Probabilistic Inverse Theory, doi: 10.1016/S0065-2687(10)52001-6

Interactive comment on Solid Earth Discuss., doi:10.5194/se-2016-67, 2016.