In the following the authors provide a point-by-point answer which was essentially submitted in this form in the public discussion. However, we improved the manuscript further after the discussion.

Answer letter to referee #1 S. Langer

Response prepared by C. Pelties, A.-A. Gabriel, and J.-P. Ampuero.
The authors thank the referee for providing a thorough review. We prepared the following changes and clarifications to improve the manuscript accordingly. In the following, we address all suggestions and comments of the review in detail. Each of our response items starts with the corresponding quote from the review. A revised version of the manuscript can be found in the supplement.

"General comments
The paper "Verification of an ADER-DG method for complex dynamic rupture problems" by Pelties et al. attempts to verify an arbitrary high-order derivative Discontinuous Galerkin (ADER-DG) method by comparing dynamic rupture simulation results to results of Finite Element simulations. As the authors have published papers using ADER-DG in the past, benchmarking their implementation against other models used in the community is desirable and worth publishing.
The paper is structured by different areas of interest in the realm of dynamic rupture simulations. The authors try to reproduce a large number of test cases (bi-material interfaces, branching, supershear rupture) which is impressing. By covering a large variety of topics the authors show good insight into the field of dynamic rupture simulations. They've also dealt sufficiently with the literature of each of their subtopics. The authors use the SCEC test cases for comparison of their simulation results which allows for comparison with other existing and future approaches to dynamic rupture simulations. Their choice to compare against a FaultMod instance for their verification exercises is a reasonable one, as FaultMod has been used by various researchers to model dynamic rupture. The authors themselves use SeisSol, a software that I was not aware of yet.
In general their results for SeisSol (ADER-DG) and FaultMod (FEM) are in good agreement. This indicates, that both codes can be used to model a similar subset of earthquake physics. A nice side effect of this benchmarking study is that the reader can learn about the capabilities of FaultMod.

Specific comments
The paper shows an impressive number of SCEC test cases where the ADER-DG and the FEM results are in good agreement. However it would be nice to see where agreement is not as great – if you came across any such cases. Current limitations of ADER-DG and SeisSol in terms of applicability for dynamic rupture simulations would be valuable information for a reader who is about to choose their method and software to conduct dynamic rupture studies. This could be part of a small "open questions" paragraph near the end of the paper. Adding such a paragraph is at the authors' discretion, however the authors should address their code's limitations and I think a paragraph dedicated to this would substantially add value to the paper."

We found that the differences in the results obtained with the two methods are minor and largely attributable to the different accuracies of the methods and different resolutions of the simulations. Significant differences were addressed in the manuscript. They include differences in shear stresses near the branching point in the TPV15 benchmark problem and a sharper resolution by ADER-DG of the secondary rupture front in TPV17.
The advantages of SeisSol are best exploited on a certain class of problems. The usage of unstructured meshes is best justified for geometrically complicated problems, but leads to relatively costly simulations in the geometrically simple examples presented here. Another advantage of SeisSol is the fast tetrahedral mesh generation that can be fully automated. A technical difficulty in achieving complete automation, which is currently being addressed, is the need for appropriate CAD models. However, this difficulty is common to other methods on unstructured meshes (SEM

and FEM). We think our last sentence in the conclusion addresses the optimal scope of SeisSol sufficiently.

“I’d be interested to see if the simulation results have a similar effect in a spatially larger geological context: The interseismic and final slip of an earthquake are important for studying static stress triggering, rupture jumps at step-overs and post-seismic stress transfer. Could you show to what extend potency and final slip profiles are in agreement?”

SeisSol has been applied to realistic large-scale scenarios (e.g., Gabriel et al., 2013) and the accuracy of stresses computed with SeisSol in wave propagation simulations was demonstrated by Dumbser & Käser (2006). Off-fault stresses and final slip over the whole fault were not requested in the SCEC benchmark descriptions and are not available for comparisons. However, the slip time series for particular fault positions are available on the SCEC website http://scecdata.usc.edu/cvws. For instance, Figure 1 shows a representative example of slip time series computed by SeisSol and FaultMod for TPV11. The receiver position is 0 km along-strike and 1.5 km down-dip distance. The differences in slip at the end of the simulation are minor.

Gabriel, A.-A., Pelties, C., Atanasov, A., Sachdeva, V., Passone, L., Jordan, K. E., Ely, G., Mai, P. M., Large-Scale Earthquake Dynamic Rupture Simulations with the ADER-DG Method: Towards simulation based seismic hazard assessment, Oral presentation, 2013 SIAM Conference on Mathematical and Computational Issues in the Geosciences, Padova, Italy, 2013.

“P5987, L20ff: As the meshes have different edge lengths, you have apparently not used the same mesh in both methods. Have you used the same mesh generation software or algorithms for the ADER-DG and FEM simulations?”

The meshes are different: FaultMod uses hexahedral elements, whereas ADER-DG uses tetrahedral elements. We added a clarification in Section 2 of the manuscript.

“P5987, L21ff: Did you only coarsen the mesh for ADER-DG or for both methods?”

FaultMod used a grid doubling technique in every benchmark presented here, which is a form of mesh coarsening. We added a corresponding sentence in Section 2 of the manuscript.

“P5987, L23: Does Figure 1a (P6018) show the mesh for ADER-DG or for FEM? Please extend the caption.”

We only show meshes for ADER-DG. We added a clarification in Section 2 of the manuscript.

“P5987, L20: Could you briefly elucidate why ADER-DG,O5,200m and FEM,O2,100m are comparable, as they have a different order of accuracy and mesh element size?”

The resolutions are not exactly comparable, but stem from our best attempt to provide a fair comparison. A thorough comparison between two different numerical methods should include element size and order, but also computational runtime and achieved accuracy, which is difficult to determine without an exact reference solution. No equivalent mesh spacing metric is consensually accepted in the community. The problem is also discussed in Pelties et al. (2012) and Galis et al. (2014). In these publications we tested two different resolution parameters delta_x and v, depending on element edge length h and order of accuracy O or degrees of freedom dof = O(O+1)(O+2)/6.

$$\text{delta\_x} = h/O \quad (1)$$

$$v = h/(dof)^{(1/3)} \quad (2)$$

Eq. (1) was used in Pelties et al. (2012), except for low order methods. It gives delta_x = 40 m for ADER-DG O5 h = 200 m and delta_x = 100 m for FaultMod. Eq. (2) gives v = 61.14 m for ADER-DG O5 h = 200 m and v = 50 m for FaultMod. The resolutions have been chosen differently depending on the specific test case.

"P5986: As both methods remove high-frequency oscillations in different ways, yet provide similar simulation results: Do the dissipation in ADER-DG and the viscous layer/Newmark damping in FEM need time-consuming adjustment (e.g. parameter sweeps) to obtain comparable simulation results, is there an analytical way to do it or did it work out of the box?"

The ADER-DG method does not require adjustment by the user. We discussed in Pelties et al. (2012), and more briefly in Section 2 of the manuscript, the advantageous numerical properties of the ADER-DG scheme for dynamic rupture. The choice of the flux function is likely the main reason for the absence of excitation of spurious modes. The applied Godunov flux is adjusted by design to the optimal amount of numerical dissipation. Numerical dissipation has a minor effect on physically meaningful long wavelengths. The dissipation of the shortest wavelengths is virtually instantaneous, such that spurious oscillations are absent in the ADER-DG solution. This is demonstrated by the slip rate time series which are directly measured on the fault. Once implemented in the code, there is no further adjustment of the numerical properties by the user possible, except perhaps through the time step length (which we fixed by CFL = 0.5 in all our simulations).
To our knowledge, other methods adjust the damping parameters (e.g. viscosity and thickness of the viscous layer) based on experience and numerical analysis (e.g. frequency-dependent quality factor of viscous attenuation). For clarification purposes, we added in Section 2: "For details about the specific tuning strategies we refer to Rojas et al. (2007) and Barall (2009).

"Technical corrections

P5983, L14ff and other long sentences throughout the text: try to make shorter sentences with less inserted auxiliary sentences to increase readability. Instead of: "The physical solution is not necessarily insensitive to the precise parametrization of the added damping, which interferes with the actual physics of interest, for example by slowing down the rupture propagation (Andrews, 2005) and smoothing out small scale features, and may also reduce the time step length and thus increase the computational effort considerably." you could make 3 sentences out of the one above:
"The physical solution is not necessarily insensitive to the precise parametrization of an added damping. The damping may interfere with the actual physics of interest, for example by slowing down the rupture propagation (Andrews, 2005) and smoothing out small scale features. The artifical damping may also reduce the time step length and thus increase the computational effort considerably.""

We agree and revised our manuscript accordingly.

"There are some issues with the consistency of style:

P5984, L24 vs. P6005, L7 and others: use either "bi-material"or "bimaterial", but not both"

Changed to bimaterial.

"P5988, L8 vs. P5983, L13: sometimes you separate authors by ";", sometimes by ",". Be consistent."

We agree and revised our manuscript accordingly. However, we found that some inconsistencies occurred during the type-setting process which will be corrected later in collaboration with the type setters.

"P5990, L4 vs. P5985, L20: "distance from the fault" vs "distance to the fault"."

Changed.

"P5997: "velocity-weakening" (L22) vs. "velocity weakening" (L24)"

Changed in the manuscript to velocity-weakening.

"P5997, L17: "Figures" instead of "Figure", but generally decide for one way throughout the document and don't mix: either "Figure/Figures" or "Fig./Figs.""

Changed to Figs. We opted for the abbreviated form throughout the paper.

"P6003, L14: Throughout the paper you've used present tense to talk about your current work. Here and in some following sentences (e.g. P6004, L11) you are using past tense."

Changed to present tense.

"P5984, L10-11: So far you've used a lot of commas for auxiliary sentences. Here you don't. Then you go back to using a lot of commas, but sometimes (P5991L15) not. Either way is fine, but try to stay consistent."

We prefer to set commas for clarification purposes and checked for sentences where commas were missing. Actually, we think the two sentences mentioned are fine as they are.

"P5987, L17: Although this sentence is correct, you break your previous style by omitting the comma before "we"."

Sentence changed. We try to be consistent with commas after "we".

"I am not a native English speaker, so take my advice on grammar with a grain of salt:

P5983, L4: "... approximated accurately ...", I know what you mean but it sounds contradicting. Please change this to your own discretion."

Replaced "approximated" by "represented" to be more clear.

"P5984, L19: "... spontaneous rupture dynamic simulations ..." should be "... spontaneous rupture dynamics simulations ...""

Changed.

"P5986, L17: Is it really an algorithm? What about "time stepping scheme"?"

An algorithm is a step-by-step procedure to solve a given problem which is the case also for time stepping schemes in numerical methods. However, we agree that the term "time stepping scheme" is more usual and changed it accordingly.

"P5986, L26ff: Can you break this sentence up? Especially the last bit (page 5987, line 1) does not seem to fit grammatically.”

Sentence was revised.

“P5987, L6: free-surface should be “free surface””

Changed.

“P5987, L17: Wouldn’t “each” be better than “every”?”

Changed.

“P5988, L6: missing “.” after “transition””

Added.

“P5989, L26: maybe use “each”, not “every”?”

Changed.

“P5991, L7: Could you remove “sometimes”? It is at an incorrect position in the sentence and does not provide additional information.”

Removed.

“P5992, L2: should be “... or by allowing the full branch to rupture””

Sentence was revised.

“P5992, L21: “continues a further”. The “a” should go.”

Removed.

“P5992, L22: “distance the main fault”, missing “along””

Added.

“P5992, L25: “starts a little bit later like” – Could you write this less colloquial?”

Changed to 'starts slightly later similar to'.

“P5992, L27: I am confused by the “as well as at the end”. Either the second “as” has to go or something is missing”

We removed the complete 'as well as'.

“P5993, L5: “we consider this differences” should be “we consider these differences””

Changed.

“P5993, L13: “are more similar” – Could you rephrase and be more specific?”

Changed to 'closer' according to previous text.

“P5993, L15: “in the direct vicinity show” would be easier to read if it was “in the direct vicinity of the branching point show””

Changed.

“P5993, L16: “At this point, we mention that” could be “We’d like to point out that””

Changed.

“P5993, L26: Change word order to “Under certain circumstances such stress perturbation could generate additional propagation modes of rupture ...””

Changed.

“P5993, L19: Change word order to “These differences can also be noted ...” “

Changed on page 5994, L19 which was probably referred to.

“page 5997, line 9: “... is a smoothly version ...” should be “... is a smooth version ...””

Sentence was removed due to comment of reviewer #2.

“P5997, L11: please add comma after “self-consistency”, as you have one before “for””

Added.

“P5997, L15: use “smoothly” instead of “smooth””

Changed.

“P5998, L20: I think, the code is not called SBIE, could this be rephrased to something like “and the three-dimensional spectral boundary integral element method implementation by Lapusta and Liu (2009)” ?”

Changed to “... and the three-dimensional spectral boundary integral element (SBIE) method implementation by Lapusta and Liu (2009)”.

“P5999, L22: Try to make the last sentence a statement that works without references to the sentences before, like: “We conclude that the advanced geometric flexibility of ADER-DG/SeisSol combined with its(?) enhanced accuracy ... in complicated setups.” Additionally, depending on your intention you might want to use “complex” instead of “complicated”.”

Changed to “We conclude that the advanced geometric flexibility of SeisSol combined with its enhanced accuracy positions it as a competitive tool to study earthquake dynamics in complicated setups.”

“P6003, L13: “experimental based law” could be “experiment-based law” or “experimentally based law””

Changed to “experiment-based law”.

Answer letter to referee #2 Anonymous

Response prepared by C. Pelties, A.-A. Gabriel, and J.-P. Ampuero.
The authors thank the referee for providing a thorough review. We prepared the following changes and clarifications to improve the manuscript accordingly. In the following we address all suggestions and comments of the review in detail. Each of our response items starts with the corresponding quote from the review. A revised version of the manuscript can be found in the supplement.

"This is a useful paper that presents thorough benchmarking of an arbitrary high-order derivative Discontinuous Galerkin (ADER-DG) method on unstructured meshes for advanced earthquake dynamic rupture problems. The authors validated the method in comparison to well-established numerical methods in a series of verification exercises and showed that the combination of meshing flexibility and high-order accuracy of the ADER-DG method makes it a competitive tool to study earthquake dynamics in complicated setups. I think that this paper is worth publishing in GMD. My minor points are the followings:
P5983: Ohnaka and Mogi (1982) is not an appropriate reference because this paper did not discuss a constitutive law. Examples of better references to cite here are Ohnaka and Kuwahara (1990) or Ohnaka and Shen (1999). Ohnaka, M., and Y. Kuwahara, Characteristic features of local breakdown near a crack-tip in the transition zone from nucleation to unstable rupture during stick-slip shear failure, Tectonophysics 175, 197-220, 1990. Ohnaka, M., and L. Shen (1999), Scaling of the shear rupture process from nucleation to dynamic propagation: Implications of geometric irregularity of the rupturing surfaces, J. Geophys. Res., 104(B1), 817–844, doi:10.1029/1998JB900007."

The reference is changed to Ohnaka and Kuwahara (1990).

"P5988 lines 3-5. "The development of a supershear daughter pulse in TPV11, caused by stress concentration ahead of the sub-shear rupture front (Dunham, 2007), is equally well captured, as shown in Fig. 2." I cannot catch well the development of a supershear daughter pulse from Figure 2. More explanations will be necessary."

Changed to "The development of a supershear rupture front in TPV11 is equally well captured, as shown in Fig. 2. The rupture time contour plot in Fig. 2 (a) captures the boost in rupture velocity after supershear transition."

"P5995, line 25; P5996, line 16; p5992, line 21 "rate-and-state dependent constitutive relationships" should be amended to "rate- and state-dependent constitutive relationships""

Changed accordingly.

"Equation (1): Definition of L should be added."

Changed accordingly. The definition of L was moved forward from being defined only after Eq. 6. Furthermore, the naming convention of L was unified between tables 5 and 6 to 'characteristic slip scale'.

"Equation (4): W and w should be defined. If spatial distribution of a is shown in Figure 15a, this equation seems to be unnecessary."

As suggested, equation (4) is removed and replaced by a reference to Fig. 15 a).

“P 5997, line 20: Since a usual rate- and state-dependent friction law is introduced, slow velocity friction seems to be better.”

We assume the reviewer refers here to p. 5996, line 1, the heading of subsection 7.1 “Slow velocity weakening ”. We followed the suggestion and changed this heading to “Slow velocity friction”.

“P5986 line 11 “ we compare our results to the well-established software FaultMod” → “we compare our results to that from the well-established software FaultMod””

Changed accordingly.

“Equation (7) Is vˆ8/v_* correct ? In the paper by Dunham et al. (2011), this term is (v/v_*)ˆ8. Also mu_s seems to be mu_w.”

Changed to (v/v_*)^8.

“P6004 lines 14. The unit of L should be added.”

Added.

“Table 1. tau_0 is used for nucleation shear stress along-dip and nucleation shear stress along-dip.”

Nucleation shear stress along-dip is now referred to as tau_nuc.

“Table 6: this is not referred to in the main text.”

A reference is added.

“Figure 8: Unit of X and Y should be added.”

Units are added.

“Figure 9: In the main fault (a), it would be better to show the location where fault branch occurs. Please add an explanation on a blue zone (concentrated blue lines).”

The branch occurs at along-strike distance 0. We added a corresponding sentence. Furthermore, we added a discussion about the 'concentrated blue lines': “The concentration of rupture fronts after along-strike distance > 0 km on the main fault for ADER-DG without gap (concentrated blue lines in Fig. 9(a))is simply the result of a smooth, spontaneous rupture arrest in the branch (as opposed to abrupt arrest by a barrier).”

“Figure 15 a: It would be better to divide this figure into two figures; initial stress and friction parameter a.”

Since the initial distribution of the parameter a only affects the domain boundaries, whereas the initial distribution of stress acts in the vicinity of the hypocenter, we believe both visualizations can be shown in one figure and would carry too little information shown separately.

“In Figure 17, the unit of L should be added. Caption “the nucleation zone. for different” -> “the nucleation zone for different””

Dot is removed and unit is added.