Parallel uncertainty quantification with fused simulations

Sebastian Wolf[1], Linus Seelinger[2], Michael Bader[1]

[1]Technical University of Munich, School of Computation, Information and Technology, Boltzmannstraße 3, 85748 Garching, Germany
[2]Interdisciplinary Center for Scientific Computing, Ruprecht-Karls-Universität Heidelberg, Berliner Str. 41-49, 69120 Heidelberg, Germany
wolf.sebastian@cit.tum.de

SeisSol[1] is a simulation software for extreme-scale simulations of earthquake source dynamics and seismic wave propagation. It features several friction laws and failure criteria on the fault as well as various material rheologies to model the solid Earth. It builds on the Discontinuous Galerkin (DG) space discretization with Arbitrary DERivative (ADER) time-stepping to achieve high-order convergence in space and time. SeisSol has been designed for the single simulation of one complex scenario, see e.g. [2]. To geoscientists, it is not only important to simulate forward models to analyze "what-if" scenarios, but they also want to study subsurface phenomena by observations at the Earth's surface. In that regard, significant fields of research are the inversion for material parameters and the inversion of earthquake sources.

MUQ [3] is a toolbox for uncertainty quantification (UQ) with a particular focus on parallel computing [4]. It features models for forward uncertainty propagation (polynomial chaos expansion) and models for Bayesian inverse problems (Monte Carlo Markov Chain). For such inverse problems, a forward model G is given, which computes synthetic observables from some input vector. We assume that the input vector is a random variable with unknown distribution. The task is to find this distribution, conditioned on given observations. In this work, we couple MUQ with SeisSol in order to find the probability distribution of an earthquake source based on surface-recordings of an earthquake.

With SeisSol, we can fuse $N \geq 1$ simulations in one run, if they share the same mesh and material model [6]. This gives us several advantages: First, we can choose the number of simulations as a multiple of the vector register length. This way, we can perfectly vectorize the code with SIMD instructions along the dimension of fused simulations. Second, in comparison to running N non-fused simulations sequentially, the workload per mesh-element is N times higher in the fused case, such that we reach the strong-scaling barrier later. Third, we only have to do the setup phase, e.g. mesh I/O, once for N simulations. Fused simulations are also a valuable tool for efficient parameter studies.

To reflect the fused simulations on the UQ side, we have considered different approaches [5]. The best choice is the generalized Metropolis-Hastings (GMH) algorithm [1]. The GMH algorithm is a natural extension to the well-known Metropolis-Hastings algorithm, used in MCMC simulations. The idea behind the standard MCMC approach is to sample a new state (in our case coordinates of the hypocenter) from a prior distribution. With this state as input, we

---

[1] `https://seissol.org`

evaluate the forward model (in our case synthetic seismograms at given receiver stations). Then, we calculate the likelihood of the state by comparing the result of the simulation and the observation. Based on the misfit, we either append the proposed state as a new entry to our Markov Chain, or we neglect it. In the GMH approach, for each step, we draw N different samples from the prior distribution at once. For all N points, we evaluate the forward model concurrently and then accept $M \leq N$ points based on their likelihoods.

On this poster, we will present scaling and performance studies that highlight the benefit of using fused simulations over subsequent single simulations. We focus in detail on a layer over halfspace scenario, which is a widely accepted benchmark in the wave-propagation community. In this scenario, a point source excites seismic waves, which are recorded at a given set of receivers. We find the probability distribution of the source location by applying the previously mentioned coupling of SeisSol and MUQ. We compare the efficiency of using different configurations of N and M.

References

[1] Calderhead, B. (2014). A general construction for parallelizing Metropolis-Hastings algorithms. Proceedings of the National Academy of Sciences, 111(49), 17408–17413.

[2] Krenz, L., Uphoff, C. et al. (2021). 3D acoustic-elastic coupling with gravity: The dynamics of the 2018 Palu, Sulawesi earthquake and tsunami. Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 1–14.

[3] Parno, M., Davis, A. et al. (2014). MIT uncertainty quantification (MUQ) library.

[4] Seelinger, L., Reinarz, A. et al. (2021). High performance uncertainty quantification with parallelized multilevel Markov chain Monte Carlo. Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 1–15.

[5] Sperling, N. (2022). Uncertainty Quantification of Seismic Simulations on High Performance Computers [Master's thesis, Technische Universität München].

[6] Uphoff, C., Bader, M. (2020). Yet Another Tensor Toolbox for Discontinuous Galerkin Methods and Other Applications. ACM Transactions on Mathematical Software, 46(4), 34:1-34:40.