

# Can we accelerate Maximum-Likelihood Computations on GPUs?

Praxis der Forschung Projekt von Christoph Stelz

Phylogenies are increasingly used in all fields of medical and biological research. Moreover, because of the next-generation sequencing revolution, datasets used for conducting phylogenetic analyses grow at an unprecedented pace.

Maximum Likelihood (ML) tree inference is a statistical method used in phylogenetics to determine the phylogenetic tree which best explains the observed data based on a given model of sequence evolution. Typically, Markov model of sequence evolution is assumed, which defines the probabilities of different types of changes (e.g., substitutions of one DNA character for another). Then, we search for the tree topology, branch lengths, and model parameters that maximize the likelihood of observing the given sequence data under this model. Topological search under the maximum likelihood criterion is NP-hard, and tree likelihood evaluation is computationally expensive. Thus, highly optimized and scalable codes are needed to analyze constantly growing empirical datasets.

RAxML-NG (Randomized Accelerated Maximum Likelihood Next Generation) is a popular program for phylogenetic analyses of large datasets under maximum likelihood. RAxML-NG currently uses three levels of parallelism: Instruction-level parallelism (SIMD), thread-level parallelism (PTHREADS), and distributed memory parallelism (MPI). Likelihood computation in RAxML-NG is implemented in computational kernels which are highly optimized for modern x86 CPUs and most common sequence data types (DNA, proteins).

In this “Praxis der Forschung” project we plan to evaluate the feasibility of parallelizing Maximum Likelihood methods on modern (2023) GPUs and possibly ARM CPUs. We plan to start with a simplified computational kernel to explore capabilities without having to deal too much with the internal workings of RAxML-NG. That is, the first step will be to develop a simple program that reads in a tree, traverses it, computes the likelihood, and optimizes the branch lengths as well as model parameters. If we can achieve promising results, the next steps towards GPU-accelerated RAxML-NG would be to perform the actual tree search and integrate advanced optimizations such as subtree site repeats.