

## **Toward Next-Generation Benchmarking of CEM Methods: Comparing Computational Costs**

Jackson W. Massey, Anton Menshov, and Ali E. Yilmaz  
Department of Electrical and Computer Engineering  
The University of Texas at Austin, Austin, TX 78712

Benchmarking in computational electromagnetics research is commonly understood as the task of using a (new) algorithm to solve sample scattering problems (by those creating and implementing the algorithm) and comparing the newly computed values of certain quantities of interest (e.g., radar cross section) to theoretical, experimental, or previously computed references. The sample scattering problems range from perfect spheres with analytical Mie series solutions of fields at all points in space and time to canonical shapes with experimentally measured far fields at sample frequencies (A. C. Woo et al., *IEEE Antennas Propag. Mag.*, Feb. 1993). The primary goal of this mode of benchmarking is to (i) obtain empirical evidence that validates (and can potentially falsify) theoretically expected features of an algorithm, e.g., its error convergence rate or asymptotic cost scaling with respect to the degrees of freedom.

Beyond this rather perfunctory role, benchmarking can also be used to (ii) quantify the performance of a computational method—more precisely, the performance of a specific software implementation of an algorithm executed on a particular hardware—(iii) inform researchers and the public about the latest problems of interest as well as the current state of computational methods for solving these problems, (iv) compare competing computational methods (algorithm, software, and hardware combinations) more objectively and judge their suitability for solving problems of interest, and (v) identify weaknesses of computational methods and inspire future research for addressing them. Such benchmarking, however, is rarer and more difficult. Indeed, existing benchmark suites rarely contain all the ingredients necessary to perform this mode of benchmarking; in particular, computational cost measures are often conspicuously absent.

This article presents the approach used in a recently introduced bioelectromagnetic benchmark suite for quantifying computational costs (J. W. Massey et al., *URSI Int. Symp. Electromagn. Theory*, Aug. 2016). Benchmark participants are asked to report wall-clock times of the methods they used to solve the benchmark problems along with the theoretical peak performance of their computer. The methods are then compared both in terms of their observed wall-clock time, which helps identify the fastest method (for a given error level, which is also found for each method as part of the benchmark), and in terms of their normalized time (wall-clock time \* computer's theoretical peak floating point operations/s), which helps identify the opportunity cost of each method. This dual comparison highlights the two extreme modes of operation for computational methods, especially parallel ones: obtaining a fast run time (but with a high opportunity cost/low parallel efficiency) by using many threads/processes or a low opportunity cost/high parallel efficiency (but with a slow run time) by using few threads/processes. While the normalization used in the bioelectromagnetic benchmark suite is simple to implement and does enable a coarse comparison of different methods deployed on different hardware platforms, it should eventually be supplanted with more precise estimates of the actual computational power available to the methods in order to resolve differences among closer performing methods.