Extending Proto-Benchmarks to Create Benchmarks for Quantifying Modern Computational Electromagnetics Performance

Jon T. Kelley[1], Jackson W. Massey[2], and Ali E. Yılmaz*[2]

[1] Department of Physics

[2] Department of Electrical and Computer Engineering

The University of Texas at Austin, Austin, TX 78712

In software engineering, a "proto-benchmark" is defined as a set of tests used to compare the performance of alternative tools or techniques that is missing one of the components of a proper benchmark; most commonly, a performance measure (S. E. Sim *et al.*, "Using benchmarking to advance research: a challenge to software engineering," in *Proc. Int. Conf. Software. Eng.*, May 2003). Both proto-benchmarks and benchmarks can be used to demonstrate features of a new computational system (algorithm, code, supporting software, hardware), while only benchmarks can also systematically combat the ubiquity of error, ameliorate the hazards of specialization, fortify scientific integrity, and inspire research in addition to also demonstrate features of a computational system (A. E. Yılmaz "Advancing computational electromagnetics research through benchmarking," in *Proc. USNC/URSI Rad. Sci. Meet.*, July 2017).

To be used for these purposes, computational electromagnetics (CEM) benchmarks must contain precisely defined problems of increasing complexity that are representative of sub-fields of electromagnetics, specific quantities of interest (e.g., monostatic radar cross section), available reference solutions, and performance metrics that include both error and cost measures (J. W. Massey *et al.*, "Benchmarking to close the credibility gap: A computational BioEM benchmark suite," in *Proc. URSI Int. Symp. Electromagn. Theory*, Aug. 2016). Although such CEM benchmarks are currently in their infancy, a variety of verification and validation studies are available in the literature (A. Greenwood, "Electromagnetic code consortium benchmarks," Tech. Report, AFRL-DE-TR 2001-1086, Dec. 2001; A. C. Woo *et al.*, "Benchmark radar targets for the validation of computational electromagnetic programs," *IEEE Antennas Propagat. Mag.*, Feb. 1993). The tests in these studies can generally be categorized as proto-benchmarks because they do not include performance metrics that quantify the trade-off between computational error and cost.

This article proposes a method for constructing CEM benchmarks by extending existing validation and verification proto-benchmarks; specifically, it proposes to supplement the test cases in existing studies with error definitions (relative to carefully chosen reference data) and computational cost data. To demonstrate the proposed method, several proto-benchmarks commonly used to validate scattering from metallic objects are extended by: (i) defining error measures with respect to high-resolution references that are cross-validated with existing measurement data; (ii) performing empirical computational experiments using a variety of simulation methods and providing their errors and computational costs; and (iii) publishing the performance data on a website. Benchmarks resulting from these extensions can facilitate CEM benchmarking on modern computational systems.