SPECIAL SECTION ON TWO THEMES: CSE SOFTWARE AND BIG DATA IN CSE

The 2015 SIAM Conference on Computational Science and Engineering (CSE) was held March 14-18, 2015, in Salt Lake City, Utah. The *SIAM Journal on Scientific Computing* (SISC) created this special section in association with the CSE15 conference. The special section focuses on two topics that are of significant current interest to CSE researchers: CSE software and big data in CSE.

The development of efficient, robust, and sustainable software is at the core of CSE. Software itself is now broadly recognized as a key crosscutting technology that connects advances in mathematics, computer science, and domain-specific science and engineering. Software encapsulates progress in algorithms, methods, and implementations across CSE and thereby provides a foundation for long-term CSE collaboration and scientific progress.

A second topic of great current interest in CSE and beyond is how the data revolution enables new ways of generating quantitative insight for applications in science, engineering, technology, and society. The world is experiencing an explosion of digital data, and there are important challenges in designing scalable algorithms and methods to turn the data deluge into a source of scientific and engineering progress.

The special section contains 34 papers. There are 22 papers on topics related to CSE software and 10 papers on algorithms with applications in big data. The final two papers combine aspects of software and big data; this is not a surprising combination, since open source software for data analysis is of timely interest.

Among the 22 software-related papers, there are 11 papers on scientific software packages that use differential equation models for science and engineering for such diverse applications as soft materials, explosive devices, and insect flight. Aspects of multiphysics and multiresolution modeling, scalability, and mesh management are highlighted. Two of these papers discuss efficiency and performance of adaptive grid refinement packages, and a further paper presents a model order reduction package. There are four papers on notable aspects of simulation software, including algorithmic resilience, automatic simulation workflow change, data and task parallelism, and systems to guide solver choices. There are also seven papers on linear algebra software, including preconditioners, eigensolvers, and hierarchical and multilevel system solvers. Several of the software papers focus on aspects of high-level software organization, while others highlight algorithms, efficient parallel implementations, and performance tests. As such, the scope of this special section is somewhat broader, e.g., in terms of scientific software functionality, than the papers that usually appear in SISCs Software and High-Performance Computing section. Indeed, CSE software is a valuable research product and a cornerstone of CSE collaboration, and as part of the growing realization of the importance of software for the CSE endeavor, the field continues to see evolution as to where and how software contributions of various nature are best published. This special section is part of that exploration and evolution.

The special section contains 10 papers on algorithms with applications in big data. There are two papers on matrix and tensor completion, with applications in, e.g., large-scale recommendation and stochastic PDE simulation. Two papers discuss randomized algorithms for large linear algebra problems. One paper considers an adaptive metropolis approach for scalable parallel Bayesian inference, and another paper considers a multilevel method for large-scale sparse inverse covariance estimation. There are two papers on parallel algorithms for matching and partitioning on

large graphs targeting problems with up to billions of vertices and edges. There is a paper on parallel density estimation for 100 million particles on 500 billion grid points and a paper on parallel nearest neighbor search in high dimensions with scalability tests on more than 100,000 computing cores for 4 billion points in 1,000 dimensions.

Finally, the first paper that combines software and big data aspects presents an extension of the Apache Spark distributed data analytics environment to enable high-performance visual computing applications on GPU-based cluster systems. The second software-big data paper discusses a parallel open-source library for high-dimensional kernel summations with applications in machine learning and computational statistics, scaling to billions of points in hundreds of dimensions.

The guest editorial board worked very hard to ensure a rigorous peer-review process while meeting deadlines. Special thanks are due to Mitch Chernoff (SIAM Publications Manager) and Brittni Holland (SIAM Editorial Associate) for their efforts on this special section.

Guest editors
Hans De Sterck, Chris Johnson, and Lois Curfman McInnes

Guest associate editors
Evrim Acar, David Bader, Wolfgang Bangerth, Peter Bastian, Erik Boman, Jed Brown, Donna Calhoun, Hank Childs, Paul Constantine, Petros Drineas, Omar Ghattas, David Gleich, Robert Harrison, Andreas Klöckner, Tamara Kolda, Vipin Kumar, Anders Logg, Andrew Lumsdaine, Luke Olson, Jeff Phillips, Ali Pinar, Jack Poulson, Han-Wei Shen, Masha Sosonkina, Garth Wells, Tim Warburton, and Carol Woodward