Bigger, Faster, Random(ized): Computing in the Era of Big Data

Abstract:
Our capacity to produce and store large sets of data has increased exponentially over the course of the last two decades; the development of algorithms for sifting through it efficiently is somewhat lagging behind. Randomization is being recognized as a powerful tool, whether in constructing models on which algorithms can be tested, or in sampling the data reliably, or in speeding up and optimizing existing algorithms. In particular, basic results from random matrix and random graph theory are being employed at the forefront of scientific computing; often, assembling algorithms and producing theoretical guarantees for them requires a blend of probability, combinatorics, graph theory, numerical analysis, and optimization. I will speak about two results, one in which randomization is used to achieve a communication-minimizing non-symmetric eigenvalue solver, and one establishing a spectral gap in bipartite biregular graphs, with applications in areas as varied as community detection, matrix completion, and error-correcting codes. This is joint work with Jim Demmel and Grey Ballard, respectively, Gerandy Brito and Kameron Harris.

Hosts: Philip Gill & Jacques Verstrate

Monday, December 11, 2017
2:00 PM
AP&M 6402