FOPAM 2026
Proceedings of the 3rd Foundations of Process/Product Analytics and Machine Learning
Decatur, USA, July 26-30 (2026)

Data-driven Optimization: Efficient Adaptive Learning for Self-Driving Laboratories

Nick Sahinidis

Georgia Tech, School of Chemical & Biomolecular Engineering

ABSTRACT

Self-driving laboratories promise to compress materials-discovery timelines from years to weeks by replacing trial-and-error experimentation with closed-loop, algorithm-guided campaigns. Yet, despite the rapid proliferation of robotic and automation hardware, today's autonomous labs rely almost exclusively on Bayesian optimization (BO) to decide what experiment to run next. BO is a sensible approach to low-dimensional optimization problems with smooth response surfaces, but it struggles in precisely the regimes that matter most for real materials campaigns: tight experimental budgets, dozens of process parameters, mixed-integer choices, hard physical constraints, and noisy expensive measurements.

In this talk, I will show how moving from BO to partitioning-based algorithms can substantially improve data efficiency, scale gracefully to dozens of process variables, and handle the constraints and noise that characterize realistic experimental campaigns. I will summarize a recently completed large-scale black-box optimization (BBO) benchmark in which we compared 42 solvers across 502 problems ranging from one to 300 dimensions and from smooth and convex to non-smooth and nonconvex. The results overturn several community assumptions: BO solves only about 9% of problems within a 2,500-evaluationbudget, while a new branch-and-model (BAM) algorithm reaches an 81% success rate, with GLCCLUSTER, MULTIMIN, MCS, and SNOBFIT also performing strongly. A minimal, irreducible set of eight complementary solvers attains 88% solvability on the full suite.

I will then move from in-silico benchmarks to the wet lab, presenting a recent algorithm-guided experimental campaign on high-performance perovskite solar cells in which a non- BO solver was used to co-optimize six process variables spanning the perovskite, electron-transport, and hole-transport layers. Time permitting, I will also share early results from applying ensembles of BBO algorithms to digital twins of self-driving labs across additional materials systems.

I will close with a forward-looking research vision: accelerating autonomous labs by developing, benchmarking, and experimentally validating data-efficient adaptive algorithms across batteries, semiconductors, catalysts, polymeric membranes, and biomolecules. The benchmarking software will be released as open source, with BAM and most BBO software available free to academic users, so that experimental groups can deploy these tools on their own self-driving platforms.