Empirically Comparing the Finite-Time Performance of Simulation-Optimization Algorithms

Anna Dong, **David Eckman**, Xueqi Zhao, Shane Henderson
Cornell University

Matthias Poloczek
University of Arizona

Winter Simulation Conference
December 5, 2017

Simulation Optimization (SO)

Optimize a real-valued objective function, estimated via simulation, over a deterministic domain.

Challenges:

1. Error in estimating objective function.
2. Unknown topology (continuity/differentiability/convexity).

SO algorithms are designed to solve a broad class of problems.

SO Algorithm Performance

Many theoretical results for asymptotic performance,

- . . . as simulation budget approaches infinity.

Examples:

- Algorithm converges to local (global) optimizer.
- Convergence rate, once within neighborhood of optimizer.

Required budget for such results may exceed practical budget!

To decide which algorithm to use, a practitioner cares about **finite-time** performance.

Evaluating Algorithms

SO community lags behind other optimization communities:

- Established testbed of problems for benchmarking.
- Metrics for empirical finite-time performance.
- Comparison of algorithms on large testbed.

We implement several popular SO algorithms and test them on a subset of problems from the SimOpt library (`www.simopt.org`).

Objectives

Near-term:

1. Comparison of finite-time performance of different algorithms.
2. Insights on the types of problems on which certain algorithms work well.

Long-term:

1. More contributions to SimOpt library.
   - Problems and/or algorithms.
2. Development of finite-time performance metrics.
   - E.g., adapting performance profiles.
3. Motivate others to do similar comparisons.
4. Development of algorithms with strong finite-time performance.

Evaluating Finite-Time Performance

In deterministic optimization:

- Measure computational effort needed to get to optimal solution (or within specified tolerance).
  - Number of function evaluations or wall clock time.

Doesn't work so well for SO.

- Optimal solution is often unknown.
- Often no certificate of optimality.
- Estimation error makes it hard to check tolerance condition.

Evaluating Finite-Time Performance

Idea

1. Fix a simulation budget.
2. Evaluate the objective function at the estimated best solution found within the budget.

We measure the budget in number of objective function evaluations.

Evaluating Finite-Time Performance

Let $Z(n)$ be the **true** objective function value of the **estimated** best solution visited in the first n objective function evaluations.

- $Z(n)$ is a random variable, because the estimated best solution, $X(n)$, is random.
- Conditional on $X(n)$, $Z(n)$ is fixed, but needs to be estimated.

In our experiments, we estimate $Z(n)$ (conditional on $X(n)$) in a *post-processing* step.

Evaluating Finite-Time Performance

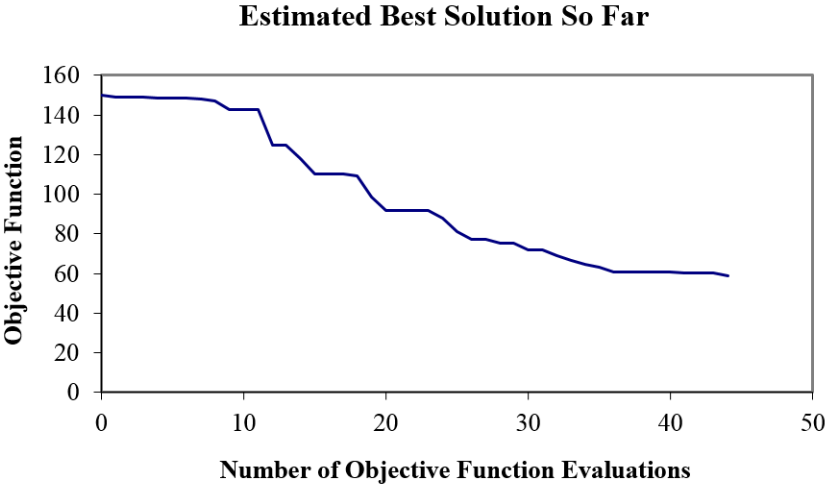


Figure: Pasupathy and Henderson (2006).

Evaluating Finite-Time Performance

Can obtain $Z(n)$ curve from single macroreplication of an algorithm.

- Unless the algorithm uses the budget in setting parameters.

Location of $Z(n)$ curve is random, so take several macroreplications and look at:

- mean,
- median/quantile,
- empirical cdf *(hard to show on one plot)*.

Algorithms

1. Random Search (non-adaptive)
2. Gradient Search with Random Restarts
   - Central finite differences for gradient estimate.
3. Simultaneous Perturbation Stochastic Approximation (SPSA)
   - Uses budget to set gain sequence.
4. Stochastic Trust-Region Response-Surface Method (STRONG)
   - Didn't use design of experiments for fitting.
   - Central finite differences for gradient estimate.
     a. BFGS estimate of Hessian. (STRONG)
     b. No second-order model. (STRONG-Stage1)
5. Nelder-Mead
   - Simplicial method that doesn't use gradient information.

Benchmark Problems

Table: SimOpt benchmark problems and their characteristics.

Name on SimOpt	Dimension	Optimal Solution
A Multimodal Function	2	Known
Ambulances in a Square	6	Unknown
Continuous Newsvendor	1	Known
Dual Sourcing	2	Unknown
Economic-Order-Quantity	1	Known
Facility Location	4	Unknown
GI/G/1 Queue	1	Unknown
M/M/1 Metamodel	3	Known
Optimal Controller for a POMDP	10	Unknown
Optimization of a Production Line	3	Unknown
Parameter Estimation: 2D Gamma	2	Known
Rosenbrock's Function	40	Known
Route Prices for Mobility-on-Demand	12	Unknown
SAN Duration	13	Unknown
Toll Road Improvements	12	Unknown

Problems

Properties of all problems:

- Continuous decision variables.
- Deterministic (box) constraints or unbounded.

Initial solution is drawn from probability distribution over domain.

- Uniform distribution for bounded variables.
- Exponential/Laplace distribution for unbounded variables.

Took 30 replications at given solution to estimate its objective value.

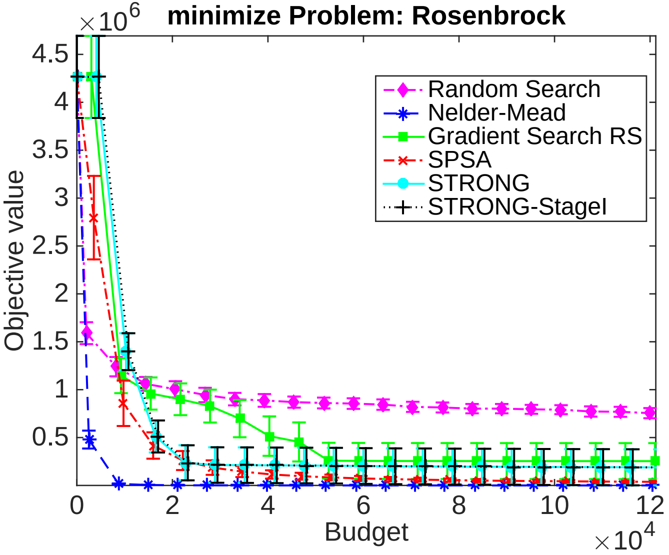
- Used common random numbers (CRN) across solutions.

Experiments

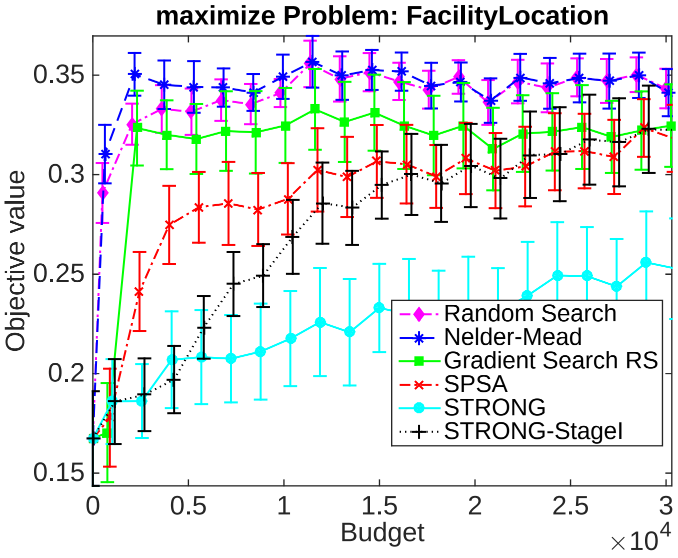
For every problem:

- Ran 30 macroreplications of each algorithm.
- Recorded estimated best solution $X(n)$ for range of n values.
- Ran 30 function evaluations at each $X(n)$ to estimate $Z(n)$ conditional on $X(n)$.
- Averaged 30 estimates of $Z(n)$.
- Constructed 95% normal confidence intervals.

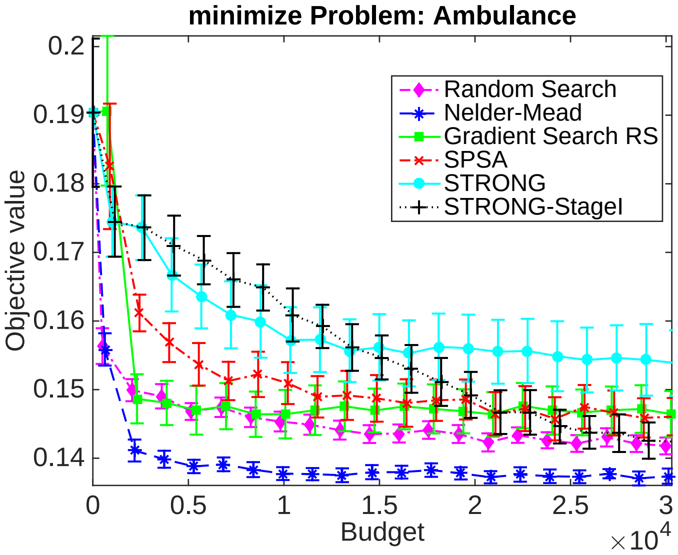
Rosenbrock ($dim = 40$)



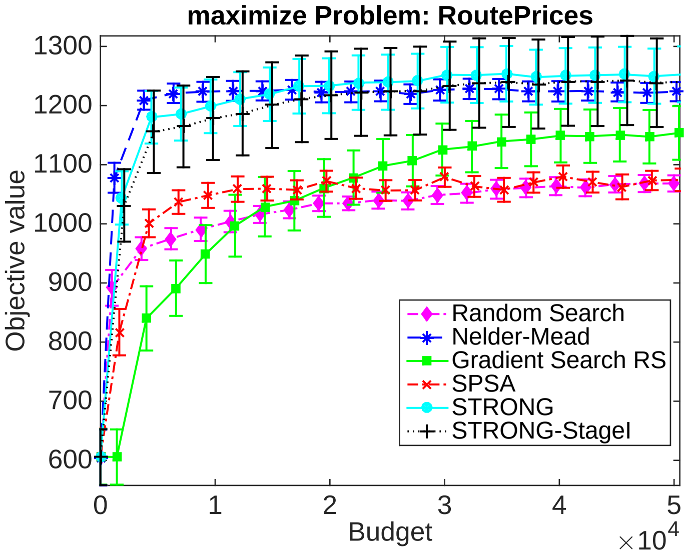
Facility Location ($dim = 4$)



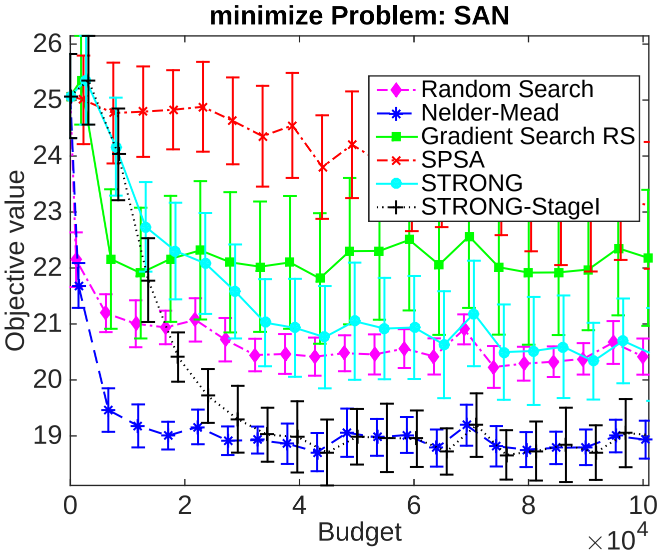
Ambulance ($dim = 6$)



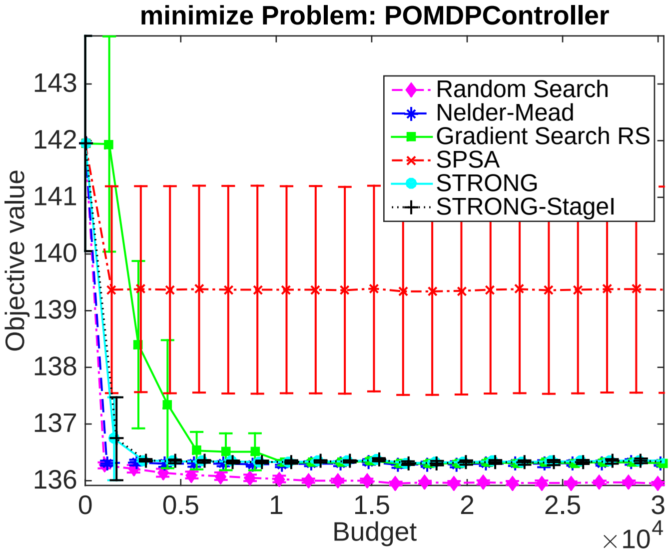
Route Prices ($dim = 12$)



Stochastic Activity Network ($dim = 13$)



POMDP Controller ($dim = 10$)



Conclusions

Takeaways

1. Robust performance of Nelder-Mead across problems.
2. STRONG-Stage1 did as well as (or better than) STRONG.
3. Random Search did better than expected.
4. Performance of SPSA was sometimes highly variable.

Future Work

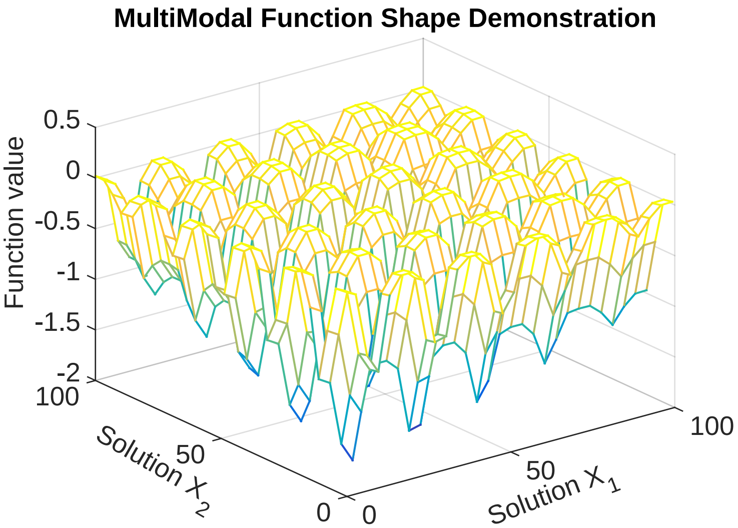
Similar comparisons for:

- High-dimensional problems.
- Discrete/integer-ordered variables.
- Stochastic constraints.

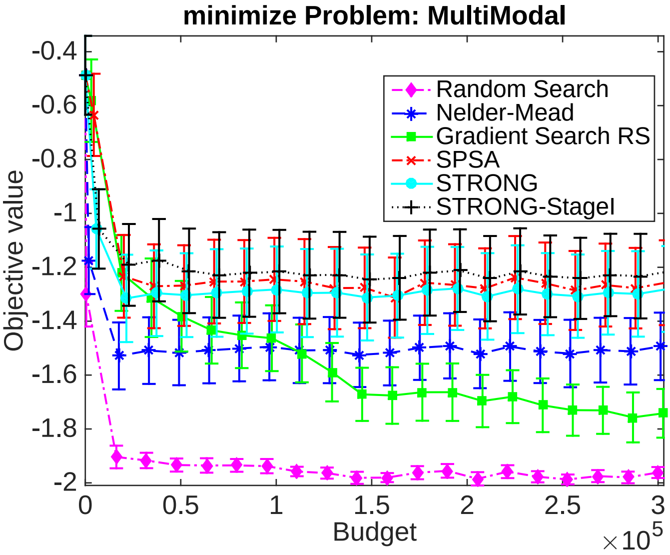
Acknowledgments

This material is based upon work supported by the National Science Foundation under grants CMMI–1537394, CMMI–1254298, CMMI–1536895, and IIS–1247696, by the Air Force Office of Scientific Research under grants FA9550–12–1–0200, FA9550–15–1–0038, and FA9550–16–1–0046 and by the Army Research Office under grant W911NF–17–1–0094.

MultiModal ($dim = 2$)



MultiModal ($dim = 2$)



Ambulance ($dim = 6$)