Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

OPTIMIZING STOCHASTIC SYSTEMS USING STREAMING SIMULATION

Robert J. Lambert[1]

[1]STOR-i Centre for Doctoral Training, Lancaster University, Lancaster, UK

ABSTRACT

Sequential optimization of stochastic systems is an increasingly important area of research. With modern systems generating continuous streams of data, decision-making policies must be able to adapt in real time to incorporate new information as it arrives. My PhD research develops theory and methods for online sequential decision-making with streaming observations. In particular, it addresses the key challenges of providing convergence guarantees for simulation optimization procedures, whilst maintaining model accuracy within computational limits. Initial work focused on optimizing M/G/c queueing systems with unknown arrival processes, deriving asymptotic convergence results for a Monte Carlo-based algorithm. Subsequent research has extended these ideas to finite-time performance evaluation, decision-dependent observations, and adaptive policies guided by approximate dynamic programming principles. These results provide a foundation for designing adaptive, data-driven policies in complex stochastic systems, enabling more reliable real-time decision-making.

1 STREAMING SIMULATION OPTIMIZATION

Across a number of domains, practitioners face the challenge of optimizing stochastic systems. In settings where analytical models are unavailable or intractable, simulation optimization (SO) methodologies are a natural approach to adopt. Traditionally, SO is viewed as an offline procedure, where models are constructed according to some fixed collection of historical observations about the target system. With recent technological advances, we increasingly encounter data-rich environments in which system observations can be collected sequentially, creating an ever-growing pool of information. This information structure is often referred to as *streaming observations* or *streaming data*.

We consider a SO problem of the form

$$\min_{x\in\mathscr{X}} \left\{\mathbb{E}_{\xi}\left[F(x;\theta,\xi)\right]\right\},$$

where $x \in \mathscr{X}$ denotes an action within a finite decision set, θ denotes an unknown system parameter, $F(x;\theta,\xi)$ is a performance measure of interest, and the randomness is fully characterized by a random vector ξ.

The difficulty in this problem lies in the fact that action $x \in \mathscr{X}$ that minimizes the expectation depends directly on θ; however, we assume no prior knowledge of the true parameter θ. Instead, we rely on streaming system observations to sequentially obtain estimates of θ, denoted by $\hat{\theta}_n$ for $n \in \mathbb{N}$, and use these within the simulation model. This problem naturally aligns with a sequential decision-making framework, where iterative decisions are made in order to balance exploration of the parameter space against optimal system operation in terms of x. As such, a number of questions arise:

1. What statistical convergence guarantees can be established for SO methods under streaming observations?
2. In online settings, how can we maintain reliable model accuracy in real time without incurring unnecessarily large computational costs?
3. How can we improve the estimation efficiency using structural/shape constraints, i.e., monotonicity, convexity of $\mathbb{E}_{\xi}\left[F(x;\theta,\xi)\right]$ in θ?

These questions have formed the motivation for my doctoral research.

2 ASYMPTOTIC RESULTS

The canonical example of a stochastic system is a queueing system. These often feature in decision-making problems for staffing, routing, or scheduling, and align well with the streaming SO framework. For example, in domains such as retail operations or airport management, the same system must be optimized repeatedly whilst new information becomes available.

We recently developed theoretical results characterizing the asymptotic behavior and computational limitations of simulation-based decision-making algorithms in streaming data settings (Lambert et al. 2025). The motivating problem involved adaptively selecting the number of servers to implement within an $M/G/c$ queueing system with an unknown exogenous arrival process. By formulating the problem to account for a generalizable class of objective functions, the work provided new theoretical tools to evaluate SO algorithms in a streaming setting, thus contributing to the growing literature on SO under input uncertainty (He and Song 2024).

Our work focuses on evaluating the performance of a greedy Monte Carlo-based algorithm in the presence of two sources of uncertainty:

- *Input uncertainty:* uncertainty introduced by approximating an unknown quantity with a sample average - driven by the number of observation periods, n.
- *Simulation uncertainty:* uncertainty introduced by running m_n simulation replications in each period.

Taking into account each of these, we derived a Lindeberg-Feller Central Limit Theorem, which we used to derive an asymptotic bound on the probability of incorrect selection for the policy, namely $O(\max\{n^{-1}, m_n^{-1}\})$. We further derived bounds on the expected cumulative regret of the algorithm R_n, providing conditions for optimal computational budget allocation during the process, namely

$$R_n := \begin{cases} O(\log n), & \text{if } m_n = \Omega(n) \\ O(\sqrt{n}), & \text{if } m_n = o(n). \end{cases}$$

3 CURRENT RESEARCH DIRECTIONS

Recently, my research has sought to extend these ideas, with a focus on applications and real-world considerations. One such extension concerns generalizing to endogenous (decision-dependent) systems. Unlike the exogenous setting, the endogenous data generation process cannot be treated as independent of the decision-making process. This coupling of the two procedures compounds the difficulty of both estimation and optimization, motivating the development of robust policies that can adapt to decision-dependent observations. Another natural extension is to consider the exploration vs. exploitation trade-off that plays a key role in more-general online learning and multi-armed bandits theory. By viewing the sequential optimization problem as an approximate dynamic programming problem, it becomes possible to quantify the sub-optimality gap between simple greedy policies and alternative/heuristic policies that actively explore the decision space. My current work develops methods that optimally address these trade-offs in systems with multiple conflicting sources of uncertainty—specifically, queueing systems with multiple forms of customer impatience.

REFERENCES

He, L., and E. Song. 2024. "Introductory Tutorial: Simulation Optimization Under Input Uncertainty". In *2024 Winter Simulation Conference (WSC)*, 1338–1352. IEEE.

Lambert, R. J., J. A. Grant, R. Shone, and R. Szechtman. 2025. "Optimization of Queueing Systems Using Streaming Simulation". In *Proceedings of 2025 Winter Simulation Conference (WSC)*, To Appear. IEEE.