Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

SEQUENTIAL SIMULATION OPTIMIZATION WITH CENSORING: AN APPLICATION TO BIKE SHARING SYSTEMS

Cedric Gibbons

Directorate of Programs, Research and Development
Armada de Chile
Gral. Del Canto 398
Valparaíso, CHILE

James A. Grant

Department of Mathematics and Statistics
Lancaster University
Lancaster, LA1 4YW, UK

Roberto Szechtman

Operations Research
Naval Postgraduate School
1411 Cunningham Road
Monterey, CA 93943, USA

ABSTRACT

Sequential Simulation Optimization is an online optimization framework where an operator iterates periodically between collecting data from a real-world system, using stochastic simulation to approximate the optimal values of some operational variables, and setting some choice of variables in the system for the next period. The aim is to converge to an optimum efficiently, as uncertainty due to finite data and finitely many simulations eventually reduces. Using Bike Sharing Systems (BSS) as a motivating example, we analyze a variant where data from the real-world system is subject to censoring, whose nature depends on the system variables selected by the operator. In the BSS setting, censoring is of customer demand, to pick up or drop off bikes. We show that a method built upon Sample Average Approximation attains asymptotically vanishing error in its parameter estimates and specification of the optimal operational variables.

1 INTRODUCTION

Simulation Optimization is classically conceived as an offline methodology, wherein a finite volume of data is collected, and used to estimate the input processes of a simulation system aiming to identify an optimal action. Recent work (Wu et al. 2022; Song and Shanbhag 2019) has explored practical settings where input data arrives sequentially. Such frameworks are relevant in modern, data-driven applications of simulation where data may be collected alongside the simulation optimization process, and where decisions in the real-world may be taken and updated as the simulation evolves.

We consider such an example in the context of *Bike-Sharing Systems* (BSS): popular initiatives in urban areas where customers can collect a bike from a rack, use it to travel around a city and return it later, potentially to an alternative rack located elsewhere, to be used by other customers. We consider the important challenge of stock management, a critical concern for operators of BSSs: if racks empty during periods of high demand for collections, or reach capacity while there high demand for returns, it will be impossible to service the demand, and customers will be dissatisfied. Choosing optimal stock levels (i.e. the number of bikes on a rack) at the start of the day can mitigate some of this dissatisfaction. Placing fewer

979-8-3503-6966-3/23/$31.00 ©2023 IEEE

bikes on a rack at the start of the day reduces the chance of it reaching capacity later, and starting the day with more reduces the risk of it becoming empty. In particular, we consider how the operator *learns* an optimal initial daily stock-level through sequential simulation optimization (using sample average approximation), adapting as observations of input processes gradually arrive. This approach is applicable where operators introduce a new rack in an area where demand rates are unknown, and can only be determined through experimentation and data collection.

Achieving the optimal stocking of even a single rack presents a non-trivial optimization problem, particularly without precise knowledge of the rates of demand for collections and returns. Following Raviv and Kolka (2013) and as described fully in Section 3, we model a single BSS rack as a Markovian queue-like system, with demands for collection and return arising according to separate non-homogeneous Poisson processes, analogous to services and arrivals in the queueing context. We approach the challenge of determining (through sequential experimentation and data collection) an optimal number of bikes to initialise the system with, with a view to minimising the number of dissatisfied customers (those who wish to collect but find an empty rack, or wish to return but find a full rack). The objective of this problem is intractable, even for short running times or simple, known, non-homogeneous patterns. Thus effective solution requires the assistance of stochastic approximation techniques, and consideration of the unknown rate parameters.

Our BSS setting brings two important additional challenges, which have not been considered in contemporaneous work on sequential simulation optimization: 1) that dissatisfied customers are unobserved - the system only records serviced demand (collections and returns), and not those where a customer's demand could not be satisfied and they went elsewhere. This censoring of the system in its two boundary states (empty or full), mean that the quantity the operator seeks to minimise (the number of dissatisfied customers) is not explicitly observed, and must be inferred approximately via inference on the demand rates during the times in boundary states. 2) that the decisions the BSS operator takes in pursuit of optimal deployment of resource, while the simulation proceeds, impact the distribution of the sequentially observed data (specifically changing the probability of being in a censoring state at a given time). This creates the potential for biased feedback loops and slow convergence or freezing in local optima, if not addressed appropriately. This could occur, for instance, if the demand for drop-offs on a rack is initially (and incorrectly) estimated to be very low, the operator could favour starting the day with a full rack, leading to much of the day being spent in a censored, full state. As the censoring prevents observation of unsatisfied drop-off demand, the operator will have limited chance to collect data which improves their demand rate estimate and is unknowingly stuck in the situation of not realising they are creating unsatisfied customers.

In response to these challenges we, propose a sequential method based on Sample Average Approximation. The method discards simulations between periods so as to avoid bias, and can be shown to guarantee convergence of both the parameter estimates and the decision variable to their true or optimal values. The remainder of the paper is organised as follows: we review related literature in Section 2 before defining our framework, key assumptions, and solution method in Section 3. Our theoretical analysis is in Section 4, and a numerical study and conclusions follow in Sections 5 and 6 respectively.

2 RELATED LITERATURE

Raviv and Kolka (2013) and Freund et al. (2019) introduce and employ the same Markovian model of the BSS process we utilise. Herein, the decision-maker seeks to minimise the expected number of unsatisfied customers as a function of the number of bikes on the rack at the start of a period. These works show that in the setting where the demand rates are known and piece-wise constant, the optimal solution can be well-approximated by stochastic simulation. In particular, Raviv and Kolka (2013) also prove that the objective is (discretely) convex, which is fundamental to our theoretical contributions. Schuijbroek et al. (2017) extend to consider a system-wide measure of performance, encompassing multiple racks, using heuristic optimisation to tackle a more challenging inventory management task.

Alternative models are considered by George and Xia (2011) and Leurent (2012). The former cast the BSS as a closed queuing network and characterise its asymptotic behaviour. The latter models a single station, and considers a situation without balking, and with homogenoeous service rates, where customers queue until their demand is serviced.

The aforementioned approaches are offline, in the sense that they do not seek to integrate streaming data as it becomes available. The more general simulation optimisation literature has begun to consider the streaming setting.

Wu et al. (2022), address the interplay of simulation uncertainty and input model uncertainty in the Ranking and Selection setting. They design variants of Sequential Elimination algorithms which fixed confidence guarantees, in a setting where one wishes to learn an optimal action within fixed computational budget, which can be spent on data collection or simulating from the system given particular parameters. Wu and Zhou (2017) and Wang and Zhou (2022) study a similar problem where the allocation of budget to data collection and simulation must be declared a priori. Song and Shanbhag (2019) study a Multi-stage Simulation Optimization setting closer to ours where data is collected at regular intervals, independent of operational decisions. They consider an optimization problem over a compact, convex decision space and focus on the design of gradient based optimizers which achieve optimal scaling of the cumulative regret.

At the methodological level, our work is similar to Wu et al. (2022) in that it considers optimisation over a discrete set, and Song and Shanbhag (2019) in that it is focused on minimization of the cumulative regret when data arises in a streaming fashion. A notable difference from both frameworks is that in ours, the actions chosen in the optimization step affect the distribution of the data observed from the real-world process. This presents a greater challenge in the theoretical analysis of the system, since data collected at different epochs cannot be considered independent. We overcome this via careful analysis recognising the cumulative error as a martingale difference sum.

3 MODEL

Consider customers who arrive to a bike rack of capacity B to drop off or pick up bikes, where the drop offs (or arrivals) and pick ups (or departures) are Poisson processes. Customers who attempt to pick up a bike when the rack is empty or to drop off a bike when the rack leave unsatisfied, so the operator's goal is to minimize the expected number of unsatisfied customers. We model a situation where at the end of each period (e.g., a day), the rack operator employs the arrival and departure observations to-date to restock the rack at some desired level. If the arrival and departure process parameters were known, this could be done using, for instance, sample-average approximation (SAA) methods (Shapiro et al. 2021). In reality, however, these parameters can be estimated on a daily basis, albeit in the presence of censored arrivals, when the rack is full, or censored departures, when the rack is empty.

For simplicity, we assume both arrival and departure Poisson processes are independent of each other and i.i.d. across days, with unknown rates $\lambda(h)$ for the arrivals and $\mu(h)$ for the departures, for $h \in [0,H]$, $H \in \mathbb{N}$. The Poisson process of attempted bike drop offs is $(X_\lambda(h) : h \in [0,H])$, and the one corresponding to the pickups is $(Y_\mu(h) : h \in [0,H])$. The stochastic process $(Z_{\lambda,\mu}(h) : h \in [0,H])$ measures the number of bikes in the rack at time h, and is defined in terms of X_λ and Y_μ for dh small by

$$Z_{\lambda,\mu}(h+dh) - Z_{\lambda,\mu}(h) = 1$$

if $X_\lambda(h+dh) - X_\lambda(h) = 1, Y_\mu(h+dh) - Y_\mu(h) = 0, Z_{\lambda,\mu}(h) < B$ or $X_\lambda(h+dh) - X_\lambda(h) = 1, Z_{\lambda,\mu}(h) = 0$, and

$$Z_{\lambda,\mu}(h+dh) - Z_{\lambda,\mu}(h) = -1$$

if $X_\lambda(h+dh) - X_\lambda(h) = 0, Y_\mu(h+dh) - Y_\mu(h) = 1, Z_{\lambda,\mu}(h) > 0$ or $Y_\mu(h+dh) - Y_\mu(h) = 1, Z_{\lambda,\mu}(h) = B$. The initial number of bikes is $Z_{\lambda,\mu}(0) = b$ w.p. 1.

It follows that the (random) number of unsatisfied customers over a period when the initial number of bikes in the rack is b is

$$f(b;\lambda,\mu) := \int_0^H \mathbb{I}(Z_{\lambda,\mu}(h)=0)dY_\mu(h) + \mathbb{I}(Z_{\lambda,\mu}(h)=B)dX_\lambda(h), \tag{1}$$

with expectation (by Fubini's theorem)

$$F(b;\lambda,\mu) := \int_0^H (\mathbb{P}(Z_{\lambda,\mu}(h)=0)\mu(h) + \mathbb{P}(Z_{\lambda,\mu}(h)=B)\lambda(h))dh.$$

As mentioned earlier, if the rates λ and μ were known, the rack operator would minimize F over the possible values of $b \in [B] := \{0,1,\ldots,B\}$, and obtain an optimal daily restock level $b^* \in \arg\min_{b\in[B]} F(b;\lambda,\mu)$. Since F is convex when b is viewed as a continuous variable (Raviv and Kolka 2013), there exist at most two values of b that minimize F over $[B]$.

In this work we assume the rates are initially unknown, but can be estimated. By day t, there are rate estimates λ_t and μ_t (whose exact form is specified later), leading to the approximate problem

$$\min_{b\in[B]} F(b;\lambda_t,\mu_t),$$

with (possibly non-unique) solution b_t. The function F is not known but can be estimated by Monte Carlo sampling from the estimated arrival and departure processes, leading to the estimator

$$\bar{F}_{n_t,t}(b;\lambda_t,\mu_t) := \frac{1}{n_t}\sum_{i=1}^{n_t} f(b;X_{\lambda_t,i},Y_{\mu_t,i}), \tag{2}$$

where $(X_{\lambda_t,i},Y_{\mu_t,i})$, $i=1,\ldots,n_t$, are i.i.d. copies drawn from the distribution of $(X_{\lambda_t},Y_{\mu_t})$ (which is available to the analyst), and the sample size is n_t (which may depend on t).

In summary, the operator starts day j with b_j bikes and observes censored versions of (X_λ,Y_μ). These censored observations, along with observations collected in earlier days, are used to form estimators λ_j and μ_j. The operator then uses SAA to determine the value of b_{j+1}, and the process repeats itself in day $j+1$. This iterative process is visualised in Figure 1. The measure of performance to minimize is the expected regret, and the decisions taken by the operator each day are two: Determine the number initial number of bikes, b_t, and the number of samples used for estimation, n_t.

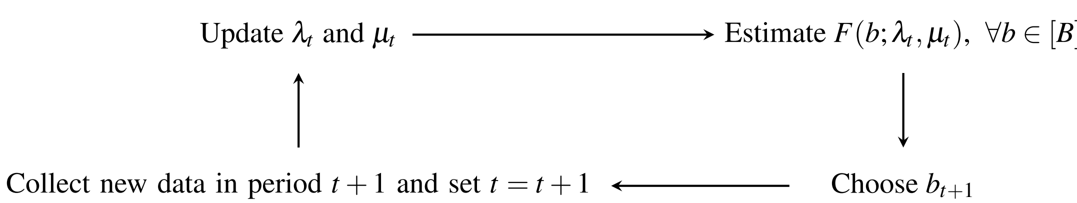


Figure 1: Overview of main steps.

To simplify the presentation, in this paper we focus our analysis on a setting where both Poisson process rate functions, λ and μ, are piecewise constant over a discretization of $[0,H]$, expressed through the key assumption below.

Assumption A1 (Piecewise constant rates): The arrival and departure rate functions are piecewise constant, with

$$\lambda(s) = \lambda_h \le \lambda_{max}, \quad \mu(s) = \mu_h \le \mu_{max},$$

for all $s \in [h-1,h)$, and each $h \in [H]$, where λ_{max}, μ_{max} are finite, known positive constants. Also, both arrival and departure rates are strictly positive for all $h \in [H]$.

Assumption A1 is commonly employed by operators, and has the benefit of restricting the estimation problem to finite dimensions, so that the ensuing analysis is simplified. Under Assumption A1, the expected number of unsatisfied customers over a period is,

$$F(b;\lambda,\mu) = \sum_{h=1}^{H} \left(\mu_h \int_h \mathbb{P}(Z_{\lambda,\mu}(u) = 0) du + \lambda_h \int_h \mathbb{P}(Z_{\lambda,\mu}(u) = B) du \right). \tag{3}$$

The smoothness of F is guaranteed by Assumption A1, as stated in the next Lemma.

Lemma 1 (Lipschitz Objective) Under Assumption A1, F is Lipschitz continuous in all rate parameters λ_h, μ_h, $h \in [H]$. Specifically, there exist finite coefficients $L_{\lambda_1}, \ldots L_{\lambda_H}, L_{\mu_1}, \ldots, L_{\mu_H}$ such that

$$\frac{\partial F}{\partial \lambda_h} \le L_{\lambda_h} \le 1 + \sum_{i=h}^{H} (\mu_i + \lambda_i), \quad \frac{\partial F}{\partial \mu_h} \le L_{\mu_h} \le 1 + \sum_{i=h}^{H} (\mu_i + \lambda_i),$$

for $h \in [H]$.

The full proof of Lemma 1 is omitted here for conciseness. It exploits Assumption A1 to relate F to various Poisson probabilities and expectations, which are continuous functions of the positive rates.

The main steps of the proposed approach are shown in Algorithm 1, which consists of two nested iterations – the outer one over the physical realizations (and thus unavoidable) and the inner one over the simulation replications. The rate functions are estimated in Step 4. Each simulation replication produces a number of unsatisfied customers for each possible initial number of bikes. Finally, in Step 10 the initial number of bikes in period $j+1$, b_{j+1}, is chosen greedily with respect to the estimator of F. Other possible selection rules of b_{j+1} are studied in Grant and Szechtman (2023).

We conclude this section with some notation and definitions. When the context is clear, we write $F(b)$, $F_t(b)$, and $\bar{F}_t(b)$ in lieu of $F(b;\lambda,\mu)$, $F(b;\lambda_t,\mu_t)$, and $\bar{F}_{n_t,t}(b;\lambda_t,\mu_t)$, respectively. Likewise, we often write Z in lieu of $Z_{\lambda,\mu}$. The index i is used to iterate over the simulation replications $1,\ldots,n_t$, and the index j indicates the days $1,\ldots,t$. In this way, $Z_{\lambda,\mu,j}(\cdot)$ is the process for the number of bikes on day j, which depends on the initial number of bikes b_j. The number of detected arrivals (departures) during period h on day j is $C_{h,j}$ (respectively $D_{h,j}$), and $\mathscr{F}_j$ is σ-algebra generated by the history of the bike level process $\{Z_1,\ldots,Z_j\}$ up to day j.

4 THEORETICAL ANALYSIS

This section analyses the properties of Algorithm 1 and is split in three parts, we first motivate the rate estimators and give consistency results for the rates, objective function, and optimal action. The second section goes further to provide central limit theorems for estimators of the rates and objectives, and the third applies these results to give a guarantee on the asymptotic regret of Algorithm 1.

4.1 Rates Estimation

In this section we consider the problem of estimating the Poisson rates over each $h \in [H]$. We present the analysis just for the drop-offs rate estimator since the one corresponding to the pick-ups rate estimator is analogous.

Assumption A1 ensures that $\int_h P(Z(u) < B) du > 0$ for all $h \in [H]$. Proposition 4.10.1 of Resnick (1992) implies that the number of detected arrivals, $C_{h,j}$, has distribution

$$C_{h,j} | \mathscr{F}_{j-1} \sim \text{Poisson}\left(\lambda_h \int_h P(Z_j(u) < B) du \right), \tag{4}$$

Algorithm 1

Initialization: Set $b_1 = \lfloor B/2 \rfloor$.
for $j = 2$ **to** t **do**
  For each $h \in [H]$, collect $(C_{h,j}, D_{h,j})$, corresponding to the drop-offs and pick-ups, respectively.
  Update $\lambda_{h,j}$ and $\mu_{h,j}$ as

$$\lambda_{h,j} = \frac{\sum_{j=1}^{t} C_{h,j}}{\sum_{j=1}^{t} \int_h I(Z_j(u) < B) du}, \text{ and } \mu_{h,j} = \frac{\sum_{j=1}^{t} D_{h,j}}{\sum_{j=1}^{t} \int_h I(Z_j(u) > 0) du}.$$

  for $i = 1$ **to** n_j **do**
    For each $h \in [H]$, draw homogeneous Poisson processes $(X_{i,j}, Y_{i,j})$ independent of each other with rates $(\lambda_{h,j}, \mu_{h,j})$, corresponding to the drop-offs and pick-ups, respectively.
    For each $b \in [B]$, obtain the number of unhappy customers (1)

$$f_{i,j}(b) := \int_0^H \mathbb{I}(Z_{i,j}(h) = 0) dY_{i,j}(h) + \mathbb{I}(Z_{i,j}(h) = B) dX_{i,j}(h),$$

  end for
  Produce the estimator $\bar{F}_j(b)$ as defined in (2).
  Set $b_{j+1} \in \arg\min \bar{F}_j(b)$.
end for

for $j = 1, \ldots, t$. Since $E[C_{h,j} - \lambda_{h,j} \int_h P(Z_j(u) < B) du | \mathscr{F}_{j-1}] = 0$ for $j > 1$, it follows that

$$\sum_{j=1}^{t} \left(C_{h,j} - \lambda_h \int_h P(Z_j(u) < B) du \right)$$

is a martingale difference sum, with

$$E\left[\left(C_{h,j} - \lambda_h \int_h P(Z_j(u) < B) du \right)^2 | \mathscr{F}_{j-1} \right] = \lambda_h \int_h P(Z_j(u) < B) du, \text{ a.s.} \tag{5}$$

Theorem 2.15 of Hall and Heyde (2014) then implies,

$$\frac{1}{t} \sum_{j=1}^{t} \left(C_{h,j} - \lambda_h \int_h P(Z_j(u) < B) du \right) \to 0, \text{ a.s.} \tag{6}$$

as $t \to \infty$. Likewise, using the same result leads to

$$\frac{1}{t} \sum_{j=1}^{t} \int_h (I(Z_j(u) < B) - P(Z_j(u) < B)) du \to 0, \text{ a.s.} \tag{7}$$

as $t \to \infty$.

The rate estimator inflates the cumulative number of drop-offs in period h by the total amount of time uncensored for drop-offs,

$$\lambda_{h,t} = \frac{\sum_{j=1}^{t} C_{h,j}}{\sum_{j=1}^{t} \int_h I(Z_j(u) < B) du}, \tag{8}$$

where $0/0 = 0$.

Lemma 2 (Consistency of Rate Estimators) Under Assumption A1,

$$\lambda_{h,t} \to \lambda_h, \tag{9}$$

a.s. as $t \to \infty$.

Proof: Using (6) – (7), the fact that $\sum_{j=1}^{t} \int_h I(Z_j(u) < B)du > 0$ a.s. as $t \to \infty$, and continuous mapping yield (9). ■

Lemma 3 (Consistency of Objective Estimators) Under Algorithm 1, if $n_t \to \infty$ then $\bar{F}_t(b) \to F(b)$ a.s. as $t \to \infty$ for each $b \in [B]$.

Proof: Write,

$$|\bar{F}_t(b) - F(b)| \leq |F_t(b) - F(b)| + |\bar{F}_t(b) - F_t(b)|. \tag{10}$$

Since F is continuous in λ and μ by Lemma 1, continuous mapping and Lemma 2 lead to $F_t(b) \to F(b)$ a.s. as $t \to \infty$ for each $b \in [B]$.

Regarding the second term in the right-hand side of (10), remark that

$$\sum_{i=1}^{n_t} (f_{t,i}(b) - F_t(b))$$

is a square integrable martingale difference sum. Therefore, for each $b \in [B]$, Theorem 2.15 of Hall and Heyde (2014) results in $|\bar{F}_t(b) - F_t(b)| \to 0$ a.s. as $t \to \infty$. The claim now follows from (10). ■

Corollary 1 (Consistency of Action Selection) $b_t \to b^*$ a.s. as $n_t \to \infty$.

Proof: By Lemma 3, for $\varepsilon = \min_{b \neq b^*}\{F(b) - F(b^*)\}$, there exists $\ell(\varepsilon)$ such that for all $t \geq \ell(\varepsilon)$, $\bar{F}_t(b) - \bar{F}_t(b^*) \geq \varepsilon/2$ w.p.1 for each $b \neq b^*$. Step 10 of Algorithm 1 then chooses $b_t = b^*$ for all $t \geq \ell(\varepsilon)$, also w.p.1. ■

4.2 Central Limit Theorems

To get a CLT for $\lambda_{h,t}$, arbitrary oscillations of the sequence $(b_t)_{t \in N}$ should not occur. More precisely, $t^{-1} \sum_{j=1}^{t} \int_h P(Z_j(u) < B)du$ needs to converge a.s.

Theorem 2 (CLT for Rate Estimators) Under Assumption A1,

$$\sqrt{t}\lambda_{h,t} \Rightarrow \lambda_h + \left(\frac{\lambda_h}{\int_h P(Z^*(u) < B)du} + \left(\sigma_{I_h} \frac{\lambda_h}{\int_h P(Z^*(u) < B)du} \right)^2 \right)^{1/2} N(0,1).$$

as $t \to \infty$, where $\sigma_{I_h}^2 = \mathrm{var}(\int_h I(Z^*(u) < B)du)$.

Proof sketch: Corollary 1 and continuous mapping lead to,

$$t^{-1} \sum_{j=1}^{t} \int_h P(Z_j(u) < B)du \to \int_h P(Z^*(u) < B)du, \tag{11}$$

a.s., where Z^* is the rack level process with b^* initial bikes. A similar argument leads to,

$$t^{-1} \sum_{j=1}^{t} \int_h I(Z_j(u) < B)du \to \int_h P(Z^*(u) < B)du, \tag{12}$$

a.s.

Define the averages,

$$\bar{C}_{h,t} = t^{-1}\sum_{j=1}^{t} C_{h,j}, \text{ and } \int \bar{I}(Z_j(u) < B)du = t^{-1}\sum_{j=1}^{t}\int I(Z_j(u) < B)du.$$

Eqs. (11)–(12) and p. 58 of Hall and Heyde (2014) result in

$$\sqrt{t}(\bar{C}_{h,t}, \int_h \bar{I}(Z_t(u) < B)du) \Rightarrow N\left((\lambda_h \int_h P(Z^*(u) < B)du, \int_h P(Z^*(u) < B)du), \Sigma\right),$$

where Σ is the covariance matrix. The two variables $(C_{h,t}, \int_h I(Z_t(u) < B)du)$ are independent due to the Poisson assumption, so that Σ is the matrix with main diagonal $(\sigma^2_{C_h}, \sigma^2_{I_h})$, where

$$\sigma^2_{C_h} = \lambda_h \int_h P(Z^*(u) < B)du, \text{ and } \sigma^2_{I_h} = \text{var}\left(\int_h I(Z^*(u) < B)du\right).$$

The key is to use the multivariate Delta method (see Serfling (2009)), whereby for a function g with continuous first partial derivatives,

$$\sqrt{t}g(\bar{C}_{h,t}, \int_h \bar{I}(Z_t(u) < B)du) \Rightarrow N\left(g(\lambda_h \int_h P(Z^*(u) < B)du, \int_h P(Z^*(u) < B)du), \nabla_g^T \Sigma \nabla_g\right), \tag{13}$$

where ∇_g is gradient of g evaluated at $(\lambda_h \int_h P(Z^*(u) < B)du, \int_h P(Z^*(u) < B)du)$. In our case

$$g\left(\bar{C}_{h,t}, \int_h \bar{I}(Z_t(u) < B)du\right) = \frac{\bar{C}_{h,t}}{\int_h \bar{I}(Z_t(u) < B)du} = \lambda_{h,t}, \tag{14}$$

and

$$\nabla_g^T = \left(\frac{1}{\int_h P(Z^*(u) < B)du}, -\frac{\lambda_h}{\int_h P(Z^*(u) < B)du}\right). \tag{15}$$

The result now follows from Eqs. (13)–(15). ■

The CLT for the rate estimator leads to an analogue result for $\bar{F}_t(b)$. The main result is stated next.

Theorem 3 (CLT for Objective Function) Under Assumption A1,

$$\sqrt{n_t}(\bar{F}_t(b) - F(b)) \Rightarrow \sigma(b)N(0,1),$$

as $t \to \infty$, for $\sigma(b) > 0$.

The proof will be presented in Grant and Szechtman (2023).

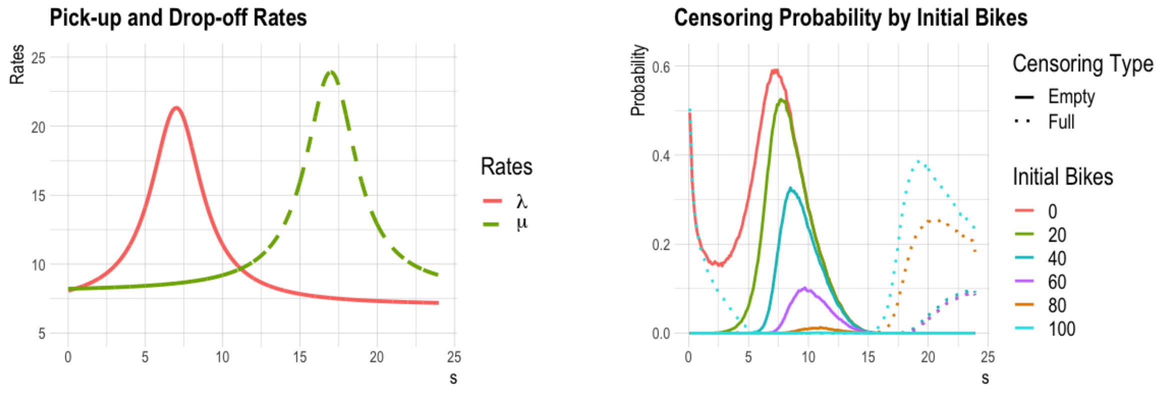
4.3 Regret Analysis

Combining the results on the concentration of the parameter estimates and sample average approximation leads us to the theorem below, characterising the asymptotic regret of Algorithm 1. We defer its proof to the Appendix.

Theorem 4 (Asymptotic Regret) Suppose Assumption A1 holds. The per-period regret of the actions, b_t, selected by UncenSAA, with $n_t = \Omega(t)$, satisfies

$$\mathbb{E}\left[f(b_t;\lambda,\mu) - f(b^*;\lambda,\mu)\right] = O\left(t^{-1/2}\right), \quad \text{as} \quad t \to \infty.$$

We note that while Theorem 2 holds for all simulation budget sequences $n_t = \Omega(t)$, the final stage of the proof of the Theorem illustrates why $n_t = \Theta(t)$ is the more practical choice. If n_t is chosen as a super-linear sequence, the input uncertainty term will dominate the regret regardless, and no further improvement to the *rate* of convergence can be anticipated.



(a) Arrival and Departure Rate Functions (b) Probability of Censoring

Figure 2: Underlying True Functions for the Numerical Study.

5 NUMERICAL STUDY

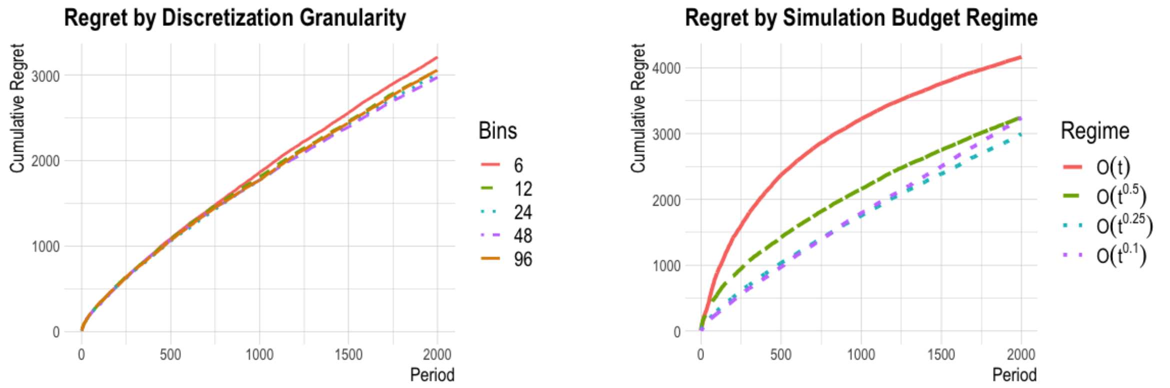
This section provides an empirical assessment of Algorithm 1. With inspiration from Raviv and Kolka (2013), we experiment in a setting with rack size $B = 100$, period length $H = 24$, a fixed horizon of $T = 2000$ periods, and continuous rate functions, $\lambda(s) = 7 + 180/(\pi(s-7)^2 + 4\pi)$, and $\mu(s) = 8 + 200/(\pi(s-17)^2 + 4\pi)$ for $s \in [0, 24]$ as plotted in Figure 2a. These unimodal rate functions yield a convex (in a discrete sense) objective function F with $\arg\min_b F(b) = 62$, and $\min_b F(b) \approx 20$. Thus, even with an optimal initialization, this problem is such that some censoring of demand is inevitable. Figure 2b plots the probability of censoring (due to the empty or full racks separately) against in-period time for various choices of initialization, $b \in \{0, 20, 40, 60, 80, 100\}$. This shows that while the problem is constructed so that some censoring is unavoidable in expectation, where, in the period, the censoring is likely to occur can vary substantially with b, indicating that we can expect parameter estimates to concentrate at quite different rates.

Both functions are unimodal, and their peaks are such that some censoring of both pick-ups and drop-offs is likely within a single period for most choices of b, but the objective function can vary substantially.

We apply Algorithm 1 to make decisions as to the initial number of bikes over $T = 2000$ periods, and present results on the cumulative regret, $\sum_{t=1}^{T} F(b_t) - F(b^*)$, averaged over 20 replications. Our numerical study focuses on the sensitivity to two algorithmic choices: the granularity of the discrete approximation to the rate functions, and the simulation budget.

In Figure 3a we show the effect of the granularity of the discrete approximation: using 6, 12, 14, 48, and 96 bins. A sublinear growth is seen in each case, but the final regret decreases as bin number grows from 6 to 48. However, the largest number of bins, 98, is associated with a larger regret, suggesting there comes a point where the problem is over-parameterized: i.e. the capacity for finer-grained approximation is not of sufficient benefit to counteract the reduction in inference quality.

In Figure 3b we see the results of experimenting with different regimes for allocating a fixed simulation budget over many periods, using bins of width 1, and a total of 20000 SAA simulations throughout. We compare the linear spread ($O(t)$) recommended by the theory to various more slowly increasing functions which allocate more simulation weight to earlier periods, via functions of order $t^{0.1}$, $t^{0.25}$, and $t^{0.5}$. The results suggest that the linear spread is suboptimal for finite horizons, as it incurs a higher regret than all other options, which yield broadly similar performance, with $t^{0.25}$ having the best among those trialed.



(a) Effect of Discretization Granularity on Regret (b) Effect of Simulation Budget Regime on Regret

Figure 3: Sensitivity Analyses of Regret.

6 CONCLUSIONS

Sequential Simulation Optimization has the capacity to allow more efficient, targeted data collection, and faster convergence towards optimal system performance compared to offline simulation optimization. This work has presented a first foray into the challenging but pertinent setting where decisions enacted during the sequential iteration between simulation and data collection directly impact the distribution of the observed data. Using the Bike Sharing System as a central example, we have shown that despite the interplay between censoring and decision-making, straightforward SAA-based algorithms can see their regret vanish asymptotically, with a reasonable level of robustness to key model and budget parameters.

Our work raises several questions and inspires directions for future research. Algorithm 1 does not explicitly consider the trade-off between the long-term benefits of minimising censoring and the short-term benefits of minimising expected regret. It is of interest to determine whether algorithms that do so, e.g. lower confidence bound approaches, outperform Algorithm 1. Furthermore, Algorithm 1 does not re-use any simulation samples across periods. This is a useful construction for the theoretical analysis and weakens the dependence across periods, but may be sub-optimal in the later stages when input parameters show little variation from period to period. Finally, we have centred our work on the BSS example with a single rack, but the challenges of Sequential Simulation Optimization in the face of censoring may be encountered in larger BSS networks and in many other applications. There is scope to broaden the analysis and methodology to a wider range of Simulation Optimization problems. The forthcoming work Grant and Szechtman (2023) aims to extend the contributions of the present manuscript to consider these challenges.

A APPENDIX

Proof of Theorem 4 The expected regret in period t may be bounded with input and simulation uncertainty terms, as follows,

$$
\begin{aligned}
\mathbb{E}[f(b_t;\lambda,\mu)-f(b^*;\lambda,\mu)] &= F(b_t;\lambda,\mu)-\bar{F}_t(b_t;\lambda_t,\mu_t)+\bar{F}_t(b_t;\lambda_t,\mu_t)-F(b^*;\lambda,\mu) \\
&\leq F(b_t;\lambda,\mu)-\bar{F}_t(b_t;\lambda_t,\mu_t)+\bar{F}_t(b^*;\lambda_t,\mu_t)-F(b^*;\lambda,\mu) \\
&= [F(b_t;\lambda,\mu)-F(b_t;\lambda_t,\mu_t)]+[F(b^*;\lambda_t,\mu_t)-F(b^*;\lambda,\mu)] \\
&\quad +[F(b_t;\lambda_t,\mu_t)-\bar{F}(b_t;\lambda_t,\mu_t)]+[\bar{F}(b^*;\lambda_t,\mu_t)-F(b^*;\lambda_t,\mu_t)]\,. \qquad (16)
\end{aligned}
$$

The inequality uses the fact that b_t minimises $\bar{F}_t$, so $\bar{F}_t(b^*;\lambda_t,\mu_t) \geq \bar{F}_t(b_t;\lambda_t,\mu_t)$. At (16), the first line's terms are driven by the uncertainty in the inputs, and the second's by uncertainty due to the SAA. The concentration of the simulation uncertainty is characterised by Theorem 3. For the input uncertainty, the following Corollary of Theorem 2 extends the CLT on the rate estimators to a CLT on the function $F(;\lambda_t,\mu_t)$, thus characterising the concentration of the input uncertainty.

Corollary 5 Suppose Assumption A1 holds. For each $b \in \{0,1,...,B\}$, and as $t \to \infty$, the expected input uncertainty $|F(b;\lambda_t,\mu_t) - F(b;\lambda,\mu)|$ is bounded by a random variable whose distribution approaches

$$t^{-1/2}\sum_{h=1}^{H}\left(\sqrt{\frac{L^2_{\lambda_h}\lambda_h}{\int_h P(Z^*(u)<B)du}}+\sqrt{\frac{L^2_{\mu_h}\mu_h}{\int_h P(Z^*(u)>0)du}}\right)N(0,1).$$

Corollary 5's proof uses the Lipschitz continuity of Lemma 1 to decompose $|F(b;\lambda_t,\mu_t) - F(b;\lambda,\mu)|$ in terms of biases in individual rates, before replacing those with their asymptotic distributions.

Combining Corollary 5, Theorem 3, and Equation (16), we have that the per-period regret, $F(b_t;\lambda,\mu) - F(b^*;\lambda,\mu)$ is bounded by a random variable whose distribution approaches that of

$$\left|2t^{-1/2}\sum_{h=1}^{H}\left(\sqrt{\frac{L^2_{\lambda_h}\lambda_h}{\int_h P(Z^*(u)<B)du}}+\sqrt{\frac{L^2_{\mu_h}\mu_h}{\int_h P(Z^*(u)>0)du}}\right)N(0,1)+2n_t^{-1/2}\max(\sigma(b_t),\sigma(b^*))N(0,1)\right|$$

as $t \to \infty$. Thus, if $n_t = \Omega(t)$, we have $F(b_t;\lambda,\mu) - F(b^*;\lambda,\mu) = o(t^{-1/2})$. ■

REFERENCES

Freund, D., S. G. Henderson, E. OMahony, and D. B. Shmoys. 2019. "Analytics and Bikes: Riding Tandem with Motivate to Improve Mobility". *INFORMS Journal on Applied Analytics* 49(5):310–323.

George, D. K., and C. H. Xia. 2011. "Fleet-Sizing and Service Availability for a Vehicle rental System via Closed Queueing Networks". *European Journal of Operational Research* 211(1):198–207.

Grant, J. A., and R. Szechtman. 2023. "Sequential Simulation Optimization of Bike Sharing Systems with Censored Data – in progress".

Hall, P., and C. C. Heyde. 2014. *Martingale Limit Theory and its Application*. Academic press.

Leurent, F. 2012. "Modelling a Vehicle-Sharing Station as a Dual Waiting System: Stochastic Framework and Stationary Analysis". *Working Paper. LVMT - Laboratoire Ville, Mobilité, Transport*.

Raviv, T., and O. Kolka. 2013, 10. "Optimal Inventory Management of a Bike-Sharing Station". *IIE Transactions* 45:1077–1093.

Resnick, S. I. 1992. *Adventures in Stochastic Processes*. Boston: Birkhäuser.

Schuijbroek, J., R. Hampshire, and W.-J. van Hoeve. 2017. "Inventory Rebalancing and Vehicle Routing in Bike Sharing Systems". *European Journal of Operational Research* 257(3):992–1004.

Serfling, R. J. 2009. *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons.

Shapiro, A., D. Dentcheva, and A. Ruszczynski. 2021. *Lectures on Stochastic Programming: Modeling and Theory*. SIAM.

Song, E., and V. Shanbhag, Uday. 2019. "Stochastic Approximation for Simulation Optimization under Input Uncertainty with Streaming Data". In *Proceedings of the 2019 Winter Simulation Conference*, edited by N. Mustafee, K.-H. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, 3597–3608. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Wang, Y., and E. Zhou. 2022. "Fixed Budget Ranking and Selection with Streaming Input Data". In *Proceedings of the 2022 Winter Simulation Conference*, edited by B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C. Corlu, L. Lee, E. Chew, T. Roeder, and P. Lendermann, 3027–3038. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Wu, D., Y. Wang, and E. Zhou. 2022. "Data-driven R&S Under Input Uncertainty". *Operations Research* In Press.

Wu, D., and E. Zhou. 2017. "Ranking and Selection Under Input Uncertainty: A Budget Allocation Formulation". In *Proceedings of the 2017 Winter Simulation Conference*, edited by W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, 2245–2256. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

CEDRIC GIBBONS is an Officer in the Chilean Navy. He received an M.Sc. degree in Operations Research from the Naval Postgraduate School in 2021, and won the Prof. Donald P. Gaver, Jr. thesis research award in the same year. His email is cedric.gibbons@gmail.com.

JAMES A GRANT is a Lecturer (equiv. Assistant Professor) in Statistics in the Department of Mathematics and Statistics at Lancaster University, UK. His primary research interests are in online optimisation and multi-armed bandits. His email is j.grant@lancaster.ac.uk and his homepage is https://www.lancaster.ac.uk/pg/grantj/.

ROBERTO SZECHTMAN received his Ph.D. from Stanford University. Currently, he is Professor in the Operations Research Department at the Naval Postgraduate School. His research interests include applied probability and military operations research. His email address is rszechtm@nps.edu.