R-SPLINE FOR LOCAL INTEGER-ORDERED SIMULATION OPTIMIZATION PROBLEMS WITH STOCHASTIC CONSTRAINTS

Kalyani Nagaraj
Raghu Pasupathy

Industrial and Systems Engineering
Virginia Tech
Blacksburg, VA 24061, USA

ABSTRACT

R-SPLINE is a recently proposed competitor to the popular COMPASS algorithm for solving local integer-ordered simulation optimization problems that have either an unconstrained or a deterministically-constrained feasible region. R-SPLINE is a refined sample-average approximation algorithm with a structure that is particularly conducive to the inclusion of stochastic constraints. In this paper we consider one such trivial adaptation of R-SPLINE. Our aim is narrow in that we wish only to investigate the asymptotic behavior of the resulting iterates. Accordingly, we demonstrate sufficient conditions under which the proposed adaptation's iterates match the consistency and convergence rate qualities of the iterates from the originally proposed R-SPLINE. Ongoing numerical experiments show much promise but raise important questions about the choice of algorithm parameters when the adaptation is executed on problems where one or more of the constraints are binding.

1 INTRODUCTION

We consider the context of simulation optimization (SO) problems over a set of discrete decision variables, where both the objective function and constraint functions can be observed implicitly via a simulation model. Such integer-ordered SO problems seem quite prevalent, as evidenced by entries in the SO library www.simopt.org (Pasupathy and Henderson 2006, Pasupathy and Henderson 2011). Despite this seeming popularity, there has been relatively little in the literature in terms of methods for solution. The main available algorithms for this context are COMPASS (Xu, Hong, and Nelson 2010, Hong and Nelson 2006), R-SPLINE (Wang, Pasupathy, and Schmeiser 2013), and more recently, Discrete Stochastic Approximation (Lim 2013). While COMPASS and R-SPLINE have been shown to perform competitively, they are explicitly designed for contexts where the feasible region is integer-ordered and *unconstrained or deterministically constrained*.

In this paper, we ask whether R-SPLINE can be adapted to contexts where the local integer-ordered SO problem includes stochastic constraints. We believe that this is a useful question to ask for two reasons: (i) R-SPLINE, being "gradient based," has the ingredients to be successful in the context of problems requiring local solutions, and is hence a natural candidate to consider for extensions to include stochastic constraints; and (ii) R-SPLINE, due to the use of a sample-path framework, lends itself to the incorporation of constraints that can be estimated. The latter point makes R-SPLINE's adaptation to include stochastic constraints almost trivial, although consequent questions about the asymptotic behavior of the resulting iterates, and the choice of algorithm parameters to guarantee good finite-time performance, are challenging.

Our goal in this paper is narrow. We consider the asymptotic behavior of the iterates that result from incorporating estimated constraints into R-SPLINE's algorithmic framework. Specifically, we wish to investigate if the iterates from the adapted version of R-SPLINE are consistent from the standpoint of

returning a locally optimal solution. For instance, we demonstrate sufficient conditions that ensure that the adapted R-SPLINE's iterates converge with probability one (wp1) to a locally optimal solution, when the feasible region of the underlying problem is bounded or unbounded. Furthermore, we demonstrate that the original R-SPLINE's results on the exponential decay rate (of selecting a suboptimal solution) transfer over to the adapted version modulo certain restrictions on sampling and structure. Of course, the results that we present, while reassuring to some degree, say little about whether the adapted version of R-SPLINE performs well in finite-time. Unfortunately, such finite-time performance is still a subject of ongoing investigation, particularly in terms of how certain algorithmic parameters should be chosen.

The work that is closest in spirit to what we propose here is Lim (2013) and Luo and Lim (2013), where the stochastic approximation (Kushner and Yin 2003) recursion is extended to work on discrete sets and in the presence of stochastic constraints. The current paper should be seen as an analogous adaptation of a sample-average approximation (Shapiro, Dentcheva, and Ruszczynski 2009) method to solve local integer-ordered SO problems with constraints. It is also worth mentioning that methods to solve more generic SO problems obviously apply to the problem class under consideration in this paper. For example, Li, Sava, and Xie (2009) and Park and Kim (2011) consider SO problems having stochastic constraints on discrete sets. Likewise, the entire recent literature on ranking and selection in the presence of stochastic constraints (Andradóttir and Kim 2010, Andradóttir, Goldsman, and Kim 2005, Batur and Kim 2005, Hunter and Pasupathy 2013, Hunter, Pujowidianto, Lee, Chen, and Pasupathy 2011) applies to the class of integer-ordered SO problems as long as the feasible set is finite. However, we believe that such methods are disadvantaged because they do not have access to the immense structure that is usually present in integer-ordered SO problems. Furthermore, such generic methods tend to seek globally optimal solutions due to the lack of an assumed neighborhood structure.

In what follows, we present a formal problem statement and detail a simple adaptation of the R-SPLINE method to include stochastic constraints. This is followed by a section that lists and proves the main results characterizing the behavior of the iterates generated by the proposed adaptation. We then conclude the paper with some final remarks.

2 PROBLEM STATEMENT AND SOLUTION METHOD

Given a neighborhood definition N, our goal is to find an N-local minimum of the real-valued function $g : \mathbb{X} \to \mathbb{R}$, where $\mathbb{X} \subseteq \mathbb{Z}^d$ and the feasible region is defined in terms of inequalities involving real-valued functions $h_i : \mathbb{X} \to \mathbb{R}$, $i = 1, \cdots, c$. For each $x \in \mathbb{X}$, $g(x)$ and $h_i(x)$ are unknown functions whose point estimators are observed via a given simulation oracle. The optimization problem is stated as follows.

$$\tilde{P}_1 : \text{minimize } g(x)$$
$$\text{subject to } h_i(x) \leqslant 0, \quad i = 1, \cdots, c,$$
$$x \in \mathbb{X}.$$

The point estimators $\hat{g}_m(x)$ and $\hat{h}_{i,m}(x)$, $1 \leqslant i \leqslant c$, of the unknown functions $g(x)$ and $h_i(x)$ are "observed simultaneously" via a computer simulation, where m is some measure of simulation effort. Also, for each $x \in \mathbb{X}$ we assume that $\lim_{m \to \infty} \hat{g}_m(x) = g(x)$ wp1 and $\lim_{m \to \infty} \hat{h}_{i,m}(x) = h_i(x)$ wp1, $1 \leqslant i \leqslant c$. Let $\mathcal{F}$ denote the feasible region of $\tilde{P}_1$, that is, $\mathcal{F} = \{x \in \mathbb{X} : h_i(x) \leqslant 0, i = 1, \cdots, c\}$. Also, we define $\mathcal{F}^c = \mathbb{X} \backslash \mathcal{F}$ and the interior $\mathcal{F}^\circ = \{x \in \mathbb{X} : h_i(x) < 0, i = 1, \cdots, c\}$. Given a neighborhood definition N, let $M^*(N)$ denote the set of N-local minimizers of $g(x)$ over $\mathcal{F}$:

$$M^*(N) = \{x^* \in \mathcal{F} : g(x^*) \leqslant g(x), \forall x \in N(x^*) \cap \mathcal{F}\}.$$

Instead of $\tilde{P}_1$, consider the following problem.

$$\tilde{P}_2 : \text{minimize } \tilde{g}(x)$$
$$\text{subject to } x \in \mathbb{X}$$

where $\tilde{g} : \mathbb{X} \to \mathbb{R} \cup \{\infty\}$ is defined as

$$\tilde{g}(x) = \begin{cases} g(x) & \text{if } x \in \mathcal{F}, \\ \\ \infty & \text{otherwise.} \end{cases}$$

Problems $\tilde{P}_1$ and $\tilde{P}_2$ are not equivalent, but it is intuitively clear that they have the same solution. As stated, the problem in $\tilde{P}_2$ is unconstrained. However, the local SO algorithms COMPASS (Xu, Hong, and Nelson 2010, Hong and Nelson 2006) and R-SPLINE (Wang, Pasupathy, and Schmeiser 2013) cannot be used on $\tilde{P}_2$ directly because of the nature of the feasible region $\mathbb{X}$. Specifically, since $\mathbb{X}$ is defined through functions that can only be observed using a simulation oracle, the feasible region of the problem $\tilde{P}_2$ is not "deterministic" in the sense of knowing whether or not a point x is feasible with no uncertainty.

We propose solving $\tilde{P}_2$ (instead of $\tilde{P}_1$) using a straightforward adaptation of R-SPLINE. Like R-SPLINE, the framework we propose relies on two ideas: (i) a retrospective approximation (RA) framework (Kim, Pasupathy, and Henderson 2012, Pasupathy and Kim 2011, Pasupathy 2010) that implicitly generates a sequence of increasingly accurate sample-path problems to approximate the true problem $\tilde{P}_2$; and (ii) a local-search method called SPLINE explicitly devised to solve the generated sample-path problems through the use of "phantom gradients" on the integer-ordered set. The sample-path problem in the adapted version of R-SPLINE is formally stated as

$$P_k : \text{ Find } X_k^* \text{ such that } X_k^* \in \{x' \in \mathcal{F}_k : \hat{\tilde{g}}_{m_k}(x') \leqslant \hat{\tilde{g}}_{m_k}(x), \ \forall x \in \mathcal{F}_k \cap N(x')\},$$

where $\mathcal{F}_k = \{x \in \mathbb{X} : \hat{\tilde{g}}_{m_k}(x) < \infty\}$ and

$$\hat{\tilde{g}}_{m_k}(x) = \begin{cases} \hat{g}_{m_k}(x) & \text{if } \hat{h}_{i,m_k}(x) < \epsilon_{i,k}, \ i = 1, \cdots, c, \\ \\ \infty & \text{otherwise.} \end{cases}$$

The sample-path problem P_k is thus obtained by replacing the objective function g and the constraints functions $h_i, i = 1, 2, \ldots, c$ by their corresponding estimators $\hat{g}_{m_k}$ and $\hat{h}_{m_k}$ obtained using a sample size m_k. Importantly, the constraints appearing in the original problem $\tilde{P}_2$ are relaxed by an amount $\epsilon_{i,k} > 0$ such that $\lim_{k \to \infty} \epsilon_{i,k} = 0$, $i = 1, 2, \ldots, c$. The local-search method, SPLINE, that is mentioned in (ii) is an integer-ordered gradient-based search routine introduced in R-SPLINE. See Figure 1 for a sense of SPLINE'S functioning, and Wang, Pasupathy, and Schmeiser (2013) for complete details.

Remark 1 We note that in the above framework, SPLINE can be replaced with any locally convergent algorithm with no corresponding change in the results that we are about to present. SPLINE, however, is a tailored local search algorithm that has been shown to perform well on integer-ordered problems.

Algorithm 1 R-SPLINE

Require: initial solution $x_0 \in \mathbb{X}$; sample sizes $\{m_k\}$; limits on simulation calls $\{b_k\}$
Ensure: sequence of solutions $\{X_k^*\}$

1: Set $X_0^* = x_0$.
2: **for** $k = 1, 2, \ldots$
3:     $X_k^* = \text{SPLINE}(X_{k-1}^*, m_k, b_k)$ {obtain a solution to sample-path problem P_k}
4: **end for**

Remark 2 The advantages of the RA framework have been noted in numerous earlier papers. First, RA naturally allows the use of common random numbers (Law 2007). Specifically, during the kth iteration of the procedure, the sample-path function $\hat{\tilde{g}}_{m_k}(x)$ is generated using the *same* sample size m_k at each point
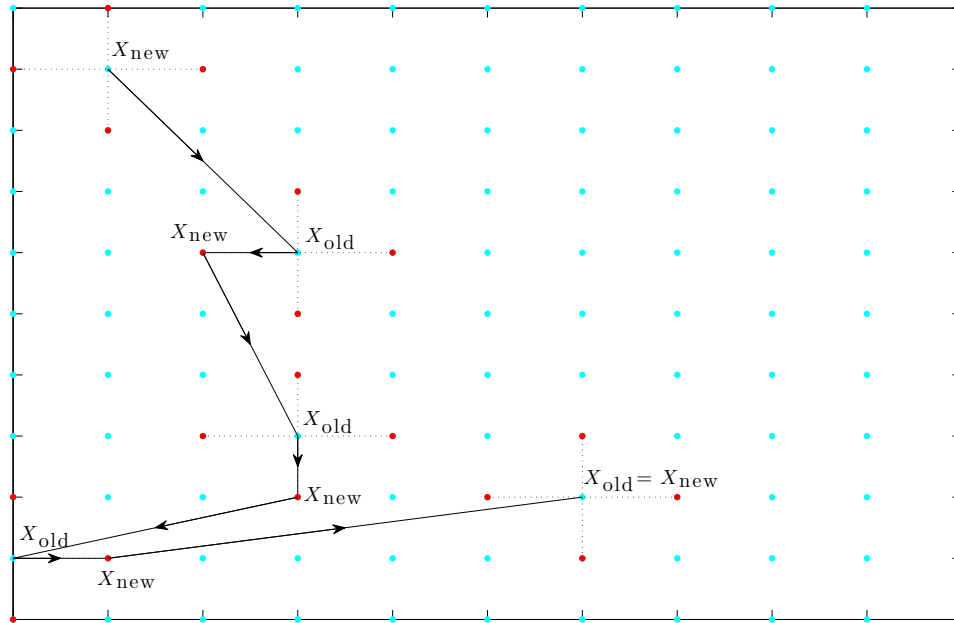
Figure 1: (Above figure adapted from Wang, Pasupathy, and Schmeiser (2013).) A depiction of the trajectory of SPLINE, the algorithm used to solve the sample-path problems generated within R-SPLINE. SPLINE consists of a gradient search routine and a neighborhood enumeration routine. The gradient search constructs "phantom gradient" directions (shown as direct lines) using an appropriate extension of the sample-path function onto continuous space. At each step, a neighborhood enumeration is performed to check if a local solution has been attained. The routine terminates if a local solution is attained; otherwise, the search restarts from the superior solution identified during the neighborhood search.

$x \in \mathbb{X}$, thereby allowing any exploitable structure in the underlying true objective function to percolate into the function $\hat{\bar{g}}_{m_k}(x)$. Second, the progressive solution framework in RA allows using the solution X^*_{k-1} from the most recent $(k-1)$th iteration as a "warm start" for solving the sample-path problem in the subsequent kth iteration.

As was noted in the introduction, the scope of this paper is narrow. Our goal is to understand the theoretical aspects of the framework that we have outlined. In what ensues, we limit ourselves to presenting results that highlight important asymptotic aspects of the sample-path problem P_k, and the sequence of solutions generated by the proposed framework. We present no numerical experience.

3   MAIN RESULTS

If X^*_k is the solution returned by SPLINE to problem P_k in the kth iteration of R-SPLINE, we would like to show that the sequence of sample path solutions $\{X^*_k\}$ converges to the set of locally optimal solutions of $\tilde{P}_1$. We emphasize that we do not need to modify the SPLINE algorithm described in Wang, Pasupathy, and Schmeiser (2013) for solving the "generated" sample-path problems P_k. This is essentially because the objective function and constraints appearing in P_k are "observable." Consequently, as in R-SPLINE, the following statements are true about the sequence of solutions $\{X^*_k\}$.

1. Since SPLINE returns a solution X^*_k in finite time, $\{X^*_k\}$ is an infinite sequence of solutions returned by the adapted version of R-SPLINE.
2. $\hat{\bar{g}}_{m_k}(X^*_k) \leqslant \hat{\bar{g}}_{m_k}(X^*_{k-1})$

Under certain assumptions on the rate of increase of the sample sizes $\{m_k\}$, and the quality of the constraint estimators $\hat{h}_{i,m_k}, i = 1, 2, \ldots, c$, the feasible and the infeasible regions for the sample-path problem P_k and the true problem $\tilde{P}_2$ coincide in a certain precise sense. We note this in Lemma 1 after stating assumptions on the sample sizes $\{m_k\}$ and the constraint estimators $\hat{h}_{i,m_k}, i = 1, 2, \ldots, c$.

Assumption 1 The sequence $\{m_k\}$ satisfies $\limsup_{k\to\infty} k/m_k = 0$.

Assumption 2 The sequence of random variables $\{h_i(x) - \hat{h}_{i,m_k}(x)\}$ is governed by a large-deviation principle with rate function $I_{i,x}(s)$, $i \in \{1, \cdots, c\}$, such that for any $\epsilon > 0$, $\inf_{x\in\mathbb{X}} \min(I_{i,x}(-\epsilon), I_{i,x}(\epsilon)) = \eta_i > 0$.

Lemma 1 Suppose Assumption 1 and Assumption 2 hold. Then $\Pr\{\mathcal{F}_k \nsubseteq \mathcal{F} \ i.o.\} = 0$ and $\Pr\{\mathcal{F} \nsubseteq \mathcal{F}_k \ i.o.\} = 0$.

Proof. Under Assumption 2 and by the Gartner-Ellis theorem (Dembo and Zeitouni 1998), for any $\epsilon > 0$ and $x \in \mathbb{X}$ there exists K (independent of x and ϵ) $\in \mathbb{N}$ such that if $k \geqslant K$ then

$$\frac{1}{m_k} \log \Pr\{|h_i(x) - \hat{h}_{i,m_k}(x)| > \epsilon\} \leqslant -\min(I_{i,x}(\epsilon), I_{i,x}(-\epsilon)) \leqslant -\eta_i < 0,$$

$i = 1, \cdots, c$. Suppose $\eta = \min_{1\leqslant i\leqslant c} \eta_i$.

Let $x \in \mathcal{F}$. Then for $\epsilon < \epsilon_k$ and $k \geqslant K'$ (independent of x and ϵ)

$$Pr\{x \in \mathcal{F}_k^c\} \leqslant Pr\{\bigcup_{i=1}^{c} \hat{h}_{i,m_k}(x) \geqslant \epsilon_k\}$$

$$\leqslant \sum_{i=1}^{c} Pr\{\hat{h}_{i,m_k}(x) \geqslant \epsilon_k\}$$

$$= \sum_{i=1}^{c} Pr\{\hat{h}_{i,m_k}(x) \geqslant \epsilon_k, \hat{h}_{i,m_k}(x) \in (h_i(x) - \epsilon, h_i(x) + \epsilon)\}$$

$$+ Pr\{\hat{h}_{i,m_k}(x) \geqslant \epsilon_k, \hat{h}_{i,m_k}(x) \notin (h_i(x) - \epsilon, h_i(x) + \epsilon)\}$$

$$\leqslant ce^{-m_k\eta}. \tag{1}$$

Since $\limsup_{k\to\infty} k/m_k = 0$, $\Pr\{x \in \mathcal{F}_k^c \ i.o.\} = 0$.

Let $y \in \mathcal{F}^c$. Suppose $h_{j,m_k}(y) > 0$ for some $j \in \{1, \cdots, c\}$. Then for $\epsilon > \epsilon_1$ and $k \geqslant K''$ (independent of y and ϵ)

$$Pr\{y \in \mathcal{F}_k\} \leqslant Pr\{\bigcap_{i=1}^{c} \hat{h}_{i,m_k}(y) < \epsilon_k\}$$

$$\leqslant Pr\{\hat{h}_{j,m_k}(y) - \epsilon_k < 0\}$$

$$= Pr\{\hat{h}_{j,m_k}(y) - \epsilon_k < 0, \hat{h}_{j,m_k}(y) \in (h_j(y) - \epsilon, h_j(y) + \epsilon)\}$$

$$+ Pr\{\hat{h}_{j,m_k}(y) - \epsilon_k < 0\, \hat{h}_{j,m_k}(y) \notin (h_j(y) - \epsilon, h_j(y) + \epsilon)\}$$

$$\leqslant Pr\{\hat{h}_{j,m_k}(y) \notin (h_j(y) - \epsilon, h_j(y) + \epsilon)\}$$

$$\leqslant e^{-m_k\eta}. \tag{2}$$

Since $\limsup_{k\to\infty} k/m_k = 0$, $\Pr\{y \in \mathcal{F}_k \ i.o.\} = 0$.

And since x and y were arbitrarily chosen from $\mathcal{F}$ and $\mathcal{F}^c$, respectively, $\mathcal{F}_k = \mathcal{F}$ wp1 for $k \geqslant K = \max(K', K'')$. □

Given Lemma 1, and assuming that the feasible region is finite, it seems intuitively clear that R-SPLINE should return a correct local solution wp1, in the limit as $k \to \infty$. This is because, every pair of points in $\mathcal{F}$

orders itself (when measured with respect to the true objective g) correctly as $k \to \infty$. However, since $\mathcal{F}$ is assumed to be finite, there are only a finite number of such pairs, yielding the correct overall ordering. This is proved rigorously in the result that follows.

Theorem 1 Suppose $\mathcal{F}$ is finite. Then under Assumptions 1 and 2 and given a neighborhood definition N, R-SPLINE returns a sequence of sample path solutions $\{X_k^*\}$ that converges wp1 to the set of true local minima $M^*(N)$ of problem $\tilde{P}_2$.

Proof. Under Assumptions 1 and 2, $\Pr\{\mathcal{F}_k \neq \mathcal{F} \ i.o.\} = 0$. Since $X_k^* \in \mathcal{F}_k$, $\Pr\{X_k^* \notin \mathcal{F} \ i.o.\} = 0$. That is, there exists $K_1 \in \mathbb{N}$ such that $X_k^* \in \mathcal{F}$ with probability 1 if $k \geqslant K_1$. $\mathcal{F}$ is bounded. Thus R-SPLINE returns a sequence of solutions that are bounded with probability 1. Also, $\Pr\{X_k^* \in \mathcal{F}^c \ i.o.\} = 0$.

Let $\epsilon = \min \{|g(x) - g(x')| : (x, x') \in \mathcal{F}, g(x) \neq g(x')\}$. Then since $\mathcal{F}$ is finite, $\hat{\bar{g}}_{m_k}$ converges uniformly to g with probability 1 on the set $\mathcal{F}$ as $k \to \infty$. Thus there exists $K_2(\epsilon) \in \mathbb{N}$ such that $|\hat{\bar{g}}_{m_k}(x) - g(x)| < \epsilon/2$ with probability 1 if $k \geqslant K_2$ for all $x \in \mathcal{F}$. So if $g(x') < g(x)$ then with probability 1, $\hat{\bar{g}}_{m_k}(x') < \hat{\bar{g}}_{m_k}(x)$ for all x, $x' \in \mathcal{F}$ if $k \geqslant K_2$. This in turn implies that with probability 1 if $\hat{\bar{g}}_{m_k}(x') \geqslant \hat{\bar{g}}_{m_k}(x)$ then $g(x') \geqslant g(x)$ for all x, $x' \in \mathcal{F}$, $k \geqslant K_2$. Thus with probability 1, $X_k^* \in \mathcal{F}$ bounded and $g(X_k^*) \leqslant g(X_{k-1}^*)$ for all $k \geqslant K = \max(K_1, K_2)$, and hence $\Pr\{X_k^* \notin M^*(N) \ i.o.\} = 0$. $\square$

If one is to relax Theorem 1 to include the context of unbounded feasible regions, more needs to be assumed about the behavior of the objective function g to prevent "chase-offs" to infinity both due to the structure of the recursion, and due to mischance. Once such assumptions are made, results analogous to Lemma 1 and Theorem 1 can be proved for unbounded sets $\mathcal{F}$. In the results that follow, we let $S_k(x)$ denote the set $\{x' \in \mathbb{X} : \hat{\bar{g}}_{m_k}(x') \leqslant \hat{\bar{g}}_{m_k}(x), x' \in \mathcal{F}_k\}$ for each $x \in \mathcal{F}$.

Assumption 3 Let the sequence of random variables $\{g(x) - \hat{g}_{m_k}(x)\}$ be governed by a large-deviation principle with rate function $I_x(s)$ such that for any $\epsilon > 0$, $\inf_{x \in \mathbb{X}} \min(I_x(-\epsilon), I_x(\epsilon)) = \eta_g > 0$.

Assumption 4 For each $x \in \mathbb{X}$ there exists $\delta > 0$ such that the level set $S(x, \delta) = \{x' \in \mathbb{X} : g(x') \leqslant g(x) + \delta\}$ is finite.

Lemma 2 Suppose $x_0 \in \mathcal{F}$, and $\delta > 0$ is such that $S(x_0, \delta)$ is finite. Then under Assumptions 1–4, $\Pr\{S_k(x_0) \nsubseteq \mathcal{F} \cap S(x_0, \delta) i.o.\} = 0$.

Proof. Pick $y \in (\mathcal{F} \cap S(x_0, \delta))^c$. Then either $y \in \mathcal{F}^c$ or $y \in S(x_0, \delta)^c \cap \mathcal{F}$. Suppose $y \in \mathcal{F}^c$. Then for $k \geqslant K_1$(independent on x_0), $Pr\{y \in S_k(x_0)\} \leqslant Pr\{y \in \mathcal{F}_k\} \leqslant e^{-m_k \eta}$ (from (2)).

If $y \in S(x_0, \delta)^c \cap \mathcal{F}$ then $g(y) > g(x_0) + \delta$. Then for $0 < 2\epsilon < \delta$ and $k \geqslant K_2$(independent on x_0 and δ),

$$
\begin{aligned}
Pr\{y \in S_k(x_0)\} &= Pr\{\hat{g}_{m_k}(y) \leqslant \hat{g}_{m_k}(x_0)\} \\
&\leqslant Pr\{\hat{g}_{m_k}(y) \leqslant \hat{g}_{m_k}(x_0), \hat{g}_{m_k}(x_0) \in (g(x_0) - \epsilon, g(x_0) + \epsilon)\} \\
&\quad + Pr\{\hat{g}_{m_k}(y) \leqslant \hat{g}_{m_k}(x_0), \hat{g}_{m_k}(x_0) \notin (g(x_0) - \epsilon, g(x_0) + \epsilon)\} \\
&\leqslant Pr\{\hat{g}_{m_k}(y) \notin (g(y) - \epsilon, g(y) + \epsilon)\} \\
&\quad + Pr\{\hat{g}_{m_k}(x_0) \notin (g(x_0) - \epsilon, g(x_0) + \epsilon)\} \\
&\leqslant 2e^{-m_k \eta_g} \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (3)
\end{aligned}
$$

Let $p_k = \max(2e^{-m_k \eta_g}, e^{-m_k \eta})$. Then for any $y \in (\mathcal{F} \cap S(x_0, \delta))^c$, $\Pr\{y \in S_k(x_0)\} \leqslant p_k$ if $k \geqslant K = \max(K_1, K_2)$. And since $\limsup_{k \to \infty} k/m_k = 0$, $\Pr\{y \in S_k(x_0) \ i.o.\} = 0$. Thus $S_k(x_0)$ does not contain any point in $(\mathcal{F} \cap S(x_0))^c$ with probability 1 for $k \geqslant K$. In other words $\Pr\{S_k(x_0) \nsubseteq \mathcal{F} \cap S(x_0), \delta) i.o.\} = 0$. $\square$

We are now ready to present the analogue to Theorem 1 for the context of unbounded feasible regions $\mathcal{F}$.

Theorem 2 Given a starting solution $x_0 \in \mathcal{F}$ and a neighborhood definition N, and under Assumptions 1–4, R-SPLINE returns a sequence of sample path solutions $\{X_k^*\}$ that converges wp1 to the set of true local minima $M^*(N)$ of problem $\tilde{P}_2$.

Proof.   Suppose $\delta > 0$ is such that $S(x_0, \delta)$ is finite. Then $\Pr\{S_k(x_0) \nsubseteq \mathcal{F} \cap S(x_0, \delta) \, i.o.\} = 0$. Since $X_k^* \in S_k(x_0)$, $\Pr\{X_k^* \notin S_k(x_0) \, i.o.\} = 0$. Hence there exists K_1(independent on x_0 and δ) $\in \mathbb{N}$ such that $X_k^* \in \mathcal{F} \cap S(x_0, \delta)$ with probability 1 if $k \geqslant K_1$. And since $\mathcal{F} \cap S(x_0, \delta)$ is finite, R-SPLINE returns a sequence of solutions that are bounded with probability 1. We also note (from to Lemma 1) that $\Pr\{X_k^* \in \mathcal{F}^c \, i.o.\} = 0$.

Let $\epsilon = \min \{|g(x) - g(x')| : (x, x') \in \mathcal{F} \cap S(x_0, \delta), g(x) \neq g(x')\}$. Then since $\mathcal{F} \cap S(x_0, \delta)$ is finite and contained in $\mathcal{F}$, $\hat{\bar{g}}_{m_k}$ converges uniformly to g with probability 1 on the set $\mathcal{F} \cap S(x_0, \delta)$ as $k \to \infty$. Thus there exists $K_2(\epsilon) \in \mathbb{N}$ such that $|\hat{\bar{g}}_{m_k}(x) - g(x)| < \epsilon/2$ with probability 1 if $k \geqslant K_2$ for all $x \in \mathcal{F} \cap S(x_0, \delta)$. So if $g(x') < g(x)$ then with probability 1, $\hat{\bar{g}}_{m_k}(x') < \hat{\bar{g}}_{m_k}(x)$ for all $x, x' \in \mathcal{F} \cap S(x_0, \delta)$ if $k \geqslant K_2$. This in turn implies that with probability 1 if $\hat{\bar{g}}_{m_k}(x') \geqslant \hat{\bar{g}}_{m_k}(x)$ then $g(x') \geqslant g(x)$ for all $x, x' \in \mathcal{F} \cap S(x_0, \delta)$, $k \geqslant K_2$. Thus with probability 1, $X_k^* \in \mathcal{F} \cap S(x_0, \delta)$ bounded and $g(X_k^*) \leqslant g(X_{k-1}^*)$ for all $k \geqslant K = \max(K_1, K_2)$, and hence $\Pr\{X_k^* \notin M^*(N) \, i.o.\} = 0$.   □

Theorems 1 and 2 prove the almost sure convergence of R-SPLINE's iterates $\{X_k^*\}$ to a true local minimum. How fast does such convergence happen? In other words, can anything be said about the rate at which the probability of adapted R-SPLINE returning a point other than a true local minimum decays to zero? What is the corresponding rate for returning a truly feasible solution. The following two results assert that these rates are exponential when the sample sizes are increased at a fast enough rate.

Theorem 3 If $\mathbb{X}$ is finite then under Assumptions 1 and 2, the probability that R-SPLINE returns an infeasible solution decreases exponentially in k, i.e., $\Pr\{X_k^* \notin \mathcal{F}\} = O(e^{-k\eta})$ for some $\eta > 0$.

Proof.   Let $\eta = \min_{1 \leqslant i \leqslant c} \eta_i$. Then

$$\Pr\{\mathcal{F}_k \nsubseteq \mathcal{F}\} = \Pr\{\bigcup_{y \in \mathbb{X} \setminus \mathcal{F}} \{y \in \mathcal{F}_k\}\}$$
$$\leqslant \sum_{y \in \mathbb{X} \setminus \mathcal{F}} \Pr\{y \in \mathcal{F}_k\}$$
$$\leqslant \sum_{y \in \mathbb{X} \setminus \mathcal{F}} e^{-m_k \eta}$$
$$\leqslant |\mathbb{X}| e^{-k\eta} \tag{4}$$

where the last inequality of (4) follows from Assumption 1. Thus, $\Pr\{X_k^* \notin \mathcal{F}\} \leqslant \Pr\{\mathcal{F}_k \nsubseteq \mathcal{F}\} \leqslant |\mathbb{X}| e^{-k\eta}$.   □

Theorem 4 If $\mathbb{X}$ is finite then given a neighborhood definition N and under Assumptions 1, 2 and 3, the probability that R-SPLINE returns a locally suboptimal solution decreases exponentially in k, i.e., $\Pr\{X_k^* \notin M^*(N)\} = O(e^{-k\eta'})$ for some $\eta' > 0$.

Proof.   Let $W^*(\mathbb{X})$ and $W_k^*(\mathbb{X})$ denote the set of global minima for problems $\tilde{P}_1$ and P_k, respectively. Both $W^*(\mathbb{X})$ and $W_k^*(\mathbb{X})$ are nonempty since $\mathbb{X}$ is finite. Then let $\epsilon > 0$ be such that $2\epsilon < \min_{y \in \mathcal{F} \setminus W^*(\mathbb{X})}(g(y) - g(x^*))$,

where $x^* \in W^*(\mathbb{X})$. Then for some $y \in \mathcal{F} \setminus W^*(\mathbb{X})$ and k large enough,

$$
\begin{aligned}
\Pr\{y \in W_k^*(\mathbb{X})\} &\leqslant \Pr\{\hat{g}_{m_k}(y) \leqslant \hat{g}_{m_k}(x^*)\} \\
&= \Pr\{\hat{g}_{m_k}(y) \leqslant \hat{g}_{m_k}(x^*),\ \hat{g}_{m_k}(x^*) \in (g(x^*) - \epsilon, g(x^*) + \epsilon)\} \\
&\quad + \Pr\{\hat{g}_{m_k}(y) \leqslant \hat{g}_{m_k}(x^*),\ \hat{g}_{m_k}(x^*) \notin (g(x^*) - \epsilon, g(x^*) + \epsilon)\} \\
&\leqslant \Pr\{\hat{g}_{m_k}(y) \leqslant g(y) - \epsilon\} + \Pr\{\hat{g}_{m_k}(x^*) \notin (g(x^*) - \epsilon, g(x^*) + \epsilon)\} \\
&\leqslant \Pr\{\hat{g}_{m_k}(y) \notin (g(y) - \epsilon, g(y) + \epsilon)\} + \Pr\{\hat{g}_{m_k}(x^*) \notin (g(x^*) - \epsilon, g(x^*) + \epsilon)\} \\
&\leqslant 2e^{-k\eta_g}
\end{aligned}
\tag{5}
$$

The last inequality of (5) follows from Assumptions 1 and 3.

Thus for large enough k,

$$
\begin{aligned}
\Pr\{M_k^*(N) \nsubseteq M^*(N)\} &\leqslant \sum_{y \in \mathbb{X} \setminus M^*(N)} \Pr\{y \in M_k^*(N)\} \\
&= \sum_{y \in \mathbb{X} \setminus \mathcal{F}} \Pr\{y \in M_k^*(N)\} + \sum_{y \in \mathcal{F} \setminus M^*(N)} \Pr\{y \in M_k^*(N)\} \\
&\leqslant \sum_{y \in \mathbb{X} \setminus \mathcal{F}} \Pr\{y \in \mathcal{F}_k\} + \sum_{y \in \mathcal{F} \setminus M^*(N)} \Pr\{y \in W_k^*(N(y))\} \\
&\leqslant \sum_{y \in \mathbb{X} \setminus \mathcal{F}} e^{-k\eta} + \sum_{y \in \mathcal{F} \setminus M^*(N)} 2e^{-k\eta_g} \\
&\leqslant 3|\mathbb{X}| e^{-k\eta'}
\end{aligned}
\tag{6}
$$

where the third inequality of (6) follows from (2) and (5), and $\eta' = \min\{\eta_g, \eta\}$. $\qquad\square$

4 CONCLUDING REMARKS

We consider the relatively unexplored problem of local integer-ordered SO with stochastic constraints. For solving this problem, we adapt the recently proposed R-SPLINE algorithm (for unconstrained or deterministically-constrained contexts) by including a sampled and relaxed version of the constraints from the true problem. It so happens that the asymptotics of the adapted version resembles the originally proposed R-SPLINE, in terms of the consistency and convergence rate of the returned solutions with respect to optimality and feasibility. When the true solutions lie on the boundary of the feasible region, interesting questions relating to how one should choose the extent of relaxation of the constraints remains an unexplored but important question. When all local solutions lie on the interior of the feasible region, the proposed adaptation presents no surprises. While we have assumed that the SPLINE algorithm is used within the retrospective approximation context to create the adapted version of R-SPLINE, any deterministic locally minimizing algorithm for integer ordered spaces can be used SPLINE's stead, with no change in the results we have presented.

ACKNOWLEDGMENTS

REFERENCES

Andradóttir, S., D. Goldsman, and S. H. Kim. 2005. "Finding the best in the presence of a stochastic constraint". In *Proceedings of the 2005 Winter Simulation Conference*, edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 732–738: Institute of Electrical and Electronics Engineers: Piscataway, New Jersey.

Andradóttir, S., and S. H. Kim. 2010. "Fully sequential procedures for comparing constrained systems via simulation". *Naval Research Logistics* (57): 403–421.

Batur, D., and S. H. Kim. 2005. "Procedures for feasibility detection in the presence of multiple constraints". In *Proceedings of the 2005 Winter Simulation Conference*, edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 692–698: Institute of Electrical and Electronics Engineers: Piscataway, New Jersey.

Dembo, A., and O. Zeitouni. 1998. *Large Deviations Techniques and Applications*. New York, NY.: Springer-Verlag.

Hong, J., and B. L. Nelson. 2006. "Discrete Optimization Via Simulation Using COMPASS". *Operations Research* 54 (1): 115–129.

Hunter, S. R., and R. Pasupathy. 2013. "Optimal Sampling Laws for Stochastically Constrained Simulation Optimization". *INFORMS Journal on Computing*. To appear.

Hunter, S. R., N. A. Pujowidianto, L. H. Lee, C. H. Chen, and R. Pasupathy. 2011. "Optimal Sampling Laws for Constrained Simulation Optimization on Finite Sets: The Bivariate Normal Case". In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu: Institute of Electrical and Electronics Engineers: Piscataway, New Jersey.

Kim, S., R. Pasupathy, and S. G. Henderson. 2012. "A Guide to SAA". Frederick Hilliers OR Series. Elsevier.

Kushner, H. J., and G. G. Yin. 2003. *Stochastic Approximation and Recursive Algorithms and Applications*. New York, NY.: Springer-Verlag.

Law, A. M. 2007. *Simulation Modeling and Analysis*. New York, NY.: McGraw-Hill.

Li, J., S. Sava, and X. Xie. 2009. "Simulation-based Discrete Optimization of Stochastic Discrete Event Systems Subject to Non Closed-form Constraints". 54:2900–2904.

Lim, E. 2013. "Stochastic Approximation Over Multidimensional Discrete Sets with Applications to Inventory Systems and Admission Control of Queueing Networks". *ACM TOMACS*. To appear.

Luo, Y., and E. Lim. 2013. "Simulation-based optimization over discrete sets with noisy constraints". *IIE Transactions* 45 (7): 699–715.

Park, C., and S. H. Kim. 2011. "Handling Stochastic Constraints in Discrete Optimization via Simulation". In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu: Institute of Electrical and Electronics Engineers: Piscataway, New Jersey.

Pasupathy, R. 2010. "On choosing parameters in retrospective-approximation algorithms for stochastic root finding and simulation optimization". *Operations Research* 58:889–901.

Pasupathy, R., and S. Henderson. 2006. "A Testbed of Simulation-Optimization Problems". In *Proceedings of the 2006 Winter Simulation Conference*, edited by L. Perrone, F. Wieland, J. Liu, B. Lawson, D. Nicol, and R. Fujimoto: Institute of Electrical and Electronics Engineers: Piscataway, New Jersey.

Pasupathy, R., and S. G. Henderson. 2011. "SimOpt: A Library of Simulation Optimization Problems". In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. Fu: Institute of Electrical and Electronics Engineers: Piscataway, New Jersey.

Pasupathy, R., and S. Kim. 2011. "The stochastic root-finding problem: overview, solutions, and open questions". *ACM TOMACS* 21 (3).

Shapiro, A., D. Dentcheva, and A. Ruszczynski. 2009. *Lectures on Stochastic Programming: Modeling and Theory*. Philadelphia, PA: SIAM.

Wang, H., R. Pasupathy, and B. W. Schmeiser. 2013. "Integer-Ordered Simulation Optimization using R-SPLINE: Retrospective Search using Piecewise-Linear Interpolation and Neighborhood Enumeration". *ACM TOMACS*. To appear.

Xu, J., L. J. Hong, and B. L. Nelson. 2010. "Industrial Strength COMPASS: A Comprehensive Algorithm and Software for Optimization via Simulation". *ACM TOMACS* (20): 1–29.

AUTHOR BIOGRAPHIES

KALYANI NAGARAJ is a PhD student in the Grado Department of Industrial and Systems Engineering at Virginia Tech. Her research interests include analysis and methodology for Monte Carlo simulations, and specifically simulation optimization and input model uncertainty. Her e-mail address is kalyanin@vt.edu.

RAGHU PASUPATHY is an Associate Professor in the Grado Department of Industrial and Systems Engineering at Virginia Tech. His research interests lie broadly in Monte Carlo methods with a specific focus on simulation optimization. He is a member of INFORMS, IIE, and ASA, and serves as Department Editor for the simulation desk at IIE Transactions and Associate Editor for *Operations Research*, *ACM TOMACS* and *INFORMS Journal on Computing*. His email address is pasupath@vt.edu and his web page is https://filebox.vt.edu/users/pasupath/pasupath.htm.