Proceedings of the 2019 Winter Simulation Conference
N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, eds.

BALANCING OPTIMAL LARGE DEVIATIONS IN RANKING AND SELECTION

Ye Chen

Statistical Sciences & Operations Research
Virginia Commonwealth University
1015 Floyd Avenue
Richmond, VA 23284, USA

Ilya O. Ryzhov

Robert H. Smith School of Business &
Institute for Systems Research
University of Maryland
7699 Mowatt Lane
College Park, MD 20742, USA

ABSTRACT

The ranking and selection problem deals with the optimal allocation of a simulation budget to efficiently identify the best among a finite set of unknown values. The large deviations approach to this problem provides very strong performance guarantees for static (non-adaptive) budget allocations. Using this approach, one can describe the optimal static allocation with a set of highly nonlinear, distribution-dependent optimality conditions whose solution depends on the unknown parameters of the output distribution. We propose a new methodology that provably learns this solution (asymptotically) and is very computationally efficient, has no tunable parameters, and works for a wide variety of output distributions.

1 INTRODUCTION

Ranking and selection is a mathematical framework for understanding and studying the optimal allocation of resources for the purpose of information collection. In the classical statement of the problem, there is a finite set of "designs" or " alternatives" (for example, competing simulation models or differently-tuned versions of the same model) whose performance can be estimated through individual simulations. Assigning more simulation replications to one alternative produces more accurate estimates of that particular performance value, but reduces the budget available for learning about other alternatives, creating a tradeoff. Applications where information is collected through simulation include manufacturing (Boesel et al. 2003) and inventory management (Xu et al. 2010); more generally, a "replication" may represent any expensive experiment, such as a fundraising campaign (Han et al. 2016) or a clinical trial (Chick et al. 2017).

The methodological research on this problem focuses on the development of algorithms for efficiently allocating the budget. Allocations can be *static*, which means that they are determined once at the beginning, before any information is collected; this type of approach is studied in the literature on statistical design of experiments (see, e.g., Qian et al. 2008; Zhang and Qian 2013). Alternately, allocations can be *adaptive*, meaning that simulations are assigned sequentially and each assignment takes into account the result of previous experiments. Most recent research in the simulation community has focused on adaptive allocations, and can be divided into several broad algorithmic classes. Each of these classes is based on a particular algorithmic concept: for example, indifference-zone methods (Kim and Nelson 2001; Andradóttir and Kim 2010) successively screen out alternatives after comparing their estimated values in groups of two or three (Kim and Dieker 2011); expected improvement methods (Chick et al. 2010; Salemi et al. 2014) use Bayesian predictions of the potential (variously defined) of the next simulation to improve the current estimate of the best value; posterior sampling methods (Russo and Van Roy 2014) assign simulations in a randomized, but guided manner based on the uncertainty in a set of Bayesian posteriors; and upper confidence bound methods (Bubeck et al. 2011) invert various concentration inequalities to create criteria

for adaptive assignments. Each of these schools of thought has also developed its own analytical techniques that produce certain types of guarantees.

Our work focuses on one specific school of thought, which originates from the seminal paper by Glynn and Juneja (2004) and is based on the idea of using optimal static allocations to guide suboptimal adaptive ones. Using large deviations theory, this approach proves very strong optimality guarantees for static allocations under general distributional assumptions. It was shown that a static allocation causes the probability of incorrect selection to converge to zero at an exponential rate, with an exponent that depends on the allocation. It follows that the *optimal* allocation is the one that optimizes this exponent; one can derive a set of highly nonlinear optimality conditions whose solution will depend on the underlying distributional parameters. Of course, in practice one does not know these parameters, but one can plug in a set of estimates and thus create heuristics that are inspired by the optimality conditions. There is an entire stream of papers using precisely this approach, with some representative examples being Pasupathy et al. (2014), Hunter and McClosky (2016), and Gao et al. (2017). A very similar principle is used by the literature on optimal computing budget allocation (Chen and Lee 2010, Chen et al. 2015, Zhang et al. 2016, which considers various tractable approximations of the optimality conditions. In general, these papers focus on the derivation of the optimality conditions for the setting of interest, and then implement various heuristics that aim to learn or approximate the solution.

In this paper, we summarize a new methodology (laid out in detail in Chen and Ryzhov 2019a) which is guaranteed to learn the solution to the optimality conditions under a wide variety of sampling distributions. This approach, which we call Balancing Optimal Large Deviations or BOLD, does not need to solve systems of nonlinear equations, as does the vast majority of heuristics for large deviations-based allocations. Rather, it adaptively learns the solution by iteratively "balancing" the two sides of each equation, which only requires us to separately evaluate these sides for a given set of parameters, rather than to solve for the value that makes them equal. This can be done very quickly, making BOLD highly computationally efficient.

Another major advantage of BOLD is that it does not use any tunable parameters. This differentiates our approach from those of Russo (2019) and Qin et al. (2017), which likewise aim to learn the solution to the large deviations optimality conditions. These "top-two" methods rely on the observation, also made by Glynn and Juneja (2004), that the decision to allocate a simulation to the *optimal* alternative is based on different logic from the choice between *suboptimal* alternatives. Top-two methods essentially pass the first of these decisions back to the decision-maker: one assigns the next simulation to the best alternative (or to one's best guess of it) with some pre-specified probability, and if this does not occur, then the algorithm chooses between the remaining alternatives in a provably optimal fashion. Therefore, convergence to the correct solution requires prior knowledge of the correct proportion of the budget to assign to the optimal alternative. Our BOLD procedure eliminates this requirement and is guaranteed to satisfy all of the optimality conditions.

We view BOLD as the fundamental algorithmic paradigm for large deviations-based methods. This literature has mainly focused on the theoretical optimality conditions, but thus far has not agreed on any particular decision criterion that could be implemented in a practical setting while remaining compatible with the theoretical result. BOLD provides an exceptionally convenient and computationally efficient criterion, which is easily seen to apply to practically any standard sampling distribution in ranking and selection, and thus may also be applicable to more complex simulation optimization problems where the large deviations approach can be leveraged to derive optimality conditions.

2 PROBLEM BACKGROUND

We first define some formal notation for the problem. Let there be $M > 2$ alternatives; for any alternative $x \in \{1,...,M\}$, we may collect independent random samples from the distribution F_x with mean μ_x, assuming for convenience that $\mu_x \neq \mu_y$ for any $x \neq y$. We wish to identify $x^* = \arg\max_x \mu_x$ with high probability. Given a budget of N simulations, we let $\{x^n\}_{n=0}^{N-1}$ be a sequence of alternatives chosen for sampling. In time stage $n = 0,...,N-1$, we observe $W_{x^n}^{n+1} \sim F_{x^n}$. Let $\mathscr{F}^n$ be the sigma-algebra generated by $x^0, W_{x^0}^1, ..., x^{n-1}, W_{x^{n-1}}^n$.

The allocation $\{x^n\}$ is adaptive if $x^n \in \mathscr{F}^n$ for all n, and static if $x^n \in \mathscr{F}^0$ for all n; in words, a static allocation is deterministic and does not depend on any observed values, whereas an adaptive allocation does and may adjust itself sequentially.

Now let $N_x^n = \sum_{m=0}^{n-1} 1_{\{x^m=x\}}$ be the number of simulations assigned to alternative x in the first n stages (sample size), and let $\theta_x^n = \frac{1}{N_x^n} \sum_{m=0}^{n-1} 1_{\{x^m=x\}} W_x^{m+1}$ be the sample mean at time n, with θ_x^0 chosen arbitrarily. Denote by $x^{*,n} = \arg\max_x \theta_x^n$ the index of the alternative with the highest sample mean; if this argmax is not unique, let $x^{*,n}$ be the alternative that has the smallest sample size N_y^n among $y \in \arg\max_x \theta_x^n$, with further ties broken in any arbitrary fashion. *Correct selection* is said to occur if $x^{*,n} = x^*$.

Glynn and Juneja (2004) studied static allocations that satisfy $\lim_{n\to\infty} \frac{N_x^n}{n} = \alpha_x$ with $\alpha_x > 0$ for all x. The main result was that such allocations satisfy

$$\lim_{n\to\infty} \frac{1}{n} \log P(x^{*,n} \neq x^*) = -\Gamma(\alpha;\mu), \tag{1}$$

where Γ is a function of the allocation $\alpha = (\alpha_1, ..., \alpha_M)$ as well as the population parameters. That is, the probability of *incorrect* selection vanishes to zero at an exponential rate, but the exponent Γ can be manipulated by changing the allocation. The convex program

$$\begin{array}{rll} \max_{\alpha\in\mathbb{R}^M} & \Gamma(\alpha;\mu) & \\ \text{s.t.} & \sum_{x=1}^M \alpha_x & = 1, \\ & \alpha & \geq 0. \end{array} \tag{2}$$

yields the optimal allocation α^*, which can be further characterized using the KKT conditions. Clearly, α^* depends on μ and therefore cannot be computed in any practical application, but we can already see that it might be approximated, for example by plugging $\theta^n = (\theta_1^n, ..., \theta_M^n)$ into the KKT conditions instead of μ.

We now give a very brief summary of the large deviations analysis used by Glynn and Juneja (2004) to derive (1). The following discussion assumes that $N_x^n = \alpha_x n$ for all n, which will not be true in finite time, but which does not matter for the asymptotic result (only minor technical nuisances are needed to handle the case where $\frac{N_x^n}{n} \to \alpha_x$).

First, it was shown that $\Gamma(\alpha;\mu) = \min_{x\neq x^*} \Gamma_x(\alpha_{x^*}, \alpha_x; \mu)$, where

$$\Gamma_x(\alpha_{x^*}, \alpha_x; \mu) = -\lim_{n\to\infty} \frac{1}{n} \log P(\theta_{x^*}^n < \theta_x^n), \tag{3}$$

provided that the limit on the RHS of (3) exists. This means that the convergence rate of the probability of incorrect selection is related to the convergence rate of the probability of making an error in the pairwise comparison between x^* and some $x \neq x^*$. In fact, the overall error probability is governed by the *slowest* rate among the pairwise comparisons. Intuitively, the rate for the comparison between x^* and x speeds up if we increase α_x. However, doing so will reduce α_y for other $y \neq x$, slowing down the rates for those comparisons. Therefore, the proportions α_x should be chosen in such a way as to make the error rates for all the pairwise comparisons equal ("balanced"). This will be expressed more rigorously further down.

To characterize (3) more precisely, let us take a fixed x and define $\Psi_x(\gamma) = \log \mathbb{E}\left(e^{\gamma W_x}\right)$, where $W_x \sim F_x$ represents a generic sample. The domain of Ψ_x is taken to be the set of all γ for which $\mathbb{E}\left(e^{\gamma W_x}\right) < \infty$. Let

$$I_x(u;\mu_x) = \sup_\gamma \gamma u - \Psi_x(\gamma)$$

be the Fenchel-Legendre transform of Ψ_x. By observing that

$$\log \mathbb{E}\left(e^{\gamma_1 \theta_{x^*}^n + \gamma_2 \theta_x^n}\right) = \log \mathbb{E}\left(e^{\gamma_1 \theta_{x^*}^n}\right) + \log \mathbb{E}\left(e^{\gamma_2 \theta_x^n}\right)$$

and applying the Gärtner-Ellis theorem (Dembo and Zeitouni 2009), it is shown that

$$\Gamma_x(\alpha_{x^*}, \alpha_x; \mu) = \inf_{u \in [\mu_x, \mu_{x^*}]} \alpha_{x^*} I_{x^*}(u; \mu_{x^*}) + \alpha_x I_x(u; \mu_x), \tag{4}$$

so the convergence rate of the probability of error in the pairwise comparison is governed by the individual rate functions I_{x^*}, I_x. It turns out that the infimum in (4) is achieved at a unique u; after some algebra, one finds that the KKT conditions of problem (2) consist of two parts:

- Total balance condition:
$$\sum_{x \neq x^*} \frac{I_{x^*}(u_{x^*,x}(\alpha_{x^*}, \alpha_x); \mu_{x^*})}{I_x(u_{x^*,x}(\alpha_{x^*}, \alpha_x); \mu_x)} = 1. \tag{5}$$
- Individual balance conditions:
$$\Gamma_x(\alpha_{x^*}, \alpha_x; \mu) = \Gamma_y(\alpha_{x^*}, \alpha_y; \mu), \qquad \text{for all } x, y \neq x^*. \tag{6}$$

Condition (5) arises due to the equality constraint in (2), while (6) is in line with our intuition that the error rates for the pairwise comparisons should all be equal under the optimal allocation.

Equations (5)-(6) have much more explicit forms if one makes additional distributional assumptions. Chen and Ryzhov (2019a) gives four examples; here, in the interests of space, we only show one example for the very common setting of normal distributions. Suppose that F_x is $\mathcal{N}(\mu_x, \lambda_x^2)$; then, $I_x(u; \mu_x) = \frac{(u-\mu_x)^2}{2\lambda_x^2}$ and, as shown in Example 1 of Glynn and Juneja (2004), (5) becomes

$$\left(\frac{\alpha_{x^*}}{\lambda_{x^*}}\right)^2 = \sum_{x \neq x^*} \left(\frac{\alpha_x}{\lambda_x}\right)^2, \tag{7}$$

while (6) becomes

$$\frac{(\mu_x - \mu_{x^*})^2}{\frac{\lambda_{x^*}^2}{\alpha_{x^*}} + \frac{\lambda_x^2}{\alpha_x}} = \frac{(\mu_y - \mu_{x^*})^2}{\frac{\lambda_{x^*}^2}{\alpha_{x^*}} + \frac{\lambda_y^2}{\alpha_y}}, \qquad x, y \neq x^*. \tag{8}$$

Even in this special case, it is clear that (7)-(8) are quite difficult to solve even for given μ. A typical heuristic might plug θ^n into these equations, solve to get an approximate allocation α^n, use this allocation to assign a portion of the budget, then update the sample means and repeat. This would involve solving many systems of nonlinear equations, which is quite time-consuming.

3 BALANCING OPTIMAL LARGE DEVIATIONS

The main insight behind our Balancing Optimal Large Deviations (BOLD) method is that evaluating the two sides of (6) for given values of α and μ is much easier than solving for the value of α that makes these two sides equal. We will evaluate the individual rate functions under the most recent set of sample means, then balance them iteratively by assigning simulations to alternatives whose estimated rates appear to be too small.

Define the shorthand $I_x^n(\cdot) = I_x(\cdot; \theta_x^n)$. Similarly, for $D_x(u; \mu) = \frac{d}{du} I_x(u; \mu)$, denote $D_x^n(\cdot) = D_x(\cdot; \theta_x^n)$. The functions I_x^n and D_x^n are plug-in estimates of the large deviations rate function and its derivative. Recalling that the solution to (4) is governed by the equation

$$\alpha_{x^*} D_{x^*}(u; \mu_{x^*}) + \alpha_x D_x(u; \mu_x) = 0, \tag{9}$$

let us instead consider a similar equation

$$N_x^n D_x^n(u) + N_y^n D_y^n(u) = 0 \tag{10}$$

for arbitrary $x \neq y$. If either x or y is equal to x^*, (10) is a version of (9) in which the true means are replaced by sample means, and the proportions α are replaced by time-n sample sizes. The solution to (10), which we denote by u^n_{xy}, is unaffected if we divide by n, so it does not matter whether we work with the sample sizes or the empirical proportions. We now define a kind of rescaled, approximate rate function

$$\Gamma^n_{x,y} = N^n_x I^n_x \left(u^n_{x,y}\right) + N^n_y I^n_y \left(u^n_{x,y}\right), \tag{11}$$

with the specific case $\Gamma^n_{x^*,x}$ being analogous to Γ_x.

We now present some informal intuition for our algorithmic approach. Suppose for the sake of argument that $N^n_x \to \infty$ for all x (as will be formally proved later), and that n is large enough such that $x^{*,n} = x^*$. Then (11) suggests that $\Gamma^n_{x^*,x}$ will also increase to infinity. This, in turn, suggests that we can make $\Gamma^n_{x^*,x}$ converge to infinity at the same rate, for all $x \neq x^*$, by assigning simulations to whichever such x has the *smallest* value of $\Gamma^n_{x^*,x}$ at time n. Similarly, we can approximate the left-hand side of the total balance condition (5) and assign simulations to $x^{*,n}$ depending on whether this approximation is above or below 1. This is exactly the structure of the BOLD algorithm, which is formally stated in Figure 1 below.

Step 0: Initialize $n = 0$ and $N^n_x = 0$ for all x.

Step 1: If $\arg\max_x \theta^n_x$ is not unique, assign $x^n = x^{*,n}$ and proceed directly to Step 4.

Step 2: If $\arg\max_x \theta^n_x$ is unique, check whether

$$\sum_{x \neq x^{*,n}} \frac{I^n_{x^{*,n}}\left(u^n_{x^{*,n},x}\right)}{I^n_x\left(u^n_{x^{*,n},x}\right)} > 1. \tag{12}$$

Step 3: If (12) holds, assign $x^n = x^{*,n}$. Otherwise, assign

$$x^n = \arg\min_{x \neq x^{*,n}} \Gamma^n_{x^{*,n},x}. \tag{13}$$

Step 4: Collect new information $W^{n+1}_{x^n}$, update sample means. Increment n by 1 and return to step 1.

Figure 1: Description of BOLD algorithm.

Step 1 of the algorithm is needed to accommodate discrete sampling distributions (for example, if F_x is Bernoulli), under which one could have two sample means with identical values at some finite n. However, since we assume that $\mu_x \neq \mu_y$ for any $x \neq y$, one can show that this will happen at most finitely many times as $n \to \infty$, so Step 1 does not play a major role in the asymptotic analysis of the procedure. From Steps 2 and 3, we see that the key to the procedure is the idea of evaluating (but not solving) approximate versions of the optimality equations and assigning simulations in a way that seeks to balance them.

Let us see how BOLD works in our example of normal distributions. In this setting, (12) becomes

$$\left(\frac{N^n_{x^{*,n}}}{\lambda_{x^{*,n}}}\right)^2 < \sum_{x \neq x^{*,n}} \left(\frac{N^n_x}{\lambda_x}\right)^2, \tag{14}$$

so BOLD will assign $x^n = x^{*,n}$ if (14) holds. If (14) does not hold, the calculation in (13) becomes

$$x^n = \arg\min_{x \neq x^{*,n}} \frac{(\theta^n_x - \theta^n_{x^{*,n}})^2}{\frac{\lambda^2_{x^{*,n}}}{N^n_{x^{*,n}}} + \frac{\lambda^2_x}{N^n_x}}, \tag{15}$$

and if the variances are unknown, one could simply plug in estimators of these quantities where necessary. If F_x belongs to a different distributional family, one should derive I_x in that setting and obtain an explicit form for (12)-(13); this form will be different from (14)-(15), but will likewise be straightforward to compute. Chen and Ryzhov (2019a) gives examples for three other distributional families.

From this we see that BOLD can be straightforwardly adapted to many distributions, but we require knowledge of the distributional family in order to do so. If we cannot evaluate $I(u;\theta)$ for arbitrary u,θ, implementing BOLD will become more difficult. A possible direction for future work is to consider a version of BOLD in which the rate function itself is estimated, but this case is not considered here. However, we note that the present setting is already comparable to, or more general than, the assumptions underlying many existing methods in the simulation literature (many of which require normal distributions).

We briefly note that there have been a few recent papers proposing algorithms with BOLD-like structure, most notably Shin et al. (2018), which proposes algorithm (14)-(15) for solving the optimality conditions in the special case of normality, and Gao et al. (2017), which uses a more general form closer to (12)-(13). However, these papers treat these algorithms as heuristics; the proof of convergence for general sampling distributions is new to our work.

4 SUMMARY OF THEORETICAL ANALYSIS

We now summarize the theoretical analysis of Chen and Ryzhov (2019a). The main result is that BOLD learns the solution to (5)-(6), with probability 1, as $n \to \infty$. The proof of this result relies on a number of more fundamental properties and intermediate steps. In this section, we give a high-level discussion of the proof structure, but defer to Chen and Ryzhov (2019a) for the technical details.

First, however, we should make sure that the main result is properly interpreted. Although Glynn and Juneja (2004) explicitly characterizes the convergence rate of the optimal static allocation, an adaptive allocation that *converges* to that static allocation as $n \to \infty$ is not guaranteed to achieve the same convergence rate. In fact, as discussed in Glynn and Juneja (2011) and Wu and Zhou (2018), it may not achieve an exponential convergence rate at all. However, the large deviations optimality conditions still play an important role in the performance of adaptive procedures. For example, Russo (2019) shows that these same conditions are required for the rate-optimality of a different type of performance criterion presented in that paper. Thus, although the static framework of Glynn and Juneja (2004) may have certain limitations, nonetheless, as of this writing this framework continues to serve as the foundation for virtually all of the work on convergence rates for ranking and selection, and furthermore can be observed to be a kind of unifying link for algorithmic classes such as expected improvement (Ryzhov 2016) and optimal computing budget allocation (Chen and Lee 2010). For these reasons, the optimal static allocation is an important object of study, and our contribution with this work is to show that it can provably be learned by a computationally efficient and tuning-free sequential algorithm.

We now state four major assumptions of our analysis; all four can be shown to hold on a variety of useful output distributions. Assumptions 1-2 are regularity conditions on the rate functions I_x. They ensure that $I_x(u;\theta)$ has a unique zero at $u=\theta$ and is well-behaved around this value.

Assumption 1 For any x and arbitrarily small $\varepsilon > 0$,

$$\inf_{\theta\in H, |u-\theta|>\varepsilon} I_x(u;\theta) > 0,$$
$$\inf_{\theta\in H, |u-\theta|>\varepsilon} |D_x(u;\theta)| > 0,$$

where H is any closed interval with non-empty interior containing μ_x.

Assumption 2 Both $I_x(u;\theta)$ and $D_x(u;\theta)$ are continuous at every pair (u,θ) with $\theta \in H$, where H is as in Assumption 1. Furthermore, for any x, and any $\varepsilon > 0$, there exists some $\delta > 0$ such that, for all pairs (u,θ) satisfying $\theta \in H$ and $|u-\theta| < \delta$, we have $I_x(u;\theta) < \varepsilon$ and $|D_x(u;\theta)| < \varepsilon$.

Assumption 3 ensures that x and y cannot be "mistaken" for each other if the sample mean for x and the true value of y are close to each other. This assumption is mainly needed for the pathological situation where $\mu_x = \theta_y^n$ for $x \neq y$, which cannot happen with non-zero probability when the sampling distributions are continuous. In the continuous setting, Assumption 3 can be omitted entirely, but it enables us to handle discrete distributions (for example, one can verify that it holds for Bernoulli distributions).

Assumption 3 For any x, y, there exists a fixed positive constant C such that

$$\sup_{u \neq \theta} \left| \frac{D_x(u;\theta)}{D_y(\theta;u)} \right| \leq C.$$

The final assumption is a regularity condition on $I_x(u;\theta)$ and $D_x(u;\theta)$, viewed here as bivariate functions of both u and θ.

Assumption 4 For all x, both $\frac{\partial I_x}{\partial \theta}(u;\theta)$ and $\frac{\partial D_x}{\partial \theta}(u;\theta)$ are continuously differentiable in both u and θ.

We now proceed to the first major step in the analysis, namely the statistical consistency of the BOLD algorithm. A brief, high-level discussion of the proof will follow.

Theorem 1 (Chen and Ryzhov 2019a) Let Assumptions 1-3 hold. Under the BOLD algorithm, $N_x^n \to \infty$ for all x.

Since BOLD uses two distinct criteria (one for choosing $x^{*,n}$, one for choosing among $x \neq x^{*,n}$), we must examine both criteria to make sure that there is no z with $\lim_{n\to\infty} N_z^n < \infty$. It is first necessary to show that, if such a z exists, it must be the case that $x^{*,n} \neq z$ for all large enough n. Because of this, for any large enough n when condition (12) holds, the sampling decision must choose from among alternatives that are measured infinitely often. Because the law of large numbers applies to all such alternatives, it follows that, for all large enough n, $\arg\max_x \theta_x^n$ will eventually yield the same alternative index and $x^{*,n}$ will take the same value. It also follows from this that Step 1 in Figure 1 will only be invoked finitely many times, which means that condition (12) will be checked for all sufficiently large n.

One then shows that condition (12) will fail for infinitely many n. In other words, there will be infinitely many n in which the criterion in (13) is checked. From this, however, it follows that, if there exists any z satisfying $\lim_{n\to\infty} N_z^n < \infty$, we must nonetheless have had infinitely many chances to choose that z. That is, if $x^{*,n} \neq z$ for all large enough n, then z must be one of the available alternatives that we compare in (13) whenever (12) has failed. From this, the non-existence of such z can be proved: we show that $\Gamma^n_{x^{*,n},x}$ increases over time when both $x^{*,n}$ and x are sampled infinitely often, whereas $\Gamma^n_{x^{*,n},z}$ is bounded from above. This implies that, for arbitrarily large n, the argmin of (13) must correspond to an alternative that can no longer be sampled, contradicting the existence of any z satisfying $\lim_{n\to\infty} N_z^n < \infty$.

Consistency plays an important role in the rest of the analysis because it allows us to study a simpler form of the BOLD algorithm, in which Step 1 of Figure 1 is removed entirely (because, in any case, it will be invoked only finitely many times) and $x^{*,n}$ in (12)-(13) is replaced by x^*. This is because our analysis focuses on large n on a fixed sample path, and for any such sample path we can simply take n large enough that the original BOLD algorithm also has these behaviours. With these modifications, one can show the next major step, which is that the BOLD algorithm assigns $\mathcal{O}(n)$ samples to every alternative.

Theorem 2 (Chen and Ryzhov 2019a) Let Assumptions 1-3 hold. For any two alternatives x and y,

$$\limsup_{n\to\infty} \frac{N_x^n}{N_y^n} < \infty.$$

The next result states that BOLD satisfies the total balance condition, i.e., (5) is satisfied asymptotically.

Theorem 3 (Chen and Ryzhov 2019a) Let Assumptions 1-4 hold. Then,

$$\lim_{n\to\infty} \sum_{x \neq x^*} \frac{I_{x^*}^n\left(u_{x^*,x}^n\right)}{I_x^n\left(u_{x^*,x}^n\right)} = 1.$$

Again, we give some informal discussion of how the proof works. For notational convenience, let $\Delta^n = \sum_{x \neq x^*} \frac{I^n_{x^*}\left(u^n_{x^*,x}\right)}{I^n_x\left(u^n_{x^*,x}\right)}$. From previous results, we know that condition (12) both succeeds and fails infinitely many times, meaning that Δ^n crosses the level 1 infinitely often from both above and below. Let us consider some sufficiently large n at which $\Delta^n = 1$, whence the BOLD algorithm chooses $x^n = x^*$. Let us also choose m such that $n+m$ is the next time stage when $x^{n+m} = x^*$, which means that only suboptimal alternatives are sampled (and $\Delta^n \leq 1$) at times $n+s$ for $0 < s < m$. The proof of Theorem 3 first shows that:

- The quantity m, as defined above, must be bounded by $\mathscr{O}\left(\sqrt{n \log\log n}\right)$;
- For any $m \sim \mathscr{O}\left(\sqrt{n \log\log n}\right)$, it follows that

$$\left|u^n_{x^*,x} - u^{n+m}_{x^*,x}\right| = O\left(\sqrt{\frac{\log\log n}{n}}\right).$$

In words, we show that m must be "small" in a certain sense; we then show that "small" changes from time n to time $n+m$ produce "small" incremental changes from $u^n_{x^*,x}$ to $u^{n+m}_{x^*,x}$. This, in turn, is shown to lead to "small" changes from Δ^n to Δ^{n+m}, such that, for any $0 < s < m$, $\Delta^{n+s} - \Delta^n$ vanishes to zero as $n \to \infty$. A similar, symmetric argument can be made for the case where $x^n \neq x^*$ and Δ^n crosses the level 1 from below. Therefore, as Δ^n repeatedly crosses the level 1, it stays closer and closer (converges) to that level as n becomes large.

The final major result states that BOLD satisfies the individual balance conditions, i.e., (6) is satisfied asymptotically for all pairs of suboptimal alternatives.

Theorem 4 (Chen and Ryzhov 2019a) Let Assumptions 1-4 hold. For any $y, z \neq x^*$,

$$\lim_{n\to\infty} \frac{\Gamma^n_{x^*,y}}{\Gamma^n_{x^*,z}} = 1.$$

Similarly to the previous result, we require a number of intermediate steps that place bounds of $\mathscr{O}\left(\sqrt{n \log\log n}\right)$ on the number of time periods that can occur between two samples of interest. Previously we saw that such a bound holds on the number of time stages in between two samples of the optimal alternative. Now, we also establish the following:

- The number of samples of x^* that can be made in between two samples of some fixed $z \neq x^*$ is $\mathscr{O}\left(\sqrt{n \log\log n}\right)$;
- Again taking a fixed $z \neq x^*$, the number of samples of any $y \neq z$ in between two samples of z is $\mathscr{O}\left(\sqrt{n \log\log n}\right)$.

Analogously to the proof of the total balance condition, we use these facts to show that the ratios $\frac{\Gamma^n_{x^*,y}}{\Gamma^n_{x^*,z}}$ can only change by "small" increments (that vanish to zero) around the level 1. This completes the proof that BOLD is guaranteed to learn the solution to the large deviations optimality conditions as $n \to \infty$.

The overall sequence of arguments used in this proof builds on Chen and Ryzhov (2019b), which studied a BOLD-like algorithm in the special case of normal distributions. In that context, this earlier paper also proved convergence to the optimal allocation using a number of intermediate $\mathscr{O}\left(\sqrt{n \log\log n}\right)$ bounds. However, the normal case is much easier to handle: for example, the total balance condition (7) has no explicit dependence on the sample means and thus is much more tractable. Many of the intermediate steps are easier to show, or may even be omitted entirely, when restricted to the normal case.

5 NUMERICAL ILLUSTRATION

We present a numerical illustration on a synthetic example with $M = 30$ alternatives and normal sampling distributions. Since the true values μ_x are generated artificially, the optimality conditions (5)-(6) may be solved by brute force. We compare the empirical allocation made by BOLD after $N = 10^5$ samples with the optimal allocation in order to illustrate our theoretical analysis and also highlight how this ability is unique to BOLD as compared to benchmark methods.

Suppose that F_x is $\mathcal{N}\left(\mu_x, \lambda_x^2\right)$, with the parameters (μ_x, λ_x) instantiated according to the rules in Sec. 5.1 of Ryzhov (2018). Two instances (i.e., two sets of population means generated according to the same rules) are considered. The sampling variances are known to all methods, while the means are unknown. We implemented the standard expected improvement (EI) algorithm (Ryzhov 2016), the Thompson sampling algorithm (Russo and Van Roy 2014) with normal priors, and the top-two expected improvement method of Qin et al. (2017). It is known that EI and Thompson sampling satisfy $\frac{N_{x^*}^n}{n} \to 1$, meaning that they do not sample suboptimal alternatives as often as prescribed by large deviations theory. Top-two EI assigns a flat proportion $\frac{1}{2}$ of the budget to $x^{*,n}$ and uses a variant of expected improvement (specifically, the complete EI method of Salemi et al. 2014) to choose among the remaining alternatives whenever $x^{*,n}$ is not sampled. This flat proportion can be viewed as a tunable parameter, and if it is set to the optimal value of α_{x^*} as prescribed by large deviations theory, the remaining proportions will automatically be optimized. However, Qin et al. (2017) specifically recommends using the value $\frac{1}{2}$, so we implement the method in this way.

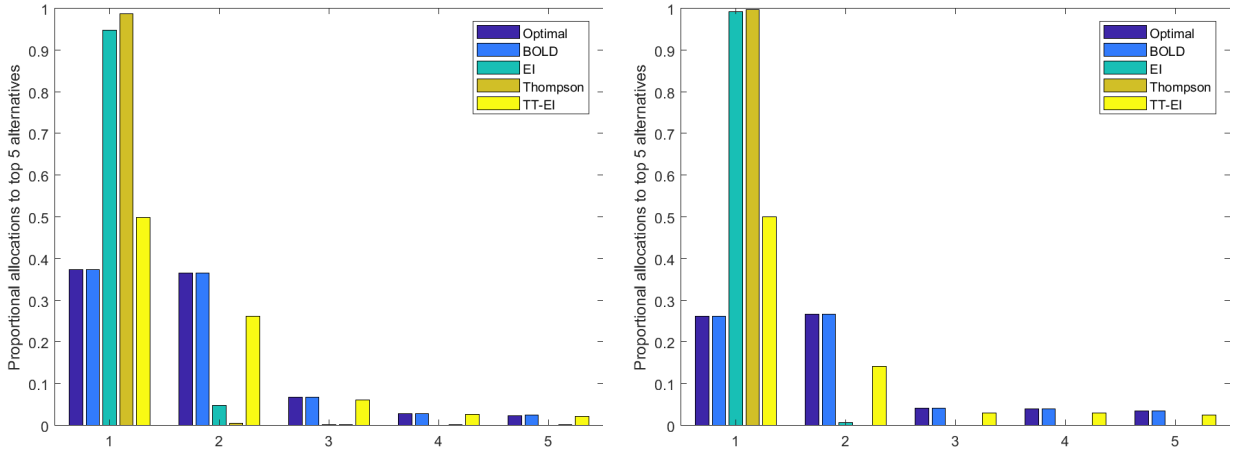
Figure 2 compares the empirical allocations (averaged over 1000 macro-replications) achieved by these methods in the test instances. For readability, we only show allocations to the five alternatives with the highest *optimal* proportions; it is easily seen that these proportions quickly become very small for suboptimal alternatives. As expected, BOLD learns the optimal proportions with high precision, while EI and Thompson sampling spend virtually the entire budget on x^*. Top-two EI is forced to over-sample x^*, and therefore underestimates the other alternatives, most notably the second-best.

In Figure 3, we show the average value of $\mu_{x^{*,n}}$, as a function of n, under each procedure. Overall, the differences between methods are small, reflecting the maturity of the algorithmic literature on ranking and selection with normal distributions; most modern methods will perform similarly well in this setting, and from Ryzhov (2016) and Russo (2019) we know that the large deviations optimality conditions provide a kind of link between multiple algorithmic methodologies (that is, many of these methods, even though they may be derived from different principles, can somehow be related to the optimality conditions or to some asymptotic approximation of them). In general, BOLD has an advantage for small simulation budgets on the order of 10-100 samples. When the budget is around 10^3 samples, Thompson sampling and top-two EI have a slight advantage, followed by a regime where BOLD catches up to these methods and the differences between them are no longer visible.

While these examples cannot be used to draw far-reaching conclusions about finite-time empirical performance, we can see that, in these instances, BOLD is quite competitive against the state of the art; also, from a computational point of view, BOLD and Thompson sampling were the fastest methods. In our view, a major advantage of BOLD is its generality – we can easily implement a version of BOLD for non-normal distributions (some examples are given in Chen and Ryzhov 2019a).

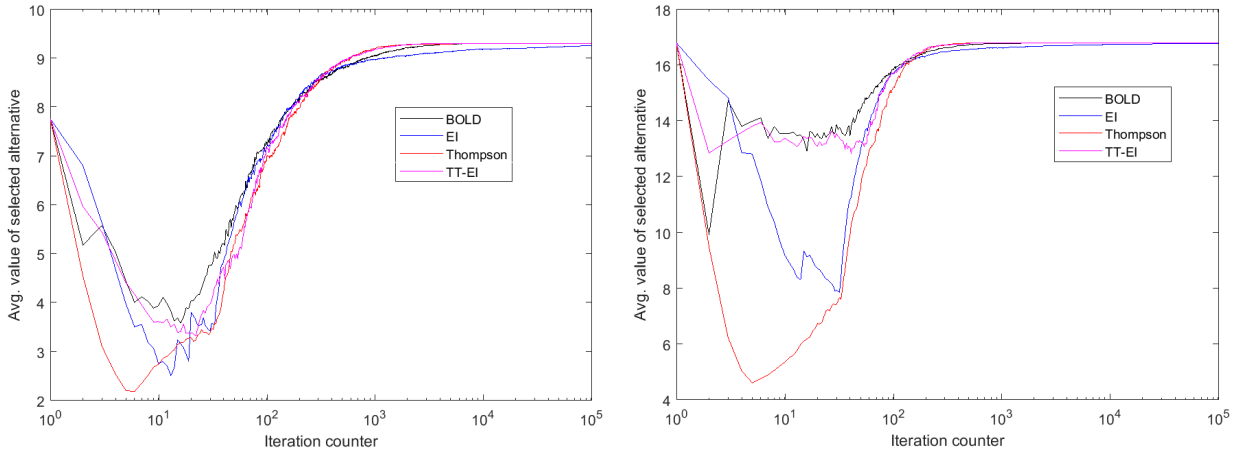
6 CONCLUSION

We have presented a simple, fast, and general algorithm for learning optimal allocations (as prescribed by large deviations theory) in ranking and selection. We view this work as building a computational foundation for the broader stream of literature (starting with Glynn and Juneja 2004) on large deviations-based allocations for simulation optimization problems. The general philosophy adopted by this literature is to characterize optimal static allocations, and then to use various adaptive heuristics to solve for or approximate the optimal allocation. Our approach provides a rigorous and fast computational principle for achieving this, under a variety of output distributions, without the need for tunable parameters. We hope



(a) Instance 1. (b) Instance 2.

Figure 2: Empirical allocations, $N = 10^5$.



(a) Instance 1. (b) Instance 2.

Figure 3: Average value of $\mu_{x^*,n}$ under each procedure.

that BOLD will come to be viewed as the go-to algorithmic approach for large deviations-based methods in ranking and selection as well as other problem classes where such methods have been developed.

REFERENCES

Andradóttir, S., and S.-H. Kim. 2010. "Fully Sequential Procedures for Comparing Constrained Systems Via Simulation". *Naval Research Logistics* 57(5):403–421.

Boesel, J., B. L. Nelson, and N. Ishii. 2003. "A Framework for Simulation-Optimization Software". *IIE Transactions* 35(3):221–229.

Bubeck, S., R. Munos, and G. Stoltz. 2011. "Pure Exploration in Finitely-Armed and Continuous-Armed Bandits". *Theoretical Computer Science* 412(19):1832–1852.

Chen, C.-H., S. E. Chick, L. H. Lee, and N. A. Pujowidianto. 2015. "Ranking and Selection: Efficient Simulation Budget Allocation". In *Handbook of Simulation Optimization*, edited by M. C. Fu, 45–80. New York: Springer-Verlag.

Chen, C.-H., and L. H. Lee. 2010. *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*. Singapore: World Scientific.

Chen, Y., and I. O. Ryzhov. 2019a. "Balancing Optimal Large Deviations in Sequential Selection". Technical report, University of Maryland.

Chen, Y., and I. O. Ryzhov. 2019b. "Complete Expected Improvement Converges to an Optimal Budget Allocation". *Advances in Applied Probability* 51(1):209–235.

Chick, S., M. Forster, and P. Pertile. 2017. "A Bayesian Decision Theoretic Model of Sequential Experimentation With Delayed Response". *Journal of the Royal Statistical Society* B79(5):1439–1462.

Chick, S. E., J. Branke, and C. Schmidt. 2010. "Sequential Sampling to Myopically Maximize the Expected Value of Information". *INFORMS Journal on Computing* 22(1):71–80.

Dembo, A., and O. Zeitouni. 2009. *Large Deviations Techniques and Applications (2nd ed.)*. New York: Springer-Verlag.

Gao, S., W. Chen, and L. Shi. 2017. "A New Budget Allocation Framework for the Expected Opportunity Cost". *Operations Research* 65(3):787–803.

Glynn, P. W., and S. Juneja. 2004. "A Large Deviations Perspective on Ordinal Optimization". In *Proceedings of the 2004 Winter Simulation Conference*, edited by R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 577–585. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Glynn, P. W., and S. Juneja. 2011. "Ordinal Optimization: A Nonparametric Framework". In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. C. Fu, 4062–4069. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Han, B., I. O. Ryzhov, and B. Defourny. 2016. "Optimal Learning in Linear Regression With Combinatorial Feature Selection". *INFORMS Journal on Computing* 28(4):721–735.

Hunter, S. R., and B. McClosky. 2016. "Maximizing Quantitative Traits in the Mating Design Problem via Simulation-Based Pareto Estimation". *IIE Transactions* 48(6):565–578.

Kim, S.-H., and A. B. Dieker. 2011. "Selecting the Best by Comparing Simulated Systems in a Group of Three". In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R. R. Creasey, J. Himmelspach, K. P. White, and M. C. Fu, 3987–3997. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Kim, S.-H., and B. L. Nelson. 2001. "A Fully Sequential Procedure For Indifference-Zone Selection in Simulation". *ACM Transactions on Modeling and Computer Simulation* 11(3):251–273.

Pasupathy, R., S. R. Hunter, N. A. Pujowidianto, L. H. Lee, and C.-H. Chen. 2014. "Stochastically Constrained Ranking and Selection via SCORE". *ACM Transactions on Modeling and Computer Simulation* 25(1):1:1–1:26.

Qian, P. Z. G., H. Wu, and C. F. J. Wu. 2008. "Gaussian Process Models for Computer Experiments With Qualitative and Quantitative Factors". *Technometrics* 50(3):383–396.

Qin, C., D. Klabjan, and D. Russo. 2017. "Improving the Expected Improvement Algorithm". In *Advances in Neural Information Processing Systems*, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Volume 30, 5381–5391. Red Hook, New York: Curran Associates, Inc.

Russo, D. 2019. "Simple Bayesian Algorithms for Best Arm Identification". *arXiv preprint arXiv:1602.08448*.

Russo, D., and B. Van Roy. 2014. "Learning to Optimize via Posterior Sampling". *Mathematics of Operations Research* 39(4):1221–1243.

Ryzhov, I. O. 2016. "On the Convergence Rates of Expected Improvement Methods". *Operations Research* 64(6):1515–1528.

Ryzhov, I. O. 2018. "The Local Time Method for Targeting and Selection". *Operations Research* 66(5):1406–1422.

Salemi, P., B. L. Nelson, and J. Staum. 2014. "Discrete Optimization via Simulation Using Gaussian Markov Random Fields". In *Proceedings of the 2014 Winter Simulation conference*, edited by A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 3809–3820. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Shin, D., M. Broadie, and A. Zeevi. 2018. "Tractable Sampling Strategies for Ordinal Optimization". *Operations Research* 66(6):1693–1712.

Wu, D., and E. Zhou. 2018. "Analyzing and Provably Improving Fixed Budget Ranking and Selection Algorithms". *arXiv preprint arXiv:1811.12183*.

Xu, J., B. L. Nelson, and L. J. Hong. 2010. "Industrial Strength COMPASS: A Comprehensive Algorithm and Software for Optimization via Simulation". *ACM Transactions on Modeling and Computer Simulation* 20(1):3:1–3:29.

Zhang, Q., and P. Z. G. Qian. 2013. "Designs for Crossvalidating Approximation Models". *Biometrika* 100(4):997–1004.

Zhang, S., L. H. Lee, E. P. Chew, J. Xu, and C.-H. Chen. 2016. "A Simulation Budget Allocation Procedure for Enhancing the Efficiency of Optimal Subset Selection". *IEEE Transactions on Automatic Control* 61(1):62–75.

AUTHOR BIOGRAPHIES

YE CHEN is an Assistant Professor of Statistical Sciences and Operations Research at Virginia Commonwealth University. He received a Ph.D. in Statistics from the University of Maryland in 2018. His research interests include applied probability, statistical learning, and stochastic optimization. He was a finalist in the Best Theoretical Paper Award competition at the 2016

Winter Simulation Conference. His email address is ychen24@vcu.edu.

ILYA O. RYZHOV is an Associate Professor of Operations Management in the Decision, Operations and Information Technologies department of the Robert H. School of Business, with a joint appointment in the Institute for Systems Research, all at the University of Maryland. His research primarily focuses on simulation optimization and statistical learning, with applications in business analytics, revenue management, and nonprofit/humanitarian operations. He is a coauthor of the book Optimal Learning (Wiley, 2012). His work was recognized in WSC's Best Theoretical Paper Award competition on three separate occasions (winner in 2012, finalist in 2009 and 2016), and he received I-SIM's Outstanding Publication Award in 2017. His email address is iryzhov@rhsmith.umd.edu.