Proceedings of the 2015 Winter Simulation Conference
L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, eds.

EXPECTED IMPROVEMENT IS EQUIVALENT TO OCBA

Ilya O. Ryzhov

Robert H. Smith School of Business
University of Maryland
College Park, MD 20742, USA

ABSTRACT

This paper summarizes new theoretical results on the asymptotic sampling rates of expected improvement (EI) methods in fully sequential ranking and selection (R&S). These methods have been widely observed to perform well in practice, and often have asymptotic consistency properties, but rate results are generally difficult to obtain when observations are subject to stochastic noise. We find that, in one general R&S problem, variants of EI produce simulation allocations that are virtually identical to the rate-optimal allocations calculated by the optimal computing budget allocation (OCBA) methodology. This result provides new insight into the good empirical performance of EI under normality assumptions.

1 INTRODUCTION

We summarize recent results by Ryzhov (2015). This work demonstrates the asymptotic equivalence of two methodological approaches in a certain version of the ranking and selection (R&S) problem. R&S is a fundamental problem in simulation, dealing with the efficient use of a limited information budget to identify the highest-valued element of a finite set of design alternatives (e.g., competing simulation models). See Kim (2013) or Chen et al. (2015) for an overview. Aside from its immediate applicability, R&S offers an analytical framework for modeling the exploration/exploitation tradeoff, that is, the value of experimenting with a seemingly suboptimal design in order to learn new information that may change the decision-maker's selection decision.

R&S is a well-established area of research, and a variety of methodological approaches is available; see, e.g., Chau et al. (2014) and the references cited therein for an overview. In the following, we will focus on two methodologies, namely the optimal computing budget allocation (OCBA) approach (surveyed in Chen and Lee 2010), and the expected improvement (EI) approach (surveyed in Powell and Ryzhov 2012), and establish a new equivalence result between them in one specific (but widely-studied) R&S model. This result is fairly surprising, since the two methodologies proceed from completely different principles.

The OCBA approach (Chen, Fu, and Shi 2008) can be thought of as calculating an optimal deterministic simulation allocation. Given M designs and a budget of N simulation runs (where only one design can be implemented in a single run), we solve a single optimization problem to determine the number N_x of runs to allocate to each design $x = 1,...,M$ in a way that optimizes a desired metric. Usually, the metric is the probability of correct selection (Chen et al. 2000), that is, the probability that the design with the highest expected performance will also have the highest sample mean after the output of the allocated runs has been collected and averaged. The approach can also be extended to the metric of expected opportunity cost (He, Chick, and Chen 2007). When $N \to \infty$, the optimal allocations N_x admit a closed-form expression (Glynn and Juneja 2004); however, this expression is a function of the true performance values of the alternatives, which are unknown. In practice, it is estimated iteratively using sample means.

The EI approach (Chick, Branke, and Schmidt 2010) allocates simulations one at a time in an adaptive manner. One typically adopts a Bayesian view of the unknown values of the designs, and models them as random variables. The posterior distributions of these quantities can be updated recursively after each

978-1-4673-9743-8/15/$31.00 ©2015 IEEE

simulation run. Furthermore, using Bayesian predictive distributions, we can make a probabilistic forecast of the outcome of the next simulation. We then allocate that simulation to the design which is expected (on average) to make the greatest improvement in the desired metric (usually expected opportunity cost). The simulation output is observed, the posterior distributions are updated, and the expected improvement calculation can be repeated with a new set of posterior parameters. Thus, while this method is adaptive, in the sense that the allocation decisions are improved over time as new information is collected, it is also myopic, in the sense that the forecast only looks ahead to the outcome of the next simulation.

To provide context for these methods, consider a standard R&S model where the simulation output is normally distributed and independent across alternatives. In this setting, both EI and OCBA methods are asymptotically consistent, in the sense that they learn the exact value of every design as $N \to \infty$. Furthermore, the construction of OCBA naturally characterizes the rate at which each design is sampled; each design receives a non-zero proportion of the budget, and the relative frequency of simulation is represented by the *sampling ratios* $\frac{N_x}{N_y}$ for $x \neq y$. By contrast, the sampling rates of EI methods are more difficult to characterize due to the myopic nature of the allocation decisions, and the fact that the EI calculations involve complicated functions whose asymptotic behaviour is not immediately evident. As of this writing, the state of the art in this theory is the work by Bull (2011), which provides convergence rate results for EI in global optimization (also known as the EGO method; see Jones, Schonlau, and Welch 1998), but assumes *zero* variance in the simulation output. The proof technique cannot be easily adapted to handle stochastic noise.

Ryzhov (2015) derives asymptotic sampling ratios for variants of EI in independent normal R&S. Surprisingly, these ratios are virtually identical to the theoretically optimal values derived for OCBA in Glynn and Juneja (2004) and Chen and Lee (2010), with some minor variations depending on the precise version of EI used. Although the two methods optimize different criteria (opportunity cost for EI, probability of correct selection for OCBA), recent work (Gao and Shi 2014) has observed that these criteria converge to their limiting values at the same rate, providing intuition for the idea that the same policy may asymptotically optimize both of them in some cases. Even so, the new results demonstrate that this optimal allocation is asymptotically obtained by a *myopic* policy, which to our knowledge is the strongest such result currently available for EI in problems with noise.

It is worth pointing out that asymptotic convergence to the optimal sampling ratios is not enough to guarantee an optimal rate of convergence for the algorithm (an issue discussed in Glynn and Juneja 2011). However, most practical implementations of OCBA would not be able to guarantee this either; rather, they would use the sample means in place of the true values in the OCBA ratios, and use those estimated ratios to drive a randomized sequential policy, thus obtaining the optimal ratios asymptotically. Nonetheless, this convergence is an important regularity property for such algorithms, and in our view provides insight into the good empirical performance of EI methods that has widely been observed under normality assumptions (see, e.g., Frazier, Powell, and Dayanik 2008 or Ryzhov, Powell, and Frazier 2012).

This paper provides an informal overview of the result; for full proofs and technical details, please see Ryzhov (2015). Here, we state the main results and provide intuition. Section 2 lays out the R&S model and EI algorithms that will be studied. Section 3 describes the analysis. Section 4 gives a numerical example, and Section 5 concludes.

2 INDEPENDENT NORMAL RANKING AND SELECTION

We consider a version of the ranking and selection problem with independent normal simulation output with known variance. The model is Bayesian, since EI methods are based on Bayesian arguments, but the learning mechanism is identical to that used in the frequentist analog of this problem. Although there exist more powerful learning models that exploit correlations between alternatives (Qu, Ryzhov, and Fu 2012), the independent normal model continues to be widely studied, e.g. in the literature on indifference-zone methods (Kim and Nelson 2006); the method of batch means is suggested as a way to render the output approximately normal (Kim and Nelson 2007).

Let μ_x be the true average performance of design x. Since this quantity is unknown, we use the Bayesian model $\mu_x \sim \mathcal{N}\left(\theta_x^0, \left(\sigma_x^0\right)^2\right)$, where the parameters θ_x^0, σ_x^0 are user-specified. The simulation output for design x is modeled as $Y_x \sim \mathcal{N}\left(\mu_x, \lambda_x^2\right)$, where λ_x is presumed known. Let $x^n \in \{1,...,M\}$ be the index of the design chosen for the $(n+1)$st simulation, with $Y_{x^n}^{n+1}$ being the output. Denote by $\mathcal{F}^n$ the sigma-algebra generated by $x^0, Y_{x^0}^1, ..., x^{n-1}, Y_{x^{n-1}}^n$. It is well-known (DeGroot 1970) that the conditional distribution of μ_x given $\mathcal{F}^n$ is $\mathcal{N}\left(\theta_x^n, (\sigma_x^n)^2\right)$, where the posterior parameters follow the recursive update

$$\theta_x^{n+1} = \begin{cases} \frac{(\sigma_x^n)^{-2}\theta_x^n + \lambda_x^{-2} Y_x^{n+1}}{(\sigma_x^n)^{-2} + \lambda_x^{-2}} & x^n = x \\ \theta_x^n & x^n \neq x, \end{cases} \tag{1}$$

$$\left(\sigma_x^{n+1}\right)^2 = \begin{cases} \left((\sigma_x^n)^{-2} + \lambda_x^{-2}\right)^{-1} & x^n = x \\ (\sigma_x^n)^2 & x^n \neq x. \end{cases} \tag{2}$$

To simplify the model, we suppose that $\sigma_x^0 = \infty$, whence (1)-(2) become

$$\theta_x^n = \frac{1}{N_x^n} \sum_{n'=0}^{n-1} 1_{\{x^{n'}=x\}} Y_{x^{n'}}^{n'+1}, \qquad (\sigma_x^n)^2 = \frac{\lambda_x^2}{N_x^n}, \tag{3}$$

where $N_x^n = \sum_{n'=0}^{n-1} 1_{\{x^{n'}=x\}}$ is the number of times x has been simulated up to time n. This is simply the usual frequentist mean and its variance.

The Bayesian model also enables us to make probabilistic forecasts about the future. Given $\mathcal{F}^n$ and $x^n = x$, the conditional distribution of θ_x^{n+1} is $\mathcal{N}\left(\theta_x^n, (\tilde{\sigma}_x^n)^2\right)$ where

$$(\tilde{\sigma}_x^n)^2 = (\sigma_x^n)^2 - \left(\sigma_x^{n+1}\right)^2.$$

EI methods use this fact to evaluate the potential improvement offered by running one additional simulation of design x.

Different definitions of "improvement" are possible. For example, the EI criterion of Jones, Schonlau, and Welch (1998) can be defined in our setting as

$$v_x^{EI,n} = \mathbb{E}\left[\max\left\{\mu_x - \max_y \theta_y^n, 0\right\} \mid \mathcal{F}^n, x^n = x\right].$$

To evaluate the potential of design x, we consider the possibility that its true value μ_x will be greater than (improve over) the current estimate $\max_y \theta_y^n$ of the best value. We calculate the expected value of this improvement (if $\mu_x < \max_y \theta_y^n$, there is no improvement, so we only integrate over the positive tail). Under the Bayesian assumptions, the expectation can be calculated in closed form as

$$v_x^{EI,n} = \sigma_x^n f\left(-\frac{\left|\theta_x^n - \max_y \theta_y^n\right|}{\sigma_x^n}\right), \tag{4}$$

where the *information valuation* function $f(z) = z\Phi(z) + \phi(z)$ and ϕ, Φ are the standard normal pdf and cdf. The method then chooses the design $x^n = \arg\max_x v_x^{EI,n}$ for simulation.

The "knowledge gradient" or KG criterion (Frazier, Powell, and Dayanik 2008) defines the expected improvement as

$$\begin{aligned} v_x^{KG,n} &= \mathbb{E}\left[\max_y \theta_y^{n+1} - \max_y \theta_y^n \mid \mathcal{F}^n, x^n = x\right] \\ &= \tilde{\sigma}_x^n f\left(-\frac{\left|\theta_x^n - \max_{y\neq x} \theta_y^n\right|}{\tilde{\sigma}_x^n}\right). \end{aligned}$$

In this case we look ahead only to the outcome of the next measurement. However, the result is quite similar to (4). Both methods use the information valuation function and compare θ^n_x, the current estimated value of design x, with a reference value (either $\max_y \theta^n_y$ or $\max_{y\neq x} \theta^n_y$). The KG policy likewise chooses design $x^n = \arg\max_x \nu^{KG,n}_x$ for simulation.

It is easy to show that, as $n \to \infty$, we have $\nu^{EI,n}_x \to 0$ and $\nu^{KG,n}_x \to 0$ for the two respective policies. Consequently, $N^n_x \to \infty$ and $\theta^n_x \to \mu_x$ almost surely under both policies. However, the growth rate of N^n_x is important to the finite-time performance of the procedures. We will now discuss how these rates can be characterized.

3 SAMPLING RATIOS FOR EXPECTED IMPROVEMENT METHODS

In the following, we focus on the EI policy (our analysis can later be applied to the KG policy with very minor changes). To characterize N^n_x, it is necessary to understand the rate at which $\nu^{EI,n}_x$ converges to zero, which depends on the behaviour of the information valuation function f. The main difficulty in understanding this behaviour is due to the stochastic nature of the process (θ^n) in (4). We eliminate this difficulty by considering a modification of the EI policy that is purely *deterministic*.

3.1 Deterministic Modification of the EI Policy

For simplicity, suppose that the true means μ_x are *fixed* for all x. Although the derivation of (4) is based on Bayesian arguments, the computational formula can still be applied in a frequentist setting. (Alternately, one can replace μ_x by $\mu_x(\omega)$ for a fixed sample path ω in the following discussion.) Define the "modified EI criterion" as

$$\bar{\nu}^{EI,n}_x = \sigma^n_x f\left(-\frac{|\mu_x - \max_y \mu_y|}{\sigma^n_x}\right).$$

From (3), it follows that

$$\bar{\nu}^{EI,n}_x = \frac{\lambda_x}{\sqrt{N^n_x}} f\left(-c_x\sqrt{N^n_x}\right) \tag{5}$$

where $c_x = \frac{|\mu_x - \max_y \mu_y|}{\lambda_x}$ is now a constant (independent of time); we assume that $c_x \neq c_y$ for $x \neq y$. Thus, the decisions $x^n = \arg\max_x \bar{\nu}^{EI,n}_x$ do not depend on the simulation output at all. The continuous function

$$\bar{\nu}_x(z) = \frac{\lambda_x}{\sqrt{z}} f\left(-c_x\sqrt{z}\right) \tag{6}$$

can be viewed as an interpolation of (5), where z is a continuous analog of N^n_x. This representation can provide insight into the behaviour of f.

Proposition 1 (Ryzhov 2015) Let $c_1, c_2 > 0$. Then,

$$\lim_{z\to\infty} \frac{f(-c_1\sqrt{z})}{f(-c_2\sqrt{z})} = \begin{cases} \infty & c_1 < c_2 \\ 1 & c_1 = c_2 \\ 0 & c_1 > c_2. \end{cases}$$

We see that the (modified) information valuation of two designs, given comparable values of z, either converges at exactly the same rate, or vanishes an order of magnitude faster for one of the designs. This suggests the following supposition: if the modified EI policy is used to select systems, then $f(-c_x\sqrt{N^n_x})$ and $f\left(-c_y\sqrt{N^n_y}\right)$ must converge to zero at the same rate for *any* $x \neq y$.

The informal intuition behind this idea is as follows. Suppose that N^n_x and N^n_y are comparable in magnitude for some sufficiently large n. If $c_x > c_y$, then $f(-c_x\sqrt{N^n_x})$ must vanish to zero faster than

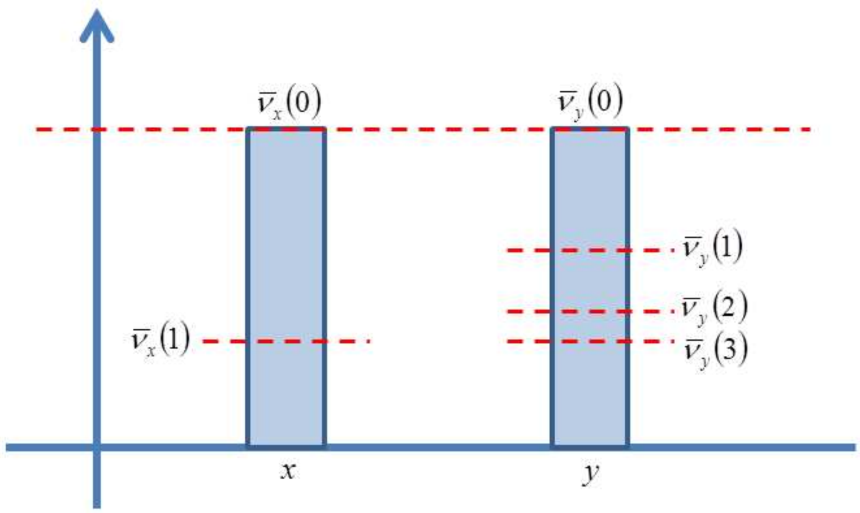

Figure 1: Let x,y be two systems with $c_x > c_y$. Then, $\bar{v}_x$ decreases faster than $\bar{v}_y$, and the modified EI method has to measure y more times in order to reduce the EI quantity by the same amount.

$f\left(-c_y\sqrt{N_y^n}\right)$. However, in that case, the modified EI method will consistently prefer design y to design x, since it always chooses the system with the highest EI quantity. It follows that N_y^n should grow faster than N_x^n in order for $\bar{v}_y^{EI,n}$ to catch up with $\bar{v}_x^{EI,n}$. The policy will adjust the sampling rates as necessary in order to ensure that the EI quantities vanish at the same rate. Figure 1 gives a visual illustration.

This intuition is formalized in the next result. Under the modified EI method, none of the individual EI quantities can shrink substantially faster (or slower) than the others.

Proposition 2 (Ryzhov 2015) For $x \neq y$, $\limsup_{n\to\infty} \frac{\bar{v}_x^{EI,n}}{\bar{v}_y^{EI,n}} < \infty$.

We then connect this result back to the sampling rates. Recall that, for the modified EI policy, the constant c_x in (5) depends on the distance between μ_x and the reference value $\max_y \mu_y$. Then, in the special case of the design $x^* = \arg\max_x \mu_x$ (and not for any other design), we always have $c_x = 0$. It follows that $\bar{v}_{x^*}^{EI,n} = \frac{\lambda_{x^*}}{N_{x^*}^n} f(0)$, and the declining behaviour of this quantity does not depend on f. The design x^* is the only design for which this is the case. As a consequence, under the modified EI policy, the sampling rate is always higher for x^* than for any other alternative.

Proposition 3 (Ryzhov 2015) For $x \neq x^*$, $\lim_{n\to\infty} \frac{N_x^n}{n} = 0$.

To prove this result, it is sufficient to consider an example with two designs, x and x^*. The convergence rate of $\bar{v}_{x^*}^{EI,n}$ does not depend on f, but the rate of $\bar{v}_x^{EI,n}$ does. Consequently, $\bar{v}_x(z)$ vanishes faster than $\bar{v}_{x^*}(z)$ for comparable values of z, so we must make substantially more measurements of x^* in order for the two EI criteria to decline at the same rate. By contrast, for any two suboptimal designs $x,y \neq x^*$, this is never the case.

Proposition 4 (Ryzhov 2015) For $x,y \neq x^*$, $\liminf_{n\to\infty} \frac{N_x^n}{N_y^n} > 0$.

Thus, any accumulation point of the sequence $\left(\frac{N_x^n}{N_y^n}\right)$ must be strictly positive and finite. In fact, there is only one accumulation point, i.e., the sampling ratio has a limit. Again, we do not give the proof here, but we provide the intuition. If $c_x \neq c_y$, but the EI criteria for x and y decline at the same rate, the relationship between N_x^n and N_y^n should satisfy

$$c_x\sqrt{N_x^n} \approx c_y\sqrt{N_y^n}.$$

The theoretical limit of the sampling ratio confirms this intuition.

Theorem 1 (Ryzhov 2015) For $x,y \neq x^*$, $\lim_{n\to\infty} \frac{N_x^n}{N_y^n} = \left(\frac{c_y}{c_x}\right)^2$.

3.2 Deterministic Modification of the KG Policy

The extension of our analysis to the KG policy is straightforward. First, we define a modified KG criterion

$$\bar{v}_x^{KG,n} = \tilde{\sigma}_x^n f\left(-\frac{\left|\mu_x - \max_{y\neq x}\mu_y\right|}{\tilde{\sigma}_x^n}\right), \tag{7}$$

by analogy with (5). From here we proceed as before, but with two main differences. First, $\tilde{\sigma}_x^n$ is $\mathscr{O}\left(\frac{1}{N_x^n}\right)$, whereas σ_x^n in the EI policy is $\mathscr{O}\left(\frac{1}{\sqrt{N_x^n}}\right)$. Second, the reference value in (7) is now $\max_{y\neq x}\mu_y$, rather than the maximum over all y as in the EI policy. Consequently, Proposition 3 no longer holds, and the limit of the sampling ratio, for any two designs x,y, becomes

$$\lim_{n\to\infty}\frac{N_x^n}{N_y^n} = \frac{c_y}{c_x},$$

where

$$c_x = \frac{\left|\mu_x - \max_{y\neq x}\mu_y\right|}{\lambda_x}. \tag{8}$$

Let x' be the design with the second-highest value. Both x^* and x' have the same numerator in (8), so

$$\lim_{n\to\infty}\frac{N_{x^*}^n}{N_{x'}^n} = \frac{\lambda_{x^*}}{\lambda_{x'}}. \tag{9}$$

3.3 Connection to OCBA

We are now able to make a surprising connection between EI and OCBA methods. First, consider the modified EI policy of Section 3.1. From Theorem 1, we find that, for $x,y \neq x^*$, we have

$$\frac{N_x^n}{N_y^n} \to \frac{\lambda_x^2\left(\mu_y - \mu_*\right)^2}{\lambda_y^2\left(\mu_x - \mu_*\right)^2}, \tag{10}$$

where $\mu_* = \max_y \mu_y$ is the value of design x^*. The limit in (10) is identical to the sampling ratio for two suboptimal designs used by OCBA in independent normal R&S (see Thm. 3.2 of Chen and Lee 2010). The only difference between the two methods lies in their treatment of x^*: EI asymptotically allocates almost every simulation to x^* (Proposition 3), whereas OCBA uses a computational formula for $N_{x^*}^n$ in terms of the other N_x^n. In practice, however, OCBA also tends to assign a much larger portion of the budget to x^*.

The OCBA calculations are based on an approximation of the performance metric (probability of correct selection). However, Glynn and Juneja (2004) derives the exact optimal allocation for independent normal R&S. Letting $p_x = \lim_{n\to\infty}\frac{N_x^n}{n}$ be the limiting proportion of the budget allocated to design x, the optimal proportions for $x,y \neq x^*$ satisfy

$$\frac{\left(\mu_x - \mu_*\right)^2}{\left(\frac{\lambda_{x^*}^2}{p_{x^*}} + \frac{\lambda_x^2}{p_x}\right)} = \frac{\left(\mu_y - \mu_*\right)^2}{\left(\frac{\lambda_{x^*}^2}{p_{x^*}} + \frac{\lambda_y^2}{p_y}\right)}.$$

An expression is then derived for p_{x^*} in terms of the other p_x. However, when $p_{x^*} \gg p_x$ for all $x \neq x^*$, we again obtain the sampling ratios

$$\frac{p_x^*}{p_y^*} \approx \frac{\lambda_x^2\left(\mu_y - \mu_*\right)^2}{\lambda_y^2\left(\mu_x - \mu_*\right)^2}.$$

In our case, Proposition 3 means precisely that x^* receives a much greater proportion of the budget than any other design, thus explaining the sampling ratios for suboptimal alternatives. Pasupathy et al. (2014) discusses conditions under which it is optimal to allocate most of the budget to x^*, and finds that this can happen when the number of systems becomes large.

Next, we discuss the modified KG policy of Section 3.2. The sampling ratios for this policy do not match the OCBA ratios in general. However, consider a special case with only two designs, x and y. From (9), we have

$$\frac{N_x^n}{N_y^n} \to \frac{\lambda_x}{\lambda_y},$$

which is precisely the OCBA ratio for independent normal R&S with two designs (see Remark 3.1 in Chen and Lee 2010). To see the intuition behind this, note that the original form of the KG policy defined in (5) is provably optimal when there are only two designs (see Thm. 7.2 in Frazier, Powell, and Dayanik 2008). We would therefore expect it to recover the optimal sampling ratio in that special case. This has also been observed by Chick, Branke, and Schmidt (2010).

To summarize, we have shown connections between the OCBA methodology and two (modified) variants of expected improvement. The KG version is identical to OCBA when there are only two designs, matching what we know about the optimality of the KG method in that setting. The EI version asymptotically allocates most of the budget to the best alternative, and recovers the optimal sampling ratios for all suboptimal alternatives.

3.4 Returning to the Stochastic Case

It remains to connect these results back to the original forms of the methods, which make calculations based on the sample means θ_x^n rather than the true values. Return to the original form of the EI criterion in (4), and fix a sample path ω. For convenience, we will continue to view μ_x as a fixed constant (though it can be made to explicitly depend on ω). The quantity c_x is defined as in (5). Fix $\varepsilon > 0$ and define

$$\nu_x(z,\varepsilon) = \frac{\lambda_x}{\sqrt{z}} f\left(-(c_x+\varepsilon)\sqrt{z}\right),$$

a perturbed version of (6). For sufficiently small ε, we can find some N_ω such that

$$\nu_x(N_x^n(\omega),\varepsilon) \leq \nu_x^{EI,n}(\omega) \leq \nu_x(N_x^n(\omega),-\varepsilon) \tag{11}$$

for all $n \geq N_\omega$. This follows from the almost sure convergence of θ_x^n to μ_x under the EI policy. Thus, on the sample path ω, there is a point in time after which the stochastic EI criterion $\nu_x^{EI,n}(\omega)$ is bounded above and below by quantities that do not directly depend on $\theta_x^n(\omega)$, only through $N_x^n(\omega)$.

We can then develop similar bounds for the ratio $\frac{N_x^n(\omega)}{N_y^n(\omega)}$ for $x,y \neq x^*$. Recall that x tends to be chosen for simulation *more* frequently if its EI criterion declines *slower*. Suppose that the simulation output behaves in a way that will cause x to be chosen as frequently as possible. This will happen if $\nu_x^{EI,n}(\omega)$ is always equal to its upper bound in (11), while $\nu_y^{EI,n}(\omega)$ for any $y \neq x$ is always equal to its lower bound. Similarly, x would be chosen as infrequently as possible if $\nu_x^{EI,n}(\omega)$ would always equal its lower bound in (11), while $\nu_y^{EI,n}(\omega)$ would always equal its upper bound.

However, in these extreme cases, the EI criteria would decline in the same way as the modified EI criteria from (5). In the case where x is measured as frequently as possible, we would repeat the analysis of Section 3.1, but replace c_x and c_y by $c_x - \varepsilon$ and $c_y + \varepsilon$, respectively. When x is measured as infrequently as possible, we use $c_x + \varepsilon$ and $c_y - \varepsilon$ instead of c_x and c_y. For large enough n, we obtain the bounds

$$\left(\frac{c_y-\varepsilon}{c_x+\varepsilon}\right)^2 \leq \frac{N_x^n(\omega)}{N_y^n(\omega)} \leq \left(\frac{c_y+\varepsilon}{c_x-\varepsilon}\right)^2.$$

Taking $\varepsilon \to 0$ recovers the OCBA ratios.

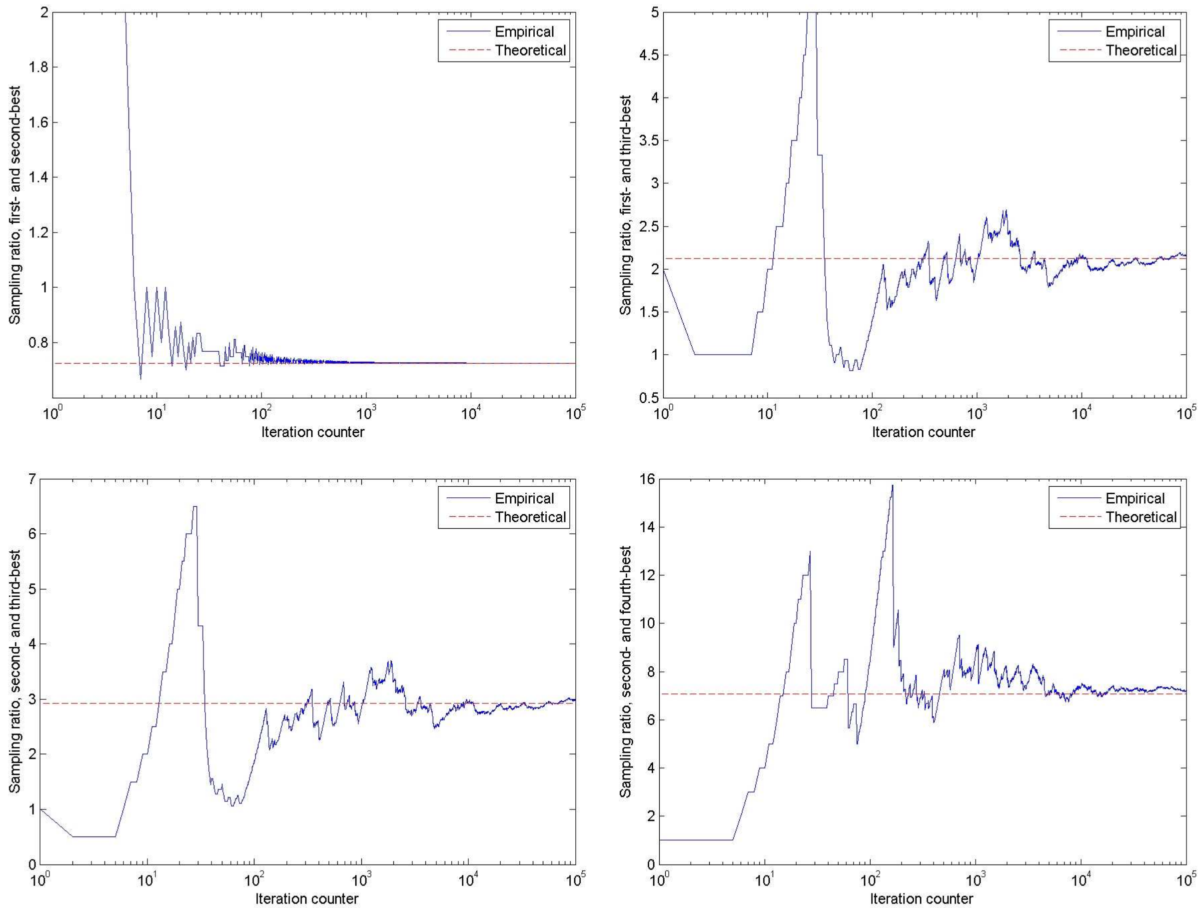

Figure 2: Empirical sampling ratios for different alternatives over 10^5 iterations.

4 NUMERICAL EXAMPLE

We give a numerical example to illustrate our results. Consider a problem with 10 alternatives where $\theta_x^0 = 0$ for every x. The prior variances $\left(\sigma_x^0\right)^2$ and the noise variances λ_x^2 are generated independently from a uniform distribution on the interval $[1,2]$. At the beginning of the experiment, a single value μ_x is generated from the distribution $\mathcal{N}\left(\theta_x^0, \left(\sigma_x^0\right)^2\right)$ for each x. These values are fixed for the remainder of the experiment; we use them to generate the simulation output, but they are not revealed to the simulation allocation policy.

We implement the KG policy defined in (5), which uses sample means to calculate expected improvement. Since KG asymptotically allocates a non-zero proportion of the budget to every alternative (unlike EI), the convergence of the sampling ratios can be observed more easily in 10^5 simulations. Figure 2 shows the empirical convergence of four ratios involving various combinations of the four highest-valued designs.

In this example, we have simulated only a single sample path to illustrate that convergence to the theoretical limiting ratios holds almost surely. However, the speed with which this occurs is highly problem-dependent. Table 1 gives the number of simulations allocated to each design in the numerical example. The four best designs were 10, 7, 9, and 1. Together, these four designs received over 85% of the budget. For some pairs of alternatives (such as 2 and 7), the limiting ratios are quite large, and it may take longer for the corresponding empirical ratios to converge.

Table 1: Number of simulations allocated to each design in the numerical example.

x	N_x	x	N_x
1	5391	6	2716
2	1425	7	38792
3	2435	8	1845
4	3254	9	13053
5	2943	10	28146

5 CONCLUSION

We have presented new theoretical results demonstrating a connection between expected improvement and OCBA methods. One well-known variant of EI asymptotically achieves the OCBA sampling ratios for all suboptimal alternatives, while a second variant does the same in the special case with only two designs. We believe that these results provide insight into the good practical performance of EI methods under normality assumptions, since the theoretical OCBA ratios optimize the rate of convergence in R&S.

However, some caveats are in order. First, the rate of convergence is optimized when the exact OCBA ratios are used, but this is not guaranteed if those ratios are only achieved asymptotically. Second, it is likely that the normality assumptions play a key role in the equivalence, since EI-type methods may not even be consistent in non-normal settings (Ding and Ryzhov 2015). Nonetheless, normality assumptions remain widely used in the literature and in practice, and it may be possible to obtain similar results for other learning problems where such assumptions are made. One avenue for future research may be to consider, e.g., subset selection problems, where OCBA methods are computationally tractable (Chen et al. 2008).

REFERENCES

Bull, A. D. 2011. "Convergence rates of efficient global optimization algorithms". *Journal of Machine Learning Research* 12:2879–2904.

Chau, M., M. C. Fu, H. Qu, and I. O. Ryzhov. 2014. "Simulation optimization: a tutorial overview and recent developments in gradient-based methods". 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, 21–35. Piscataway, NJ: IEEE, Inc.

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. Springer.

Chen, C.-H., M. C. Fu, and L. Shi. 2008. "Simulation and optimization". In *INFORMS TutORials in Operations Research*, edited by Z.-L. Chen and S. Raghavan, 247–260. INFORMS Hanover, MD.

Chen, C.-H., D. He, M. C. Fu, and L. H. Lee. 2008. "Efficient simulation budget allocation for selecting an optimal subset". *INFORMS Journal on Computing* 20 (4): 579–595.

Chen, C.-H., and L. H. Lee. 2010. *Stochastic simulation optimization: an optimal computing budget allocation*. World Scientific.

Chen, C.-H., J. Lin, E. Yücesan, and S. E. Chick. 2000. "Simulation budget allocation for further enhancing the efficiency of ordinal optimization". *Discrete Event Dynamic Systems* 10 (3): 251–270.

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.

DeGroot, M. H. 1970. *Optimal statistical decisions*. John Wiley and Sons.

Ding, Z., and I. O. Ryzhov. 2015. "Optimal learning with non-Gaussian rewards". *Advances in Applied Probability (to appear)*.

Frazier, P. I., W. B. Powell, and S. Dayanik. 2008. "A knowledge gradient policy for sequential information collection". *SIAM Journal on Control and Optimization* 47 (5): 2410–2439.

Gao, S., and L. Shi. 2014. "An optimal opportunity cost selection procedure for a fixed number of designs". 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, 3952–3958. Piscataway, NJ: IEEE, Inc.

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, NJ: IEEE, 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. Fu, 4062–4069. Piscataway, NJ: IEEE, Inc.

He, D., S. E. Chick, and C.-H. Chen. 2007. "Opportunity cost and OCBA selection procedures in ordinal optimization for a fixed number of alternative systems". *IEEE Transactions on Systems, Man, and Cybernetics* C37 (5): 951–961.

Jones, D. R., M. Schonlau, and W. J. Welch. 1998. "Efficient global optimization of expensive black-box functions". *Journal of Global Optimization* 13 (4): 455–492.

Kim, S.-H. 2013. "Statistical Ranking and Selection". In *Encyclopedia of Operations Research and Management Science*, edited by S. I. Gass and M. C. Fu, 1459–1469. Springer.

Kim, S.-H., and B. L. Nelson. 2006. "Selecting the best system". In *Handbooks of Operations Research and Management Science, vol. 13: Simulation*, edited by S. G. Henderson and B. L. Nelson, 501–534. North-Holland Publishing, Amsterdam.

Kim, S.-H., and B. L. Nelson. 2007. "Recent advances in ranking and selection". In *Proceedings of the 2007 Winter Simulation Conference*, edited by S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 162–172.

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.

Powell, W. B., and I. O. Ryzhov. 2012. *Optimal learning*. John Wiley and Sons.

Qu, H., I. O. Ryzhov, and M. C. Fu. 2012. "Ranking and selection with unknown correlation structures". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, 144–155.

Ryzhov, I. O. 2015. "On the convergence rates of expected improvement methods". *Working paper, University of Maryland*.

Ryzhov, I. O., W. B. Powell, and P. I. Frazier. 2012. "The knowledge gradient algorithm for a general class of online learning problems". *Operations Research* 60 (1): 180–195.

AUTHOR BIOGRAPHY

ILYA O. RYZHOV is an Assistant Professor of Operations Management and Management Science in the Decision, Operations and Information Technologies department of the Robert H. School of Business at the University of Maryland. He received a Ph.D. in Operations Research and Financial Engineering from Princeton University in 2011. His research mostly focuses on simulation optimization and optimal learning, with applications in business analytics. He coauthored the book *Optimal Learning* (Wiley, 2012), received the Best Theoretical Paper Award at the 2012 Winter Simulation Conference, and was a finalist in the 2014 INFORMS Junior Faculty Forum paper competition.