Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

CONSTRUCTING SIMULATION OUTPUT INTERVALS UNDER INPUT UNCERTAINTY VIA DATA SECTIONING

Peter W. Glynn

Department of MS&E
Stanford University
Huang Engineering Center 359A
Stanford, CA 94305, USA

Henry Lam

Department of IEOR
Columbia University
500 W. 120th Street
New York, NY 10027, USA

ABSTRACT

We study the problem of constructing confidence intervals (CIs) for simulation outputs in the presence of input uncertainty, where the constructed CIs capture both the statistical noises from the simulation replications and the input data. We present a simple technique based on sectioning input data that provides exact asymptotic confidence guarantees. Unlike some existing approaches, our technique bypasses the need to consistently estimate variances that could be computationally demanding. It can be flexibly applied to dependent data and to both parametric and nonparametric input models.

1 INTRODUCTION

Stochastic simulation consists of repeated generation of random variates from input models, fed through the system logic, to estimate system outputs. When the input models are not fully known but estimated from relevant historical data (the "input data"), valid output analysis needs to account for the input statistical noise, in addition to the Monte Carlo noise from the random variate generation. This issue, sometimes known as the input uncertainty, has been actively studied in recent years (see, e.g., the surveys Barton et al. 2002; Henderson 2003; Chick 2006; Barton 2012; Song et al. 2014; Lam 2016).

This paper focuses on a standard output-analytic task of constructing confidence intervals (CIs) for target performance measures under input uncertainty. When only simulation noise is considered, these CIs can be constructed by conventional techniques based on, e.g., asymptotic normality (Asmussen and Glynn 2007 Chapter 3). When input noise is substantial, this conventional construction will under-cover the true measure, and the CIs need correspondingly inflated to retain enough coverage. In the literature, several approaches have been studied to approximate this inflation. First is to utilize the delta method, which requires estimating the sensitivities of the performance measure with respect to the input parameters and combining them with the parameter sampling variance (e.g., Cheng and Holland 2004; Lin et al. 2015). Second is to bootstrap the input data and use them to drive simulation runs to obtain bootstrapped quantiles (Barton and Schruben 1993; Barton and Schruben 2001) or variance (Cheng and Holland 1997). The third approach follows a Bayesian perspective that samples from the posterior distributions of the input parameters (e.g., Glynn 1986; Chick 2001; Zouaoui and Wilson 2004; Xie et al. 2014; Biller and Corlu 2011). These approaches are closely related and some of them can be integrated with metamodeling to speed up computation (Barton et al. 2013).

While these methods can be effective in some situations, they face computational challenges in terms of the required simulation efforts, unless more restrictive assumptions are imposed. The root of these challenges lies in the need to consistently estimate key quantities in the CI, such as the parameter sensitivities in the delta method, and the variance or quantile in the bootstrap. This estimation typically requires heavy simulation effort to "outwash" the Monte Carlo noise from the input noise. Moreover, along this effort comes the

need to devise simulation allocation rules or procedural tuning, e.g., the distribution of nested simulation efforts to the resampling of the input data and the simulation replications per resample in the bootstrap, and the selection of the finite-difference perturbation parameter in the delta method. These procedural configurations can determine the performance of the CIs in a delicate fashion and are not always easy to optimize. Approaches that avoid these challenges, on the other hand, are based on additional, parametric assumptions. These include, e.g., Gaussian assumptions on the response surface in metamodel-assisted bootstrap and on the simulation outputs in Bayesian approaches.

Motivated by these challenges, this paper proposes a simple alternative to construct CIs under minimal simulation requirements and structural assumptions. Our method is based on sectioning the input data, where each input section drives a batch of simulation runs which are then integrated via a t-statistic to construct the CI. Our distinction is that, since using the involved t-statistic does not hinge on a consistent estimation of variance, we no longer need an overwhelming simulation effort to deconvolute the Monte Carlo noise from the input noise. As long as the output estimate in each section is approximately normal (with a variance including both sources of noise which could be unknown), we can construct a pivotal t-statistic that leads to a valid CI. This applies to both parametric and nonparametric input regimes, and to dependent data under suitable mixing conditions.

Sectioning and related techniques (on output samples) have been widely used in simulation output analysis, such as in steady-state estimation (e.g., Glynn and Iglehart 1990) and estimating nonlinear performance measures, e.g., quantiles, conditional value-at-risk (Asmussen and Glynn 2007; Muñoz and Glynn 1997; Nakayama 2014). In the input uncertainty literature, these are related to the so-called direct resampling in Barton and Schruben (2001) that divides input data and uses the quantiles of the resulting simulation runs to construct CIs. However, this latter procedure is heuristic and serves to motivate the bootstrap schemes suggested by the authors.

In the remainder of this paper, we will give an overview of our approach (Section 2), present our main assumptions to justify a joint central limit theorem that underpins our t-asymptotic (Section 3), and discuss our main performance guarantees (Section 4). Due to space limit, we only give brief discussion on some of our results, and leave further details and other generalizations, including a variant of our approach using "sectioned jackknife", to a full journal paper.

2 OVERVIEW OF OUR APPROACH

Consider estimating a simulatable performance measure driven by s input models. Notationally, we denote this performance measure as $\psi = \psi(\boldsymbol{\nu})$ where $\boldsymbol{\nu} = (\nu^1, \ldots, \nu^s)$, with ν^j each representing the "parameter" of the j-th input model. Given any $\tilde{\boldsymbol{\nu}}$, $\psi(\tilde{\boldsymbol{\nu}})$ can be approximated through simulation runs. As a generic example, ν^1 and ν^2 can denote the rates of exponential interarrival and service time distributions respectively, and $\psi(\nu^1, \nu^2)$ the expected average queue length over some time horizon of a queue that can be simulated. In this example, $\boldsymbol{\nu}$ denote real-valued parameters. However, our framework is equally applicable in the nonparametric regime, in which case $\boldsymbol{\nu}$ can denote the collection of the whole probability distributions for the input models, e.g., the interarrival and service time distributions, without restricting to any parametric family.

We consider the situations where $\boldsymbol{\nu}$ needs to be estimated from input data, say the estimate is $\widehat{\boldsymbol{\nu}}$. Assuming ψ is sufficiently smooth, the delta method implies that

$$\psi(\widehat{\boldsymbol{\nu}}) - \psi(\boldsymbol{\nu}) \approx N\left(0, \frac{\sigma^2}{n}\right) \tag{1}$$

where n is some scaling of the overall sample size (the reason for using a "scaling" is that there can be different sample sizes for different input models). The quantity σ^2/n is the variance contributed from all the input noises.

On top of these input noises, the output bears the Monte Carlo noise incurred in simulating ψ. We denote $\hat{\psi}(\widehat{\boldsymbol{\nu}})$ as a simulation-based estimate of $\psi(\widehat{\boldsymbol{\nu}})$ using input models calibrated with $\widehat{\boldsymbol{\nu}}$. We focus on the typical

setting where the simulation replications are conditionally unbiased given $\boldsymbol{\nu}$, with $\hat{\psi}(\widehat{\boldsymbol{\nu}}) = (1/r)\sum_{l=1}^{r} \hat{\psi}_l(\widehat{\boldsymbol{\nu}})$ and $\hat{\psi}_l(\widehat{\boldsymbol{\nu}}), l = 1,\ldots,r$ being r i.i.d. copies of simulation replications. Under a central limit theorem (CLT), we have

$$\hat{\psi}(\widehat{\boldsymbol{\nu}}) - \psi(\widehat{\boldsymbol{\nu}}) \approx N\left(0, \frac{\tau^2}{r}\right) \tag{2}$$

where τ^2/r is the variance from the simulation noise, using the implicit property that $\widehat{\boldsymbol{\nu}}$ well approximates $\boldsymbol{\nu}$ in some probabilistic sense.

We then combine (1) and (2) to obtain

$$\hat{\psi}(\widehat{\boldsymbol{\nu}}) - \psi(\boldsymbol{\nu}) \approx N\left(0, \frac{\sigma^2}{n} + \frac{\tau^2}{r}\right) \tag{3}$$

where the variance in the normal approximation is the sum of variances contributed from the input noise and the simulation noise.

Existing approaches in constructing CIs rely on estimating σ^2/n, either by bootstrapping or by direct estimation (They also estimate τ^2/r, which is more straightforward). When using the bootstrap, consistent estimation of σ^2/n typically requires simulation load that is of higher order than n (this can be seen by tracing the mean square error of the bootstrapped variance using, e.g., the formula in Sun et al. 2011). Direct estimation of σ^2 via the delta method, on the other hand, would require estimating sensitivity coefficients whose involved effort grows with the number of parameters (e.g., Cheng and Holland 1997), unless one uses conservative approximations (e.g., Cheng and Holland 2004). The challenge becomes more substantial in the nonparametric case.

The key of our idea is to bypass the consistent estimation of σ^2/n by sectioning the input data. Namely, we divide the input data into k sections (of equal length). Using each section, we estimate $\widehat{\boldsymbol{\nu}}$ and run independent simulation replications to obtain an estimate $\hat{\psi}(\widehat{\boldsymbol{\nu}})$, say they are $Y_1,\ldots,Y_k$. We then construct a pivotal statistic

$$\frac{\bar{Y} - \psi(\boldsymbol{\nu})}{S/\sqrt{k}} \tag{4}$$

where $\bar{Y}$ and S^2 are the sample mean and variance of $Y_1,\ldots,Y_k$. If each section contains enough data and drives enough simulation replications, the Y_i's are approximately normal according to (3) and moreover are i.i.d. (exactly if the data are i.i.d., and approximately if the data are serially dependent). Thus, by using a standard relation of normal variables with t-statistics, (4) is distributed approximately as t_{k-1}, the t-distribution with degree of freedom $k-1$. This implies that

$$\left[\bar{Y} - t_{k-1,1-\alpha/2}\frac{S}{\sqrt{k}}, \bar{Y} + t_{k-1,1-\alpha/2}\frac{S}{\sqrt{k}}\right]$$

is a $(1-\alpha)$-level CI, where $t_{k-1,1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of t_{k-1}. Constructing this CI does not require any knowledge about variances σ^2/n and τ^2/r, because the term $\sigma^2/n + \tau^2/r$ is "canceled out" when forming the pivotal statistic (4).

Regarding the assumptions needed in the above derivation, we essentially only need the normal approximation (3), a property satisfied for a wide range of settings when n and r are sufficiently large, including cases where the input data are dependent under suitable mixing conditions. We note that approximation (3) has been used in some previous work (e.g., Cheng and Holland 1997). Since it is the basis for our sectioning method, we first study in detail the conditions that guarantee such an approximation.

3 CENTRAL LIMIT THEOREM FOR JOINT INPUT AND SIMULATION NOISES

This section presents our assumptions and justifies in more details the central limit theorem (3). We divide our discussion into parametric and nonparametric regimes, which are conceptually similar but involve different sets of assumptions/notations.

We recall that there are s input models. For each input model j, we have an available data set. Our approach involves dividing each of these data sets into k sections. In other words, letting $\{X_l^j\}_{l=1,\ldots,kn^j}$ (where $X_l^j \in \mathscr{X}^j$ for some space $\mathscr{X}^j$) be the full data set for input model j, we batch the data into $\{X_l^j\}_{l=(i-1)n^j+1,\ldots,in^j},\ i=1,\ldots,k$ each of size n^j. We call the collection of data $\{X_l^j\}_{l=(i-1)n^j+1,\ldots,in^j},\ j=1,\ldots,s$ the i-th section. Throughout our exposition, we ignore the issue of divisibility and assume for convenience that the data size for each input model is divisible by k.

3.1 Conditions in the Parametric Regime

We first discuss the parametric case. Assume a parametric family for each input model, with a parameter vector $\theta^j \in \mathbb{R}^{d^j}, j=1,\ldots,s$. Denote $d=\sum_{j=1}^s d^j$ as the dimension of the concatenated parameter vector. Consider a simulatable performance measure $\psi=\psi(\boldsymbol{\theta})$ where $\boldsymbol{\theta}=(\theta^1,\ldots,\theta^s)$. We make the following assumption regarding the parameter estimators:

Assumption 1 (Normality of parameter estimators) We can construct an estimator $\hat{\boldsymbol{\theta}}_i$ for $\boldsymbol{\theta}$ using section i. There is a scaling parameter m for the section sample size such that as m goes to ∞, we have

$$(\sqrt{m}(\hat{\boldsymbol{\theta}}_i-\boldsymbol{\theta}))_{i=1,\ldots,k} \Rightarrow (W_i)_{i=1,\ldots,k}$$

where $W_i \overset{i.i.d.}{\sim} N(0,\Sigma)$ for some covariance matrix $\Sigma \in \mathbb{R}^d$. □

To get a sense of the meaning of m and the quantity Σ, consider the simple example where each independent input model j is estimated from i.i.d. data of size $kn^j=kmp^j$, so that each section consists of $n^j=mp^j$ observations. Suppose each parameter vector $\theta^j \in \mathbb{R}^{d^j}$ is estimated with $\hat{\theta}_i^j$ such that $\sqrt{n^j}(\hat{\theta}_i^j-\theta^j) \Rightarrow N(0,\Sigma^j)$. This holds if we use, e.g., maximum likelihood estimator (MLE) under standard conditions, in which case Σ^j is the inverse of the Fisher information matrix (see, e.g., Serfling 2009 Section 4.2.2). In this case, if we fix p^j and let $m\to\infty$, Assumption 1 holds with $\Sigma=\mathrm{diag}((1/p^j)\Sigma^j, j=1,\ldots,s)$ where $\mathrm{diag}(\cdot)$ denotes the block diagonal matrix with entries $(\cdot)$. In other words, we have $\hat{\boldsymbol{\theta}}_i \overset{\text{approx.}}{\sim} N(\boldsymbol{\theta},\mathrm{diag}((1/n^j)\Sigma^j, j=1,\ldots,k))$ as n^j grow large for all j.

Assumption 1 holds more generally for data exhibiting cross-sectional dependence (where one can simply think of the collection of input models as a single giant input model), serially dependent data under stationarity and suitable mixing conditions, and combinations of these aforementioned situations.

Next we make an assumption on the function $\psi(\cdot)$:

Assumption 2 (Parametric smoothness of ψ) The function $\psi(\cdot)$ has a non-zero differential at $\boldsymbol{\theta}$. □

The differential defined for ψ in Assumption 2 means the function

$$\nabla\psi(\boldsymbol{\theta})'\mathbf{t},\ \mathbf{t}\in\mathbb{R}^d$$

where $\nabla\psi$ is the gradient of ψ, given that the partial derivatives of ψ all exist. As a direct consequence of Serfling (2009) Section 1.12.2, the following is a handy sufficient condition for Assumption 2:

Lemma 3 Suppose all the partial derivatives of ψ exist in a neighborhood of $\boldsymbol{\theta}$ and are continuous at $\boldsymbol{\theta}$. Moreover, $\nabla\psi(\boldsymbol{\theta})$ is not a zero vector. Then Assumption 2 holds. □

Roughly speaking, the existence and continuity of partial derivatives in Lemma 3 hold when the input densities are differentiable with respect to $\boldsymbol{\theta}$, and that an exchange of derivative and expectation (appearing typically in defining ψ) is valid; general conditions of the latter can be found in, e.g., L'Ecuyer (1995).

Assumptions 1 and 2 together imply asymptotic normality via the delta method (see, e.g., Serfling 2009 Section 3.3 Theorem A):

Lemma 4 Under Assumptions 1 and 2, we have

$$(\sqrt{m}(\psi(\hat{\boldsymbol{\theta}}_i)-\psi(\boldsymbol{\theta})))_{i=1,\ldots,k} \Rightarrow (Z_i)_{i=1,\ldots,k}$$

where $Z_i \overset{i.i.d.}{\sim} N(0,\sigma^2)$ and $\sigma^2 = \nabla\psi(\boldsymbol{\theta})'\Sigma\nabla\psi(\boldsymbol{\theta})$, with $\nabla\psi$ being the gradient of ψ that exists under Assumption 2. □

Next we state some assumptions about the simulation noise. Let $\hat{\psi}_l(\boldsymbol{\beta})$ be an unbiased simulation run (l is used as an index when running multiple replications) so that $E[\hat{\psi}_l(\boldsymbol{\beta})|\boldsymbol{\beta}] = \psi(\boldsymbol{\beta})$. Denote $\tau^2(\boldsymbol{\beta}) = Var(\hat{\psi}_l(\boldsymbol{\beta})|\boldsymbol{\beta})$ as the simulation variance for each replication conditional on the input parameter. Similarly, denote $\kappa_3(\boldsymbol{\beta}) = E[|\hat{\psi}_l(\boldsymbol{\beta}) - \psi(\boldsymbol{\beta})|^3|\boldsymbol{\beta}]$ as the conditional third-order absolute central moment. Let $\mathscr{N}_\eta(\boldsymbol{\theta}) = \{\boldsymbol{\beta} \in \mathbb{R}^d : \|\boldsymbol{\beta} - \boldsymbol{\theta}\|_2 < \eta\}$ be an η-neighborhood of $\boldsymbol{\theta}$.

Assumption 5 (Parametric simulation noise) We have

1. $\tau^2(\cdot)$ is continuous at $\boldsymbol{\theta}$ with $0 < \tau^2(\boldsymbol{\theta}) < \infty$.
2. $\kappa_3(\boldsymbol{\beta})$ is uniformly bounded for any $\boldsymbol{\beta} \in \mathscr{N}_\eta(\boldsymbol{\theta})$ for some $\eta > 0$. □

Sufficient conditions guaranteeing Assumption 5 are similar to the ones in Lemma 3, but needing a higher order moment condition on the performance function that guarantees the existence of τ^2 and κ_3.

3.2 Conditions in the Nonparametric Regime

We now turn to the nonparametric regime. Instead of using/assuming parametric families for the input distributions, we utilize the empirical distributions of the data as the input estimates. In this case, consider a performance measure $\psi = \psi(\mathbf{F})$ where $\mathbf{F} = (F^1, \ldots, F^s)$ denotes the vector of distribution functions, each for one of the s independent input models.

With collected data divided into k sections each of size n^j for model j (like before), we construct $\hat{\mathbf{F}}_i = (\hat{F}_i^1, \ldots, \hat{F}_i^s)$, the vector of empirical distributions obtained from section i, where

$$\hat{F}_i^j(\cdot) = \frac{1}{n^j} \sum_{l=(i-1)n^j+1}^{in^j} \delta_{X_l^j}(\cdot)$$

and $\delta_{X_l^j}(\cdot)$ denotes the delta measure at X_l^j.

For any linear map $\mathscr{L} = (L^1, \ldots, L^s)$ where $L^j : \mathscr{X}^j \to \mathbb{R}$, and any $\mathbf{G} = (G^1, \ldots, G^s)$ where each $G^j : \mathscr{X}^j \to \mathbb{R}$ is a probability distribution, we define

$$\langle \mathscr{L}, \mathbf{G} \rangle = \sum_{j=1}^{s} \int L^j dG^j.$$

Hence we have

$$\langle \mathscr{L}, \hat{\mathbf{F}}_i \rangle = \sum_{j=1}^{s} \frac{1}{n^j} \sum_{l=(i-1)n^j+1}^{in^j} L^j(X_l^j).$$

In contrast to the parametric case, here in the nonparametric setting we depict local assumptions (i.e., differentiability) in a probabilistic way. In other words, instead of separating the input estimation errors (similar to Assumption 1) and the smoothness property of ψ (similar to Assumption 2), we impose an assumption that enforces ψ applied on $\hat{\mathbf{F}}_i$ is linearizable with an error that is statistically small. This is described as follows:

Assumption 6 (Nonparametric input noise) Assume ψ is differentiable in the sense that

$$\psi(\hat{\mathbf{F}}_i) = \psi(\mathbf{F}) + \langle D\psi, \hat{\mathbf{F}}_i - \mathbf{F} \rangle + \xi_i$$

where $D\psi$ is a linear map that could depend on $\hat{\mathbf{F}}_i$ and $\mathbf{F}$, such that there exists a scaling parameter m with the property that as $m \to \infty$,

1.
$$(\sqrt{m}\langle D\psi, \hat{\mathbf{F}}_i - \mathbf{F}\rangle)_{i=1,\ldots,k} \Rightarrow (Z_i)_{i=1,\ldots,k}$$
where $Z_i \overset{i.i.d.}{\sim} N(0,\sigma^2)$ for some $\sigma^2 < \infty$.

2.
$$\sqrt{m}\xi_i \xrightarrow{p} 0 \text{ for all } i = 1,\ldots,k.$$

□

The linear map $D\psi$ in Assumption 6 can be viewed as a derivative with respect to the collection of distribution functions (hence the notation). The reason why we use such a probabilistic assumption is to avoid the introduction of a metric in defining the error in using $\hat{\mathbf{F}}_i$ to approximate $\mathbf{F}$, which can lead to the need of more regularity condition on ψ. Note that, in using the latter approach, the error term ξ_i depends on both the chosen metric and the continuity of ψ with respect to this metric. When a weak norm is used then $\hat{\mathbf{F}}_i \to \mathbf{F}$ (e.g., the supremum norm, in which case $\hat{\mathbf{F}}_i \to \mathbf{F}$ as a consequence of the Glivenko-Cantelli lemma), but more conditions are needed for the function ψ to be continuous (e.g., if $\psi = E[h(X)]$, then h needs to have bounded variation). The imposition of a probabilistic statement like Assumption 6 avoids this dilemma, and in many situations a direct verification of this assumption can be done.

For example, consider a finite-horizon performance measure $\psi(\mathbf{F}) = E_{\mathbf{F}}[h(X^j(t),\ j = 1,\ldots,s,\ t = 1,\ldots,T^j)]$ where h is some performance function, T^j is the time horizon for each input sequence, $X^j(t)$ are i.i.d. replications generated from independent input model j and $E_{\mathbf{F}}[\cdot]$ denotes the expectation under the collection of input distributions $\mathbf{F}$. Suppose that

$$E[h(X^j(l_t^j),\ j = 1,\ldots,s,\ t = 1,\ldots,T^j)^2] < \infty \tag{5}$$

where $l_t^j \geq 1$ can be any integers. Suppose also that the data size of each section for model j satisfies $n^j = mp^j$ for a scaling parameter m and $p^j > 0$ are fixed constants. Then one can verify directly that Assumption 6 holds.

Note that the parametric conditions in Section 3.1 can be translated into the nonparametric conditions in Assumption 6 under appropriate modifications. For instance, if $\psi(\boldsymbol{\theta})$ satisfies Assumption 2 and $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathbf{F})$ is differentiable with respect to $\mathbf{F}$ in a similar sense as Assumption 6, then ψ also satisfies Assumption 6 when viewed as a function of $\mathbf{F}$.

Next we state some assumptions about the simulation noise. Let $\hat{\psi}_l(\cdot)$ be a conditionally unbiased simulation replication for $\psi(\cdot)$, such that $E[\hat{\psi}_l(\mathbf{G})|\mathbf{G}] = \psi(\mathbf{G})$. Denote $\tau^2(\mathbf{G}) = Var(\hat{\psi}_l(\mathbf{G})|\mathbf{G})$ as the conditional variance, and for convenience we denote $\tau^2 = \tau^2(\mathbf{F})$. Similarly, denote $\kappa_3(\mathbf{G}) = E[|\hat{\psi}_l(\mathbf{G}) - \psi(\mathbf{G})|^3|\mathbf{G}]$ as the conditional third-order absolute central moment. We make the following assumption on the simulation noise:

Assumption 7 (Nonparametric simulation noise) We have

1. $\tau^2(\hat{\mathbf{F}}_i) \xrightarrow{p} \tau^2$ with $0 < \tau^2 < \infty$.
2. $P(\kappa_3(\hat{\mathbf{F}}_i) \leq M) \to 1$ for some fixed $M > 0$. □

Assumption 7 holds for, e.g., the finite-horizon performance measure discussed above under a higher order moment condition than in (5).

3.3 A Joint Central Limit Theorem

The following proposition describes a CLT held jointly between the input noise and the simulation noise:

Proposition 8 Consider a collection of estimators $\hat{\boldsymbol{\nu}}_i$, $i = 1,\ldots,k$ for $\boldsymbol{\nu}$ that satisfies

$$(\sqrt{m}(\gamma(\hat{\boldsymbol{\nu}}_i) - \gamma(\boldsymbol{\nu})))_{i=1,\ldots,k} \Rightarrow (Z_i)_{i=1,\ldots,k}$$

where $Z_i \overset{i.i.d.}{\sim} N(0,\sigma^2)$ for some function γ and $\sigma^2 > 0$. Consider another (possibly dependent) collection of estimators $\hat{\boldsymbol{\nu}}'_u$, $u = 1,\ldots,v$. Let

$$\hat{\psi}(\hat{\boldsymbol{\nu}}'_u) = \frac{1}{r_u}\sum_{l=1}^{r_u} \hat{\psi}_l(\hat{\boldsymbol{\nu}}'_u)$$

where $r_u = mq_u$ for some fixed $q_u > 0$, $\hat{\psi}_l(\hat{\boldsymbol{\nu}}'_u)$ are i.i.d. unbiased simulation runs for $\psi(\hat{\boldsymbol{\nu}}'_u)$ given $\hat{\boldsymbol{\nu}}'_u$, and are conditionally independent across $u = 1,\ldots,v$. Denote $\tau^2(\boldsymbol{\beta}) = Var(\hat{\psi}_l(\boldsymbol{\beta})|\boldsymbol{\beta})$ as the conditional simulation variance for each replication, and for convenience denote $\tau^2 = \tau^2(\boldsymbol{\nu})$. Similarly, denote $\kappa_3(\boldsymbol{\beta}) = E[|\hat{\psi}_l(\boldsymbol{\beta}) - \psi(\boldsymbol{\beta})|^3|\boldsymbol{\beta}]$ as the conditional third-order absolute central moment. Assume that $\tau^2(\hat{\boldsymbol{\nu}}'_u) \xrightarrow{p} \tau^2$ and $P(\kappa_3(\hat{\boldsymbol{\nu}}'_u) \le M) \to 1$ for $u = 1,\ldots,v$, as $m \to \infty$.

Then,

$$\left((\sqrt{m}(\gamma(\hat{\boldsymbol{\nu}}_i) - \gamma(\boldsymbol{\nu}))_{i=1,\ldots,k}, (\sqrt{m}(\hat{\psi}(\hat{\boldsymbol{\nu}}'_u) - \psi(\hat{\boldsymbol{\nu}}'_u)))_{u=1,\ldots,v}\right) \Rightarrow ((Z_i)_{i=1,\ldots,k}, (W_u)_{u=1,\ldots,v}) \tag{6}$$

where $Z_i \overset{i.i.d.}{\sim} N(0,\sigma^2)$ and $W_u \overset{i.i.d.}{\sim} N(0,\tau^2/q_u)$ are independent. □

The proposition applies whether $\boldsymbol{\nu}$ is a real-valued parameter vector $\boldsymbol{\theta}$ (parametric case) or the collection of probability distributions $\mathbf{F}$ (nonparametric case), or a mix of both. To apply Proposition 8 to justify our sectioning scheme, we will see that the function $\gamma(\cdot)$ can be taken either as the performance function $\psi(\cdot)$ or its linearization. The assumption on the simulation replications, namely $\tau^2(\hat{\boldsymbol{\nu}}'_u) \xrightarrow{p} \tau^2$ and $P(\kappa_3(\hat{\boldsymbol{\nu}}'_u) \le M) \to 1$, can be satisfied by using a law of large numbers under the presented assumptions in the parametric case and Assumption 7 directly in the nonparametric case.

Note that similar results hold if r_u is not of the same order as m. Suppose $r_u = \omega(m)$, then, under (otherwise) the same assumptions as Proposition 8, the limit in (6) can be replaced by $((Z_i)_{i=1,\ldots,k}, (0)_{u=1,\ldots,v})$. On the other hand, if $r_u = mpq_u$ where $p = o(m)$, then (6) is replaced by

$$\left((\sqrt{mp}(\gamma(\hat{\boldsymbol{\nu}}_i) - \gamma(\boldsymbol{\nu}))_{i=1,\ldots,k}, (\sqrt{mp}(\hat{\psi}(\hat{\boldsymbol{\nu}}'_u) - \psi(\hat{\boldsymbol{\nu}}'_u)))_{u=1,\ldots,v}\right) \Rightarrow ((0)_{i=1,\ldots,k}, (W_u)_{u=1,\ldots,v}).$$

The sectioning scheme we depict next works for these imbalanced cases with little modification.

4 SECTONING INPUT DATA FOR CONFIDENCE INTERVAL ESTIMATION

This section presents our sectioning approach to construct CIs. Section 4.1 states our procedure and shows the asymptotic confidence guarantee. Section 4.2 discusses other performance aspects of our proposed procedure.

4.1 Procedure and Main Guarantee

We revisit our procedure discussed in Section 2 using our introduced notations. To estimate $\psi = \psi(\boldsymbol{\nu})$ (where $\boldsymbol{\nu}$ can be $\boldsymbol{\theta}$ or $\mathbf{F}$), we divide the data set of each input model into k sections of equal length. For section i that consists of $\{X_l^j\}_{l=(i-1)n^j+1,\ldots,in^j}$, $j = 1,\ldots,s$, we estimate $\hat{\boldsymbol{\nu}}_i$ (can be $\hat{\boldsymbol{\theta}}_i$ or $\hat{\mathbf{F}}_i$), and then $Y_i = \hat{\psi}(\hat{\boldsymbol{\nu}}_i)$ where

$$\hat{\psi}(\hat{\boldsymbol{\nu}}_i) = \frac{1}{r}\sum_{l=1}^{r} \hat{\psi}_l(\hat{\boldsymbol{\nu}}_i).$$

Here $\hat{\psi}_l(\hat{\boldsymbol{\nu}}_i)$ are conditionally unbiased i.i.d. simulation replications for $\psi(\hat{\boldsymbol{\nu}}_i)$, and independent across i. Let $\bar{Y} = (1/k)\sum_{i=1}^k Y_i$ and $S^2 = (1/(k-1))\sum_{i=1}^s (Y_i - \bar{Y})^2$ be the sample mean and variance of the k sectioned estimates. We propose to use the CI

$$\mathscr{I}_{sec} = \left[\bar{Y} - t_{k-1,1-\alpha/2}\frac{S}{\sqrt{k}}, \bar{Y} + t_{k-1,1-\alpha/2}\frac{S}{\sqrt{k}}\right] \tag{7}$$

where $t_{k-1,1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of the t_{k-1} distribution. This construction holds as long as $k \geq 2$.

We have the following guarantees for (7). For this, we introduce the scaling $r = mq$, where q is fixed while m goes to ∞. We have:

Theorem 9 Under either:

1. (parametric case) Assumptions 1, 2 and 5; or
2. (nonparametric case) Assumptions 6 and 7,

we have

$$\frac{\bar{Y} - \psi}{S/\sqrt{k}} \Rightarrow t_{k-1} \tag{8}$$

for any fixed $k \geq 2$. Consequently, $\mathscr{I}_{sec}$ is an asymptotically exact $(1-\alpha)$-level CI for ψ, i.e.,

$$P(\psi \in \mathscr{I}_{sec}) \to 1-\alpha \quad \text{as } m \to \infty.$$

Proof. We first develop a limit theorem for $\sqrt{m}(Y_i - \psi)$ for both the parametric and the nonparametric case. Then we proceed to justify the t-asymptotic (8) and conclude the theorem.

Parametric case: We can write

$$Y_i = \psi + U_i + V_i$$

where $U_i = \psi(\hat{\boldsymbol{\theta}}_i) - \psi$ and $V_i = \hat{\psi}(\hat{\boldsymbol{\theta}}_i) - \psi(\hat{\boldsymbol{\theta}}_i)$. Under Assumptions 1 and 2, we use Lemma 4 to deduce that $\sqrt{m}(U_i)_{i=1,\ldots,k} \Rightarrow (Q_i)_{i=1,\ldots,k}$ where $Q_i \overset{i.i.d.}{\sim} N(0,\sigma^2)$ with $\sigma^2 = \nabla\psi(\boldsymbol{\theta})'\Sigma\nabla\psi(\boldsymbol{\theta})$. Using the notations in Assumption 5, a law of large numbers for $\hat{\boldsymbol{\theta}}_i$ implied from Assumption 1 and the local properties in Assumption 5 conclude that $\tau^2(\hat{\boldsymbol{\theta}}_i) \xrightarrow{p} \tau^2(\boldsymbol{\theta})$ and $P(\kappa_3(\hat{\boldsymbol{\theta}}_i) \leq M) \to 1$ for some $M > 0$. Hence Proposition 8 implies that$\sqrt{m}((U_i)_{i=1,\ldots,k}, (V_i)_{i=1,\ldots,k}) \Rightarrow ((Q_i)_{i=1,\ldots,k}, (W_i)_{i=1,\ldots,k})$ where $W_i \overset{i.i.d.}{\sim} N(0,\tau^2/q)$ and Q_i and W_i are all independent. Therefore,

$$\sqrt{m}(Y_i - \psi)_{i=1,\ldots,k} \Rightarrow (Z_i)_{i=1,\ldots,k}$$

where $Z_i \overset{i.i.d.}{\sim} N(0,\sigma^2+\tau^2/q)$.

Nonparametric case: Under Assumption 6, we write

$$Y_i = \psi + U_i + \xi_i + V_i$$

where $U_i = \langle D\psi, \hat{\mathbf{F}}_i - \mathbf{F}\rangle$, ξ_i defined as in Assumption 6, and $V_i = \hat{\psi}(\hat{\mathbf{F}}_i) - \psi(\hat{\mathbf{F}}_i)$. Assumption 6 states that $(\sqrt{m}U_i)_{i=1,\ldots,k} \Rightarrow (Q_i)_{i=1,\ldots,k}$ where $Q_i \overset{i.i.d.}{\sim} N(0,\sigma^2)$ for some $\sigma^2 < \infty$ and $\sqrt{m}\xi_i \xrightarrow{p} 0$ for all $i = 1,\ldots,k$. Note that Assumption 7 enforces $\tau^2(\hat{\mathbf{F}}_i) \xrightarrow{p} \tau^2(\mathbf{F})$ and $P(\kappa_3(\hat{\mathbf{F}}_i) \leq M) \to 1$ for some $M > 0$. Hence, Proposition 8 implies that$\sqrt{m}((U_i)_{i=1,\ldots,k}, (V_i)_{i=1,\ldots,k}) \Rightarrow ((Q_i)_{i=1,\ldots,k}, (W_i)_{i=1,\ldots,k})$ where $W_i \overset{i.i.d.}{\sim} N(0,\tau^2/q)$ and Q_i and W_i are all independent. Therefore, by Slutsky's theorem,

$$\sqrt{m}(Y_i - \psi)_{i=1,\ldots,k} \Rightarrow (Z_i)_{i=1,\ldots,k}$$

where $Z_i \overset{i.i.d.}{\sim} N(0,\sigma^2+\tau^2/q)$.

Finally, in both cases

$$\frac{\bar{Y} - \psi}{S/\sqrt{k}} \Rightarrow \frac{\bar{Z}}{\sqrt{(1/(k-1))\sum_{i=1}^k (Z_i - \bar{Z})^2/k}} \tag{9}$$

where $\bar{Z} = (1/k)\sum_{i=1}^{k} Z_i$, by the continuous mapping theorem. Note that, since Z_i are Gaussian, $\bar{Z}$ is independent of $Z_i - \bar{Z}$, and $\sum_{i=1}^{k}(Z_i - \bar{Z})^2 \sim \chi^2_{k-1}$ where χ^2_{k-1} is a χ^2 random variable with degree of freedom $k-1$. This shows that the limit in (9) is a t_{k-1} random variable. The last assertion in the theorem follows by a standard pivotal argument. □

4.2 Conservativeness and Variability of the Confidence Intervals

We discuss the performance of our sectioning-based CIs. Ideally, one would use as many sections as possible, as this would improve the length of the interval (in terms of reducing the average length and also its variability); in the limiting case, this would give rise to a CI constructed from normality asymptotic, which can be viewed as the best possible. However, with fixed data and simulation budgets, using more sections means a smaller data size and simulation replications per section, so that the quality of the Gaussian approximation within each section, and consequently the performance of the pivotal t-statistic, deteriorates. The number of sections and the size per section thus result in a tradeoff between desirable half-width properties and the statistical validity of the CIs.

We use the mean half-width to measure the conservativeness of the CI, and its standard deviation to measure its variability. In the following, we do a rough-cut analysis on the effect of section numbers on these two measurements, which reveals the minimal section numbers needed to achieve a certain performance compared to the benchmark normality CI. A full investigation of the optimal section number, however, would require analyzing the approximation error of the involved asymptotic t-limit.

Suppose that we use the full data set and all the simulation budget to get an estimate $\hat{\psi}(\hat{\boldsymbol{\nu}})$ where $\hat{\boldsymbol{\nu}}$ is the input estimate and $\hat{\psi}$ denotes the average of all simulation runs. Then, under the same assumption as in Theorem 9 (with k set to 1), we have, asymptotically, $\hat{\psi}(\hat{\boldsymbol{\nu}}) \sim N(\psi, \sigma_0^2 + \tau_0^2)$ for some σ_0^2 and τ_0^2, where σ_0^2 denotes the variance from the input noise and τ_0^2 the simulation noise. For convenience, we denote $\lambda^2 = \sigma_0^2 + \tau_0^2$, so that $\hat{\psi}(\hat{\boldsymbol{\nu}}) \sim N(\psi, \lambda^2)$ approximately. If we instead divide the data and the simulation budget into k sections, then for each section i our estimate becomes $\hat{\psi}(\hat{\boldsymbol{\nu}}_i)$, and they are asymptotically $\overset{i.i.d.}{\sim} N(\psi, k\lambda^2)$.

The half-width of $\mathscr{I}_{sec}$ in (7) is $t_{k-1,1-\alpha/2}S/\sqrt{k}$. Here S^2, the sample variance of the $\hat{\psi}(\hat{\boldsymbol{\nu}}_i)$'s, follows $k\lambda^2 \cdot \chi^2_{k-1}/(k-1)$ approximately. Therefore, the half-width H of $\mathscr{I}_{sec}$ is

$$H \approx \frac{t_{k-1,1-\alpha/2}}{\sqrt{k}} \cdot \sqrt{k\lambda^2 \frac{\chi^2_{k-1}}{k-1}} = t_{k-1,1-\alpha/2}\lambda \sqrt{\frac{\chi^2_{k-1}}{k-1}}.$$

Using the fact that $E\sqrt{\chi^2_{k-1}} = \sqrt{2}\Gamma(k/2)/\Gamma((k-1)/2)$, the expected half-width satisfies

$$EH \approx t_{k-1,1-\alpha/2}\lambda \frac{\Gamma(k/2)}{\Gamma((k-1)/2)} \sqrt{\frac{2}{k-1}}.$$

On the other hand, using $E\chi^2_{k-1} = k-1$, the second moment of the half-width satisfies

$$EH^2 \approx t^2_{k-1,1-\alpha/2}\lambda^2$$

so that the variance is

$$Var(H) \approx t^2_{k-1,1-\alpha/2}\lambda^2 \left(1 - \left(\frac{\Gamma(k/2)}{\Gamma((k-1)/2)}\right)^2 \frac{2}{k-1}\right).$$

It can be shown that $\Gamma(k/2)/\Gamma((k-1)/2)\sqrt{2/(k-1)} \to 1$ as $k \to \infty$. Combined with the fact that $t_{k-1} \Rightarrow N(0,1)$, the approximate $EH \to z_{1-\alpha/2}\lambda$, where $z_{1-\alpha/2}$ is the $1-\alpha/2$ standard normal quantile,

and the approximate $Var(H) \to 0$. For example, when $\alpha = 5\%$ (so that we are looking at the 95% CI), $EH/\lambda \approx 1.98$ at $k = 100$, compared with $z_{1-\alpha/2} = 1.96$. Note that the quantity $z_{1-\alpha/2}\lambda$ is the half-width of a normality CI. Therefore, as discussed above, the normality CI can be viewed as the best possible (Of course, to get this, one would need information about λ, which as discussed in the introduction can be challenging to obtain in the input uncertainty setting).

The approximate EH, EH^2 and $Var(H)$ all monotonically decrease as k increases. Fixing $\alpha = 5\%$, the decreases of all the considered quantities occur rapidly from $k = 2$ to 6 ($EH/\lambda \approx 10.1, 3.8, 2.9, 2.6, 2.4$ and $\sqrt{Var(H)}/\lambda \approx 7.7, 2.0, 1.2, 0.9, 0.8$ at $k = 2$ to 6), continues to decrease at a slower rate from $k = 6$ to 10 (e.g., $EH/\lambda \approx 2.2$ and $\sqrt{Var(H)}/\lambda \approx 0.5$ at $k = 10$), and afterwards the gain becomes quite negligible (e.g., $EH/\lambda \approx 2.0$ and $\sqrt{Var(H)}/\lambda \approx 0.2$ at $k = 50$). These numbers suggest that increasing k from 2 to 3 leads to a significant gain, whereas increasing k further to 4 to 6 can improve sensibly further. Depending on the data/simulation size, a reasonable k could be chosen in the range 3 to 6.

5 A NUMERICAL EXAMPLE

We test our approach with a numerical example of $M/M/1$ queue. In this queue, the true arrival rate is 1 and service rate is 1.2. We initiate from an empty system. We consider two performance measures: the expected waiting time averaged over the first 20 customers, denoted as $E_{waittime}$, and the expected time-averaged queue length over time period $[0,5]$, denoted as $E_{queuelength}$. The true values of $E_{waittime} = 1.53$ and $E_{queuelength} = 0.63$.

We use a data size 100 on both the interarrival and service times (i.e., these data are sampled from the exponential distributions with rates 1 and 1.2 respectively). We execute our sectioning scheme to construct 95% CIs for measures $E_{waittime}$ and $E_{queuelength}$, using various numbers of sections, but under a total simulation budget of 100 runs. In each setting, we repeat the calculation 1000 times, where each time we sample a new set of observations and construct a CI. From these repetitions, we estimate the coverage probability and the average CI length. We consider the nonparametric setting (i.e., using the empirical distributions for the interarrival and service times).

Table 1 shows the estimated coverage probabilities (both the point estimate and the 95% confidence interval) and the average CI lengths in estimating $E_{waittime}$ and $E_{queuelength}$ as the number of sections k varies (e.g., when $k = 2$, each section constitutes 50 observations for both the interarrival and service times and 50 replications). The coverages of both performance measures seem to be valid and quite stable until about $k = 14$, after which they start to deteriorate from the nominal 95%. This shows, at least in this example, that the coverage probability is relatively robust to the data size in each section. On the other hand, the average CI length drops significantly from $k = 2$ to 3, continues to drop from $k = 3$ to 6, and afterwards levels off. This shows a huge gain in reducing conservativeness from 2 to 3 sections, and further gains till about 6 sections. These observations reconcile with our discussion in Section 4.2.

ACKNOWLEDGMENTS

We gratefully acknowledge support from the National Science Foundation under grants CMMI-1542020 and CAREER CMMI-1653339/1834710.

REFERENCES

Asmussen, S., and P. W. Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis*, Volume 57. Springer Science & Business Media, New York.

Barton, R. R. 2012. "Tutorial: Input Uncertainty in Output Analysis". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque et al., 1–12. Piscataway, New Jersey: IEEE.

Barton, R. R., S. E. Chick, R. C. Cheng, S. G. Henderson, A. M. Law, B. W. Schmeiser, L. M. Leemis, L. W. Schruben, and J. R. Wilson. 2002. "Panel Discussion on Current Issues in Input Modeling".

Table 1: Nonparametric CIs using 100 observations on the interarrival times and 100 observations on the service times, with total simulation budget 100.

Section #	Coverage (wait time)	CI length (wait time)	Coverage (queue length)	CI length (queue length)
2	0.940 (0.925, 0.955)	10.024	0.936 (0.921, 0.951)	3.704
3	0.943 (0.929, 0.957)	3.995	0.947 (0.933, 0.961)	1.485
4	0.928 (0.912, 0.944)	3.034	0.946 (0.932, 0.960)	1.136
5	0.949 (0.935, 0.963)	2.745	0.952 (0.939, 0.965)	1.046
6	0.930 (0.914, 0.946)	2.558	0.945 (0.931, 0.959)	1.011
7	0.941 (0.926, 0.956)	2.553	0.941 (0.926, 0.956)	0.973
8	0.942 (0.927, 0.957)	2.451	0.960 (0.948, 0.972)	0.995
9	0.936 (0.921, 0.951)	2.423	0.939 (0.924, 0.954)	0.980
10	0.953 (0.940, 0.966)	2.430	0.949 (0.935, 0.963)	0.973
11	0.947 (0.933, 0.961)	2.362	0.950 (0.936, 0.964)	0.989
12	0.971 (0.961, 0.981)	2.465	0.963 (0.951, 0.975)	1.019
13	0.947 (0.933, 0.961)	2.513	0.943 (0.929, 0.957)	1.098
14	0.947 (0.933, 0.961)	2.397	0.949 (0.935, 0.963)	1.006
15	0.945 (0.931, 0.959)	2.551	0.936 (0.921, 0.951)	1.122
16	0.924 (0.908, 0.940)	2.429	0.928 (0.912, 0.944)	1.080
17	0.936 (0.921, 0.951)	2.628	0.918 (0.901, 0.935)	1.164
18	0.928 (0.912, 0.944)	2.512	0.908 (0.890, 0.926)	1.153
19	0.915 (0.898, 0.932)	2.454	0.900 (0.881, 0.919)	1.100
20	0.922 (0.905, 0.939)	2.398	0.900 (0.881, 0.919)	1.080

In *Proceedings of the 2002 Winter Simulation Conference*, edited by E. Yücesan et al., 353–369. Piscataway, New Jersey: IEEE.

Barton, R. R., B. L. Nelson, and W. Xie. 2013. "Quantifying Input Uncertainty via Simulation Confidence Intervals". *INFORMS Journal on Computing* 26(1):74–87.

Barton, R. R., and L. W. Schruben. 1993. "Uniform and Bootstrap Resampling of Empirical Distributions". In *Proceedings of the 1993 Winter Simulation Conference*, edited by G. W. Evans et al., 503–508. Piscataway, New Jersey: IEEE.

Barton, R. R., and L. W. Schruben. 2001. "Resampling Methods for Input Modeling". In *Proceedings of the 2001 Winter Simulation Conference*, edited by B. A. Peters et al., Volume 1, 372–378. Piscataway, New Jersey: IEEE.

Biller, B., and C. G. Corlu. 2011. "Accounting for Parameter Uncertainty in Large-Scale Stochastic Simulations with Correlated Inputs". *Operations Research* 59(3):661–673.

Cheng, R. C., and W. Holland. 1997. "Sensitivity of Computer Simulation Experiments to Errors in Input Data". *Journal of Statistical Computation and Simulation* 57(1-4):219–241.

Cheng, R. C., and W. Holland. 2004. "Calculation of Confidence Intervals for Simulation Output". *ACM Transactions on Modeling and Computer Simulation* 14(4):344–362.

Chick, S. E. 2001. "Input Distribution Selection for Simulation Experiments: Accounting for Input Uncertainty". *Operations Research* 49(5):744–758.

Chick, S. E. 2006. "Bayesian Ideas and Discrete Event Simulation: Why, What and How". In *Proceedings of the 2006 Winter Simulation Conference*, edited by L. F. Perrone et al., 96–106. Piscataway, New Jersey: IEEE.

Glynn, P. W. 1986. "Problems in Bayesian Analysis of Stochastic Simulation". In *Proceedings of the 1986 Winter Simulation Conference*, edited by J. Wilson et al., 376–379. New York, New York: ACM.

Glynn, P. W., and D. L. Iglehart. 1990. "Simulation Output Analysis Using Standardized Time Series". *Mathematics of Operations Research* 15(1):1–16.

Henderson, S. G. 2003. "Input Modeling: Input Model Uncertainty: Why Do We Care and What Should We Do about It?". In *Proceedings of the 2003 Winter Simulation Conference*, edited by S. Chick et al., 90–100. Piscataway, New Jersey: IEEE.

Lam, H. 2016. "Advanced Tutorial: Input Uncertainty and Robust Analysis in Stochastic Simulation". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder et al., 178–192. Piscataway, New Jersey: IEEE.

L'Ecuyer, P. 1995. "Note: On the Interchange of Derivative and Expectation for Likelihood Ratio Derivative Estimators". *Management Science* 41(4):738–747.

Lin, Y., E. Song, and B. Nelson. 2015. "Single-Experiment Input Uncertainty". *Journal of Simulation* 9(3):249–259.

Muñoz, D. F., and P. W. Glynn. 1997. "A Batch Means Methodology for Estimation of a Nonlinear Function of a Steady-State Mean". *Management Science* 43(8):1121–1135.

Nakayama, M. K. 2014. "Confidence Intervals for Quantiles Using Sectioning When Applying Variance-Reduction Techniques". *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 24(4):19.

Serfling, R. J. 2009. *Approximation Theorems of Mathematical Statistics*, Volume 162. John Wiley & Sons.

Song, E., B. L. Nelson, and C. D. Pegden. 2014. "Advanced Tutorial: Input Uncertainty Quantification". In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk et al., 162–176. Piscataway, New Jersey: IEEE.

Sun, Y., D. W. Apley, and J. Staum. 2011. "Efficient Nested Simulation for Estimating the Variance of a Conditional Expectation". *Operations Research* 59(4):998–1007.

Xie, W., B. L. Nelson, and R. R. Barton. 2014. "A Bayesian Framework for Quantifying Uncertainty in Stochastic Simulation". *Operations Research* 62(6):1439–1452.

Zouaoui, F., and J. R. Wilson. 2004. "Accounting for Input-Model and Input-Parameter Uncertainties in Simulation". *IIE Transactions* 36(11):1135–1151.

AUTHOR BIOGRAPHIES

PETER W. GLYNN is the Thomas Ford Professor in the Department of Management Science and Engineering (MS&E) at Stanford University. He is a Fellow of INFORMS and of the Institute of Mathematical Statistics, has been co-winner of Best Publication Awards from the INFORMS Simulation Society in 1993, 2008, and 2016, and was the co-winner of the John von Neumann Theory Prize from INFORMS in 2010. In 2012, he was elected to the National Academy of Engineering. His research interests lie in stochastic simulation, queueing theory, and statistical inference for stochastic processes. His email address is glynn@stanford.edu.

HENRY LAM is an Associate Professor in the Department of Industrial Engineering and Operations Research at Columbia University. His research focuses on Monte Carlo simulation, risk and uncertainty quantification, and stochastic optimization. His email address is khl2114@columbia.edu.