# Comparing Regenerative-Simulation-Based Estimators of the Distribution of the Hitting Time to a Rarely Visited Set

Peter Glynn, Marvin Nakayama, Bruno Tuffin

# Comparing Regenerative-Simulation-Based Estimators of the Distribution of the Hitting Time to a Rarely Visited Set

Peter W. Glynn

Department of Management Science and Engineering
Stanford University
475 Via Ortega
Stanford, CA 94305, USA

Marvin K. Nakayama

Department of Computer Science
New Jersey Institute of Technology
Newark, NJ 07102, USA

Bruno Tuffin

Inria, Univ Rennes, CNRS, IRISA
Campus de Beaulieu
35042 Rennes, FRANCE

## ABSTRACT

We consider the estimation of the distribution of the hitting time to a rarely visited set of states for a regenerative process. In a previous paper, we provided two estimators that exploited the weak convergence of the hitting time divided by its expectation to an exponential as the rare set becomes rarer. We now add three new estimators, based on a corrected exponential, a gamma, and a bootstrap approach, the last possibly providing less biased estimators when the rare set is only moderately rare. Numerical results illustrate that all of the estimators perform similarly. Although the paper focuses on estimating a distribution, the ideas may also be applied to estimate risk measures, such as a quantile or conditional tail expectation.

## 1 INTRODUCTION

The hitting time of a rarely visited set $\mathscr{A}$ of states is of interest in many areas, such as dependability and finance. While rare-event simulation is often used to study the hitting time (Rubino and Tuffin 2009b), most of the literature focuses on estimating its mean. But its quantiles or entire distribution can often provide more useful information. For example, if the hitting time represents the failure time of a product that a company plans to sell, a quantile of its distribution may be employed to define an appropriate warranty length. The time to ruin of an insurance company may be modeled as a hitting time, and we may be interested in its distribution; e.g., see Section 1.3.5 of Kalashnikov (1997). But the study of rare-event-simulation techniques for computing the distribution of hitting times has not received a lot of attention.

We now address this problem for a stochastic process having a regenerative structure; i.e., the process "probabilistically restarts" at an increasing sequence of regeneration times (Kalashnikov 1994). We then can express the mean hitting time as a ratio, which may be utilized to design efficient simulation methods; see Goyal et al. (1992), Nicola et al. (1993), and Glynn et al. (2017), and Rubino and Tuffin (2009b).

It is often the case (Kalashnikov 1997) that when the probability $p$ of hitting $\mathscr{A}$ before regeneration goes to zero, the distribution of the hitting time divided by its mean converges to an exponential with mean 1. We can exploit this property to approximate the distribution of the hitting time by an exponential, with the problem now being reduced to estimating the mean hitting time, which may be efficiently estimated, as noted above. Glynn et al. (2018) develop this idea to provide estimators of the distribution, quantile and conditional tail expectation (CTE), each requiring only an estimator of the mean hitting time.

Glynn et al. (2018) further propose a second family of methods, resulting in so-called convolution estimators. This approach uses the fact that the scaled sum of lengths of the cycles (i.e., the process between regenerations) before the one hitting $\mathscr{A}$ converges weakly (as $p$ shrinks) to an exponential, whose mean can be estimated. We then convolve the exponential with the distribution of the time to reach $\mathscr{A}$ within a cycle, which is estimated by importance sampling (IS; Asmussen and Glynn 2007, Chapter VI).

While the above estimators are both computationally efficient and can have small variance, a drawback is that they are also biased. The weak convergence on which they are based requires $p \to 0$. But when this probability is not very small, bias could potentially surpass variance. For this reason, we propose in this paper three other estimators, which may be more accurate: i) a convolution estimator for which the exponential approximation for the sum of lengths of cycles before the one hitting $\mathscr{A}$ is replaced by a refined approximation of Blanchet and Glynn (2007); ii) an estimator that substitutes the exponential approximation with one based on a gamma, giving one extra degree of freedom in the inference of the distribution, which may be helpful for moderate $p$; iii) a bootstrap estimator that does not use any distributional approximations but rather resamples simulated lengths of independent cycles. We focus on estimating the distribution of the hitting time, but quantile and CTE estimators can be derived as well. Our numerical experiments show that all of the distribution estimators are close in terms of performance.

The rest of the paper proceeds as follows. Section 2 sets the mathematical framework and reviews the exponential and convolution estimators of Glynn et al. (2018). The next three sections develop the new estimators, with the hope that they can reduce the error when the probability $p$ is not too small. Section 3 describes the modified convolution estimator based on a refinement of the exponential approximation by Blanchet and Glynn (2007). Section 4 explains the gamma estimator, and we present the bootstrap estimator in Section 5. Section 6 compares all estimators through numerical experiments on two standard regenerative processes (a model of highly reliable systems used in dependability analysis, and an M/M/1 queue) for which the exact values are known, so we can analyze the error. Section 7 gives concluding remarks and directions for future work.

## 2 MODEL AND PREVIOUS RESULTS

Consider a stochastic process $X = [X(t) : t \geq 0]$ evolving on a state space $\mathscr{S} \subseteq \Re^d$. The goal is to estimate the distribution/quantiles of the *hitting time* $T = \inf\{t \geq 0 : X(t) \in \mathscr{A}\}$ of a subset $\mathscr{A} \subset \mathscr{S}$ of states.

Assume process $X$ is regenerative, with $0 = \Gamma_0 < \Gamma_1 < \Gamma_2 < \cdots$ as the sequence of regeneration times, so $X$ "probabilistically restarts" at each $\Gamma_i$. For $i \geq 1$, define $\tau_i = \Gamma_i - \Gamma_{i-1}$ as the time between regenerations, and let $\tau$ be a generic copy of $\tau_i$. The process $(X(\Gamma_{i-1}+s) : 0 \leq s < \tau_i)$ between regenerations $i-1$ and $i$ is called the $i$th (regenerative) *cycle* of $X$. The pairs $(\tau_i, (X(\Gamma_{i-1}+s) : 0 \leq s < \tau_i))$, $i \geq 1$, are independent and identically distributed (i.i.d.) (Kalashnikov 1994, Section 1.3). Define also $T_i = \inf\{s > 0 : X(\Gamma_{i-1}+s) \in \mathscr{A}\}$ as the first time to reach $\mathscr{A}$ after the $(i-1)$th regeneration, and let $M = \sup\{i > 0 : T_i > \tau_i\}$ be the number of cycles completed before the first in which $\mathscr{A}$ is hit. We then can express

$$T = S + V \equiv \sum_{i=1}^{M} \tau_i + T_{M+1}, \tag{1}$$

where the *geometric sum* $S$ is independent of $V$. Let $F$ be the cumulative distribution function (cdf) of $T$, denoted as $T \sim F$. Also, let $G$ and $H$ be the cdfs of $S$ and $V$, respectively. Independent of $M$, each $\tau_i$ in $S$ has a cdf $K$ that is the conditional cdf of $\tau$ given $\tau < T$; i.e., $K(x) = P(\tau \leq x \mid \tau < T)$.

We also assume that the probability $p = \mathbb{P}(T < \tau)$ to reach $\mathscr{A}$ before a regeneration is small. To study a limiting behavior, we parameterize (and index) the model by $\varepsilon$, and our goal is to study the cdf $F \equiv F_\varepsilon$ of $T \equiv T_\varepsilon$ when $p \equiv p_\varepsilon \to 0$ as $\varepsilon \to 0$. With $\mu_\varepsilon = \mathbb{E}_\varepsilon[T_\varepsilon]$, Glynn et al. (2018) note that in different contexts (Kalashnikov 1997),

$$\lim_{\varepsilon \to 0} \mathbb{P}_\varepsilon(T_\varepsilon/\mu_\varepsilon \leq t) = 1 - e^{-t}, \quad \forall t \geq 0, \tag{2}$$

so $T_\varepsilon/\mu_\varepsilon$ converges weakly to an exponential as $\varepsilon \to 0$. An interesting consequence of (2) is that the computation of the cdf $F_\varepsilon$ of hitting times reduces asymptotically (as $\varepsilon \to 0$) to computing its mean.

We next describe two standard examples of asymptotic regimes in which $p_\varepsilon \to 0$ as $\varepsilon \to 0$.

- For the process with $X(t)$ representing the number of customers at time $t$ in a GI/GI/1 queue with first-in-first-out discipline, we may be interested in computing the distribution of the hitting time $T = T_\varepsilon$ of the set $\mathscr{A} \equiv \mathscr{A}_\varepsilon = \{\lfloor 1/\varepsilon \rfloor, \lfloor 1/\varepsilon \rfloor + 1, \dots\}$ of states. Regenerations occur at the beginnings of each busy period (Kalashnikov 1994, Example 1.2.2), and it is clear that $p_\varepsilon \to 0$ as $\varepsilon \to 0$.
- For a highly reliable Markovian system (HRMS), studied among others in Shahabuddin (1994), Nakayama (1996), L'Ecuyer and Tuffin (2012), Rubino and Tuffin (2009a), let $c$ be the number of types of components, where each type $l = 1, 2, \dots, c$, has a given redundancy $n_l$. Each component of each type is subjected to failures and repairs, with failure and repair times exponentially distributed. We can then define $X$ as a continuous-time Markov chain on a state space $\mathscr{S}$, where a state in $\mathscr{S}$ includes information on the number of operational components of each type (along with any necessary queueing information about failed components waiting for repair). The set $\mathscr{S}$ of states is decomposed into a set of states in which the system is defined as working and the set $\mathscr{A}$ of failed states. Regenerations occur on returns to the fully operational state, and $T$ represents the hitting time to $\mathscr{A}$. Failure rates are small, of order $O(\varepsilon)$, with respect to repair rates, of order $\Theta(1)$, with a repair possible from any non-fully operational state. (Recall that a function $f(\varepsilon)$ is $O(g(\varepsilon))$ if $|f(\varepsilon)/g(\varepsilon)|$ remains bounded when $\varepsilon \to 0$, and it is $\Theta(g(\varepsilon))$ if $|f(\varepsilon)/g(\varepsilon)|$ is bounded and also bounded away from 0, when $\varepsilon \to 0$.) Again, we have $p = p_\varepsilon \to 0$ as $\varepsilon \to 0$ (Shahabuddin 1994).

We adopt the following notational convention in the rest of the paper. For an unknown quantity such as $F$, we use a tilde, as in $\widetilde{F}$, to denote a (non-simulation) approximation to $F$, typically derived from a weak-convergence result, such as (2). A simulation-based estimator of $F$ has a hat, as in $\widehat{F}$.

### 2.1 Exponential Approximation Estimator

By (2), the computation of the hitting time's cdf reduces asymptotically (as $\varepsilon \to 0$, where we often omit the subscript $\varepsilon$ to simplify notation) to estimating its mean $\mu = \mathbb{E}[T]$: $\forall t \geq 0$,

$$F(t) = \mathbb{P}(T \leq t) \approx \widetilde{F}_{\exp}(t) \equiv 1 - e^{-t/\mu}. \tag{3}$$

For regenerative systems, we can express $\mu$ as a ratio (Goyal et al. 1992)

$$\mu = \frac{\mathbb{E}[T \wedge \tau]}{\mathbb{E}[\mathscr{I}(T < \tau)]} \equiv \frac{\zeta}{p} \tag{4}$$

with $x \wedge y = \min(x, y)$ and $\mathscr{I}(\cdot)$ the indicator function. This allows estimating $\mu$ by *measure-specific importance sampling* (MSIS; Goyal et al. 1992), which independently estimates $\zeta$ and $p$ in (4). To do this, we specify a proportion $0 < \gamma < 1$ used to allocate a total of $n$ independent simulated cycles as follows.

- Generate $n_{\text{CS}} \equiv \gamma n$ independent cycles by *crude simulation* (CS; i.e., no IS), giving i.i.d. observations $T_i \wedge \tau_i$, $i = 1, 2, \dots, n_{\text{CS}}$, used to estimate the numerator $\zeta$ in (4) by $\widehat{\zeta}_n = (1/n_{\text{CS}}) \sum_{i=1}^{n_{\text{CS}}} T_i \wedge \tau_i$.
- Because $\{T < \tau\}$ in the denominator of (4) is a rare event (Glynn et al. 2018), we employ IS to estimate $p = \mathbb{E}[\mathscr{I}(T < \tau)] = \int \mathscr{I}(T < \tau)\, \mathrm{d}\mathbb{P}$. Specifically, rather than sampling using the original probabilistic dynamics of $\mathbb{P}$, IS instead simulates under another probability measure $\mathbb{P}'$, and we can apply a change of measure to express $p = \int \mathscr{I}(T < \tau) L\, \mathrm{d}\mathbb{P}' = \mathbb{E}'[\mathscr{I}(T < \tau) L]$, with $L = \mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{P}'$ as the likelihood ratio and $\mathbb{E}'$ the expectation operator under $\mathbb{P}'$. This motivates estimating $p$ by

$$\widehat{p}_n = \frac{1}{n_{\text{IS}}} \sum_{i=1}^{n_{\text{IS}}} \mathscr{I}(T_i' < \tau_i') L_i', \tag{5}$$

where $(\mathscr{I}(T_i' < \tau_i'), L_i')$, $i = 1,2,\ldots,n_{\mathrm{IS}} \equiv (1-\gamma)n$, are i.i.d. copies of $(\mathscr{I}(T < \tau), L)$ under $\mathbb{P}'$.

Applying MSIS leads to the following estimator (Glynn et al. 2018).

**Estimator 1** The *exponential estimator* of the cdf $F(t)$ of $T$ is

$$\widehat{F}_{\exp,n}(t) = 1 - e^{-t/\widehat{\mu}_n}, \tag{6}$$

where $\widehat{\mu}_n = \widehat{\zeta}_n/\widehat{p}_n$ is the MSIS estimator of $\mu$.

As in Goyal et al. (1992), we may select the proportion $\gamma$ to minimize the variance per unit of computational budget (or work-normalized variance) of the estimator $\widehat{\mu}_n$. Also, Nakayama and Tuffin (2019) develop another regenerative estimator of $\mu$ based on the expression $\mu = \frac{\mathbb{E}[T;T<\tau]+\mathbb{E}[\tau;T>\tau]}{\mathbb{E}[\tau]}$, which can alternatively be used and could yield more accurate results, where $\mathbb{E}[X;A] = \mathbb{E}[X\mathscr{I}(A)]$ for an event $A$.

From the cdf estimator (6), Glynn et al. (2018) further estimate the $q$-quantile $\xi = F^{-1}(q)$, $0 < q < 1$, and the CTE $\gamma = \mathbb{E}[T|T > \xi]$ by

$$\widehat{\xi}_{\exp,n} = \widehat{F}^{-1}_{\exp,n}(q) = -\ln(1-q)\widehat{\mu}_n \quad \text{and} \quad \widehat{\chi}_{\exp,n} = (1-\ln(1-q))\widehat{\mu}_n, \tag{7}$$

respectively, where the CTE estimator uses the fact that when $T$ is exactly exponential, the memoryless property implies its CTE is $\mu + \xi$. Moreover, given a confidence interval (CI) for $\mu$ based on $\widehat{\mu}_n$, we can easily obtain CIs for $\xi$ and $\gamma$ through (7), which are linear transformations of $\widehat{\mu}_n$ (Glynn et al. 2018). In the following, we focus on estimating only the cdf $F$ of $T$ and for space reasons will not further consider estimators for quantiles and CTE; see Glynn et al. (2018) for numerical results on the latter two.

### 2.2 Convolution Estimator

Because $T = S + V$ in (1) has the geometric sum $S \sim G$ independent of $V \sim H$, the cdf $F$ of $T$ satisfies

$$F(t) = G \star H(t) = \int H(t-s)\,\mathrm{d}G(s) \tag{8}$$

where $\star$ denotes the convolution operator. This suggests convolving estimators of $G$ and $H$ to obtain an estimator of $F$, which Glynn et al. (2018) develop as follows.

- For each $t \geq 0$, we often have (Kalashnikov 1997) that $G(t) \approx \widetilde{G}_{\exp}(t) = 1 - e^{-t/\eta}$ for $p \approx 0$, where $\eta = \mathbb{E}[S] = \mathbb{E}[M] \cdot \mathbb{E}[\tau \mid \tau < T]$. As $M$ is geometric with probability mass function (pmf) $P(M = k) = (1-p)^k p$ for $k \geq 0$, its mean is $\mathbb{E}[M] = (1-p)/p$, which we estimate by $(1-\widehat{p}_n)/\widehat{p}_n$, where $\widehat{p}_n$ is from (5). As $\mathbb{E}[\tau \mid \tau < T] = \mathbb{E}[\tau\mathscr{I}(\tau < T)]/(1-p)$ is unknown, we estimate it via $(1/((1-\widehat{p}_n)n_{\mathrm{CS}}))\sum_{i=1}^{n_{\mathrm{CS}}} \tau_i \mathscr{I}(\tau_i < T_i)$ using the $n_{\mathrm{CS}}$ i.i.d. observations of $(\tau_i, \mathscr{I}(\tau_i < T_i))$ of $(\tau, \mathscr{I}(\tau < T))$ from crude simulation. We then estimate $\eta$ by

$$\widehat{\eta}_n = \frac{1}{\widehat{p}_n n_{\mathrm{CS}}} \sum_{i=1}^{n_{\mathrm{CS}}} \tau_i \mathscr{I}(\tau_i < T_i),$$

  resulting in $\widehat{G}_{\exp,n}(t) = 1 - e^{-t/\widehat{\eta}_n}$ as an exponential estimator of $G(t)$.
- The cdf $H$ of $V$ satisfies $H(x) = \mathbb{P}(V \leq x) = \mathbb{P}(T \leq x \mid T < \tau) = \mathbb{P}(T \leq x, T < \tau)/p$ by (4). We apply a change of measure to get $\mathbb{P}(T \leq x, T < \tau) = \mathbb{E}[\mathscr{I}(T \wedge \tau \leq x, T < \tau)] = \mathbb{E}'[\mathscr{I}(T \wedge \tau \leq x, T < \tau)L]$. Thus, from i.i.d. observations $(T_i' \wedge \tau_i', \mathscr{I}(T_i' < \tau_i'), L_i')$, $i = 1,2,\ldots,n_{\mathrm{IS}}$, of $(T \wedge \tau, \mathscr{I}(T < \tau), L)$ under IS measure $\mathbb{P}'$, we estimate $H(x)$ by (using (5))

$$\widehat{H}_n(x) = \frac{1}{\widehat{p}_n n_{\mathrm{IS}}} \sum_{i=1}^{n_{\mathrm{IS}}} \mathscr{I}(T_i' \wedge \tau_i' \leq x, T_i' < \tau_i')L_i' = \frac{\sum_{i=1}^{n_{\mathrm{IS}}} \mathscr{I}(T_i' \wedge \tau_i' \leq x, T_i' < \tau_i')L_i'}{\sum_{j=1}^{n_{\mathrm{IS}}} \mathscr{I}(T_j' < \tau_j')L_j'}. \tag{9}$$

Taking the convolution of $\widehat{G}_{\exp,n}$ and $\widehat{H}_n$, we get (Glynn et al. 2018)

**Estimator 2** The convolution estimator of cdf $F(t)$ of $T$ is

$$\widehat{F}_{\text{conv},n}(t) = 1 - \frac{1}{\widehat{p}_n \cdot n_{\text{IS}}} \sum_{i=1}^{n_{\text{IS}}} \mathscr{I}(T_i' < \tau_i') L_i' e^{-(t-(T_i' \wedge \tau_i'))^+/\widehat{\eta}_n}. \tag{10}$$

For a fair comparison with the exponential estimator $\widehat{F}_{\exp,n}$ in (6), we defined the convolution estimator $\widehat{F}_{\text{conv},n}$ in (10) using the same sample sizes $n_{\text{CS}} = \gamma n$ and $n_{\text{IS}} = (1-\gamma)n$ for CS and IS, respectively. But we could also choose the MSIS parameter $\gamma$ differently for the convolution estimator.

## 3 CORRECTED CONVOLUTION ESTIMATOR

While Section 2.2 obtained $\widehat{F}_{\text{conv},n}$ in (10) by convolving $\widehat{G}_{\exp,n}$ and $\widehat{H}_n$, we can construct additional estimators of $F$ by instead taking the convolution of other estimators of $G$ and $H$. In particular, we now propose to do this by exploiting an approximation, developed by Blanchet and Glynn (2007), to the cdf $G'$ of a (slightly different) geometric sum $S'$. Specifically, let $S' = \sum_{i=1}^{M'} Y_i$, where $(Y_i)_{i\geq 1}$ are i.i.d. copies of a random variable $Y$ and $M' = M+1$ for $M$ as in (1), where $M'$ is independent of the $Y_i$. Thus, $M'$ has the alternative definition of a geometric with pmf $P(M'=k) = (1-p)^{k-1}p$ for $k \geq 1$, in contrast to our geometric $M$ in (1), whose support starts from $k=0$. If $Y$ has a strongly nonlattice distribution with moment-generating function $\Phi(v) \equiv \mathbb{E}[e^{vY}] < \infty$ for some $v > 0$, then Theorem 2 of Blanchet and Glynn (2007) provides rigorous justification for the "corrected" exponential approximation

$$\mathbb{P}(S' > t) \approx c(p)e^{-\theta t} \tag{11}$$

as $p \to 0$, where $\theta$ is the unique nonnegative solution of $\Phi(\theta) = (1-p)^{-1}$ and $c(p) = \frac{p}{(1-p)^2\theta\phi(\theta)} \equiv e^{r(p)}$, with $\phi$ the derivative of $\Phi$. While the right side of (11) approaches a true tail cdf for all $t$ as $p \to 0$, for a fixed $p > 0$, it may be greater than 1 for some (small) $t$.

Blanchet and Glynn (2007) suggest two alternative options to estimate $\theta$ and $c(p)$:

1. Use a numerical root-finding method (Press et al. 2007, Chapter 9) to get $\theta$ and then estimate $c(p)$;
2. Develop expressions for $\theta$ and $r(p) = \ln(p) - \ln[(1-p)^2\theta\phi(\theta)]$ in powers of $p$:

$$\theta = \sum_{k=1}^{\infty} \chi_k p^k \quad \text{and} \quad r(p) = \sum_{k=0}^{\infty} \delta_k p^k,$$

with the $\chi_k$ and $\delta_k$ computed via the implicit-function theorem, giving as a variation from Blanchet and Glynn (2007)

- $\chi_1 = 1/\mathbb{E}[Y]$,
- $\chi_2 = \frac{2(\mathbb{E}[Y])^2 - \mathbb{E}[Y^2]}{2(\mathbb{E}[Y])^3}$
- $\chi_3 = \frac{3(\mathbb{E}[Y^2])^2 - 6\mathbb{E}[Y^2](\mathbb{E}[Y])^2 + 6(\mathbb{E}[Y])^4 - \mathbb{E}[Y]\mathbb{E}[Y^3]}{6(\mathbb{E}[Y])^5}$
- $\delta_1 = \frac{2(\mathbb{E}[Y])^2 - \mathbb{E}[Y^2]}{2(\mathbb{E}[Y])^2}$
- $\delta_2 = \frac{12(\mathbb{E}[Y])^4 - 12\mathbb{E}[Y^2](\mathbb{E}[Y])^2 + 15(\mathbb{E}[Y^2])^2 - 8\mathbb{E}[Y^3]\mathbb{E}[Y]}{24(\mathbb{E}[Y])^4}$
- ...

We now would like to use (11) to approximate the cdf $G$ of $S = \sum_{i=1}^{M} \tau_i$ from (1) (i.e., $Y_i$ is replaced with $\tau_i$ given $\tau_i < T_i$) and convolve it with the cdf $H$ of $V$ to approximate the cdf $F$ of $T = S + V$. However, we cannot directly apply (11) because our $S$ sums $M$ terms, whereas $S'$ in (11) has $M' = M+1$ summands. To account for this difference, note that (also see p. 12 of Kalashnikov 1997) for $t \geq 0$,

$$\begin{aligned}\mathbb{P}(S > t) &= \mathbb{P}(S > t \mid M \geq 1)\mathbb{P}(M \geq 1) + \mathbb{P}(S > t \mid M = 0)\mathbb{P}(M = 0) \\ &= (1-p)\mathbb{P}(S > t \mid M \geq 1) = (1-p)\mathbb{P}(S' > t) \approx (1-p)c(p)e^{-\theta t}.\end{aligned}$$

To estimate $\theta$ and $c(p)$, we use the above second option as the moments of $Y$ are easily estimated. Specifically, the $k$th moment is $\mathbb{E}[Y^k] = \mathbb{E}[\tau^k|\tau < T] = \mathbb{E}[\tau^k \mathscr{I}(\tau < T)]/(1-p)$, which MSIS estimates by

$$\frac{1}{(1-\widehat{p}_n)\, n_{\mathrm{CS}}} \sum_{i=1}^{n_{\mathrm{CS}}} \tau_i^k \mathscr{I}(\tau_i < T_i).$$

This yields estimators $\widehat{\chi}_{k,n}$ and $\widehat{\delta}_{k,n}$ of $\chi_k$ and $\delta_k$, respectively, so we estimate $\theta$ and $r(p)$ by

$$\widehat{\theta}_n = \widehat{\chi}_{1,n}\widehat{p}_n + \widehat{\chi}_{2,n}\widehat{p}_n^2 + \widehat{\chi}_{3,n}\widehat{p}_n^3 \quad \text{and} \quad \widehat{r}_n(\widehat{p}_n) = \widehat{\delta}_{1,n}\widehat{p}_n + \widehat{\delta}_{2,n}\widehat{p}_n^2,$$

respectively, leading to the corrected exponential estimator of $G(t)$ as

$$\widehat{G}_{\mathrm{corr},n}(t) = 1 - (1-\widehat{p}_n)e^{-\widehat{\theta}_n t + \widehat{r}_n(\widehat{p}_n)}.$$

We also estimate the cdf $H$ of $V$ by $\widehat{H}_n$ in (9). As (8) shows that $F$ can be expressed as a convolution of $G$ and $H$, we then estimate $F$ by (with $A_i' = T_i' \wedge \tau_i'$)

$$\begin{aligned}
\widehat{F}_{\mathrm{corr},n}(t) &= \frac{1}{\widehat{p}_n\, n_{\mathrm{IS}}} \sum_{i=1}^{n_{\mathrm{IS}}} \mathscr{I}(T_i' < \tau_i') L_i' \int_0^t \mathscr{I}(A_i' \le t-x)\, \mathrm{d}\widehat{G}_{\mathrm{corr},n}(x) \\
&= \frac{1}{\widehat{p}_n\, n_{\mathrm{IS}}} \sum_{i=1}^{n_{\mathrm{IS}}} \mathscr{I}(T_i' < \tau_i') L_i' \int_0^{(t-A_i')^+} \mathrm{d}\widehat{G}_{\mathrm{corr},n}(x) \\
&= \frac{1}{\widehat{p}_n\, n_{\mathrm{IS}}} \sum_{i=1}^{n_{\mathrm{IS}}} \mathscr{I}(T_i' < \tau_i') L_i' \widehat{G}_{\mathrm{corr},n}((t-A_i')^+).
\end{aligned}$$

Finally, we get the following expression:

**Estimator 3** The corrected convolution estimator of cdf $F(t)$ of $T$ is

$$\widehat{F}_{\mathrm{corr},n}(t) = \frac{1}{\widehat{p}_n\, n_{\mathrm{IS}}} \sum_{i=1}^{n_{\mathrm{IS}}} \mathscr{I}(T_i' < \tau_i') L_i' (1 - (1-\widehat{p}_n) e^{-(t-A_i')^+ \widehat{\theta}_n + \widehat{r}_n(\widehat{p}_n)}).$$

Again, for sake of comparison and as the optimal MSIS parameter $\gamma$ is unknown in this setting, our numerical studies (Section 6) take $n_{\mathrm{CS}} = \gamma n$ and $n_{\mathrm{IS}} = (1-\gamma)n$ with $\gamma$ optimized for estimator $\widehat{\mu}_n$ in (6).

## 4 GAMMA-BASED ESTIMATOR

We next propose to extend the exponential approximation (3) of cdf $F$ of $T$ by a gamma approximation since a gamma distribution generalizes an exponential. Instead of estimating only the rate of the exponential, we now have to estimate two parameters (shape and rate) for a gamma. Using this generalized family leads to two degrees of freedom instead of one and may lead to smaller bias when $p$ is not very small.

Recall that the cdf $F_\Gamma$ of a gamma distribution with shape parameter $\alpha > 0$ and rate $\beta > 0$ is

$$F_\Gamma(t;\alpha,\beta) = \int_0^t \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\, \mathrm{d}x,$$

where $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}\,\mathrm{d}x$ is the gamma function. We will approximate the cdf $F$ of $T$ by $F_\Gamma$, but as $\alpha$ and $\beta$ in $F_\Gamma$ are unknown, we need to estimate them. Section 7.2 of Casella and Berger (2002) describes several methods for doing this, including maximum likelihood and Bayesian estimators, but we restrict ourselves to the *method of moments* (MoM). Indeed, the first two techniques require i.i.d. copies $T_i$ of $T$, which may not be possible to obtain in a reasonable amount of time in a rare-event context because each $T_i$ typically requires a long simulation. MoM can avoid this issue, as we will explain below.

As $F_\Gamma$ has $v = 2$ unknown parameters, MoM starts by expressing the first $v$ true (central) moments of $F_\Gamma$ as functions of the distribution's $v$ parameters, and then solves for the parameters in the resulting $v$ simultaneous equations in terms of the moments. Finally, in the expressions for the parameters, replace the true moments with their estimators to obtain estimators of the parameters. Specifically, the mean and variance of $T \sim F_\Gamma$ are $\mu = \mathbb{E}[T] = \alpha/\beta$ and $\sigma^2 = \mathbb{E}[T^2] - \mu^2 = \alpha/\beta^2$, and solving for $\alpha$ and $\beta$ yields

$$\alpha = \frac{\mu^2}{\sigma^2} \quad \text{and} \quad \beta = \frac{\mu}{\sigma^2}. \tag{12}$$

(Note that if the scaled $T$ converges in distribution to an exponential as $p \to 0$, then $\alpha = \mu^2/\sigma^2 \to 1$ as expected (characterizing an exponential distribution).)

We next provide representations for $\mu$ and $\sigma^2$ that MSIS can efficiently estimate. As in (4), the mean satisfies $\mu = \frac{\mathbb{E}[T \wedge \tau]}{\mathbb{P}(T<\tau)} = \frac{\zeta}{p}$, which Section 2.1 describes how to estimate by MSIS. For the variance $\sigma^2 = \mathbb{E}[T^2] - \mu^2$, the only term left to handle is $\mathbb{E}[T^2]$, and the regenerative property implies

$$\begin{aligned}
\mathbb{E}[T^2] &= \mathbb{E}[T^2; T < \tau] + \mathbb{E}[(\tau + T - \tau)^2; T > \tau] \\
&= \mathbb{E}[T^2; T < \tau] + \mathbb{E}[\tau^2; T > \tau] + \mathbb{E}[(T-\tau)^2; T > \tau] + 2\mathbb{E}[\tau(T-\tau); T > \tau] \\
&= \mathbb{E}[(T \wedge \tau)^2; T < \tau] + \mathbb{E}[(T \wedge \tau)^2; T > \tau] + \mathbb{E}[(T-\tau)^2 \mid T > \tau]\,\mathbb{P}(T > \tau) \\
&\quad + 2\mathbb{E}[\tau(T-\tau) \mid T > \tau]\,\mathbb{P}(T > \tau) \\
&= \mathbb{E}[(T \wedge \tau)^2] + \mathbb{E}[T^2](1-p) + 2\mathbb{E}[\tau \mid T > \tau]\,\mathbb{E}[T - \tau \mid T > \tau](1-p) \\
&= \mathbb{E}[(T \wedge \tau)^2] + \mathbb{E}[T^2](1-p) + 2\mathbb{E}[\tau; T > \tau]\,\mu
\end{aligned}$$

as $\mathbb{E}[(T-\tau)^k \mid T > \tau] = \mathbb{E}[T^k]$ by the regenerative structure. Solving for $\mathbb{E}[T^2]$ and using (4) then yield

**Proposition 1** If the process $X$ is regenerative, then

$$\mathbb{E}[T^2] = \frac{1}{p}\left(\mathbb{E}[(T \wedge \tau)^2] + 2\mathbb{E}[\tau \mathscr{I}(T > \tau)]\frac{\mathbb{E}[T \wedge \tau]}{p}\right). \tag{13}$$

From (4) and (13), we then can obtain expressions for $\mu$ and $\sigma^2$ as functions of cycle expectations, so we can apply MSIS to estimate them and substitute into (12) to obtain our MoM estimators of $\alpha$ and $\beta$. Specifically, for a fixed total number $n$ of cycles to simulate,

- use $n_{\mathrm{CS}} = \gamma n$ cycles with CS to estimate $\mathbb{E}[T \wedge \tau]$, $\mathbb{E}[(T \wedge \tau)^2]$, and $\mathbb{E}[\tau \mathscr{I}(T > \tau)]$;
- use $n_{\mathrm{IS}} = (1-\gamma)n$ cycles with IS to estimate $p$ by $\widehat{p}_n$ in (5).

Our numerical studies (Section 6) again will use the MSIS parameter $\gamma$ optimized for estimator $\widehat{\mu}_n$ in (6).

Having now estimated the moments $\mu$ and $\sigma^2$, we estimate the shape and rate parameters of the gamma distribution through (12). We may then compute the resulting cdf (resp., quantile) estimator via efficient algorithms (Press et al. 2007, Chapter 6) for gamma cdf evaluation (resp., inversion). In summary:

**Estimator 4** The gamma-based estimator of the cdf $F(t)$ of $T$ is $\widehat{F}_{\Gamma,n}(t) = F_\Gamma(t; \widehat{\alpha}_n, \widehat{\beta}_n)$, with $\widehat{\alpha}_n$ and $\widehat{\beta}_n$ as estimators of the shape and rate parameters, as described above.

## 5 BOOTSTRAP ESTIMATOR

All of the previous estimators are based on limit theorems that hold as $p \to 0$. But in an actual system, we do not have $p \to 0$ but rather a fixed $p > 0$, so the resulting estimators have bias, which does not vanish even as $n \to \infty$. To address these issues, we next consider another approach that applies a type of bootstrap (Casella and Berger 2002, Section 10.1.4). The method is computationally more demanding but may be worthwhile for moderate values of $p$ or when the model is time-consuming to simulate.

The algorithm estimates cdf $F$ by resampling generated observations of $T = S + V$, where we recall from (1) that $S = \sum_{i=1}^{M} \tau_i$, $M$ is geometric with parameter $p$ (with support starting from 0), and $S \sim G$ and $V \sim H$ are independent. Also, $M, \tau_1, \tau_2, \ldots$ are mutually independent in the geometric sum $S$. Recall that each summand $\tau_i$ in $S$ has cdf $K$, where $K(x) = \mathbb{P}(\tau < x \mid \tau < T)$. Our algorithm generates observations of $T$ in two steps. The first step (setup) applies MSIS to estimate the geometric parameter $p$ and the cdfs $K$ and $H$. The second step (bootstrap) generates observations of $T$ as follows: initially sample a geometric $M^*$ with the estimated parameter $p$, then sample $M^*$ i.i.d. observations from the empirical estimate of the cdf $K$, and finally add their sum to a sample from the empirical estimate of $H$. Next are the details.

**Estimator 5** The bootstrap estimator $\widehat{F}_{\text{boot},n,n'}$ of the cdf $F$ of $T$ is constructed as follows.

**Step 1: Set-up** Apply MSIS with a total sample size $n = n_{\text{CS}} + n_{\text{IS}}$ to estimate $p$, $K$, and $H$.

- Use CS to generate $n_{\text{CS}}$ i.i.d. pairs $(\tau_i, \mathscr{I}(\tau_i < T_i))$, $i = 1, 2, \ldots, n_{\text{CS}}$, of $(\tau, \mathscr{I}(\tau < T))$ under the original measure $\mathbb{P}$, and estimate $K(x)$ by $\widehat{K}_n(x) = \sum_{i=1}^{n_{\text{CS}}} \mathscr{I}(\tau_i < x, \tau_i < T_i) / [\sum_{j=1}^{n_{\text{CS}}} \mathscr{I}(\tau_j < T_j)]$.
- Generate $n_{\text{IS}}$ cycles by IS to estimate $p$ by $\widehat{p}_n$ in (5) and the cdf $H$ by $\widehat{H}_n$ in (9).

**Step 2: Semi-Parametric Bootstrap** Repeat the following resampling procedure $n'$ independent times to obtain $n'$ conditionally i.i.d. observations $T_l^*$, $l = 1, 2, \ldots, n'$, of $T$, given the data from Step 1; i.e., in each iteration $l = 1, 2, \ldots, n'$, do the following:

- Generate $M_l^* \sim$ geometric$(\widehat{p}_n)$, i.e., $M_l^*$ has pmf $P(M_l^* = k) = (1 - \widehat{p}_n)^k \widehat{p}_n$, for $k \geq 0$.
- Generate $M_l^*$ i.i.d. observations $\tau_{l,1}^*, \tau_{l,2}^*, \ldots, \tau_{l,M_l^*}^*$ from the empirical distribution $\widehat{K}_n$, and let $S_l^* = \sum_{j=1}^{M_l^*} \tau_{l,j}^*$ be their sum.
- Generate $V_l^* \sim \widehat{H}_n$, conditionally independently of $S_l^*$.
- Return $T_l^* = S_l^* + V_l^*$.

After completing all $n'$ iterations, construct $\widehat{F}_{\text{boot},n,n'}$ with $\widehat{F}_{\text{boot},n,n'}(t) = (1/n') \sum_{l=1}^{n'} \mathscr{I}(T_l^* \leq t)$.

Recall that $K$ is the conditional cdf of $\tau$ given $\tau < T$, but Step 1 generates (via CS) each $(\tau_i, \mathscr{I}(\tau_i < T_i))$ *not* conditional on $\tau_i < T_i$. Hence, to estimate $K$ from our data, we use only those $\tau_i$ with $\tau_i < T_i$, so the empirical cdf $\widehat{K}_n$ assigns probability $\mathscr{I}(\tau_i < T_i) / [\sum_{j=1}^{n_{\text{CS}}} \mathscr{I}(\tau_j < T_j)]$ to each $\tau_i$ sampled in Step 1. Also, $\widehat{H}_n$ in (9) gives each IS-generated $T_i' \wedge \tau_i'$ a mass $\mathscr{I}(T_i' < \tau_i') L_i' / [\sum_{j=1}^{n_{\text{IS}}} \mathscr{I}(T_j' < \tau_j') L_j']$. Thus, although Step 1 entails simulating the complete stochastic model via CS and IS, Step 2 resamples the generated values from Step 1, providing a computational savings. Also, the bootstrap estimator is perhaps mainly appropriate when $p$ is not too small; otherwise, the geometric $M_l^*$ will typically take on huge values, so generating $S_l^*$ in Step 2 incurs large cost. But the situation when $p$ is not very small is exactly when the exponential approximations to $F$ or $G$ may not be so accurate. While the bootstrap estimator avoids the bias that arises from a weak convergence, as in (2), not holding for a fixed $p > 0$, the bootstrap incurs another type of bias. As ratio estimators, $\widehat{K}_n$ and $\widehat{H}_n$ are typically biased, although this bias vanishes as $n \to \infty$, in contrast to exponential-type estimators, whose bias does not change as $n$ grows. For our numerical experiments in Section 6, we will again use $n_{\text{CS}} = \gamma n$ and $n_{\text{IS}} = (1 - \gamma) n$, and take $n' = n$ in the second step.

## 6 NUMERICAL COMPARISONS

We ran numerical experiments to compare our estimators of the cdf $F$ of the hitting time $T$ of $\mathscr{A}$ for two stochastic models. For the exponential, convolution, corrected convolution, gamma, and bootstrap estimators, we simulated a total of $n = 10^4$ independent cycles for MSIS, where the allocation parameter $\gamma$ (proportion of CS cycles) minimizes (Goyal et al. 1992) the work-normalized variance of the estimator $\widehat{\mu}_n$ of $\mu$ used in (6). For comparison, we also constructed an empirical estimator of $F$ from directly generating $10^4$ i.i.d. values of $T$. When possible, we also computed the exact cdf $F$ numerically (i.e., no simulation).

Our goal is to analyze the bias and error of the estimators. To do this, we generated $m = 10^5$ independent copies $\widehat{F}_{j,n}(t_k)$, $j = 1, 2, \ldots, m$, of an estimator $\widehat{F}_n(t_k)$ at several points $t_k$. Then we estimated the bias,

$\mathbb{E}[\widehat{F}_n(t_k)] - F(t_k)$, and the mean square error (MSE), $\mathbb{E}[(\widehat{F}_n(t_k) - F(t_k))^2]$, by

$$\frac{1}{m}\sum_{j=1}^{m} \widehat{F}_{j,n}(t_k) - F(t_k) \quad \text{and} \quad \frac{1}{m}\sum_{j=1}^{m} \left(\widehat{F}_{j,n}(t_k) - F(t_k)\right)^2, \tag{14}$$

respectively, where we recall that MSE decomposes as the sum of variance and squared bias. In our experiments, except for the empirical and bootstrap estimators, the differences in the computation times for all simulation methods are negligible.

### 6.1 Highly Reliable Markovian Systems

We first consider an HRMS, as discussed in Section 2. The system has $c = 3$ types of components with $n_l = 3$ components of type $l$ ($1 \le l \le c$). Each component has an exponentially distributed time to failure with rate $\varepsilon$. Any failed component has an exponentially distributed repair time with rate 1. There are $\sum_{l=1}^{c} n_l$ repairmen, so failed components never queue for repair. The system is considered down whenever fewer than two components of any one type are operational, and $T$ is the system's time to failure (TTF). We implemented IS using the so-called zero-variance approximation (L'Ecuyer and Tuffin 2012).

We first examine plots of the estimated cdf $F$ of $T$ obtained by the various simulation methods for $n = 10^4$. Figure 1 shows the results for $\varepsilon = 10^{-4}$ on the left and for $\varepsilon = 10^{-2}$ on the right, while Figure 2 is for $\varepsilon = 10^{-1}$. For $\varepsilon = 10^{-4}$, we do not show the exact value because it requires too much time to compute; we also omit the empirical (it takes more than one hour to reach the first failure) and the bootstrap (which is potentially of interest only for moderate values of $p_\varepsilon$).

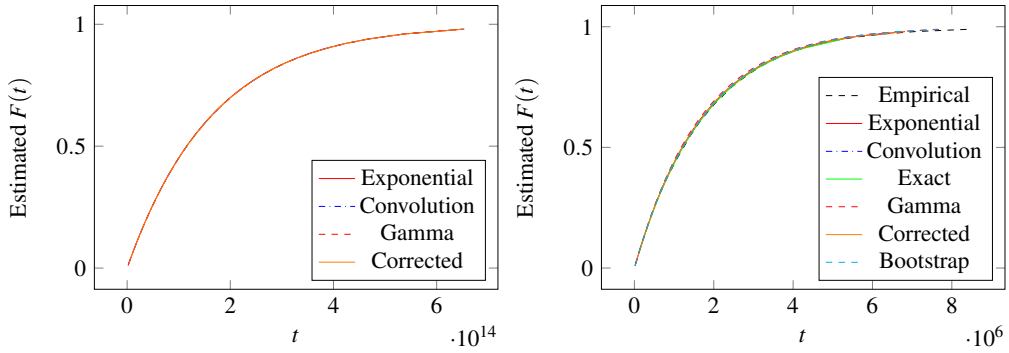


Figure 1: Plots of the estimated and exact cdf of the TTF for (left) $\varepsilon = 10^{-4}$ with MSIS allocation parameter $\gamma = 0.75370$, and for (right) $\varepsilon = 10^{-2}$ with $\gamma = 0.26390$.

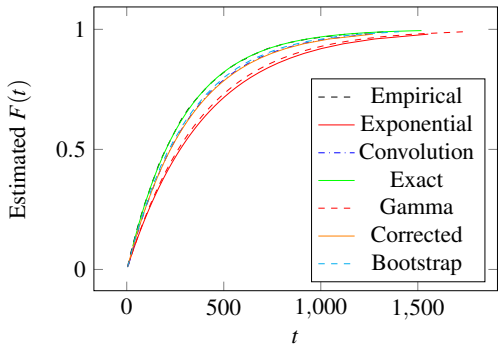


Figure 2: Plots of the estimated and exact cdf of the TTF when $\varepsilon = 10^{-1}$, with $\gamma = 0.36670$.

For the gamma estimator, the estimated shape and rate parameters are $\alpha = 1$ (resp., 0.99999 and 1.0108) and $\beta = 6.0292\text{e-}15$ (resp., 5.7437e-07 and 3.2142e-03) for $\varepsilon = 10^{-4}$ (resp., $10^{-2}$ and $10^{-1}$).

For $\varepsilon = 10^{-2}$ and $\varepsilon = 10^{-4}$, the curves overlap, showing the good behavior of all algorithms. For $\varepsilon = 0.1$ we can see slight differences: the exponential estimator seems less accurate, but we have to check whether it is a general characteristic or if it is due to random noise.

To better analyze the bias and errors, Figure 3 plots the bias and MSE, estimated using (14) from $m = 10^5$ independent experiments, when $\varepsilon = 0.1$, where $(\mathrm{L}, \mathrm{U})$ is a 95% CI for the true value $F(t)$. It took 40 minutes to get the curves. As the CIs largely overlap, the differences appear to be very limited.

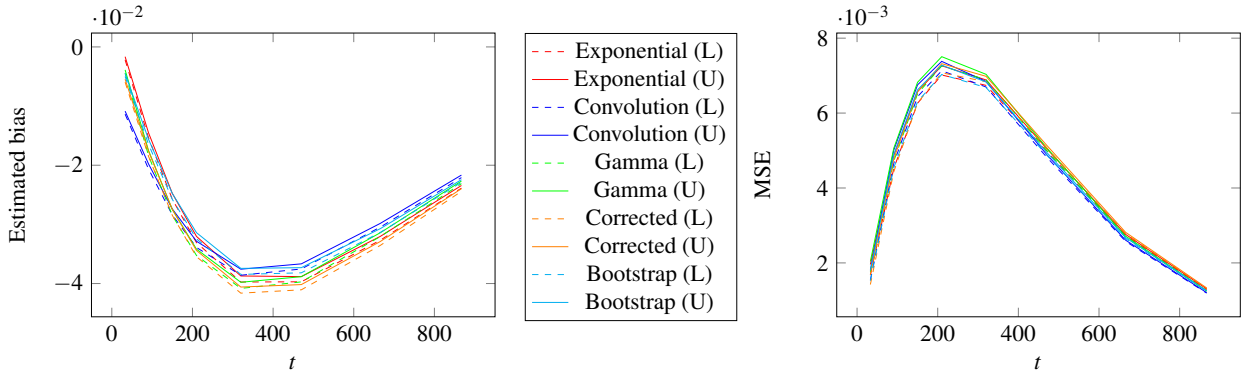


Figure 3: Bias (left) and MSE (right) at various points $t_k$ for the HRMS with $\varepsilon = 10^{-1}$ when $m = 10^5$

We do the same when $\varepsilon = 10^{-2}$ in Figure 4, which do not show the empirical and bootstrap estimators because they take too much time to compute. All the methods give equivalent results in terms of bias and

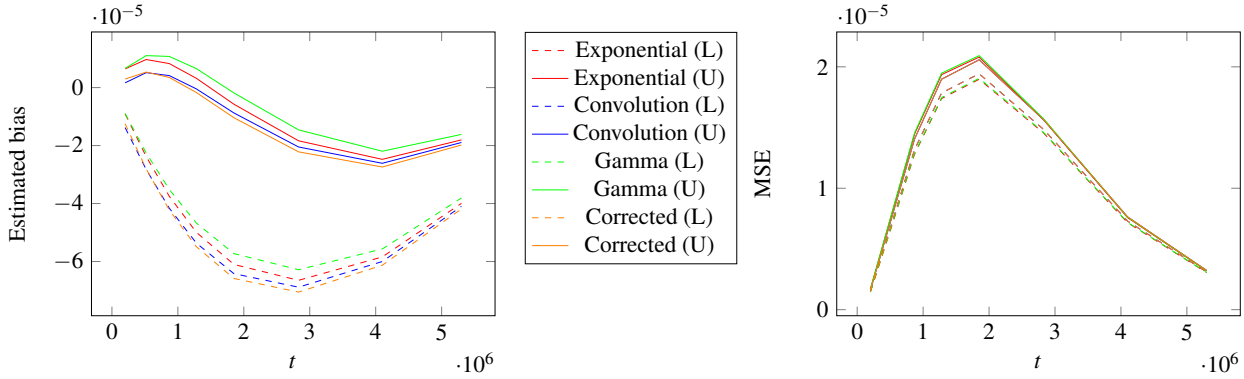


Figure 4: Bias (left) and MSE (right) at various points $t_k$ for the HRMS with $\varepsilon = 10^{-2}$ when $m = 10^5$

MSE. The squared bias is much smaller than the MSE, meaning that, at least for $n = 10^4$, bias may not be much of an issue for the CIs, which are based only on variance and do not directly account for the bias.

### 6.2 M/M/1 Queue

For a second example, we consider the process where $X(t)$ denotes the total number of customers at time $t$ in an M/M/1 queue, a special case of the GI/GI/1 queue described in Section 2, with arrival rate $\lambda = 0.5$ and service rate $\mu = 1$. We want to estimate the cdf of the hitting time of $\mathscr{A}_\varepsilon = \{\lfloor 1/\varepsilon \rfloor, \lfloor 1/\varepsilon \rfloor + 1, \dots\}$. For IS, we swap the arrival and service rates (Parekh and Walrand 1989).

Figures 5 and 6 display estimated bias and MSE from $m = 10^5$ independent replications of the estimators for both $N \equiv \lfloor 1/\varepsilon \rfloor = 10$ and $N = 20$, respectively (with respective MSIS parameter values $\gamma = 0.5612$ and

$\gamma = 0.6601$). All algorithms except the bootstrap give very similar responses. Interestingly, the bootstrap,

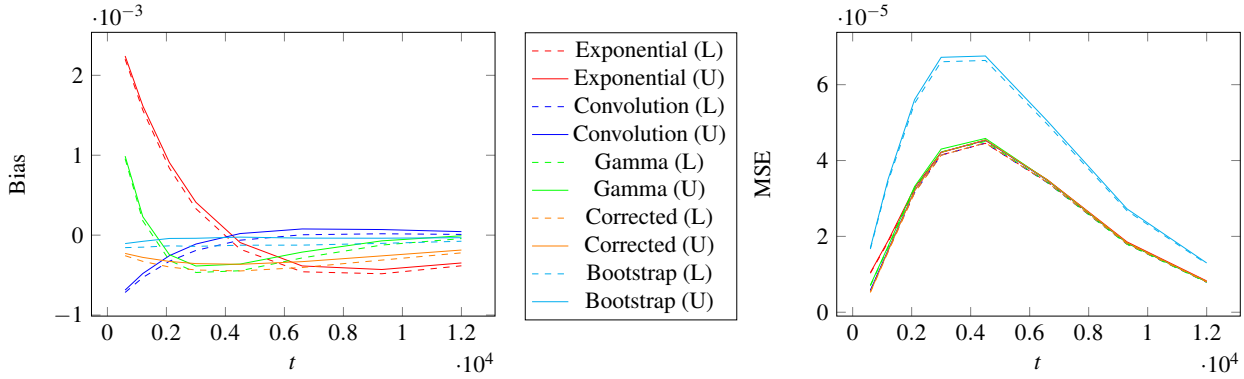


Figure 5: Bias (left) and MSE (right) at various points $t_k$ for the M/M/1 queue with $N = 10$ and $m = 10^5$ independent replications

which takes longer to compute, has much smaller bias but larger MSE due to a larger variance than the other methods. Compared to the HRMS, the M/M/1 queue has more variable cycle lengths. (Indeed, HRMS has very short cycles as returns to the regenerative state are very likely, which is less true for M/M/1.) It is also worth noting that again, even for this moderately rare case, the squared bias contributes little to the MSE.

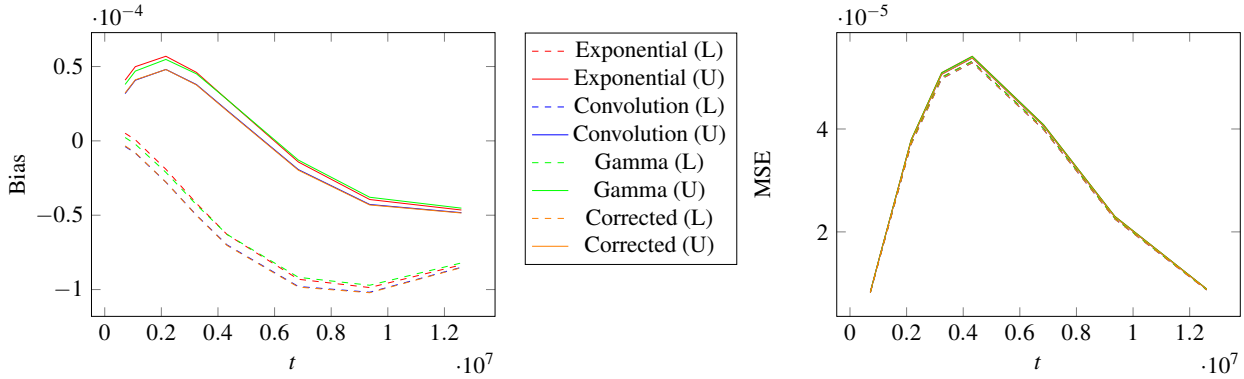


Figure 6: Bias (left) and MSE (right) at various points $t_k$ for the M/M/1 queue with $N = 20$ and $m = 10^5$ independent replications

For $N = 20$, we do not display the results for the bootstrap estimator because it is too computationally demanding. All algorithms are almost identical in terms of bias and MSE.

## 7 CONCLUDING REMARKS

We have introduced in this paper new cdf estimators of the hitting time to a rarely visited set of states for a regenerative process. We have numerically observed that the estimators are very close in performance even for a moderately rare situation. It is also interesting to observe that bias is not a significant part of the mean squared error. All this emphasizes the interest of the exponential approximation estimator. As a next step, we plan to apply the same techniques to estimate risk measures, such as quantiles and CTE.

## REFERENCES

Asmussen, S., and P. Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis*. New York: Springer.

Blanchet, J. H., and P. W. Glynn. 2007. "Uniform Renewal Theory with Applications to Expansions of Random Geometric Sums". *Advances in Applied Probability* 39:1070–1097.

Casella, G., and R. L. Berger. 2002. *Statistical Inference*. 2nd ed. Pacific Grove, Calif.: Duxbury.

Glynn, P. W., M. K. Nakayama, and B. Tuffin. 2017. "On the Estimation of the Mean Time to Failure by Simulation". In *Proceedings of the 2017 Winter Simulation Conference*, edited by W. K. V. Chan, A.D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, 1844–1855. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Glynn, P. W., M. K. Nakayama, and B. Tuffin. 2018. "Using Simulation to Calibrate Exponential Approximations to Tail-Distribution Measures of Hitting Times to Rarely Visited Sets". In *Proceedings of the 2018 Winter Simulation Conference*, edited by M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, 1802–1813. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Goyal, A., P. Shahabuddin, P. Heidelberger, V. Nicola, and P. W. Glynn. 1992. "A Unified Framework for Simulating Markovian Models of Highly Dependable Systems". *IEEE Transactions on Computers* C-41(1):36–51.

Kalashnikov, V. 1994. *Topics on Regenerative Processes*. Boca Raton: CRC Press.

Kalashnikov, V. 1997. *Geometric Sums: Bounds for Rare Events with Applications*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

L'Ecuyer, P., and B. Tuffin. 2012. "Approximating Zero-Variance Importance Sampling in a Reliability Setting". *Annals of Operations Research* 189(1):277–297.

Nakayama, M. K. 1996. "General Conditions for Bounded Relative Error in Simulations of Highly Reliable Markovian Systems". *Advances in Applied Probability* 28:687–727.

Nakayama, M. K., and B. Tuffin. 2019. "Efficient Estimation of the Mean Hitting Time to a Set of a Regenerative System". In *Proceedings of the 2019 Winter Simulation Conference*, edited by N. Mustafee, K.-H. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, 416–427. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Nicola, V. F., M. K. Nakayama, P. Heidelberger, and A. Goyal. 1993. "Fast simulation of dependability models with general failure and repair processes". *IEEE Transactions on Computers* 42:1440–1452.

Parekh, S., and J. Walrand. 1989. "A Quick Simulation Method for Excessive Backlogs in Networks of Queues". *IEEE Transactions on Automatic Control* 34:54–66.

Press, W. H., S. Teukolsky, W. Vetterling, and B. Flannery. 2007. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. 3 ed. USA: Cambridge University Press.

Rubino, G., and B. Tuffin. 2009a. "Markovian Models for Dependability Analysis". In *Rare Event Simulation using Monte Carlo Methods*, edited by G. Rubino and B. Tuffin, 125–144. Chichester, UK: John Wiley & Sons.

Rubino, G., and B. Tuffin. 2009b. *Rare Event Simulation using Monte Carlo Methods*. Wiley.

Shahabuddin, P. 1994. "Importance Sampling for Highly Reliable Markovian Systems". *Management Science* 40:333–352.