

Resampled Regenerative Estimators

JAMES M. CALVIN, New Jersey Institute of Technology
MARVIN K. NAKAYAMA, New Jersey Institute of Technology

We discuss some estimators for simulations of processes having multiple regenerative sequences. The estimators are obtained by resampling trajectories without and with replacement, which correspond to a type of U -statistic and a type of V -statistic, respectively. The U -statistic estimator turns out to be equivalent to the permuted regenerative estimator, which we previously proposed, but the V -statistic estimator is new. We compare analytically some properties of these estimators along with the semi-regenerative estimator. We show that when estimating the second moment of a cycle reward, the semi-regenerative estimator has positive bias, which is strictly larger than the (positive) bias of the V -statistic estimator. The permuted estimator is unbiased. All of the estimators have the same asymptotic central-limit behavior, with reduced asymptotic variance compared to the standard regenerative estimator. Some numerical results are included.

Categories and Subject Descriptors: I.6.6 [**Simulation and Modeling**]: Simulation Output Analysis; G.3 [**Probability and Statistics**]: Nonparametric statistics; G.3 [**Probability and Statistics**]: Probabilistic algorithms (including Monte Carlo)

General Terms: Performance, Theory

Additional Key Words and Phrases: variance reduction, regenerative method

ACM Reference Format:

James M. Calvin and Marvin K. Nakayama, 2013. Resampled regenerative estimators. *ACM Trans. Model. Comput. Simul.* V, N, Article xx (July 2013), 25 pages.
DOI: <http://dx.doi.org/10.1145/0000000.0000000>

1. INTRODUCTION

A *regenerative process* is a stochastic process having an infinite sequence of random times, known as *regeneration points*, at which times the process probabilistically restarts. The path segment of the process between two regeneration points is known as a *regenerative cycle*, and cycles are independent and identically distributed (i.i.d.). For example, successive hitting times to a fixed state for a positive-recurrent Markov chain on a discrete state space form a regeneration sequence. The standard regenerative method [Crane and Iglehart 1975] exploits the i.i.d. property of cycles to construct asymptotically valid confidence intervals.

Many stochastic processes with one regeneration sequence often actually have several such sequences. For example, for a positive-recurrent Markov chain on a discrete state space, there is a regeneration sequence corresponding to each state. Also, Moka and Juneja [2014] describe how many regeneration sequences can be identified in an

This work has been supported in part by the National Science Foundation under Grants No. DMI-9624469, DMI-9900117, CMMI-0926949, CMMI-1200065, and DMS-1331010. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Author's addresses: J. M. Calvin and M. K. Nakayama, Computer Science Department, New Jersey Institute of Technology.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.

© 2013 ACM 1049-3301/2013/07-ARTxx \$15.00

DOI: <http://dx.doi.org/10.1145/0000000.0000000>

important class of queueing network models. The standard regenerative method only uses one sequence, and we study herein methods that take advantage of more than one sequence to try to reduce variance. In this paper, we examine estimators of the second moment of a cycle reward.

Previously proposed methods that exploit multiple regeneration sequences include the semi-regenerative method [Calvin et al. 2001; 2006], permuted regenerative estimators [Calvin and Nakayama 1998; 2000a], the almost regenerative method [Gunther and Wolff 1980], and A -segments [Zhang and Ho 1992]. The almost regenerative method uses two fixed disjoint sets A_0 and A_1 (Gunther and Wolff [1980] instead call the sets U and V) of states to define *almost regeneration points* as the transition times from a state in A_0 to a state in A_1 . A generated sample path is then divided into segments demarcated by pairs of successive almost regeneration points. Both the semi-regenerative method and the A -segment approach divide a simulated sample path determined by successive visits to a fixed set A of states. To construct a permuted regenerative estimator, one starts with a simulated sample path, permutes the regenerative cycles of two regenerative sequences to construct new sample paths, and averages over all permuted paths to obtain the estimator. Computing the permuted estimator, which Calvin and Nakayama [1998] prove always has no greater and often strictly smaller (small-sample) variance than the standard regenerative estimator, actually does not require averaging over the exponential number of permuted paths since efficient closed-form formulae are derived.

In this paper, we develop two new methods based on resampling *trajectory segments* of a generated sample path. For both of these methods, we fix a subset A of states, and the trajectory segments are determined by successive entrances to the set A . One new estimator resamples segments without replacement and the other with replacement, which correspond to U - and V -statistics, respectively (e.g., see Chapters 5 and 6 of Serfling [1980]). One of the contributions of this paper is the development of a framework that allows us to express the standard estimator as a function of the generated trajectory segments, which in turn enables us to define the resampled estimators.

The U -statistic estimator turns out to be equivalent to the permuted regenerative estimator, but the V -statistic estimator is different than existing estimators. We also compare the two resampled estimators with the semi-regenerative estimator, and for the performance measure we consider, we show that they satisfy a total ordering; i.e., for any sample path, the semi-regenerative estimator is at least as large as the V -statistic estimator, and the latter is at least as large as the U -statistic estimator. Also, the U -statistic and permuted estimators are unbiased, and the other two are biased high, with the semi-regenerative estimator having the larger bias. Asymptotically (as the run length grows large), all four estimators are equivalent and have the same central-limit behavior, but as the bias and ordering results show, the estimators differ in the small-sample context. (We develop all of our results for the setting of a discrete-time Markov chain living on a discrete state space; it would be interesting to explore the applicability of the methods to more general regenerative processes [Henderson and Glynn 2001] that also have multiple regeneration sequences.)

The main contributions of our paper are as follows. Our developed framework allows one to define resampled estimators. While only the U -statistic has been proven to have smaller variance than the standard regenerative estimator, our numerical experiments seem to indicate that the V -statistic and semiregenerative estimator also can often have reduced variance (relative to the standard estimator). Another contribution is that while the resampled estimators can be defined in terms of averaging over an exponentially growing number of resampled paths, we derive efficient closed-form formulae for the estimators that can be computed very quickly. (In our numerical experiments, computing the resampled estimators incurred less than 1% overhead compared

to the standard estimator.) Moreover, the total ordering shows an interesting relationship among the various estimators. Our initial focus is on using only two regeneration sequences, but we also extend the methods to handle more than two sequences, which can result in even more variance reduction. We could also apply the resampling approaches to derive corresponding estimators for performance measures other than the second moment of additive cumulative cycle reward considered in this work, but the calculations required to obtain efficient representations of the estimators may be complicated, as they are here. As an alternative (which we do not pursue in this paper), one may average over only a (randomly or deterministically chosen) subset of all resampled paths. This may allow simple extensions to other performance measures, but there also would likely be less variance reduction.

Our estimators share some similarities with bootstrapping and other resampling methods [Efron and Tibshirani 1993]. Also, the resampled estimators bear some resemblance to various standardized time series (STS) methods, such as folded [Antonini et al. 2008] and permuted STS [Calvin and Nakayama 2006], although these techniques only preserve the asymptotic (rather than the small-sample) distribution of the resampled paths.

The rest of the paper has the following organization. In Section 2 we describe the mathematical model and the performance measure considered. We give the semi-regenerative estimator in Section 3, and we present the permuted estimator from Calvin and Nakayama [2000b] in Section 4. Section 5 contains the resampled estimators, and we compare the four estimators in Section 6. Section 7 presents the results of numerical experiments. For the results through Section 7, we restrict ourselves to estimators exploiting only two regeneration sequences, but in Section 8, we show how to combine several two-sequence estimators to obtain a single estimator that then takes advantage of more than two sequences. The longer proofs are in the appendix and an online-only appendix. Some of the results of this paper previously appeared without proofs in Calvin and Nakayama [2002].

2. FRAMEWORK

Suppose $X = (X_j : j = 0, 1, 2, \dots)$ is an irreducible positive-recurrent discrete-time Markov chain living on a discrete state space $S = \{1, 2, 3, \dots\}$. For any state $x \in S$, we define a sequence of regeneration points $T(x) = (T_k(x) : k = 0, 1, 2, \dots)$, where $T_0(x) = \inf\{j \geq 0 : X_j = x\}$ and $T_k(x) = \inf\{j > T_{k-1}(x) : X_j = x\}$, $k \geq 1$. We call the sample-path segment $(X_j : T_{k-1}(x) \leq j < T_k(x))$ the k th $T(x)$ -cycle. Define a “reward” function $f : S \rightarrow \Re$, and for $x \in S$, define

$$\alpha(x) = E_x \left[\left(\sum_{k=0}^{\tau-1} f(X_k) \right)^2 \right], \quad (1)$$

where $\tau = \inf\{k \geq 1 : X_k = w\}$ for a fixed state $w \in S$ and E_x denotes expectation given $X_0 = x$. Our goal is to estimate $\alpha(w)$, which arises, for example, when computing the time-average variance constant of a stochastic process; e.g., see Section 4 of Calvin and Nakayama [2000a] for details. We assume $E_w[(\sum_{j=0}^{\tau-1} |f(X_j)|)^2] < \infty$, which ensures $\alpha(w)$ exists and is finite.

We will compare various estimators of $\alpha(w)$. These include the semi-regenerative estimator [Calvin et al. 2006], the permuted estimator [Calvin and Nakayama 2000b], and two new estimators based on resampling, one of which turns out to be the same as the permuted estimator. To describe the estimators, we first fix a set $A \subset S$ with $w \in A$. For this A , let $T_0 = \inf\{j : X_j \in A\}$ and $T_k = \inf\{j \geq T_{k-1} + 1 : X_j \in A\}$ for $k \geq 1$.

To simplify notation, we will develop estimators when the set $A = \{1, 2\}$ and $w = 1$. (We could always relabel the states if we want to use different states in A and for w . In addition, Section 8 discusses an approach to deal with A having arbitrary cardinality.) Also, we construct the estimators based on a single sample path $\vec{X}_m = (X_j : j = 0, 1, \dots, T_m(1))$ of a fixed number m of $T(1)$ -cycles, with $X_0 = 1$. Note that $\tau = T_1(1)$. Also, define $W = (W_k : k = 0, 1, 2, \dots)$ with $W_k = X_{T_k}$ for $k \geq 0$, which is the embedded chain of X on visits to the set A . Let $M = m + |\{T_k(2) : k \geq 0, T_k(2) < T_m(1)\}|$, which is the number of returns to the set A up to time $T_m(1)$, so $T_M = T_m(1)$. We define a *trajectory* to be a sample path segment between two successive entrances to A . Thus, $(X_j : T_{k-1} \leq j < T_k)$ is a trajectory for each $k \geq 1$.

We define the standard estimator of $\alpha(1)$ based on the sample path \vec{X}_m to be

$$\hat{\alpha}_{\text{STD}}(\vec{X}_m) = \frac{1}{m} \sum_{k=1}^m Y_k^2, \quad (2)$$

where

$$Y_k = \sum_{j=T_{k-1}(1)}^{T_k(1)-1} f(X_j)$$

for $k = 1, 2, \dots, m$. Note that $\hat{\alpha}_{\text{STD}}(\vec{X}_m)$ is unbiased since the Y_k , $k = 1, 2, \dots, m$, are i.i.d. by the regenerative property.

3. SEMI-REGENERATIVE ESTIMATOR

The semi-regenerative method is based on an alternative representation of the performance measure $\alpha(1)$ as a function of expectations of functionals of trajectories. We then construct the semi-regenerative estimator by replacing each expectation with an estimator of it from the generated sample path \vec{X}_m .

Presenting the semi-regenerative estimator requires more notation; its derivation is given in the online-only appendix. Recall that we defined $A = \{1, 2\}$ and M such that $T_M = T_m(1)$. For $x, y \in A$, let $h(x, y) = \sum_{k=0}^{M-1} I(W_k = x, W_{k+1} = y)$, which is the number of times a trajectory along the path \vec{X}_m starts in state x and ends in state y . Let $H(x) = \sum_{k=0}^{M-1} I(W_k = x)$, which is the number of trajectories in \vec{X}_m that begin in state x , and $H(1) = m$. Observe that $h(x, 1) + h(x, 2) = H(x)$ for $x \in A$. Also, $h(1, 2) = h(2, 1)$ since \vec{X}_m is a path of a fixed number of $T(1)$ -cycles.

For $x, y \in A$, define $T_1(x, y) = \inf\{T_k : k \geq 0, W_k = x, W_{k+1} = y\}$ and $T_i(x, y) = \inf\{T_k \geq T_{i-1}(x, y) + 1 : k \geq 0, W_k = x, W_{k+1} = y\}$ for $i \geq 2$. Also, let $T'_1(x, y) = \inf\{T_{k+1} : k \geq 0, W_k = x, W_{k+1} = y\}$ and $T'_i(x, y) = \inf\{T_{k+1} : k \geq 0, T_k \geq T_{i-1}(x, y) + 1, W_k = x, W_{k+1} = y\}$ for $i \geq 2$. Thus, for $k = 1, 2, \dots, h(x, y)$, the k th trajectory starting in state $x \in A$ and ending in state $y \in A$ begins at time $T_k(x, y)$ and finishes at time $T'_k(x, y) - 1$, and we define $Y_k(x, y)$ as the sum of the rewards along that trajectory; i.e.,

$$Y_k(x, y) = \sum_{j=T_k(x, y)}^{T'_k(x, y)-1} f(X_j).$$

For $\ell = 1, 2$, define the sum of the ℓ th power of the $Y_k(x, y)$ as

$$S_\ell(x, y) = \sum_{k=1}^{h(x, y)} Y_k^\ell(x, y),$$

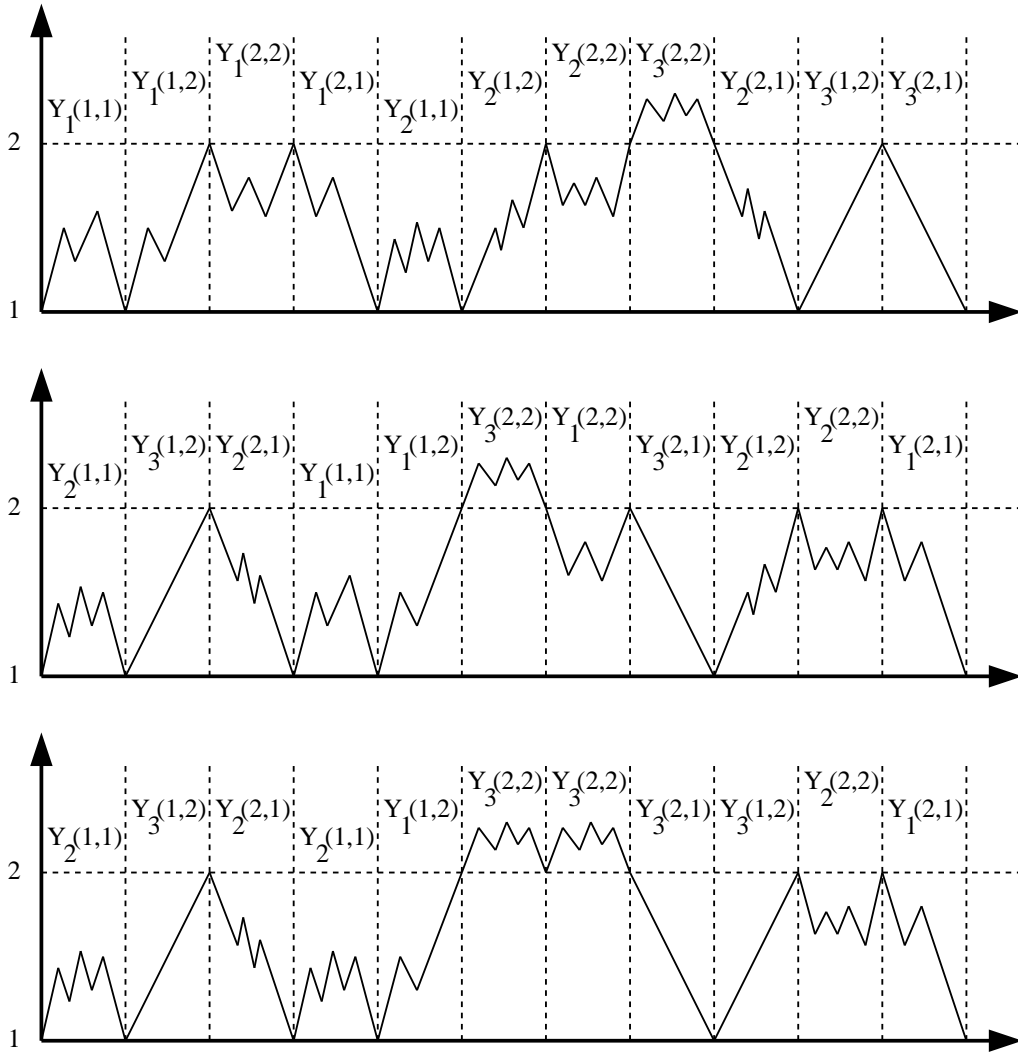


Fig. 1. The original sample path \vec{X}_m (top) and instances of resampled paths without replacement (middle) and with replacement (bottom)

where $S_\ell(x, y) = 0$ when $h(x, y) = 0$.

Figure 1 illustrates some of the notation. The top is an example of a sample path \vec{X}_m ; we later describe the other two sample paths in Section 5. State 1 corresponds to the horizontal axis, and state 2 corresponds to the dashed horizontal line. The number of $T(1)$ -cycles in path \vec{X}_m is $m = 5$. The number of $(1, 1)$ -trajectories is $h(1, 1) = 2$, the number of $(1, 2)$ -trajectories is $h(1, 2) = 3$, the number of $(2, 1)$ -trajectories is $h(2, 1) = 3$, and the number of $(2, 2)$ -trajectories is $h(2, 2) = 3$. Also, $H(1) = 5$ and $H(2) = 6$.

The semi-regenerative estimator of $\alpha(1)$ is

$$\hat{\alpha}_{\text{SR}}(\vec{X}_m) = Q + \frac{2}{mh(1, 2)} S_1^2(2, 2), \quad (3)$$

where

$$Q = \frac{1}{m} \left(S_2(1,1) + S_2(1,2) + S_2(2,1) + S_2(2,2) \right. \\ \left. + \frac{2}{h(1,2)} [S_1(1,2)S_1(2,1) + S_1(2,1)S_1(2,2) + S_1(1,2)S_1(2,2)] \right). \quad (4)$$

If $h(1,2) = 0$, then $S_\ell(1,2) = S_\ell(2,1) = S_1(2,2) = 0$ for $\ell = 1, 2$, so $Q = S_2(1,1)/m$ since we define $0/0 = 0$, and $\hat{\alpha}_{\text{SR}}(\vec{X}_m) = Q = \hat{\alpha}_{\text{STD}}(\vec{X}_m)$.

4. PERMUTED ESTIMATOR

Recall we defined the standard estimator of $\alpha(1)$ as $\hat{\alpha}_{\text{STD}}(\vec{X}_m)$ in (2). To obtain the permuted estimator of $\alpha(1)$ based on the sample path \vec{X}_m , we first permute $T(2)$ -cycles and $T(1)$ -cycles in \vec{X}_m to obtain a permuted path \vec{X}'_m , and averaging $\hat{\alpha}_{\text{STD}}(\vec{X}'_m)$ over all possible permuted paths \vec{X}'_m yields the permuted estimator. Calvin and Nakayama [2000b] showed that the permuted estimator is

$$\hat{\alpha}_{\text{P}}(\vec{X}_m) = Q + \frac{2}{m(h(1,2) + 1)} (S_1^2(2,2) - S_2(2,2)). \quad (5)$$

When $h(1,2) = 0$, we have $S_\ell(2,2) = 0$ for $\ell = 1, 2$, so $\hat{\alpha}_{\text{P}}(\vec{X}_m) = \hat{\alpha}_{\text{STD}}(\vec{X}_m)$.

5. RESAMPLED ESTIMATORS

We now present two estimators obtained via resampling the (x, y) -trajectories for $x, y \in A$. One estimator is based on resampling the (x, y) -trajectories without replacement (i.e., a permutation of trajectories), and the other resamples with replacement.

To explain how this is done, recall sample path \vec{X}_m has m $T(1)$ -cycles. We want to create a new path by resampling trajectories from \vec{X}_m so that the number of (x, y) -trajectories for each $x, y \in A$ is the same as in \vec{X}_m ; i.e., the values of $m, h(1,1), h(1,2), h(2,1), h(2,2)$ do not change. Resampling with or without replacement determines the collection and order of the $h(x, y)$ resampled (x, y) -trajectories for each $x, y \in A$ used to construct a new path. Resampling the (x, y) -trajectories without replacement results in a permutation of those trajectories. If we resample with replacement, some trajectories may appear more than once, and others may not be included. The U -statistic (resp., V -statistic) estimator results from averaging over all resampled paths without (resp., with) replacement.

We now need to understand the form of the m $T(1)$ -cycles in a resampled path. Each of the $h(1,1)$ resampled $(1,1)$ -trajectories in the constructed path is itself a $T(1)$ -cycle, and these cycles do not visit state 2. The other $T(1)$ -cycles contain a hit to state 2, so each of these cycles starts with a $(1,2)$ -trajectory and ends with a $(2,1)$ -trajectory, with zero or more $(2,2)$ -trajectories sandwiched between. Hence, constructing these $T(1)$ -cycles begins with pairing starting $(1,2)$ -trajectories with ending $(2,1)$ -trajectories. Once we have all $h(1,2) = h(2,1)$ such pairings, we then allocate the $h(2,2)$ resampled $(2,2)$ -trajectories among them.

For example, in Figure 1 the original path \vec{X}_m at the top has $m = 5$ $T(1)$ -cycles, of which $h(1,1) = 2$ (resp., $h(1,2) = h(2,1) = 3$) contain no (resp., at least one) hit to state 2. Of the $T(1)$ -cycles that visit state 2, the first (i.e., the one starting with the first $(1,2)$ -trajectory) has one $(2,2)$ -trajectory, the second has two $(2,2)$ -trajectories, and the third has none.

The middle path in Figure 1 is constructed from resampling trajectories without replacement. For each $x, y \in A$, the collection of $h(x, y)$ resampled (x, y) -trajectories is the same as in \vec{X}_m but in a different order (a permutation). The $T(1)$ -cycles with at least one hit to state 2 have different pairings of $(1, 2)$ - and $(2, 1)$ -trajectories to begin and end the cycles, and the $(2, 2)$ -trajectories are dispersed among those $T(1)$ -cycles differently than in \vec{X}_m . The first such $T(1)$ -cycle has no $(2, 2)$ -trajectories, the second has two $(2, 2)$ -trajectories, and the third has one $(2, 2)$ -trajectory.

The bottom path in Figure 1 is constructed by resampling trajectories with replacement. Some trajectories are used more than once: e.g., the second $(1, 1)$ -trajectory (from \vec{X}_m), the third $(1, 2)$ -trajectory, and the third $(2, 2)$ -trajectory. Other trajectories from \vec{X}_m do not appear in the new path, such as the first $(1, 1)$ -trajectory, the second $(1, 2)$ -trajectory, and the first $(2, 2)$ -trajectory.

We next develop notation to express these ideas precisely, which will allow us to show how to average over all resampled trajectories to obtain the resampled estimators. For $x, y \in A$, let $\Lambda(x, y)$ be the set of permutations of $(1, 2, \dots, h(x, y))$, and let $\Lambda = \times_{x, y \in A} \Lambda(x, y)$, where \times denotes Cartesian product. Define $\hat{\Lambda}(x, y) = \{(i_1, i_2, \dots, i_{h(x, y)}) : 1 \leq i_j \leq h(x, y) \text{ for } j = 1, 2, \dots, h(x, y)\}$, which is the set of all $h(x, y)$ -dimensional vectors in which the components in each vector are selected with replacement from $\{1, 2, \dots, h(x, y)\}$. Let $\hat{\Lambda} = \times_{x, y \in A} \hat{\Lambda}(x, y)$, and $\Lambda \subset \hat{\Lambda}$. Note that $\hat{\Lambda}$ consists of all possible resamplings of trajectories with replacement, where the number of resampled (x, y) -trajectories remains as $h(x, y)$ for each $x, y \in A$, and Λ is the subset of those when resampling without replacement. Define $\Delta = \{(i_1, i_2, \dots, i_{h(1,2)+1}) : i_1 = 1, i_{h(1,2)+1} = h(2, 2) + 1, i_j \leq i_{j+1} \text{ for } j = 1, 2, \dots, h(1, 2)\}$, which will be used to allocate the resampled $(2, 2)$ -trajectories to the $T(1)$ -cycles with a hit to state 2. Specifically, for a sequence $(i_1, i_2, \dots, i_{h(1,2)+1}) \in \Delta$, we will use the $(2, 2)$ -trajectories having positions $i_j, i_j + 1, i_j + 2, \dots, i_{j+1} - 1$ within $\Lambda(2, 2)$ or $\hat{\Lambda}(2, 2)$ in the j th $T(1)$ -cycle that visits state 2. If $i_j = i_{j+1}$, then the j th such $T(1)$ -cycle has no $(2, 2)$ -segments. Let $\Gamma = \Lambda \times \Delta$ and $\hat{\Gamma} = \hat{\Lambda} \times \Delta$. Consider a $K = (K(x, y) : x, y \in A) \in \hat{\Lambda}$ and $K(x, y) = (K_i(x, y) : i = 1, 2, \dots, h(x, y)) \in \hat{\Lambda}(x, y)$, so $K(x, y)$ is a particular ordered collection of $h(x, y)$ resampled (x, y) -trajectories. When constructing a resampled path using K , the order of the $(1, 2)$ -trajectories and $(2, 1)$ -trajectories in $K(1, 2)$ and $K(2, 1)$, respectively, determines how the starting $(1, 2)$ -trajectories and the ending $(2, 1)$ -trajectories are paired up in the $T(1)$ -cycles with a hit to state 2. Also, let $D = (D_1, D_2, \dots, D_{h(1,2)+1}) \in \Delta$, which is a particular allocation scheme for the $(2, 2)$ -trajectories to be distributed to the $T(1)$ -cycles visiting state 2. We now define a function $g : \hat{\Lambda} \times \Delta \rightarrow \mathfrak{R}$ to construct a sample path from \vec{X}_m for the particular resampled trajectories $K \in \hat{\Lambda}$ and allocation scheme $D \in \Delta$ as

$$\begin{aligned} g(K, D) &= \frac{1}{m} \left[\sum_{i=1}^{h(1,1)} Y_{K_i(1,1)}^2(1, 1) + \sum_{i=1}^{h(1,2)} \left(Y_{K_i(1,2)}(1, 2) + Y_{K_i(2,1)}(2, 1) + \sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2, 2) \right)^2 \right] \\ &= \frac{1}{m} \left[\sum_{i=1}^{h(1,1)} Y_{K_i(1,1)}^2(1, 1) + \sum_{i=1}^{h(1,2)} \left(Y_{K_i(1,2)}^2(1, 2) + Y_{K_i(2,1)}^2(2, 1) \right) \right] \end{aligned}$$

$$\begin{aligned}
& + \left\{ \sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2,2) \right\}^2 + 2Y_{K_i(1,2)}(1,2)Y_{K_i(2,1)}(2,1) \\
& + 2 \left\{ Y_{K_i(1,2)}(1,2) + Y_{K_i(2,1)}(2,1) \right\} \sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2,2) \Bigg]. \tag{6}
\end{aligned}$$

If $h(2,2) = 0$, then there are no $(2,2)$ -trajectories, so all of the terms in (6) that include some $Y_{K_j(2,2)}(2,2)$ are zero. If $h(1,2) = 0$, then only the first term in (6) remains.

Suppose we set $K'(x,y) = (1,2,\dots,h(x,y))$ for all $x,y \in A$, and let $K' = (K'(x,y) : x,y \in A)$. Note that $K'(x,y) \in \Lambda(x,y)$ for all $x,y \in A$, and $K' \in \Lambda$. Define $D' = (D'_1, D'_2, \dots, D'_{h(1,2)+1})$ with $D'_i = \min\{j : T_j(2,2) > T_i(1,2)\}$ for $i \geq 1$, so $D' \in \Delta$. Then

$$g(K', D') = \hat{\alpha}_{\text{STD}}(\vec{X}_m),$$

the standard estimator of $\alpha(1)$.

We now define the resampled estimators

$$\hat{\alpha}_U(\vec{X}_m) = \frac{1}{|\Gamma|} \sum_{(K,D) \in \Gamma} g(K,D) \tag{7}$$

and

$$\hat{\alpha}_V(\vec{X}_m) = \frac{1}{|\hat{\Gamma}|} \sum_{(K,D) \in \hat{\Gamma}} g(K,D), \tag{8}$$

where $\hat{\alpha}_U(\vec{X}_m)$ is the estimator based on resampling without replacement, which we can think of as a type of U -statistic, and $\hat{\alpha}_V(\vec{X}_m)$ is the estimator with replacement, a type of V -statistic.

Recall that the middle (resp., bottom) part of Figure 1 displays a resampled path sampled without (resp., with) replacement from the original sample path \vec{X}_m , on the top of the figure. The original path \vec{X}_m has $K'(1,1) = (1,2)$, $K'(1,2) = (1,2,3)$, $K'(2,1) = (1,2,3)$, $K'(2,2) = (1,2,3)$, and $D' = (1,2,4,4)$. The path shown in the middle, which is resampled without replacement, has $K(1,1) = (2,1)$, $K(1,2) = (3,1,2)$, $K(2,1) = (2,3,1)$, $K(2,2) = (3,1,2)$, and $D = (1,1,3,4)$. The bottom sample path in Figure 1, which is a resampled path with replacement, has $K(1,1) = (2,2)$, $K(1,2) = (3,1,3)$, $K(2,1) = (2,3,1)$, $K(2,2) = (3,3,2)$, and $D = (1,1,3,4)$.

The following result is established in the Appendix.

THEOREM 5.1. *For all \vec{X}_m , $\hat{\alpha}_U(\vec{X}_m) = \hat{\alpha}_P(\vec{X}_m)$ and*

$$\hat{\alpha}_V(\vec{X}_m) = Q + \frac{2(h(2,2) - 1)}{m h(2,2) (h(1,2) + 1)} S_1^2(2,2), \tag{9}$$

where Q is defined in (4).

When $h(2,2) = 0$, we have $\hat{\alpha}_V(\vec{X}_m) = Q$ since $S_1^2(2,2) = 0$, so $\hat{\alpha}_V(\vec{X}_m) = \hat{\alpha}_P(\vec{X}_m) = \hat{\alpha}_{\text{SR}}(\vec{X}_m) = Q$. Moreover, $\hat{\alpha}_V(\vec{X}_m) = \hat{\alpha}_P(\vec{X}_m) = \hat{\alpha}_{\text{SR}}(\vec{X}_m) = \hat{\alpha}_{\text{STD}}(\vec{X}_m)$ when $h(1,2) = 0$.

6. COMPARING THE ESTIMATORS

We begin by comparing the small-sample behavior of the semi-regenerative, permuted, U -statistic, and V -statistic estimators.

THEOREM 6.1. *For all \vec{X}_m and m ,*

$$\hat{\alpha}_{SR}(\vec{X}_m) \geq \hat{\alpha}_V(\vec{X}_m) \geq \hat{\alpha}_U(\vec{X}_m) = \hat{\alpha}_P(\vec{X}_m), \quad (10)$$

where

- the first inequality is strict if and only if $S_1(2, 2) \neq 0$;
- the second inequality is strict if and only if there exists $i, j \in \{1, 2, \dots, h(2, 2)\}$ such that $Y_i(2, 2) \neq Y_j(2, 2)$;
- $\hat{\alpha}_{SR}(\vec{X}_m) > \hat{\alpha}_P(\vec{X}_m)$ if and only if some $Y_k(2, 2) \neq 0$.

Moreover, if $E_1[(\sum_{j=0}^{\tau-1} |f(X_j)|)^2] < \infty$, then

$$E[\hat{\alpha}_{SR}(\vec{X}_m)] \geq E[\hat{\alpha}_V(\vec{X}_m)] \geq E[\hat{\alpha}_U(\vec{X}_m)] = E[\hat{\alpha}_P(\vec{X}_m)] = E[\hat{\alpha}_{STD}(\vec{X}_m)] = \alpha(1),$$

where

- the first inequality is strict if $P\{Y_1(2, 2) = 0\} < 1$;
- the second inequality is strict if $P\{Y_1(2, 2) = Y_2(2, 2)\} < 1$.

Thus, $\hat{\alpha}_{SR}(\vec{X}_m)$ and $\hat{\alpha}_V(\vec{X}_m)$ are biased high.

Note that the standard estimator $\hat{\alpha}_{STD}(\vec{X}_m)$ defined in (2) does not appear in the ordering given in (10). Since both $\hat{\alpha}_{STD}(\vec{X}_m)$ and $\hat{\alpha}_P(\vec{X}_m)$ are unbiased, the two estimators cannot satisfy an ordering for all \vec{X}_m .

PROOF. Calvin and Nakayama [2000b] prove that $E[\hat{\alpha}_P(\vec{X}_m)] = \alpha(1)$. (The expectation exists, as well as the others considered, under the moment condition assumed.) When $h(1, 2) > 0$,

$$\begin{aligned} D_{SR,P}(\vec{X}_m) &\equiv \hat{\alpha}_{SR}(\vec{X}_m) - \hat{\alpha}_P(\vec{X}_m) \\ &= \frac{2}{mh(1, 2)} S_1^2(2, 2) - \frac{2(S_1^2(2, 2) - S_2(2, 2))}{m(h(1, 2) + 1)} \end{aligned} \quad (11)$$

$$= \frac{2}{m(h(1, 2) + 1)} \left(\frac{S_1^2(2, 2)}{h(1, 2)} + S_2(2, 2) \right) \geq 0, \quad (12)$$

where the last inequality is strict if and only if some $Y_k(2, 2) \neq 0$. If $h(1, 2) = 0$, then trivially there does not exist $Y_k(2, 2) \neq 0$, and $\hat{\alpha}_P(\vec{X}_m) = \hat{\alpha}_{SR}(\vec{X}_m)$ and $D_{SR,P}(\vec{X}_m) = 0$. In addition, it is simple to show that $E[D_{SR,P}(\vec{X}_m)] > 0$ when $P\{Y_1(2, 2) = 0\} < 1$, so the semi-regenerative estimator is biased high.

Now if $h(1, 2) > 0$ and $h(2, 2) > 0$, then

$$\frac{2}{mh(1, 2)} > \frac{2(h(2, 2) - 1)}{mh(2, 2)(h(1, 2) + 1)},$$

so $\hat{\alpha}_{SR}(\vec{X}_m) \geq \hat{\alpha}_V(\vec{X}_m)$, with strict inequality if and only if $S_1(2, 2) \neq 0$. If $h(1, 2) = 0$ or $h(2, 2) = 0$, then $S_1(2, 2) = 0$ vacuously and $\hat{\alpha}_V(\vec{X}_m) = \hat{\alpha}_{SR}(\vec{X}_m)$. Also, comparing $\hat{\alpha}_V(\vec{X}_m)$ and $\hat{\alpha}_P(\vec{X}_m)$, we have if $h(1, 2) > 0$ and $h(2, 2) > 0$, then

$$\begin{aligned} D_{V,P}(\vec{X}_m) &\equiv \hat{\alpha}_V(\vec{X}_m) - \hat{\alpha}_P(\vec{X}_m) \\ &= \frac{2(h(2, 2) - 1)}{mh(2, 2)(h(1, 2) + 1)} S_1^2(2, 2) - \frac{2}{m(h(1, 2) + 1)} (S_1^2(2, 2) - S_2(2, 2)) \\ &= 2 \left(\frac{h(2, 2)S_2(2, 2) - S_1^2(2, 2)}{mh(2, 2)(h(1, 2) + 1)} \right) \geq 0 \end{aligned} \quad (13)$$

since Jensen's inequality implies $n \sum_{i=1}^n a_i^2 \geq (\sum_{i=1}^n a_i)^2$ for any real-valued a_i , $i = 1, 2, \dots, n$. The inequality in (13) is strict if and only if not all the $Y_k(2, 2)$ are the same. If $h(1, 2) = 0$ or $h(2, 2) = 0$, then there are no $Y_k(2, 2)$ terms, and $\hat{\alpha}_V(\vec{X}_m) = \hat{\alpha}_P(\vec{X}_m)$. Finally, it is simple to see that $E[D_{V,P}(\vec{X}_m)] > 0$ if $P\{Y_1(2, 2) = Y_2(2, 2)\} < 1$. \square

In the difference $D_{SR,P}(\vec{X}_m)$ between the semi-regenerative and permuted estimators, the second term in the right-hand side of (11), which is from the permuted estimator $\hat{\alpha}_P(\vec{X}_m)$, subtracts out the sum of the squares of the $Y_k(2, 2)$ (i.e., $S_2(2, 2)$) from the square of the sum (i.e., $S_1^2(2, 2)$). The first term of (11), which appears in the semi-regenerative estimator $\hat{\alpha}_{SR}(\vec{X}_m)$, only has the square of the sum. A similar situation occurs with the difference $D_{V,P}(\vec{X}_m)$ between the V -statistic estimator and the permuted estimator.

We now compare some asymptotic properties of the estimators. Let $\xrightarrow{\mathcal{D}}$ denote convergence in distribution and $N(a, b)$ denote a normal distribution with mean a and variance b .

THEOREM 6.2. *Suppose that $E_1[(\sum_{j=0}^{r-1} |f(X_j)|)^4] < \infty$. Then for $\hat{\alpha}(\vec{X}_m)$ defined as either $\hat{\alpha}_P(\vec{X}_m)$, $\hat{\alpha}_{SR}(\vec{X}_m)$, or $\hat{\alpha}_V(\vec{X}_m)$,*

$$\hat{\alpha}(\vec{X}_m) \rightarrow \alpha(1), \text{ a.s.},$$

and

$$\sqrt{m} [\hat{\alpha}(\vec{X}_m) - \alpha(1)] \xrightarrow{\mathcal{D}} N(0, \sigma^2)$$

as $m \rightarrow \infty$, where σ^2 is the same for all the estimators and is given in Calvin and Nakayama [2000b].

PROOF. Calvin and Nakayama [2000b] show that the results hold when $\hat{\alpha}(\vec{X}_m) = \hat{\alpha}_P(\vec{X}_m)$, so to establish our theorem, it suffices to prove that $\sqrt{m}D_{SR,P}(\vec{X}_m) \rightarrow 0$ a.s. and $\sqrt{m}D_{V,P}(\vec{X}_m) \rightarrow 0$ a.s. as $m \rightarrow \infty$. Using the representation of $D_{SR,P}(\vec{X}_m)$ in (12), we will show $\sqrt{m}D_{SR,P}(\vec{X}_m) \rightarrow 0$ a.s.; the other result can be similarly established.

For $p = 1, 2$, define

$$Z_{p,k} = \sum_{l \in B_k} Y_l^p(2, 2),$$

where B_k is the set of indices of $(2, 2)$ -trajectories that are contained in the k th $T(1)$ -cycle; i.e., $B_k = \{j = 1, 2, \dots, h(2, 2) : T_{k-1}(1) \leq T_j(2, 2) \leq T_k(1)\}$. Note that $S_p(2, 2) = \sum_{k=1}^m Z_{p,k}$ for $p = 1, 2$. Also, define $Z_{3,k} = 1$ if there exists some i such that $T_{k-1}(1) \leq T_i(2) \leq T_k(1)$, and $Z_{3,k} = 0$ otherwise. Thus, $h(1, 2) = \sum_{k=1}^m Z_{3,k}$. Because X is regenerative, for $p = 1, 2, 3$, the $Z_{p,k}$, $k = 1, 2, \dots, m$, are i.i.d., and Lemma 4 of Calvin and Nakayama [2000a] establishes that $0 < E_1[Z_{3,1}] < 1$, $|E_1[Z_{1,1}]| < \infty$, and $|E_1[Z_{2,1}]| < \infty$. Since X is irreducible and positive recurrent, $h(1, 2) \rightarrow \infty$ a.s. as $m \rightarrow \infty$. Thus, (12) implies

$$\begin{aligned} \sqrt{m}D_{SR,P}(\vec{X}_m) &= \frac{2}{\sqrt{m}} \left(\frac{m}{\sum_{k=1}^m Z_{3,k} + 1} \right) \left(\left[\frac{1}{m} \sum_{k=1}^m Z_{1,k} \right]^2 \frac{m}{\sum_{k=1}^m Z_{3,k}} + \frac{1}{m} \sum_{k=1}^m Z_{2,k} \right) \\ &\rightarrow 0 \cdot \left(\frac{1}{E_1[Z_{3,1}]} \right) \left(\frac{(E_1[Z_{1,1}])^2}{E_1[Z_{3,1}]} + E_1[Z_{2,1}] \right) = 0 \text{ a.s.} \end{aligned}$$

as $m \rightarrow \infty$. \square

Table I. Sample mean and sample variance of the estimators for the Ehrenfest model with $[B = 9, A = \{2, 4\}, w = 2]$ and $[B = 900, A = \{410, 420\}, w = 410]$.

estimator	$B = 9$		$B = 900$	
	mean ($\times 10^3$)	variance ($\times 10^4$)	mean ($\times 10^{12}$)	variance ($\times 10^{24}$)
regenerative: $\hat{\alpha}_{\text{STD}}(\vec{X}_m)$	5.1602	6.3328	7.7372	4.7519
permuted: $\hat{\alpha}_{\text{P}}(\vec{X}_m)$	5.1616	5.3448	7.8012	3.5472
V-statistic: $\hat{\alpha}_{\text{V}}(\vec{X}_m)$	5.1620	5.3508	7.8400	3.5679
semi-regenerative: $\hat{\alpha}_{\text{SR}}(\vec{X}_m)$	5.1639	5.3545	7.8986	3.6169

Table II. Sample means and sample variances of the estimators for the Ehrenfest model with $B = 9, A = \{1, v\}$, and $w = 1$. For the standard regenerative estimator the sample mean across the 10^3 replications was 5.55×10^4 and sample variance was 3.19×10^8 .

	Estimator	v							
		0	2	3	4	5	6	7	8
sample mean ($\times 10^4$)	permuted	5.55	5.49	5.52	5.48	5.56	5.50	5.44	5.47
	V-statistic	5.55	5.50	5.53	5.48	5.56	5.51	5.45	5.48
	semireg	5.55	5.58	5.61	5.57	5.66	5.60	5.52	5.51
sample variance ($\times 10^8$)	permuted	3.19	2.15	2.30	2.26	2.35	2.75	2.63	2.62
	V-statistic	3.19	2.16	2.30	2.26	2.35	2.75	2.64	2.63
	semireg	3.19	2.21	2.37	2.33	2.42	2.84	2.73	2.68

7. NUMERICAL EXPERIMENTS

We ran experiments comparing the estimators for computing $\alpha(w)$ in (1) for a discrete-time Markov chain ($X_n : n = 0, 1, 2, \dots$) with different choices for the set A and state w . The reward function for all our experiments is $f(x) = x$ for all states $x \in S$. The first model for the experiments is the Ehrenfest urn model with state space $S = \{0, 1, 2, \dots, B - 1\}$ and transition probabilities $P_{i,i+1} = (B - 1 - i)/(B - 1)$ for $0 \leq i < B - 1$, and $P_{i,i-1} = 1 - P_{i,i+1}$ for $1 \leq i < B$. We first took $B = 9, A = \{2, 4\}$, and $w = 2$, and we ran experiments to estimate $\alpha(w)$ based on 5,000 w -cycles for each of 1,000 independent replications. The second and third columns of Table I show for these parameters the sample mean and sample variance across the independent replications for each of the estimators described in this paper, and also the standard regenerative estimator. The value being estimated is 5.1616×10^3 when $w = 2$. The time to compute the resampled estimators was less than 1% more than the time to compute the standard regenerative estimator.

The last two columns of Table I show the results of similar experiments for the Ehrenfest model with $B = 900, A = \{410, 420\}$, and $w = 410$. The experiments are based on 1,000 w -cycles for each of 1,000 independent replications. Each of the non-standard estimators reduced variance by about 25% compared to the regenerative estimator.

Returning to the model with $B = 9$, Table II shows the results of an experiment with $A = \{v, w\}$ and $w = 1$ for different choices of $v \in S$. In this case the value being estimated is 5.4673×10^4 . We now used 100 w -cycles because of the long average cycle length. The point estimators now are not as close to the value being estimated as in the first case, and possible explanations for this are that in the second experiment, the estimators have much more variability and many fewer cycles were simulated. In both cases observe that the sample means of the estimators satisfy the ordering given in Theorem 6.1. Tables I and II show the choices of A and w affect the variances of and differences in the estimators. For $v = 2$ in Table II, the variance is reduced about 30% compared to the standard regenerative estimator.

We also ran experiments on the discrete-time embedded Markov chain of the M/M/1 queueing process with several values of the traffic intensity ρ (the ratio of the mean service time divided by the mean interarrival time). This chain takes values in the nonnegative integers, and has transition probabilities $P_{0,1} = 1$ and $P_{i,i+1} = \rho/(1 + \rho) =$

Table III. Comparison of estimators for M/M/1 queue with $[\rho = 0.3, A = \{0, 3\}, w = 0]$, $[\rho = 0.5, A = \{0, 3\}, w = 0]$, $[\rho = 0.9, A = \{0, 3\}, w = 0]$, and $[\rho = 0.9, A = \{0, 9\}, w = 0]$.

estimator	$\rho = 0.3$		$\rho = 0.5$		$\rho = 0.9, A = \{0, 3\}$		$\rho = 0.9, A = \{0, 9\}$	
	mean ($\times 10$)	variance ($\times 10$)	mean ($\times 10^2$)	variance ($\times 10^3$)	mean ($\times 10^6$)	variance ($\times 10^{12}$)	mean ($\times 10^6$)	variance ($\times 10^{12}$)
regenerative	3.525	2.674	3.822	7.813	3.229	4.273	3.229	4.273
permuted	3.529	2.560	3.820	7.533	3.227	4.255	3.227	4.262
V-statistic	3.531	2.573	3.824	7.563	3.231	4.266	3.240	4.307
semi-regenerative	3.536	2.580	3.827	7.569	3.231	4.266	3.243	4.309

$1 - P_{i,i-1}$ for $i > 0$. In columns 2–7 of Table III, we took $A = \{0, 3\}$ with $w = 0$ for $\rho = 0.3, 0.5$, and 0.9 . The last two columns of Table III are again for $\rho = 0.9$ but instead with $A = \{0, 9\}$ and $w = 0$. In all cases we again see that the estimators satisfy the ordering given in Theorem 6.1. The choice of A and w affects the differences in the estimators and their variances. Also, the reduction in variance is not as great for the M/M/1 experiments as for the Ehrenfest results.

8. USING MORE THAN TWO SEQUENCES

Models possessing two regeneration sequences often have more than two such sequences. For example, for a positive-recurrent Markov chain on a discrete state space S , there is a regeneration sequence corresponding to hits to each fixed state $x \in S$. We now discuss how to adapt to our setting an idea of Heidelberger [1980] (see also Calvin and Nakayama [2004]) for combining multiple regenerative estimators to obtain a new estimator.

Fix a state w , and suppose we want to estimate the performance measure $\alpha(w)$. Suppose that $x_0 = w, x_1, x_2, \dots, x_d$ are $d + 1$ distinct points in S , and let $T(x_0), T(x_1), \dots, T(x_d)$ be the corresponding regeneration sequences, with $T(x_l) = \{T_i(x_l) : i = 0, 1, 2, \dots\}$. We simulate a single sample path \vec{X}_m consisting of m $T(w)$ -cycles, and then from this path construct d (resampled, permuted, or semi-regenerative) estimators $\hat{\alpha}_l(\vec{X}_m)$ for $l = 1, 2, \dots, d$. The estimator $\hat{\alpha}_l(\vec{X}_m)$ is obtained by applying one of the methods discussed in this paper using the $T(w)$ - and $T(x_l)$ -sequences (or letting $A = \{w, x_l\}$). Then define the combined estimator

$$\hat{\alpha}_\lambda^*(\vec{X}_m) = \sum_{l=1}^d \lambda_l \hat{\alpha}_l(\vec{X}_m) \quad (14)$$

for some constants $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_d)$ such that $\sum_{l=1}^d \lambda_l = 1$. Since each $\hat{\alpha}_l(\vec{X}_m) \rightarrow \alpha(w)$ a.s. as $m \rightarrow \infty$ by Theorem 6.2, $\hat{\alpha}_\lambda^*(\vec{X}_m) \rightarrow \alpha(w)$ a.s. as $m \rightarrow \infty$.

Let $\hat{\alpha}_{U,\lambda}^*(\vec{X}_m)$, $\hat{\alpha}_{P,\lambda}^*(\vec{X}_m)$, $\hat{\alpha}_{V,\lambda}^*(\vec{X}_m)$, and $\hat{\alpha}_{SR,\lambda}^*(\vec{X}_m)$ denote the combined U -statistic, permuted, V -statistic and semi-regenerative estimators, respectively. Since each of the non-combined estimators satisfy the orderings in Theorem 6.1, we obtain the following result, which shows the combined estimators satisfy the same ordering when they use the same set of nonnegative weights. Let e denote the d -vector of all 1's, and let superscript \top denote transpose.

COROLLARY 8.1. *For all \vec{X}_m , m , and $\lambda = (\lambda_1, \dots, \lambda_d)^\top$ with each $\lambda_l \geq 0$ and $\lambda^\top e = 1$,*

$$\hat{\alpha}_{SR,\lambda}^*(\vec{X}_m) \geq \hat{\alpha}_{V,\lambda}^*(\vec{X}_m) \geq \hat{\alpha}_{U,\lambda}^*(\vec{X}_m) = \hat{\alpha}_{P,\lambda}^*(\vec{X}_m),$$

where

— the first inequality is strict if and only if $S_1(x_l, x_l) \neq 0$ for some $l = 1, \dots, d$;

- the second inequality is strict if and only if there exist $l \in \{1, \dots, d\}$ and $i, j \in \{1, 2, \dots, h(x_l, x_l)\}$ such that $Y_i(x_l, x_l) \neq Y_j(x_l, x_l)$;
- $\hat{\alpha}_{SR}(\vec{X}_m) > \hat{\alpha}_P(\vec{X}_m)$ if and only if some $Y_k(x_l, x_l) \neq 0$ for some $l = 1, \dots, d$.

Moreover, if $E_1[(\sum_{j=0}^{\tau-1} |f(X_j)|)^2] < \infty$, then

$$E[\hat{\alpha}_{SR,\lambda}^*(\vec{X}_m)] \geq E[\hat{\alpha}_{V,\lambda}^*(\vec{X}_m)] \geq E[\hat{\alpha}_{U,\lambda}^*(\vec{X}_m)] = E[\hat{\alpha}_{P,\lambda}^*(\vec{X}_m)] = \alpha(w),$$

where

- the first inequality is strict if $P\{Y_1(x_l, x_l) = 0\} < 1$ for some $l = 1, \dots, d$;
- the second inequality is strict if $P\{Y_1(x_l, x_l) = Y_2(x_l, x_l)\} < 1$ for some $l = 1, \dots, d$.

Thus, $\hat{\alpha}_{SR,\lambda}^*(\vec{X}_m)$ and $\hat{\alpha}_{V,\lambda}^*(\vec{X}_m)$ are biased high.

We now discuss how to determine the optimal value of λ to minimize the asymptotic variance of $\hat{\alpha}_\lambda^*(\vec{X}_m)$. If $E_w[(\sum_{j=0}^{\tau-1} |f(X_j)|)^4] < \infty$, then it can be shown that

$$\sqrt{m}((\hat{\alpha}_1(\vec{X}_m), \dots, \hat{\alpha}_d(\vec{X}_m)) - (\alpha(w), \dots, \alpha(w))) \xrightarrow{\mathcal{D}} N(0, C)$$

as $m \rightarrow \infty$, where $N(0, C)$ denotes a d -dimensional normal random vector with mean vector 0 and covariance matrix $C = (C_{i,j} : i, j = 1, 2, \dots, d)$. Applying the delta method (e.g., Corollary 3.3 on p. 124 of Serfling [1980]) then yields

$$\sqrt{m}(\hat{\alpha}_\lambda^*(\vec{X}_m) - \alpha(w)) \xrightarrow{\mathcal{D}} N(0, \sigma_\lambda^2)$$

as $m \rightarrow \infty$, where $\sigma_\lambda^2 = \lambda^\top C \lambda$, which is the asymptotic variance of $\hat{\alpha}_\lambda^*(\vec{X}_m)$, and superscript \top denotes transpose.

Now we want to choose λ with $\lambda^\top e = 1$ to minimize σ_λ^2 . Assuming that C is positive definite, the unique solution is

$$\lambda_* = \frac{C^{-1}e}{e^\top C^{-1}e}. \quad (15)$$

We can estimate the value of λ_* using “sectioning” as follows. The idea is to divide the sample path \vec{X}_m of m $T(w)$ -cycles into n non-overlapping sections, each consisting of p $T(w)$ -cycles, where $m = np$. Then calculate the d estimators for each section, and use these estimators to estimate C . Specifically, the k th section, $k = 1, 2, \dots, n$, starts at time $T_{(k-1)p}(w)$, ends at time $T_{kp}(w) - 1$, and consists of $T(w)$ -cycles $(k-1)p+1, (k-1)p+2, \dots, kp$. The regenerative property implies the n sections are independent. For $l = 1, 2, \dots, d$, and $k = 1, 2, \dots, n$, let $\hat{\alpha}_{l,k}(\vec{X}_m)$ be the (resampled, permuted, or semi-regenerative) estimator obtained from the k th section using sequences $T(w)$ and $T(x_l)$ (or $A = \{w, x_l\}$). Then we estimate $C_{i,j}$ by

$$\hat{C}_{i,j} = \frac{m}{n(n-1)} \sum_{k=1}^n [\hat{\alpha}_{i,k}(\vec{X}_m) - \tilde{\alpha}_i(\vec{X}_m)] [\hat{\alpha}_{j,k}(\vec{X}_m) - \tilde{\alpha}_j(\vec{X}_m)], \quad (16)$$

where $\tilde{\alpha}_l(\vec{X}_m) = \frac{1}{n} \sum_{k=1}^n \hat{\alpha}_{l,k}(\vec{X}_m)$. Substituting the estimates $\hat{C}_{i,j}$ for $C_{i,j}$ in (15) provides us with an estimate of λ_* . Asymptotically, the combined estimator with the exact λ_* will have no greater variance than the estimator for any single pair of regeneration sequences (but this may not be the case in practice with the estimated λ_*). Thus, when presented with more than two possible regeneration sequences, the user can use them all to obtain a combined estimator that is asymptotically better than the estimator from any two sequences.

8.1. Numerical Results

We ran an experiment with the combined estimator in (14) with estimates of the optimal weights in (15) using (16). The model we simulated is the Ehrenfest urn model with $B = 9$ considered in Table II. We combined permuted estimators for sets of states $A_i = \{1, i\}$ with $w = 1$ for all values of i except $i = 1$, so, using the notation in Section 8, we have $x_0 = w = 1$, $x_1 = 0$ and $x_j = j$ for $j = 2, \dots, 8$. We first estimated the covariance matrix of the permuted estimators by using (16) with 10^4 sections, each consisting of 10^3 cycles. Our estimates of the optimal weights were $(0.0648, 0.7527, 0.2072, 0.1204, 0.0197, -0.0221, -0.0129, -0.1297)$. We then ran 1,000 independent replications of 100 cycles (independent of the simulation to estimate the weights), and computed the combined estimator in each replication using the above weights. The average of the combined estimator over all of the independent replications was 5.4928×10^4 , and the sample variance across replications was 1.355×10^8 . Thus, we see about a 37% reduction in the sample variance compared to the best single permuted estimator, which is for $A = \{1, 2\}$, from Table II.

We also experimented with fixed uniform weights for the combined permuted estimator. In one case, we used sets of states $A_i = \{1, i\}$ with $w = 1$ for all values of i except $i = 1$, and the weights were $\lambda_i = 1/8$ for each i except $i = 1$. From 10^3 independent replications the resulting sample variance of the combined estimator was 0.3306×10^8 , so the variance was reduced by a factor of 6.5 compared to the single permuted estimator with $A = \{1, 2\}$.

We also tried the combined estimator with just a subset of the states, using sets of states $A_i = \{1, i\}$ with $w = 1$ for all even values of $i \leq 8$ and fixed uniform weights $\lambda_i = 1/5$ for each even i . From 10^3 independent replications the resulting sample variance of the combined estimator was 0.5159×10^8 , a reduction of about a factor of 4.2 compared to the single permuted estimator with $A = \{1, 2\}$.

The seeming discrepancy among the results for fixed uniform weights and the estimated optimal weights appears to be due to the fact there is considerable noise in the estimates of the optimal weights. This contributes to not achieving as much improvement in the variance as with the fixed uniform weights.

A. PROOF OF THEOREM 5.1: DERIVATION OF $\hat{\alpha}_U(\vec{X}_M)$

We start with $\hat{\alpha}_U(\vec{X}_m)$ in (7). Using the representation of the function g in (6), we will compute the average of each term in $g(K, D)$ over all $(K, D) \in \Gamma$. First observe that

$$\begin{aligned} \frac{1}{|\Gamma|} \sum_{(K,D) \in \Gamma} \sum_{i=1}^{h(1,1)} Y_{K_i(1,1)}^2(1,1) &= \frac{1}{|\Lambda(1,1)|} \sum_{K(1,1) \in \Lambda(1,1)} \sum_{i=1}^{h(1,1)} Y_{K_i(1,1)}^2(1,1) \\ &= \sum_{i=1}^{h(1,1)} Y_i^2(1,1) = S_2(1,1). \end{aligned} \quad (17)$$

If $h(1,2) = 0$, then all other terms in (6) vanish, so now assume $h(1,2) > 0$. Then

$$\begin{aligned} \frac{1}{|\Gamma|} \sum_{(K,D) \in \Gamma} \sum_{i=1}^{h(1,2)} \left(Y_{K_i(1,2)}^2(1,2) + Y_{K_i(2,1)}^2(2,1) \right) \\ = \sum_{i=1}^{h(1,2)} \left(Y_i^2(1,2) + Y_i^2(2,1) \right) = S_2(1,2) + S_2(2,1). \end{aligned} \quad (18)$$

Moreover, note that

$$\begin{aligned}
& \frac{1}{|\Gamma|} \sum_{(K,D) \in \Gamma} \sum_{i=1}^{h(1,2)} 2Y_{K_i(1,2)}(1,2)Y_{K_i(2,1)}(2,1) \\
&= \frac{2}{|\Lambda(1,2) \times \Lambda(2,1)|} \sum_{K(1,2) \in \Lambda(1,2)} \sum_{K(2,1) \in \Lambda(2,1)} \sum_{i=1}^{h(1,2)} Y_{K_i(1,2)}(1,2)Y_{K_i(2,1)}(2,1) \\
&= \frac{2}{h(1,2)!h(2,1)!} \sum_{i=1}^{h(1,2)} \sum_{K(1,2) \in \Lambda(1,2)} Y_{K_i(1,2)}(1,2) \sum_{K(2,1) \in \Lambda(2,1)} Y_{K_i(2,1)}(2,1) \\
&= \frac{2}{h(1,2)!h(2,1)!} \sum_{i=1}^{h(1,2)} \sum_{K(1,2) \in \Lambda(1,2)} \sum_{j=1}^{h(1,2)} Y_j(1,2)I(K_i(1,2) = j) \\
&\quad \times \sum_{K(2,1) \in \Lambda(2,1)} \sum_{\ell=1}^{h(1,2)} Y_\ell(2,1)I(K_i(2,1) = \ell) \\
&= \frac{2}{h(1,2)!h(2,1)!} \sum_{j=1}^{h(1,2)} Y_j(1,2) \sum_{\ell=1}^{h(1,2)} Y_\ell(2,1) \\
&\quad \times \sum_{i=1}^{h(1,2)} \sum_{K(1,2) \in \Lambda(1,2)} I(K_i(1,2) = j) \sum_{K(2,1) \in \Lambda(2,1)} I(K_i(2,1) = \ell) \\
&= \frac{2}{h(1,2)!h(2,1)!} \sum_{j=1}^{h(1,2)} Y_j(1,2) \sum_{\ell=1}^{h(1,2)} Y_\ell(2,1) \sum_{i=1}^{h(1,2)} (h(1,2) - 1)!(h(2,1) - 1)! \\
&= \frac{2h(1,2)(h(1,2) - 1)!(h(2,1) - 1)!}{h(1,2)!h(2,1)!} S_1(1,2)S_1(2,1) \\
&= \frac{2}{h(1,2)} S_1(1,2)S_1(2,1) \tag{19}
\end{aligned}$$

since $h(1,2) = h(2,1)$.

Also, if $h(2,2) = 0$, then any terms in (6) that include $Y_{K_j(2,2)}(2,2)$ vanish. Now assume $h(2,2) > 0$, so

$$\begin{aligned}
& \frac{1}{|\Gamma|} \sum_{(K,D) \in \Gamma} \sum_{i=1}^{h(1,2)} 2Y_{K_i(1,2)}(1,2) \sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2,2) \\
&= \frac{2}{|\Lambda(1,2) \times \Lambda(2,2) \times \Delta|} \sum_{K(1,2) \in \Lambda(1,2)} \sum_{K(2,2) \in \Lambda(2,2)} \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} Y_{K_i(1,2)}(1,2) \\
&\quad \times \sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2,2)
\end{aligned}$$

$$\begin{aligned}
&= \frac{2}{h(1,2)! |\Lambda(2,2)| |\Delta|} \sum_{i=1}^{h(1,2)} \sum_{K(1,2) \in \Lambda(1,2)} Y_{K_i(1,2)}(1,2) \\
&\times \sum_{K(2,2) \in \Lambda(2,2)} \sum_{D \in \Delta} \sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2,2) \\
&= \frac{2}{h(1,2)! |\Lambda(2,2)| |\Delta|} \sum_{\ell=1}^{h(1,2)} Y_{\ell}(1,2) (h(1,2) - 1)! \\
&\times \sum_{K(2,2) \in \Lambda(2,2)} \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} \sum_{k=1}^{h(2,2)} Y_k(2,2) I(K_j(2,2) = k) \\
&= \frac{2}{h(1,2) |\Lambda(2,2)| |\Delta|} \sum_{\ell=1}^{h(1,2)} Y_{\ell}(1,2) \sum_{K(2,2) \in \Lambda(2,2)} \sum_{D \in \Delta} \sum_{j=1}^{h(2,2)} \sum_{k=1}^{h(2,2)} Y_k(2,2) I(K_j(2,2) = k) \\
&= \frac{2}{h(1,2) |\Lambda(2,2)| |\Delta|} \sum_{\ell=1}^{h(1,2)} Y_{\ell}(1,2) \sum_{K(2,2) \in \Lambda(2,2)} \sum_{D \in \Delta} \sum_{k=1}^{h(2,2)} Y_k(2,2) \sum_{j=1}^{h(2,2)} I(K_j(2,2) = k) \\
&= \frac{2}{h(1,2) |\Lambda(2,2)| |\Delta|} \sum_{\ell=1}^{h(1,2)} Y_{\ell}(1,2) \sum_{K(2,2) \in \Lambda(2,2)} \sum_{D \in \Delta} \sum_{k=1}^{h(2,2)} Y_k(2,2) \\
&= \frac{2}{h(1,2) |\Lambda(2,2)| |\Delta|} \sum_{\ell=1}^{h(1,2)} Y_{\ell}(1,2) |\Lambda(2,2)| |\Delta| \sum_{k=1}^{h(2,2)} Y_k(2,2) \\
&= \frac{2}{h(1,2)} \sum_{\ell=1}^{h(1,2)} Y_{\ell}(1,2) \sum_{k=1}^{h(2,2)} Y_k(2,2) = \frac{2}{h(1,2)} S_1(1,2) S_1(2,2).
\end{aligned}$$

We can similarly show that

$$\begin{aligned}
&\frac{1}{|\Gamma|} \sum_{(K,D) \in \Gamma} \sum_{i=1}^{h(1,2)} 2 \{Y_{K_i(1,2)}(1,2) + Y_{K_i(2,1)}(2,1)\} \sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2,2) \\
&= \frac{2}{h(1,2)} \{S_1(1,2) + S_1(2,1)\} S_1(2,2). \tag{20}
\end{aligned}$$

Finally, letting $C_0 = 1/(|\Lambda(2,2)| \times |\Delta|)$, we get

$$\begin{aligned}
&\frac{1}{|\Gamma|} \sum_{(K,D) \in \Gamma} \sum_{i=1}^{h(1,2)} \left(\sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2,2) \right)^2 \\
&= C_0 \sum_{K(2,2) \in \Lambda(2,2)} \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \left(\sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2,2) \right) \left(\sum_{\ell=D_i}^{D_{i+1}-1} Y_{K_{\ell}(2,2)}(2,2) \right)
\end{aligned}$$

$$\begin{aligned}
&= C_0 \sum_{K(2,2) \in \Lambda(2,2)} \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \left(\sum_{j=D_i}^{D_{i+1}-1} \sum_{p=1}^{h(2,2)} Y_p(2,2) I(K_j(2,2) = p) \right) \\
&\quad \times \left(\sum_{\ell=D_i}^{D_{i+1}-1} \sum_{q=1}^{h(2,2)} Y_q(2,2) I(K_\ell(2,2) = q) \right) \\
&= C_0 \sum_{p=1}^{h(2,2)} Y_p(2,2) \sum_{q=1}^{h(2,2)} Y_q(2,2) \tag{21}
\end{aligned}$$

$$\begin{aligned}
&\times \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} \sum_{\ell=D_i}^{D_{i+1}-1} \sum_{K(2,2) \in \Lambda(2,2)} I(K_j(2,2) = p, K_\ell(2,2) = q) \\
&= C_0 \sum_{p=1}^{h(2,2)} Y_p^2(2,2) \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} \sum_{K(2,2) \in \Lambda(2,2)} I(K_j(2,2) = p) \tag{22} \\
&\quad + C_0 \sum_{p=1}^{h(2,2)} Y_p^2(2,2) \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} \sum_{\substack{\ell=D_i \\ \ell \neq j}}^{D_{i+1}-1} \sum_{K(2,2) \in \Lambda(2,2)} I(K_j(2,2) = p, K_\ell(2,2) = p) \\
&\quad + C_0 \sum_{p=1}^{h(2,2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2,2)} Y_p(2,2) Y_q(2,2) \\
&\quad \times \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} \sum_{K(2,2) \in \Lambda(2,2)} I(K_j(2,2) = p, K_j(2,2) = q). \\
&\quad + C_0 \sum_{p=1}^{h(2,2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2,2)} Y_p(2,2) Y_q(2,2) \\
&\quad \times \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} \sum_{\substack{\ell=D_i \\ \ell \neq j}}^{D_{i+1}-1} \sum_{K(2,2) \in \Lambda(2,2)} I(K_j(2,2) = p, K_\ell(2,2) = q).
\end{aligned}$$

where (22) follows by breaking up the sums in (21) indexed by p and q into cases when $p = q$ and $p \neq q$, and also doing the same for the sums indexed by j and ℓ . In (22) the first term corresponds to $p = q$ and $j = \ell$, the second term corresponds to $p = q$ and $j \neq \ell$, the third term corresponds to $p \neq q$ and $j = \ell$, and the fourth term corresponds to $p \neq q$ and $j \neq \ell$.

We now separately analyze the four terms in (22), which we denote as a_1, a_2, a_3, a_4 . For the first term, note that

$$\begin{aligned}
a_1 &= C_0 \sum_{p=1}^{h(2,2)} Y_p^2(2,2) \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} (h(2,2) - 1)! \\
&= C_0 (h(2,2) - 1)! \sum_{p=1}^{h(2,2)} Y_p^2(2,2) \sum_{D \in \Delta} \sum_{j=1}^{h(2,2)} 1
\end{aligned}$$

$$\begin{aligned}
&= C_0 h(2, 2)! |\Delta| \sum_{p=1}^{h(2,2)} Y_p^2(2, 2) \\
&= \sum_{p=1}^{h(2,2)} Y_p^2(2, 2) = S_2(2, 2). \tag{23}
\end{aligned}$$

Note that the second term in (22) is $a_2 = 0$ since $I(K_j(2, 2) = p, K_\ell(2, 2) = p) = 0$ for $j \neq \ell$ since $K(2, 2)$ is a permutation of $1, 2, \dots, h(2, 2)$. Similarly, the third term in (22) is $a_3 = 0$ since $I(K_j(2, 2) = p, K_j(2, 2) = q) = 0$ because $p \neq q$ and $K_j(2, 2)$ cannot take on two different values.

For the fourth term in (22),

$$\begin{aligned}
a_4 &= C_0 \sum_{p=1}^{h(2,2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2,2)} Y_p(2, 2) Y_q(2, 2) \sum_{i=1}^{h(1,2)} \sum_{D \in \Delta} \sum_{\substack{j=D_i \\ \ell \neq j}}^{D_{i+1}-1} \sum_{\substack{\ell=D_i \\ \ell \neq j}}^{D_{i+1}-1} (h(2, 2) - 2)! \\
&= C_0 (h(2, 2) - 2)! \sum_{p=1}^{h(2,2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2,2)} Y_p(2, 2) Y_q(2, 2) \sum_{i=1}^{h(1,2)} \sum_{D \in \Delta} 2 \binom{D_{i+1} - D_i}{2}
\end{aligned}$$

since the number of terms summed in $\sum_{j=D_i}^{D_{i+1}-1} \sum_{\substack{\ell=D_i \\ \ell \neq j}}^{D_{i+1}-1}$ is the number of ways to choose two distinct numbers j and l from among $D_i, D_i + 1, \dots, D_{i+1} - 1$. To calculate $\sum_{D \in \Delta} 2 \binom{D_{i+1} - D_i}{2}$, first recall each $D = (D_1, D_2, \dots, D_{h(1,2)}) \in \Delta$ satisfies $D_1 = 1$, $D_i \leq D_{i+1}$, and $D_{h(1,2)} \leq h(2, 2) + 1$, so to calculate $\sum_{D \in \Delta} 2 \binom{D_{i+1} - D_i}{2}$, we can think of arranging $h(2, 2) + h(1, 2)$ balls into $h(2, 2) + h(1, 2)$ boxes in a straight line, with each box receiving exactly one ball, where $h(1, 2)$ of the balls are black and identical and the rest are white and identical. We put $D_{i+1} - D_i$ white balls between the i th and $(i + 1)$ st black balls. Specifically, the first box contains a black ball since $D_1 = 1$. The next $D_2 - D_1$ boxes each contain a white ball, and a black ball is placed in box $D_2 + 1$. Then the next $D_3 - D_2$ boxes each contain a white ball, and a black ball is placed in box $D_3 + 2$. In general, after the j th black ball, the next $D_{j+1} - D_j$ boxes each contain a white ball, and a black ball is placed in box $D_{j+1} + j$. Now suppose that $D_{i+1} - D_i = k$ for some $k \geq 2$, and we now want to determine the number of $D \in \Delta$ having $D_{i+1} - D_i = k$. To calculate this, we first remove k white balls and k boxes, which corresponds to the fact that if $D_{i+1} - D_i = k$, we do not need arrange those k white balls. Then we remove one black ball and one box, which corresponds to the marker denoting the beginning of the group of k white balls. Finally, remove one more black ball and one box, which corresponds to the fact that $D_1 = 1$. The remaining $h(2, 2) - k + h(1, 2) - 2$ balls can then be arranged in any way into the $h(2, 2) - k + h(1, 2) - 2$ boxes, and the number of possible arrangements is $(h(2, 2) - k + h(1, 2) - 2)! / [(h(2, 2) - k)! (h(1, 2) - 2)!]$. Now note that $C_0 = 1 / (|\Lambda(2, 2)| \cdot |\Delta|) = 1 / (h(2, 2)! |\Delta|)$. Thus,

$$\begin{aligned}
a_4 &= \frac{1}{h(2, 2)(h(2, 2) - 1) |\Delta|} \sum_{p=1}^{h(2,2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2,2)} Y_p(2, 2) Y_q(2, 2) \\
&\quad \times \sum_{i=1}^{h(1,2)} \sum_{k=2}^{h(2,2)} 2 \binom{k}{2} \frac{(h(2, 2) - k + h(1, 2) - 2)!}{(h(2, 2) - k)! (h(1, 2) - 2)!}
\end{aligned}$$

$$\begin{aligned}
&= \frac{h(1, 2)}{h(2, 2)(h(2, 2) - 1)|\Delta|} \sum_{p=1}^{h(2, 2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2, 2)} Y_p(2, 2)Y_q(2, 2) \\
&\times \sum_{k=2}^{h(2, 2)} 2 \binom{k}{2} \frac{(h(2, 2) - k + h(1, 2) - 2)!}{(h(2, 2) - k)! (h(1, 2) - 2)!} \\
&= \frac{h(1, 2)B}{h(2, 2)(h(2, 2) - 1)|\Delta|} \sum_{p=1}^{h(2, 2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2, 2)} Y_p(2, 2)Y_q(2, 2),
\end{aligned}$$

where for any i ,

$$\begin{aligned}
B &= \sum_{D \in \Delta} 2 \binom{D_{i+1} - D_i}{2} = \sum_{k=2}^{h(2, 2)} 2 \binom{k}{2} \frac{(h(2, 2) - k + h(1, 2) - 2)!}{(h(2, 2) - k)! (h(1, 2) - 2)!} \\
&= 2 \frac{(h(2, 2) + h(1, 2) - 1)!}{(h(2, 2) - 2)! (h(1, 2) + 1)!}.
\end{aligned} \tag{24}$$

To establish the last equality, note that if $h(2, 2) < 2$, then both sides of the equality are zero. If $h(1, 2) = 1$, then both sides are equal to $h(2, 2)[h(2, 2) - 1]$. Making the substitutions $m \equiv h(2, 2) - 2$ and $n \equiv h(1, 2) - 2$, both assumed nonnegative, and removing the common factor of 2, we need to show that

$$\sum_{k=0}^m \binom{k+2}{2} \binom{m-k+n}{m-k} = \binom{m+n+3}{m} \tag{25}$$

for all $m, n \geq 0$. This follows from

$$\begin{aligned}
\sum_{k=0}^m \binom{k+2}{2} \binom{m-k+n}{m-k} &= \sum_{k=0}^m (-1)^k \binom{-3}{k} (-1)^{m-k} \binom{-n-1}{m-k} \\
&= (-1)^m \binom{-n-4}{m} = \binom{m+n+3}{m},
\end{aligned}$$

where we used ‘‘upper negation’’ for the first and last equalities (see Graham et al. [1994], p. 164), and Vandermonde’s convolution for the second equality (see Graham et al. [1994], p. 174).

Note that

$$\begin{aligned}
\sum_{p=1}^{h(2, 2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2, 2)} Y_p(2, 2)Y_q(2, 2) &= \sum_{p=1}^{h(2, 2)} Y_p(2, 2) \sum_{q=1}^{h(2, 2)} Y_q(2, 2) - \sum_{p=1}^{h(2, 2)} Y_p^2(2, 2) \\
&= S_1^2(2, 2) - S_2(2, 2).
\end{aligned} \tag{26}$$

Also,

$$|\Delta| = \frac{(h(2, 2) + h(1, 2) - 1)!}{h(2, 2)! (h(1, 2) - 1)!}, \tag{27}$$

so

$$a_4 = \frac{2h(1, 2) (h(2, 2) + h(1, 2) - 1)! h(2, 2)! (h(1, 2) - 1)! [S_1^2(2, 2) - S_2(2, 2)]}{h(2, 2)(h(2, 2) - 1) (h(2, 2) - 2)! (h(1, 2) + 1)! (h(2, 2) + h(1, 2) - 1)!}$$

$$= \frac{2}{h(1,2) + 1} [S_1^2(2,2) - S_2(2,2)]. \quad (28)$$

Therefore, from (17), (18), (19), (20), (23), and (28), we get

$$\begin{aligned} \hat{\alpha}_U(\vec{X}_m) &= \frac{1}{m} \left[S_2(1,1) + S_2(1,2) + S_2(2,1) + \frac{2}{h(1,2)} S_1(1,2)S_1(2,1) \right. \\ &\quad + \frac{2}{h(1,2)} \{S_1(1,2) + S_1(2,1)\} S_1(2,2) + S_2(2,2) \\ &\quad \left. + \frac{2}{h(1,2) + 1} [S_1^2(2,2) - S_2(2,2)] \right], \quad (29) \end{aligned}$$

which, when comparing with (5), shows that $\hat{\alpha}_U(\vec{X}_m) = \hat{\alpha}_P(\vec{X}_m)$.

B. PROOF OF THEOREM 5.1: DERIVATION OF $\hat{\alpha}_V(\vec{X}_M)$

We now show that $\hat{\alpha}_V(\vec{X}_m)$ defined in (8) can be expressed as in (9). The derivation is similar to that for $\hat{\alpha}_U(\vec{X}_m)$, but there are a few key differences since averaging is now done with replacement, and we include the derivation of $\hat{\alpha}_V(\vec{X}_m)$ for completeness.

Using the representation of the function g in (6), we will compute the average of each term in $g(K, D)$ over all $(K, D) \in \hat{\Gamma}$. First observe that

$$\begin{aligned} &\frac{1}{|\hat{\Gamma}|} \sum_{(K,D) \in \hat{\Gamma}} \sum_{i=1}^{h(1,1)} Y_{K_i(1,1)}^2(1,1) \\ &= \frac{1}{|\hat{\Lambda}(1,1)|} \sum_{i=1}^{h(1,1)} \sum_{K(1,1) \in \hat{\Lambda}(1,1)} Y_{K_i(1,1)}^2(1,1) \\ &= \frac{1}{[h(1,1)]^{h(1,1)}} \sum_{i=1}^{h(1,1)} \sum_{K(1,1) \in \hat{\Lambda}(1,1)} \sum_{j=1}^{h(1,1)} Y_j^2(1,1) I(K_i(1,1) = j) \\ &= \frac{1}{[h(1,1)]^{h(1,1)}} \sum_{j=1}^{h(1,1)} Y_j^2(1,1) \sum_{i=1}^{h(1,1)} \sum_{K(1,1) \in \hat{\Lambda}(1,1)} I(K_i(1,1) = j) \\ &= \frac{1}{[h(1,1)]^{h(1,1)}} \sum_{j=1}^{h(1,1)} Y_j^2(1,1) \sum_{i=1}^{h(1,1)} [h(1,1)]^{h(1,1)-1} \\ &= \sum_{j=1}^{h(1,1)} Y_j^2(1,1) = S_2(1,1). \quad (30) \end{aligned}$$

If $h(1, 2) = 0$, then all of the other terms in (6) vanish, so now assume $h(1, 2) > 0$. Then applying arguments as above yields

$$\begin{aligned} & \frac{1}{|\widehat{\Gamma}|} \sum_{(K,D) \in \widehat{\Gamma}} \sum_{i=1}^{h(1,2)} \left(Y_{K_i(1,2)}^2(1, 2) + Y_{K_i(2,1)}^2(2, 1) \right) \\ &= \sum_{i=1}^{h(1,2)} (Y_i^2(1, 2) + Y_i^2(2, 1)) = S_2(1, 2) + S_2(2, 1). \end{aligned} \quad (31)$$

Also, note that

$$\begin{aligned} & \frac{1}{|\widehat{\Gamma}|} \sum_{(K,D) \in \widehat{\Gamma}} \sum_{i=1}^{h(1,2)} 2Y_{K_i(1,2)}(1, 2)Y_{K_i(2,1)}(2, 1) \\ &= \frac{2}{|\widehat{\Lambda}(1, 2) \times \widehat{\Lambda}(2, 1)|} \sum_{K(1,2) \in \widehat{\Lambda}(1,2)} \sum_{K(2,1) \in \widehat{\Lambda}(2,1)} \sum_{i=1}^{h(1,2)} Y_{K_i(1,2)}(1, 2)Y_{K_i(2,1)}(2, 1) \\ &= \frac{2}{|\widehat{\Lambda}(1, 2)| \cdot |\widehat{\Lambda}(2, 1)|} \sum_{i=1}^{h(1,2)} \sum_{K(1,2) \in \widehat{\Lambda}(1,2)} Y_{K_i(1,2)}(1, 2) \sum_{K(2,1) \in \widehat{\Lambda}(2,1)} Y_{K_i(2,1)}(2, 1) \\ &= \frac{2}{[h(1, 2)]^{h(1,2)} [h(2, 1)]^{h(2,1)}} \sum_{i=1}^{h(1,2)} \left[[h(1, 2)]^{h(1,2)-1} \sum_{j=1}^{h(1,2)} Y_j(1, 2) \right] \\ & \quad \times \left[[h(2, 1)]^{h(2,1)-1} \sum_{\ell=1}^{h(2,1)} Y_\ell(2, 1) \right] \\ &= \frac{2}{h(1, 2)} \sum_{j=1}^{h(1,2)} Y_j(1, 2) \sum_{\ell=1}^{h(2,1)} Y_\ell(2, 1) = \frac{2}{h(1, 2)} S_1(1, 2) S_1(2, 1) \end{aligned} \quad (32)$$

since $h(1, 2) = h(2, 1)$.

In addition, if $h(2, 2) = 0$, then any terms in (6) that include $Y_{K_j(2,2)}(2, 2)$ vanish. Now assume $h(2, 2) > 0$, so

$$\begin{aligned} & \frac{1}{|\widehat{\Gamma}|} \sum_{(K,D) \in \widehat{\Gamma}} \sum_{i=1}^{h(1,2)} 2Y_{K_i(1,2)}(1, 2) \sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2, 2) \\ &= \frac{2}{|\widehat{\Lambda}(1, 2) \times \widehat{\Lambda}(2, 2) \times \Delta|} \sum_{K(1,2) \in \widehat{\Lambda}(1,2)} \sum_{K(2,2) \in \widehat{\Lambda}(2,2)} \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} Y_{K_i(1,2)}(1, 2) \\ & \quad \times \sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2, 2) \end{aligned}$$

$$\begin{aligned}
&= \frac{2}{[h(1,2)]^{h(1,2)} [h(2,2)]^{h(2,2)} |\Delta|} \sum_{i=1}^{h(1,2)} \sum_{K(1,2) \in \widehat{\Lambda}(1,2)} Y_{K_i(1,2)}(1,2) \\
&\times \sum_{K(2,2) \in \widehat{\Lambda}(2,2)} \sum_{D \in \Delta} \sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2,2) \\
&= \frac{2 [h(1,2)]^{h(1,2)-1}}{[h(1,2)]^{h(1,2)} [h(2,2)]^{h(2,2)} |\Delta|} \sum_{\ell=1}^{h(1,2)} Y_{\ell}(1,2) \\
&\times \sum_{K(2,2) \in \widehat{\Lambda}(2,2)} \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} \sum_{k=1}^{h(2,2)} Y_k(2,2) I(K_j(2,2) = k) \\
&= \frac{2}{h(1,2) [h(2,2)]^{h(2,2)} |\Delta|} \sum_{\ell=1}^{h(1,2)} Y_{\ell}(1,2) \\
&\times \sum_{K(2,2) \in \widehat{\Lambda}(2,2)} \sum_{D \in \Delta} \sum_{j=1}^{h(2,2)} \sum_{k=1}^{h(2,2)} Y_k(2,2) I(K_j(2,2) = k) \\
&= \frac{2}{h(1,2) [h(2,2)]^{h(2,2)} |\Delta|} \sum_{\ell=1}^{h(1,2)} Y_{\ell}(1,2) \sum_{k=1}^{h(2,2)} Y_k(2,2) \\
&\times \sum_{D \in \Delta} \sum_{j=1}^{h(2,2)} \sum_{K(2,2) \in \widehat{\Lambda}(2,2)} I(K_j(2,2) = k) \\
&= \frac{2}{h(1,2) [h(2,2)]^{h(2,2)} |\Delta|} \sum_{\ell=1}^{h(1,2)} Y_{\ell}(1,2) \sum_{k=1}^{h(2,2)} Y_k(2,2) \sum_{D \in \Delta} \sum_{j=1}^{h(2,2)} [h(2,2)]^{h(2,2)-1} \\
&= \frac{2}{h(1,2)} \sum_{\ell=1}^{h(1,2)} Y_{\ell}(1,2) \sum_{k=1}^{h(2,2)} Y_k(2,2) = \frac{2}{h(1,2)} S_1(1,2) S_1(2,2).
\end{aligned}$$

Moreover, we can similarly show that

$$\begin{aligned}
&\frac{1}{|\widehat{\Gamma}|} \sum_{(K,D) \in \widehat{\Gamma}} \sum_{i=1}^{h(1,2)} 2 \{Y_{K_i(1,2)}(1,2) + Y_{K_i(2,1)}(2,1)\} \sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2,2) \\
&= \frac{2}{h(1,2)} \{S_1(1,2) + S_1(2,1)\} S_1(2,2). \tag{33}
\end{aligned}$$

Finally, letting $\widehat{C} = 1/(|\widehat{\Lambda}(2, 2) \times \Delta|)$ and using arguments similar to those applied to establish (22), we get

$$\begin{aligned}
& \frac{1}{|\widehat{\Gamma}|} \sum_{(K,D) \in \widehat{\Gamma}} \sum_{i=1}^{h(1,2)} \left(\sum_{j=D_i}^{D_{i+1}-1} Y_{K_j(2,2)}(2, 2) \right)^2 \\
&= \widehat{C} \sum_{p=1}^{h(2,2)} Y_p^2(2, 2) \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} \sum_{K(2,2) \in \widehat{\Lambda}(2,2)} I(K_j(2, 2) = p) \\
&+ \widehat{C} \sum_{p=1}^{h(2,2)} Y_p^2(2, 2) \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} \sum_{\substack{\ell=D_i \\ \ell \neq j}}^{D_{i+1}-1} \sum_{K(2,2) \in \widehat{\Lambda}(2,2)} I(K_j(2, 2) = p, K_\ell(2, 2) = p) \\
&+ \widehat{C} \sum_{p=1}^{h(2,2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2,2)} Y_p(2, 2) Y_q(2, 2) \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} \sum_{K(2,2) \in \widehat{\Lambda}(2,2)} I(K_j(2, 2) = p, K_j(2, 2) = q) \\
&+ \widehat{C} \sum_{p=1}^{h(2,2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2,2)} Y_p(2, 2) Y_q(2, 2) \\
&\times \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} \sum_{\substack{\ell=D_i \\ \ell \neq j}}^{D_{i+1}-1} \sum_{K(2,2) \in \widehat{\Lambda}(2,2)} I(K_j(2, 2) = p, K_\ell(2, 2) = q).
\end{aligned} \tag{34}$$

We now separately analyze the four terms, in (34), which we denote as b_1, b_2, b_3, b_4 . For the first term, note that

$$\begin{aligned}
b_1 &= \widehat{C} \sum_{p=1}^{h(2,2)} Y_p^2(2, 2) \sum_{D \in \Delta} \sum_{i=1}^{h(1,2)} \sum_{j=D_i}^{D_{i+1}-1} [h(2, 2)]^{h(2,2)-1} \\
&= \widehat{C} [h(2, 2)]^{h(2,2)-1} \sum_{p=1}^{h(2,2)} Y_p^2(2, 2) \sum_{D \in \Delta} \sum_{j=1}^{h(2,2)} 1 \\
&= \widehat{C} [h(2, 2)]^{h(2,2)} |\Delta| \sum_{p=1}^{h(2,2)} Y_p^2(2, 2) \\
&= \sum_{p=1}^{h(2,2)} Y_p^2(2, 2) = S_2(2, 2).
\end{aligned} \tag{35}$$

The third term in (34) is $b_3 = 0$ since $I(K_j(2, 2) = p, K_j(2, 2) = q) = 0$ because $p \neq q$ and $K_j(2, 2)$ cannot take on 2 different values. However, contrary to what we saw in the expression (22) for the U -statistic estimator in the previous section, the second term b_2 in (34) is no longer zero since $I(K_j(2, 2) = p, K_\ell(2, 2) = p) \neq 0$ for $j \neq \ell$ since $K(2, 2)$ is an $h(2, 2)$ -dimensional vector in which the components are selected from $\{1, 2, \dots, h(2, 2)\}$ with replacement. Thus, for the second term in (34), we have

$$b_2 = \widehat{C} \sum_{p=1}^{h(2,2)} Y_p^2(2, 2) \sum_{i=1}^{h(1,2)} \sum_{D \in \Delta} \sum_{j=D_i}^{D_{i+1}-1} \sum_{\substack{\ell=D_i \\ \ell \neq j}}^{D_{i+1}-1} [h(2, 2)]^{h(2,2)-2}$$

$$\begin{aligned}
&= \widehat{C} [h(2,2)]^{h(2,2)-2} \sum_{p=1}^{h(2,2)} Y_p^2(2,2) \sum_{i=1}^{h(1,2)} \sum_{D \in \Delta} 2^{\binom{D_{i+1}-D_i}{2}} \\
&= B \widehat{C} [h(2,2)]^{h(2,2)-2} h(1,2) \sum_{p=1}^{h(2,2)} Y_p^2(2,2) \\
&= \frac{B h(1,2)}{[h(2,2)]^2 |\Delta|} S_2(2,2)
\end{aligned} \tag{36}$$

since $\widehat{C} = 1/([h(2,2)]^{h(2,2)} |\Delta|)$, where B is defined in (24).

For the fourth term in (34),

$$\begin{aligned}
b_4 &= \widehat{C} \sum_{p=1}^{h(2,2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2,2)} Y_p(2,2) Y_q(2,2) \sum_{i=1}^{h(1,2)} \sum_{D \in \Delta} \sum_{j=D_i}^{D_{i+1}-1} \sum_{\substack{\ell=D_i \\ \ell \neq j}}^{D_{i+1}-1} [h(2,2)]^{h(2,2)-2} \\
&= \widehat{C} [h(2,2)]^{h(2,2)-2} \sum_{p=1}^{h(2,2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2,2)} Y_p(2,2) Y_q(2,2) \sum_{i=1}^{h(1,2)} \sum_{D \in \Delta} 2^{\binom{D_{i+1}-D_i}{2}} \\
&= B \widehat{C} [h(2,2)]^{h(2,2)-2} h(1,2) \sum_{p=1}^{h(2,2)} \sum_{\substack{q=1 \\ q \neq p}}^{h(2,2)} Y_p(2,2) Y_q(2,2) \\
&= \frac{B h(1,2)}{[h(2,2)]^2 |\Delta|} [S_1^2(2,2) - S_2(2,2)]
\end{aligned} \tag{37}$$

by (26).

Therefore, from (24), (27), (30), (31), (32), (33), (35), (36), and (37), we get

$$\begin{aligned}
\widehat{\alpha}_V(\vec{X}_m) &= \frac{1}{m} \left[S_2(1,1) + S_2(1,2) + S_2(2,1) + \frac{2}{h(1,2)} S_1(1,2) S_1(2,1) \right. \\
&\quad + \frac{2}{h(1,2)} \{S_1(1,2) + S_1(2,1)\} S_1(2,2) + S_2(2,2) \\
&\quad \left. + \frac{2(h(2,2)-1)}{h(2,2)(h(1,2)+1)} S_1^2(2,2) \right] \\
&= Q + \frac{2(h(2,2)-1)}{m h(2,2)(h(1,2)+1)} S_1^2(2,2),
\end{aligned} \tag{38}$$

where Q is defined in (4).

ACKNOWLEDGMENTS

It is a privilege to contribute to the special issue in honor of Don Iglehart. He has made enormous contributions towards making simulation a rigorous and respected scientific method for analyzing models of complex stochastic systems. Don was a pioneer in the development of the regenerative method, and our work here continues the study of this approach, which remains as a vibrant area of research. During a graduate-student meeting with one of the authors, Prof. Iglehart posed the question of how one might exploit multiple regeneration sequences for simulation output analysis.

We thank the reviewers for providing comments that led to a better paper. One of the editors pointed out how to simplify the proof of Equation (25).

REFERENCES

- C. Antonini, C. Alexopoulos, D. Goldsman, and J. R. Wilson. 2008. Area variance estimators for simulation using folded standardized time series. *IIE Transactions* 41 (2008), 134–144.
- J. M. Calvin, P. W. Glynn, and M. K. Nakayama. 2001. Importance sampling using the semi-regenerative method. In *Proceedings of the 2001 Winter Simulation Conference*, B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer (Eds.). Institute of Electrical and Electronics Engineers, Piscataway, New Jersey, 441–450.
- J. M. Calvin, P. W. Glynn, and M. K. Nakayama. 2006. The semi-regenerative method of simulation output analysis. *ACM Transactions on Modeling and Computer Simulation* 16 (2006), 280–315.
- J. M. Calvin and M. K. Nakayama. 1998. Using permutations in regenerative simulations to reduce variance. *ACM Transactions on Modeling and Computer Simulation* 8 (1998), 153–193.
- J. M. Calvin and M. K. Nakayama. 2000a. Central limit theorems for permuted regenerative estimators. *Operations Research* 48 (2000), 776–787.
- J. M. Calvin and M. K. Nakayama. 2000b. Simulation of processes with multiple regeneration sequences. *Probability in the Engineering and Informational Sciences* 14 (2000), 179–201.
- J. M. Calvin and M. K. Nakayama. 2002. A comparison of output-analysis methods for simulations of processes with multiple regeneration sequences. In *Proceedings of the 2002 Winter Simulation Conference*, E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes (Eds.). Institute of Electrical and Electronics Engineers, Piscataway, New Jersey, 328–335.
- J. M. Calvin and M. K. Nakayama. 2004. Permuted derivative and importance-sampling estimators for regenerative simulations. *European Journal of Operational Research* 156 (2004), 390–414.
- J. M. Calvin and M. K. Nakayama. 2006. Permuted standardized time series for steady-state simulations. *Mathematics of Operations Research* 31 (2006), 351–368.
- M. Crane and D. L. Iglehart. 1975. Simulating stable stochastic systems, III: Regenerative processes and discrete-event simulations. *Operations Research* 23 (1975), 33–45.
- B. Efron and R. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman & Hall.
- R. L. Graham, D. E. Knuth, and O. Patashnik. 1994. *Concrete Mathematics: A Foundation for Computer Science* (second ed.). Addison-Wesley, Upper Saddle River, New Jersey.
- F. L. Gunther and R. W. Wolff. 1980. The almost regenerative method. *Operations Research* 28 (1980), 375–387.
- P. Heidelberger. 1980. Variance reduction techniques for the simulation of Markov processes, I: Multiple estimates. *IBM Journal of Research and Development* 24 (1980), 570–581.
- S. G. Henderson and P. W. Glynn. 2001. Regenerative steady-state simulation of discrete-event systems. *ACM Transactions on Modeling and Computer Simulation* 11 (2001), 313–345.
- S. B. Moka and S. Juneja. 2014. Regenerative Simulation for Queueing Networks with Exponential or Heavier Tail Arrival Distributions. *ACM Transactions On Modeling and Computer Simulation* (2014), conditionally accepted.
- R. J. Serfling. 1980. *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, New York.
- B. Zhang and Y.-C. Ho. 1992. Improvements in the likelihood ratio method for steady-state sensitivity analysis and simulation. *Performance Analysis* 15 (1992), 177–195.

Online Appendix to: Resampled Regenerative Estimators

JAMES M. CALVIN, New Jersey Institute of Technology
MARVIN K. NAKAYAMA, New Jersey Institute of Technology

A. DERIVATION OF SEMI-REGENERATIVE ESTIMATOR IN (3)

Let $T = T_1$, and for $x \in A$, note that

$$\begin{aligned} \alpha(x) &= E_x \left[\left(\sum_{k=0}^{T-1} f(X_k) + \sum_{y \in A} I(\tau > T, X_T = y) \sum_{k=T}^{\tau-1} f(X_k) \right)^2 \right] \\ &= E_x \left[\left(\sum_{k=0}^{T-1} f(X_k) \right)^2 \right] + 2 \sum_{y \in A} E_x \left[I(\tau > T, X_T = y) \sum_{k=T}^{\tau-1} f(X_k) \sum_{j=0}^{T-1} f(X_j) \right] \\ &\quad + \sum_{y \in A} E_x \left[I(\tau > T, X_T = y) \left(\sum_{k=T}^{\tau-1} f(X_k) \right)^2 \right], \end{aligned} \quad (39)$$

where the last term does not include another sum over $z \in A$ since $I(\tau > T, X_T = y)I(\tau > T, X_T = z) = 0$ when $y \neq z$. We now derive new expressions for the second and third terms in (39).

Each summand in the second term in (39) satisfies

$$E_x \left[I(\tau > T, X_T = y) \sum_{k=T}^{\tau-1} f(X_k) \sum_{j=0}^{T-1} f(X_j) \right] = F(x, y) e(y),$$

with

$$F(x, y) = E_x \left[I(\tau > T, X_T = y) \sum_{k=0}^{T-1} f(X_k) \right] \quad \text{and} \quad e(y) = E_y \left[\sum_{k=0}^{\tau-1} f(X_k) \right].$$

Also, note that

$$e(x) = E_x \left[\sum_{k=0}^{T-1} f(X_k) \right] + \sum_{y \in A} E_x [I(\tau > T, X_T = y)] e(y),$$

and defining

$$q(x) = E_x \left[\sum_{k=0}^{T-1} f(X_k) \right], \quad G(x, y) = E_x [I(\tau > T, X_T = y)],$$

$e = (e(x) : x \in A)$, $q = (q(x) : x \in A)$, and $G = (G(x, y) : x, y \in A)$, we get $e = q + Ge$, or $(I - G)e = q$. It is straightforward to show that X being irreducible and positive recurrent implies that $I - G$ is nonsingular, so $e = (I - G)^{-1}q$. Thus, each summand in

the second term in (39) satisfies

$$E_x \left[I(\tau > T, X_T = y) \sum_{k=T}^{\tau-1} f(X_k) \sum_{j=0}^{T-1} X_j \right] = F(x, y) ((I - G)^{-1}q)(y). \quad (40)$$

Now for each summand in the third term in (39), note that

$$E_x \left[I(\tau > T, X_T = y) \left(\sum_{k=T}^{\tau-1} f(X_k) \right)^2 \right] = G(x, y) \alpha(y). \quad (41)$$

Define

$$c(x) = E_x \left[\left(\sum_{k=0}^{T-1} f(X_k) \right)^2 \right],$$

and let $\alpha = (\alpha(x) : x \in A)$, $c = (c(x) : x \in A)$, and $F = (F(x, y) : x, y \in A)$. Then, putting together (39)–(41), we get $\alpha = c + 2F(I - G)^{-1}q + G\alpha$, so

$$\alpha = (I - G)^{-1}(c + 2F(I - G)^{-1}q)$$

since $I - G$ is nonsingular. Note that G , c , F and q are all conditional expectations, and we obtain the semi-regenerative estimator of $\alpha(1)$ by replacing each conditional expectation with a sample average obtained from a simulated sample path.

Now since $A = \{1, 2\}$ and $w = 1$, note that $G(x, 1) = F(x, 1) = 0$ for $x \in A$. Let $J = (J(x, y) : x, y \in A)$ with $J = (I - G)^{-1}$, and observe that

$$J = \begin{pmatrix} 1 & -G(1, 2) \\ 0 & 1 - G(2, 2) \end{pmatrix}^{-1} = \begin{pmatrix} 1 & G(1, 2)/(1 - G(2, 2)) \\ 0 & 1/(1 - G(2, 2)) \end{pmatrix}. \quad (42)$$

It then follows that

$$\begin{aligned} \alpha &= \begin{pmatrix} 1 & J(1, 2) \\ 0 & J(2, 2) \end{pmatrix} \left[\begin{pmatrix} c(1) \\ c(2) \end{pmatrix} + 2 \begin{pmatrix} 0 & F(1, 2) \\ 0 & F(2, 2) \end{pmatrix} \begin{pmatrix} 1 & J(1, 2) \\ 0 & J(2, 2) \end{pmatrix} \begin{pmatrix} q(1) \\ q(2) \end{pmatrix} \right] \\ &= \begin{pmatrix} c(1) + J(1, 2)c(2) \\ J(2, 2)c(2) \end{pmatrix} + 2 \begin{pmatrix} (F(1, 2) + J(1, 2)F(2, 2))J(2, 2)q(2) \\ J^2(2, 2)F(2, 2)q(2) \end{pmatrix}, \end{aligned}$$

so

$$\alpha(1) = c(1) + J(1, 2)c(2) + 2[F(1, 2) + J(1, 2)F(2, 2)]J(2, 2)q(2). \quad (43)$$

This was previously shown in Calvin and Nakayama [2002].

Now for $x \in A$, we use $(S_2(x, 1) + S_2(x, 2))/H(x)$ and $(S_1(x, 1) + S_1(x, 2))/H(x)$ as our estimators of $c(x)$ and $q(x)$, respectively. Also, we use $h(x, 2)/H(x)$ and $S_1(x, 2)/H(x)$ as estimators for $G(x, 2)$ and $F(x, 2)$, respectively. By (42) our estimator of $J(1, 2)$ is then

$$\hat{J}(1, 2) = \frac{h(1, 2)}{H(1)} \left(1 - \frac{h(2, 2)}{H(2)} \right)^{-1} = \frac{H(2)}{H(1)}$$

since $h(2, 1) + h(2, 2) = H(2)$ and $h(1, 2) = h(2, 1)$. Similarly, our estimator of $J(2, 2)$ is $\hat{J}(2, 2) = H(2)/h(1, 2)$ by (42). Thus, substituting our estimators into (43) yields (3) as the semi-regenerative estimator of $\alpha(1)$ based on the sample path \vec{X}_m with $A = \{1, 2\}$.