FOLDED STANDARDIZED TIME SERIES AREA VARIANCE ESTIMATORS FOR SIMULATION

Claudia Antonini

Departamento de Matemáticas Puras y Aplicadas
Universidad Simón Bolivar
Sartenejas, 1080, VENEZUELA

Christos Alexopoulos
David Goldsman

H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332, U.S.A.

James R. Wilson

Edward P. Fitts Department of
Industrial and Systems Engineering
Campus Box 7906
North Carolina State University
Raleigh, NC 27695-7906, U.S.A.

ABSTRACT

We estimate the variance parameter of a stationary simulation-generated process using "folded" versions of standardized time series area estimators. We formulate improved variance estimators based on the combination of multiple folding levels as well as the use of batching. The improved estimators preserve the asymptotic bias properties of their predecessors but have substantially lower variance. A Monte Carlo example demonstrates the efficacy of the new methodology.

1 INTRODUCTION

One of the most important problems in simulation output analysis is the estimation of the mean μ of a steady-state (stationary) simulation process $\{Y_i : i = 1, 2, \ldots\}$. For instance, we may be interested in determining the steady-state mean transit time in a job shop. Assuming that the simulation is indeed operating in steady state, one usually uses the sample mean of the observations, $\bar{Y}_n \equiv \sum_{i=1}^{n} Y_i/n$, as the point estimator for μ.

However, point estimation of the mean usually does not suffice, since any serious statistical analysis should also include a measure of the variability of the sample mean. One such measure is the *variance parameter*, which is defined as the sum of covariances of the process at all lags, and which can often be expressed as $\sigma^2 \equiv \lim_{n\to\infty} n\mathsf{Var}(\bar{Y}_n)$. With knowledge of such a measure in hand, we could provide, among other benefits, confidence intervals for μ, typically

of the form

$$\mu \in \bar{Y}_n \pm t\sqrt{\hat{\sigma}^2/n},$$

where t is a quantile from an appropriate distribution and $\hat{\sigma}^2$ is an estimator for σ^2.

Unfortunately, the problem of estimating the variance parameter is difficult since discrete-event simulation data are almost always serially correlated as well as non-normal, e.g., consecutive waiting times in a queueing system. Thus, traditional statistical analysis methods may be inappropriate since they may rely on the assumption of independent and identically distributed (i.i.d.) normal observations. This article gives new estimators—with pleasing properties—for the variance parameter σ^2 of a stationary simulation process.

Over the years, a number of methodologies for estimating σ^2 have been proposed in the literature (see Law 2006), e.g., the techniques of nonoverlapping batch means (NBM), overlapping batch means (OBM), standardized time series (STS), spectral analysis, and regeneration. NBM divides the data $\{Y_1, \ldots, Y_n\}$ into nonoverlapping batches, and uses the sample variance of the sample (batch) means from the batches as a foundation to estimate σ^2. OBM (Meketon and Schmeiser 1984) re-uses the data by forming *overlapping* batches, and then invokes an appropriately scaled sample variance of the resulting overlapping batch means to estimate σ^2. The OBM variance estimator has about the same bias as, but significantly lower variance than, the corresponding NBM estimator employing the same batch and total sample sizes. STS (Schruben 1983) uses a functional central limit theorem to standardize output from a

steady-state discrete-event simulation into a process that converges to a limiting Brownian bridge process as the sample size becomes large. Known properties of the Brownian bridge process are then used to obtain estimators for σ^2. Overlapping batched versions of various STS estimators have been shown to have the same bias as, but substantially lower variance than, their nonoverlapping counterparts; see Alexopoulos et al. (2007b, 2007c). Additional variance-reducing techinques in which STS reuses data can be found in Foley and Goldsman (1999), Goldsman et al. (2007), and Aktaran-Kalaycı, Goldsman, and Wilson (2007).

The current article studies the consequences of a "folding" operation on the original STS process (and its corresponding limiting Brownian bridge process). The folding operation produces multiple standardized time series processes, which in turn will ultimately allow us to re-use the original data to produce multiple estimators for σ^2—estimators that are often asymptotically independent as the sample size grows. These folded estimators will lead to combined estimators having smaller variance than existing estimators not based on the folding operation.

The article is organized as follows. Section 2 gives background material. Section 3 introduces the notion of folding a Brownian bridge, and shows that each application of folding yields a new Brownian bridge process. Section 4 gives results on the expected values, variances, and covariances of certain functionals related to the area under a folded Brownian bridge. In Section 5, we apply these results to the problem of estimating the variance parameter of a steady-state simulation process. In particular, we apply the folding operation to the original standardized time series of the simulation output process to produce new estimators that converge to the appropriate functionals of the corresponding folded Brownian bridges. We illustrate the efficacy of the folded estimators via analytical and Monte Carlo examples, and we find that the new estimators indeed reduce estimator variance at little cost in bias. Section 6 presents conclusions. For more details and proofs of the results in this paper, see Antonini (2005) and Antonini et al. (2007).

2 BACKGROUND

This section contains preliminaries on the STS methodology. We begin with some standard assumptions that we shall invoke whenever needed. These assumptions will ensure that our variance estimators work properly on a wide variety of stationary processes.

Assumptions A

1. The process $\{Y_i : i \geq 1\}$ is stationary and satisfies the following Functional Central Limit Theorem.

For $n = 1, 2, \ldots$ and $t \in [0, 1]$, the process

$$X_n(t) \equiv \frac{\lfloor nt \rfloor (\bar{Y}_{\lfloor nt \rfloor} - \mu)}{\sigma \sqrt{n}} \qquad (1)$$

satisfies $X_n \underset{n \to \infty}{\Longrightarrow} \mathcal{W}$, where: $\lfloor \cdot \rfloor$ denotes the greatest integer function; $\mathcal{W}$ is a standard Brownian motion process on $[0, 1]$; and $\underset{n \to \infty}{\Longrightarrow}$ denotes weak convergence in $D[0, 1]$, the space of functions on $[0, 1]$ that are right-continuous with left-hand limits, as $n \to \infty$. See also Billingsley (1968) and Glynn and Iglehart (1990).

2. $\sum_{i=-\infty}^{\infty} R_i = \sigma^2 \in (0, \infty)$, where $R_i \equiv \mathsf{Cov}(Y_1, Y_{1+i})$, $i = 0, 1, 2, \ldots$.
3. $\sum_{i=1}^{\infty} i^2 |R_i| < \infty$.
4. The function $f(\cdot)$, defined on $[0, 1]$, is twice continuously differentiable and satisfies the normalizing condition $\int_0^1 \int_0^1 f(s) f(t) [\min\{s, t\} - st] \, ds \, dt = 1$.

Assumptions A.1–A.3 hold for a variety of stochastic processes encountered in practice (see Glynn and Iglehart 1990).

The STS, which we define next, will form the basis of all of the estimators studied herein.

Definition 1 As in Schruben (1983), the (level-0) standardized time series of the process $\{Y_i\}$ is

$$T_{0,n}(t) \equiv \frac{\lfloor nt \rfloor (\bar{Y}_n - \bar{Y}_{\lfloor nt \rfloor})}{\sigma \sqrt{n}} \quad \text{for } t \in [0, 1]. \qquad (2)$$

The next section relates the STS (2) to a Brownian bridge process. In Section 5.2, we show how to use this process to derive estimators for σ^2.

3 FOLDED BROWNIAN BRIDGES

Our development requires some additional nomenclature. First of all, we define a Brownian bridge, which will turn out to be the limiting process of a standardized time series.

Definition 2 Suppose that $\mathcal{W}(\cdot)$ is a standard Brownian motion process. The associated level-0 Brownian bridge process is

$$\mathcal{B}_0(t) \equiv \mathcal{B}(t) \equiv t\mathcal{W}(1) - \mathcal{W}(t) \quad \text{for } t \in [0, 1].$$

In fact, a Brownian bridge $\{\mathcal{B}(t) : t \in [0, 1]\}$ is a Gaussian process with $\mathsf{E}[\mathcal{B}(t)] = 0$ and $\mathsf{Cov}(\mathcal{B}(s), \mathcal{B}(t)) = \min\{s, t\} - st$, for $s, t \in [0, 1]$. Brownian bridges are important for our purposes because under Assumptions A.1–

A.3, Schruben (1983) shows that $T_{0,n}(\cdot) \underset{n\to\infty}{\Longrightarrow} \mathcal{B}(\cdot)$ and that

$\bar{Y}_n$ and $T_{0,n}(\cdot)$ are asymptotically independent as $n \to \infty$.

We now define what we mean by the folding map, which can be applied either to a STS or a Brownian bridge.

Definition 3 The folding map $\Psi : Y \in D[0,1] \to \Psi_Y \in D[0,1]$ is defined by

$$\Psi_Y(t) \equiv Y\left(\tfrac{t}{2}\right) - Y\left(1 - \tfrac{t}{2}\right) \quad \text{for } t \in [0,1].$$

Moreover, for each nonnegative integer k, we define $\Psi^k : Y \in D[0,1] \to \Psi_Y^k \in D[0,1]$, the k-fold composition of the folding map Ψ, so that for every $t \in [0,1]$,

$$\Psi_Y^k(t) \equiv \begin{cases} Y(t), & \text{if } k = 0, \\ \Psi \circ \Psi_Y^{k-1}(t), & \text{if } k = 1, 2, \ldots. \end{cases}$$

The folding operation can be performed multiple times on the Brownian bridge process, as demonstrated by the following definition.

Definition 4 (see Shorack and Wellner 1986) For $k = 1, 2, \ldots$, the level-k folded Brownian bridge is

$$\mathcal{B}_k(t) \equiv \Psi_{\mathcal{B}_{k-1}}(t) = \mathcal{B}_{k-1}\left(\tfrac{t}{2}\right) - \mathcal{B}_{k-1}\left(1 - \tfrac{t}{2}\right),$$

so that $\mathcal{B}_k(t) = \Psi_{\mathcal{B}_0}^k(t)$ for $t \in [0,1]$.

Intuitively speaking, when the folding operator Ψ is applied to a Brownian bridge process $\{\mathcal{B}_0(t) : t \in [0,1]\}$, it does the following: (i) Ψ reflects (folds) the portion of the original process defined on the subinterval $\left[\tfrac{1}{2}, 1\right]$ about the vertical line $t = \tfrac{1}{2}$ (yielding the subprocess shown in red in Figure 1); and (ii) Ψ takes the difference between these two subprocesses defined on $\left[0, \tfrac{1}{2}\right]$ and stretches that difference over the unit interval $[0,1]$ (yielding the new process shown in Figure 2).

Lemma 1 shows that as long as we start with a Brownian bridge, folding it will produce another Brownian bridge.

Lemma 1 *For $k = 1, 2, \ldots$, the process $\{\mathcal{B}_k(t) : t \in [0,1]\}$ is a Brownian bridge.*

The next lemma relates the level-k Brownian bridge with the original (level-0) Brownian bridge and the initial Brownian motion process. These results are useful in deriving properties of certain functionals of $\mathcal{B}_k(t)$.

Lemma 2 *For $k = 1, 2, \ldots$,*

$$\mathcal{B}_k(t) = \sum_{i=1}^{2^{k-1}} \left[\mathcal{B}\left(\tfrac{i-1}{2^{k-1}} + \tfrac{t}{2^k}\right) - \mathcal{B}\left(\tfrac{i}{2^{k-1}} - \tfrac{t}{2^k}\right)\right]; \quad (3)$$
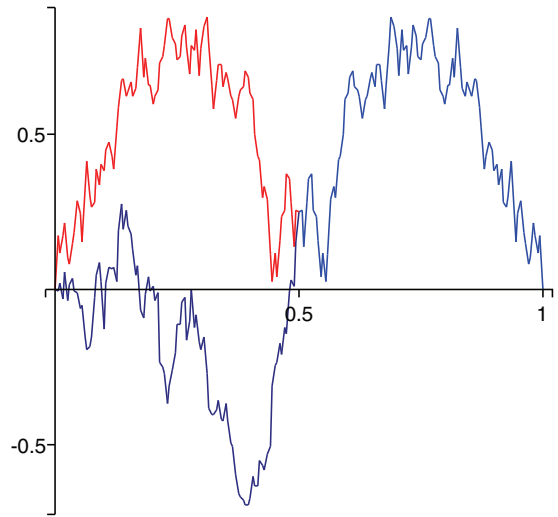

Figure 1: Illustration of a folded Brownian bridge, with reflection about $t = 1/2$ in red.

and it follows from (3) that

$$\mathcal{B}_k(t) = (t-1)\mathcal{W}(1) + \sum_{i=1}^{2^{k-1}} \left[\mathcal{W}\left(\tfrac{i}{2^{k-1}} - \tfrac{t}{2^k}\right) - \mathcal{W}\left(\tfrac{i-1}{2^{k-1}} + \tfrac{t}{2^k}\right)\right].$$

4 SOME FUNCTIONALS OF FOLDED BROWNIAN BRIDGES

This section gives results on the weighted areas under successively higher levels of folded Brownian bridges. Such functionals will be used in Section 5 to construct estimators for the variance parameter σ^2.

Definition 5 For $k = 0, 1, \ldots$, the weighted area under the level-k folded Brownian bridge is

$$N_k(f) \equiv \int_0^1 f(t)\mathcal{B}_k(t)\,dt.$$

Under simple conditions, Theorem 1 shows that $N_k(f)$ has a standard normal distribution.

Theorem 1 *For any normalized weight function $f(t)$ (satisfying Assumption A.4) and any nonnegative integer k, we have $N_k(f) \sim \mathsf{Nor}(0,1)$.*

Corollary 2 *Under the conditions of Theorem 1, we have $A_k(f) \equiv \sigma^2 N_k^2(f) \sim \sigma^2 \chi_1^2$.*

Corollary 2 shows that $\mathsf{E}[A_k(f)] = \sigma^2$, a finding that we will revisit in Theorem 8 when we develop estimators for σ^2.

Meanwhile, we proceed with several results concerning the joint distribution of the $\{N_k(f) : k = 0, 1, \ldots\}$. Our
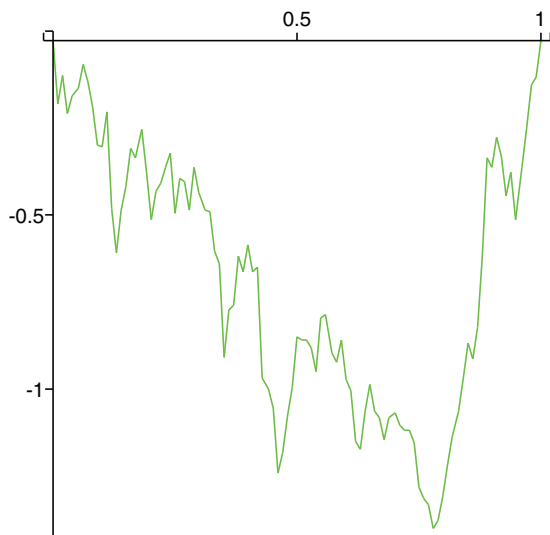
Figure 2: Illustration of a folded Brownian bridge, taking differences and stretching.

first such result gives an expression for the covariance between folded area functionals from different levels. Before stating the theorem, for any weight function $f(\cdot)$, we define $F(t) \equiv \int_0^t f(s)\, ds$, $F \equiv F(1)$, $\bar{F}(t) \equiv \int_0^t F(s)\, ds$, and $\bar{F} \equiv \bar{F}(1)$.

Theorem 3 *Let $f_1(t)$ and $f_2(t)$ be normalized weight functions. Then for $\ell = 0, 1, \ldots$ and $k = 1, 2, \ldots$, we have*

$$\mathsf{Cov}[N_\ell(f_1), N_{\ell+k}(f_2)] = -\bar{F}_1 \bar{F}_2$$
$$+ \sum_{i=1}^{2^{k-1}} \int_0^1 f_2(t) \Big[\bar{F}_1\big(\tfrac{i}{2^{k-1}} - \tfrac{t}{2^k}\big) - \bar{F}_1\big(\tfrac{i-1}{2^{k-1}} + \tfrac{t}{2^k}\big) \Big] dt.$$

The following lemma establishes the multivariate normality of the random vector $\mathbf{N}(f) \equiv [N_0(f), N_1(f), \ldots, N_k(f)]$.

Lemma 3 *If the normalized weight function satisfies $f(t) = f(1 - t)$ for all $t \in [0, 1]$, then for each positive integer k, the random vector $\mathbf{N}(f)$ has a nonsingular multivariate normal distribution.*

Theorem 3 and Lemma 3 can be used to obtain Theorem 4—the remarkable result that, under relatively simple conditions, the folded functionals $\{N_k(f) : k = 0, 1, \ldots\}$ are i.i.d. Nor$(0, 1)$.

Theorem 4 *If the normalized weight function $f(t)$ satisfies $f(t) = f(1-t)$ for all $t \in [0, 1]$, then the random variables $\{N_k(f) : k = 0, 1, \ldots\}$ are i.i.d. Nor$(0, 1)$ random variables.*

The next corollary, which follows immediately from Theorem 4, will serve as the basis for the new variance estimators to be derived in Section 5.

Corollary 5 *Under the conditions of Theorem 4, the random variables $\{A_k(f) : k = 0, 1, \ldots\}$ are i.i.d. $\sigma^2 \chi_1^2$.*

Example 1 *The following weight functions arise in simulation output analysis applications (see [Foley and Goldsman 1999](#) and Section 5 herein): $f_0(t) \equiv \sqrt{12}$, $f_2(t) \equiv \sqrt{840}(3t^2 - 3t + 1/2)$, and $f_{\cos, j}(t) \equiv \sqrt{8}\pi j \cos(2\pi j t)$, $j = 1, 2, \ldots$, all for $t \in [0, 1]$. By Theorem 4, $\{N_k(f) : k \geq 0\}$ are i.i.d. Nor$(0, 1)$, and by Corollary 5, $\{A_k(f) : k \geq 0\}$ are i.i.d. $\sigma^2 \chi_1^2$ for $f = f_0$, f_2, or $f_{\cos, j}$, $j = 1, 2, \ldots$.*

5 APPLICATION TO VARIANCE ESTIMATION

We finally show how the area functionals of folded Brownian bridges can be used in simulation output analysis. With this application in mind, we apply the folding transformation to Schruben's level-0 STS in Section 5.1, thereby obtaining several new versions of the STS. These new series are used in Section 5.2 to produce new estimators for σ^2. Section 5.3 gives obvious methods to improve the estimators, and Section 5.4 gives a Monte Carlo example showing that the estimators work as intended.

5.1 Folded Standardized Time Series

Analogous to the level-k folded Brownian bridge from Definition 4, we define the level-k folded STS.

Definition 6 *For $k = 1, 2, \ldots$, the level-k folded STS is*

$$T_{k,n}(t) \equiv \Psi_{T_{k-1,n}}(t) = T_{k-1,n}\big(\tfrac{t}{2}\big) - T_{k-1,n}\big(1 - \tfrac{t}{2}\big)$$

so that $T_{k,n}(t) = \Psi_{T_{0,n}}^k(t)$ for $t \in [0, 1]$.

The next goal is to examine the convergence of the level-k folded STS to the analogous level-k folded Brownian bridge process. The following result is an immediate consequence of the almost-sure continuity of Ψ^k on $D[0, 1]$ for $k = 0, 1, \ldots$, and the Continuous Mapping Theorem (CMT) ([Billingsley 1968](#)).

Theorem 6 *If Assumptions A.1–A.3 hold, then for any fixed nonnegative integer k, we have*

$$\big[T_{0,n}(\cdot), \ldots, T_{k,n}(\cdot)\big] \underset{n \to \infty}{\Longrightarrow} \big[\mathcal{B}_0(\cdot), \ldots, \mathcal{B}_k(\cdot)\big].$$

Moreover, $\sqrt{n}(\bar{Y}_n - \mu)$ and $[T_{0,n}(\cdot), \ldots, T_{k,n}(\cdot)]$ are asymptotically independent as $n \to \infty$.

5.2 Folded Area Estimators

We introduce folded versions of the STS area estimator for σ^2, along with their asymptotic distributions, expected values, and variances. To begin, we define our new estimators, along with their limiting Brownian bridge functionals.

Definition 7 For each nonnegative integer k, the STS level-k *folded area estimator* for σ^2 is

$$A_k(f; n) \equiv N_k^2(f; n),$$

where

$$N_k(f; n) \equiv \frac{1}{n} \sum_{j=1}^{n} f\left(\tfrac{j}{n}\right) \sigma T_{k,n}\left(\tfrac{j}{n}\right)$$

and $f(\cdot)$ is a normalized weight function (satisfying Assumption A.4). The case $k = 0$, $f = f_0$ corresponds to Schruben's original area estimator (Schruben 1983).

Definition 8 Let $\boldsymbol{A}(f; n) \equiv [A_0(f; n), A_1(f; n), \ldots, A_k(f; n)]$ and $\boldsymbol{A}(f) \equiv [A_0(f), A_1(f), \ldots, A_k(f)]$.

The following definitions provide the necessary set-up to establish in Theorem 7 below the asymptotic distribution of the random vector $\boldsymbol{A}(f; n)$ as $n \to \infty$.

Definition 9 Let Λ denote the class of strictly increasing, continuous mappings of $[0, 1]$ onto itself such that for every $\lambda \in \Lambda$, we have $\lambda(0) = 0$ and $\lambda(1) = 1$. If $X, Y \in D[0, 1]$, then the Skorohod metric $\rho(X, Y)$ defining the "distance" between X and Y in $D[0, 1]$ is the infimum of those positive ξ for which there exists a $\lambda \in \Lambda$ such that $\sup_{t \in [0,1]} |\lambda(t) - t| \leq \xi$ and $\sup_{t \in [0,1]} |X(t) - Y[\lambda(t)]| \leq \xi$. (See Billingsley 1968 for details.)

Definition 10 For each positive integer n, let $\Omega^n : Y \in D[0, 1] \to \Omega_Y^n \in D[0, 1]$ be the approximate (discrete) STS map

$$\Omega_Y^n(t) \equiv \frac{\lfloor nt \rfloor Y(1)}{n} - Y(t) \quad \text{for } t \in [0, 1].$$

Moreover, let $\Omega : Y \in D[0, 1] \to \Omega_Y \in D[0, 1]$ denote the corresponding asymptotic STS map

$$\Omega_Y(t) \equiv \lim_{n \to \infty} \Omega_Y^n(t) = tY(1) - Y(t) \quad \text{for } t \in [0, 1].$$

Note that Ω^n maps the process (1) into the corresponding STS (2) so that we have $\Omega_{X_n}^n(t) = T_{0,n}(t)$ for $t \in [0, 1]$ and $n = 1, 2, \ldots$; moreover, Ω maps a standard Brownian

motion process into a standard Brownian bridge process, $\Omega_{\mathcal{W}}(t) = t\mathcal{W}(1) - \mathcal{W}(t) \sim \mathcal{B}_0(t)$ for $t \in [0, 1]$.

Definition 11 For a given normalized weight function $f(\cdot)$, for every nonnegative integer k, and for every positive integer n, the approximate (discrete) folded area map $\Theta_n^k : Y \in D[0, 1] \to \Theta_n^k(Y) \in \mathbb{R}$ is defined by

$$\Theta_n^k(Y) \equiv \left[\frac{1}{n} \sum_{i=1}^{n} f\left(\tfrac{i}{n}\right) \sigma \Psi^k \circ \Omega_Y^n\left(\tfrac{i}{n}\right) \right]^2.$$

Moreover, the corresponding asymptotic folded area map $\Theta^k : Y \in D[0, 1] \to \Theta^k(Y) \in \mathbb{R}$ is defined by

$$\Theta^k(Y) \equiv \left[\int_0^1 f(t) \sigma \Psi^k \circ \Omega_Y(t)\, dt \right]^2.$$

In terms of Eq. (1), the definition of $A_k(f)$ from Corollary 2, and the Definitions 8–11, we see that $\Theta_n^k(X_n) = A_k(f; n)$ and $\Theta^k(\mathcal{W}) = A_k(f)$ for every nonnegative integer k. We are now ready to proceed with the main convergence theorem, which shows that the folded area estimators converge jointly to their asymptotic counterparts.

Theorem 7 *If Assumptions A hold, then*

$$\boldsymbol{A}(f, n) \underset{n \to \infty}{\Longrightarrow} \boldsymbol{A}(f).$$

Sketch of Proof. Although the proof of Theorem 7 is detailed in Antonini et al. (2007), it can be summarized as follows. Our goal is to apply the generalized CMT (i.e., Theorem 5.5 of Billingsley 1968) to prove that the $(k + 1) \times 1$ random vector with jth element $\Theta_n^j(X_n)$ converges in distribution to the $(k+1) \times 1$ random vector with jth element $\Theta^j(\mathcal{W})$ for $j = 0, 1, \ldots, k$. To establish the hypotheses of the generalized CMT, we show that if $\{x_n\} \subset D[0, 1]$ is any sequence of functions converging to a realization $\mathcal{W}$ of a standard Brownian motion process in the Skorohod metric on $D[0, 1]$, then the real-valued sequence $\Theta_n^j(x_n)$ converges to $\Theta^j(\mathcal{W})$ almost surely. First we exploit the almost-sure continuity of $\mathcal{W}(u)$ at every $u \in [0, 1]$ and the convergence of $\{x_n\}$ to $\mathcal{W}$ in $D[0, 1]$ to show that for every nonnegative integer j, with probability one (w.p.1) we have $\left| \Psi^j \circ \Omega_{x_n}^n(t) - \Psi^j \circ \Omega_{\mathcal{W}}^n(t) \right| \to 0$ uniformly for $t \in [0, 1]$ as $n \to \infty$; and it follows that

$$\lim_{n \to \infty} \left| \Theta_n^j(x_n) - \Theta_n^j(\mathcal{W}) \right| = 0 \quad \text{w.p.1}. \qquad (4)$$

Next we exploit the almost-sure convergence $\left| \Psi^j \circ \Omega_{\mathcal{W}}^n(t) - \Psi^j \circ \Omega_{\mathcal{W}}(t) \right| \to 0$ for all $t \in [0, 1]$ as $n \to \infty$ together with the almost-sure continuity and Riemann

integrability of $f(t)\Psi^j \circ \Omega_{\mathcal{W}}(t)$ for $t \in [0, 1]$ to show that

$$\lim_{n\to\infty} \left| \Theta_n^j(\mathcal{W}) - \Theta^j(\mathcal{W}) \right| = 0 \quad \text{w.p.1.} \qquad (5)$$

Combining (4) and (5) and applying the triangle inequality, we see that the corresponding vector-valued sequence $\left\{ \left[\Theta_n^0(x_n), \ldots, \Theta_n^k(x_n) \right] : n = 1, 2, \ldots \right\}$ converges to $\left[\Theta^0(\mathcal{W}), \ldots, \Theta^k(\mathcal{W}) \right]$ in $\mathbb{R}^{k+1}$ w.p.1. Thus the desired result follows from the generalized CMT. □

Remark 1 Under Assumptions A, and for the weight functions in Example 1, Theorem 7 and Corollary 2 imply that $A_0(f; n), \ldots, A_k(f; n)$ are asymptotically (as $n \to \infty$) i.i.d. $\sigma^2 \chi_1^2$ random variables.

Theorem 8 gives asymptotic expressions for the expected values and variances of the level-k area estimators.

Theorem 8 *Suppose that Assumptions A hold. Further, for fixed $k \geq 0$, suppose that the family of random variables $\left\{ A_k^2(f; n) : n \geq 1 \right\}$ is uniformly integrable (see* Billingsley 1968 *for a definition and sufficient conditions). Then*

$$\mathsf{E}[A_k(f; n)] \xrightarrow[n\to\infty]{} \mathsf{E}[A_k(f)] = \sigma^2$$

and

$$\mathsf{Var}[A_k(f; n)] \xrightarrow[n\to\infty]{} \mathsf{Var}[A_k(f)] = 2\sigma^4.$$

Remark 2 One can obtain finer-tuned results for $\mathsf{E}[A_0(f; n)]$ and $\mathsf{E}[A_1(f; n)]$. In particular, under Assumptions A, Foley and Goldsman (1999) and Goldsman, Meketon, and Schruben (1990) show that

$$\mathsf{E}[A_0(f; n)] = \sigma^2 + \frac{[(F - \bar{F})^2 + \bar{F}^2]\gamma}{2n} + o(1/n),$$

where $\gamma \equiv -2 \sum_{i=1}^{\infty} i R_i$ (cf. Song and Schmeiser 1995). In addition, Alexopoulos et al. (2007a) find that if Assumptions A hold and n is even, then

$$\mathsf{E}[A_1(f; n)] = \sigma^2 + \frac{\bar{F}^2 \gamma}{n} + o(1/n).$$

5.3 Enhanced Estimators

The individual estimators whose properties are given in Theorem 8 are all based on a single long run of n observations, and all involve a single level k of folding. We now discuss extensions of the estimators that have improved asymptotic properties—batching and combining levels.

Batching: We can organize the data by breaking the n observations into b contiguous, nonoverlapping batches,

each of size m so that $n = bm$; and then we can compute the folded variance estimators from each batch separately. As the batch size $m \to \infty$, the estimators computed from different batches are asymptotically independent under broadly applicable conditions on the original (unbatched) process $\{Y_i : i \geq 1\}$; and thus more-stable estimators can be obtained by combining the folded variance estimators computed from all available batches.

With this motivation, suppose the ith batch of size m consists of the observations $Y_{(i-1)m+1}, \ldots, Y_{im}$, for $i = 1, \ldots, b$. Using the obvious changes to the appropriate definitions, one can construct the level-k STS from the ith batch of observations; and from that, we obtain the resulting level-k area estimator from the ith batch, say $A_{k,i}(f; m)$. Finally, the level-k batched folded area estimator for σ^2 is

$$\tilde{A}_k(f; b, m) \equiv \frac{1}{b} \sum_{i=1}^{b} A_{k,i}(f; m).$$

Under the conditions of Theorem 8, we have $\lim_{m\to\infty} \mathsf{E}[\tilde{A}_k(f; b, m)] = \sigma^2$ and $\lim_{m\to\infty} \mathsf{Var}[\tilde{A}_k(f; b, m)] = 2\sigma^4/b$, where the latter result follows from the fact that the $A_{k,i}(f; m)$'s, $i = 1, \ldots, b$, are asymptotically independent as $m \to \infty$. Thus, the batched estimators have about the same expected value σ^2 as a single folded estimator arising from one long run, yet with substantially smaller variance.

Combining levels of folding: Theorem 8 shows that, for a particular weight function $f(\cdot)$, all of the estimators from different levels of folding behave about the same asymptotically in terms of their expected value and variance. We can improve upon these individual estimators by combining the different levels. To this end, denote the average of the folded area estimators from levels $0, 1, \ldots, k$ by

$$\bar{A}_k(f; n) \equiv \frac{1}{k+1} \sum_{j=0}^{k} A_j(f; n).$$

Under the conditions of Remark 1 and Theorem 8, we have $\lim_{n\to\infty} \mathsf{E}[\bar{A}_k(f; n)] = \sigma^2$ and $\lim_{n\to\infty} \mathsf{Var}[\bar{A}_k(f; n)] = 2\sigma^4/(k+1)$. Thus, we obtain combined estimators with about the same expected value σ^2 as a single folded estimator from one level, yet with significantly smaller variance.

5.4 Example

We illustrate the performance of the new folded estimator with a Monte Carlo experiment involving a stationary first-order autoregressive [AR(1)] process. The AR(1) is constructed by setting $Y_i = \phi Y_{i-1} + \epsilon_i$, $i = 1, 2, \ldots$, where the ϵ_i's are i.i.d. $\mathsf{Nor}(0, 1 - \phi^2)$, Y_0 is $\mathsf{Nor}(0,1)$ and independent of the ϵ_i's, and $-1 < \phi < 1$. For the AR(1), $R_k = \phi^k$, $k = 0, 1, 2 \ldots$, and $\sigma^2 = (1 + \phi)/(1 - \phi)$.

In the current example,

- We took $\phi = 0.9$ (so that $\sigma^2 = 19$).
- We used $b = 32$ batches of observations, various batch sizes m, and weight function $f_0(\cdot)$.
- We carried out 4096 independent realizations of the level-0 and level-1 batched folded area estimators, denoted by $\tilde{A}_0(f_0; 32, m)$ and $\tilde{A}_1(f_0; 32, m)$.
- We calculated the estimated mean and variance of the estimators, averaged over the 4096 realizations.

We demonstrate the convergence of the expected values of the estimators to σ^2 as the batch size m increases—see columns 2–5 of Table 1. In the table, a hat denotes the estimate of the analogous performance metric; we suppress the estimator notation "$(f_0; 32, m)$" for succinctness.

We also checked the performance of the batched estimators when we combine levels. In particular, we used the realizations from the individual levels 0 and 1 to calculate realizations of

$$\mathcal{A}_1(f_0; 32, m) \equiv \frac{1}{2}[\tilde{A}_0(f_0; 32, m) + \tilde{A}_1(f_0; 32, m)],$$

and we obtained the estimated performance characteristics for this combined estimator, shown in the last two columns of Table 1.

Table 1: Empirical performance of the enhanced folded area estimators.

m	$\widehat{\mathsf{E}}[\tilde{A}_0]$	$\widehat{\mathsf{Var}}[\tilde{A}_0]$	$\widehat{\mathsf{E}}[\tilde{A}_1]$	$\widehat{\mathsf{Var}}[\tilde{A}_1]$	$\widehat{\mathsf{E}}[\mathcal{A}_1]$	$\widehat{\mathsf{Var}}[\mathcal{A}_1]$
1024	18.50	21.52	18.34	21.38	18.42	10.80
2048	18.82	22.99	18.66	21.87	18.74	11.02
4096	18.78	23.33	18.85	22.61	18.82	11.40
8192	18.98	22.68	18.82	22.57	18.90	11.21
16384	18.90	22.24	18.98	22.84	18.94	11.32

We see that the estimators $\tilde{A}_0(f_0; 32, m)$ and $\tilde{A}_1(f_0; 32, m)$ perform as anticipated by the theory—for large batch size m, the estimated expected values are nearly σ^2 (≈ 19 in this example) and the variances are reasonably close to $2\sigma^4/b$ (≈ 22.563 for the AR(1) process we chose). We also observe that reasonable expected values are maintained for the combined estimator $\mathcal{A}_1(f_0; 32, m)$ as m becomes large, with the bonus of substantially ($\approx 50\%$) reduced variance.

6   CONCLUSIONS

In this paper, we introduced "folded" versions of the STS area estimator for the variance parameter arising from a stationary simulation process. We provided theoretical results showing that the folded estimators converge to appropriate functionals of Brownian motion; and these convergence results allowed us to produce asymptotically unbiased and low-variance estimators using multiple folding levels in conjunction with standard batching techniques. At each folding level, and for each weight function in Example 1, the proposed estimators can be computed in $O(n)$ time; the detailed computations will be listed in a forthcoming article.

Additional ongoing work includes the following. As in Remark 2, we can derive precise expressions for the expected values of the folded estimators—expressions that show just how quickly any estimator bias dies off as the batch size increases. We can also produce analogous folding results for other "primitive" STS variance estimators, e.g., for Cramér–von Mises estimators, as described in Antonini (2005). In any case, we have also planned a comprehensive Monte Carlo analysis to examine estimator performance over a variety of benchmark processes. Future work includes the development of overlapping versions of the folded estimators, as in Alexopoulos et al. (2007a), Alexopoulos et al. (2007b), Alexopoulos et al. (2007c) and Meketon and Schmeiser (1984).

REFERENCES

Aktaran-Kalaycı, T., D. Goldsman, and J. R. Wilson. 2007. Linear combinations of overlapping variance estimators for simulation. *Operations Research Letters* 35:439–447.

Alexopoulos, C., C. Antonini, D. Goldsman, M. Meterelliyoz, and J. R. Wilson. 2007a. Properties of folded standardized time series variance estimators for simulation. Technical report, H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

Alexopoulos, C., N. T. Argon, D. Goldsman, N. M. Steiger, G. Tokol, and J. R. Wilson. 2007b. Efficient computation of overlapping variance estimators for simulation. *INFORMS Journal on Computing* 19 (3): 314–327.

Alexopoulos, C., N. T. Argon, D. Goldsman, G. Tokol, and J. R. Wilson. 2007c. Overlapping variance estimators for simulation. *Operations Research* to appear. Available online via <ftp.ncsu.edu/pub/eos/pub/jwilson/ovestv72.pdf> [accessed June 17, 2007].

Antonini, C. 2005. *Folded variance estimators for stationary time series*. Ph.D. thesis, H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia. Available online via <hdl.handle.net/1853/6931> [accessed June 17, 2007].

Antonini, C., C. Alexopoulos, D. Goldsman, and J. R. Wilson. 2007. Area variance estimators for simulation

using folded standardized time series. Technical report, H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

Billingsley, P. 1968. *Convergence of probability measures*. New York: Wiley.

Foley, R. D., and D. Goldsman. 1999. Confidence intervals using orthonormally weighted standardized time series. *ACM Transactions on Modeling and Computer Simulation* 9:297–325.

Glynn, P. W., and D. L. Iglehart. 1990. Simulation output analysis using standardized time series. *Mathematics of Operations Research* 15:1–16.

Goldsman, D., K. Kang, S. H. Kim, A. F. Seila, and G. Tokol. 2007. Combining standardized time series area and Cramér–von Mises variance estimators. *Naval Research Logistics* 54:384–396.

Goldsman, D., M. Meketon, and L. Schruben. 1990. Properties of standardized time series weighted area variance estimators. *Management Science* 36:602–612.

Law, A. M. 2006. *Simulation modeling and analysis*. 4th ed. New York: McGraw-Hill.

Meketon, M. S., and B. W. Schmeiser. 1984. Overlapping batch means: Something for nothing? In *Proceedings of the 1984 Winter Simulation Conference*, ed. S. Sheppard, U. Pooch, and D. Pegden, 227–230. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Schruben, L. 1983. Confidence interval estimation using standardized time series. *Operations Research* 31:1090–1108.

Shorack, G. R., and J. A. Wellner. 1986. *Empirical processes with applications to statistics*. New York: Wiley.

Song, W. T., and B. W. Schmeiser. 1995. Optimal mean-squared-error batch sizes. *Management Science* 41:110–123.

AUTHOR BIOGRAPHIES

CLAUDIA ANTONINI has a tenure track position in the Department of Pure and Applied Mathematics at Simón Bolívar University in Caracas, Venezuela. Her research interests include simulation output analysis, quality control, stochastic processes, and statistics. She is a member of the AMV (Asociación Matemática Venezolana). Her e-mail address is <cfmantonini@usb.ve>.

CHRISTOS ALEXOPOULOS is an Associate Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. His research interests are in the areas of simulation, statistics, and optimization of stochastic systems. He is a member of INFORMS and an active participant in the Winter Simulation Conference, having been Proceedings Co-Editor in 1995, Asso-ciate Program Chair in 2006, and a member of the Board of Directors since 2007. He is also the Simulation Department Editor of *IIE Transactions*. His e-mail address is <christos@isye.gatech.edu>, and his web page is <http://www.isye.gatech.edu/~christos>.

DAVID GOLDSMAN is a professor in the H. Milton Stewart School of Industrial and Systems Engineering at the Georgia Institute of Technology. His research interests include simulation output analysis and ranking and selection. He is an active participant in the Winter Simulation Conference, having been Program Chair in 1995, and having served on the WSC Board of Directors since 2002. His e-mail address is <sman@isye.gatech.edu>, and his web page is <http://www.isye.gatech.edu/~sman>.

JAMES R. WILSON is professor and head of the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University. His current research interests are focused on the design and analysis of simulation experiments. He also has an active interest in applications of operations research techniques to all areas of industrial and systems engineering. He is currently the Editor-in-Chief of *ACM Transactions on Modeling and Computer Simulation*. He is a member of AAUW, ACM, and ASA, and he is a Fellow of IIE and INFORMS. His e-mail address is <jwilson@ncsu.edu>, and his web page is <www.ise.ncsu.edu/jwilson>.