Statistical Inference via Overlapping Batches (OB) and Overlapping Standardized Time Series (OSTS) t-Distributions

Dashi I. Singham[1] and Yingchieh Yeh[2†]

[1]Operations Research Department, Naval Postgraduate School, USA
[2]Institute of Industrial Management, National Central University, TAIWAN
[†]Corresponding Author, yeh@mgt.ncu.edu.tw

June 9, 2025

Abstract

Batching is a classic approach for estimating the variance of the sample mean, or other statistical functionals, obtained from a stationary data process. Nonoverlapping batching (NB) has long been the most-common foundation for statistical inference. Overlapping batching (OB), despite its inherent statistical efficiency, has not been a popular basis for statistical inference, primarily because batch dependence makes the classical χ^2 and Student's t coefficients inappropriate. Our purpose is to provide a unified foundation for OB statistical inference and overlapping standardized time series (OSTS) inference by computing the OB-t and OSTS-t distributions, analogous to the classical Student's t distribution. While past overlapping methods relied on approximating degrees of freedom using the Student's t distribution, we develop an eigenvalue-based numerical algorithm for computing the actual distributions for the limiting OB and OSTS estimators, which we use to present a table of quantile values for each distribution. The associated code delivers quantiles from these distributions for direct confidence interval generation. Finally, we employ coverage functions to assess the performance of these t-distributions compared to current approximation methods.

Keywords: Monte Carlo; output analysis; stationary processes; system simulation; confidence intervals

1 Introduction

Stochastic simulation experiments yield point estimates of performance-measure values. Estimating the standard errors (the square roots of the variances) of those point estimators is at the heart of simulation output analysis. The point and standard-error estimates can be used for statistical inference (commonly either confidence intervals or tests of hypotheses) for the performance-measure value and for the point estimator's standard error.

Statistical functionals encompass a wide range of performance estimates, including the the sample mean, quantiles, or optimal values in optimization problems. Our context for exposition will use the sample mean of a stationary stochastic process to estimate its mean, though we later will demonstrate our approach applies to a class of general statistical functionals. From Mechanic (1966a,b) through Law and Kelton (1984), the focus of standard-error estimation was on fixed sample sizes, often based on nonoverlapping batching (NB) with statistical inference based on the classical χ^2 and Student's t sampling distributions. Meketon and Schmeiser (1984) introduced overlapping batching (OB) as a more-efficient estimator of the sample mean's standard error, but overlapping batches create dependence that negate the use of the classical sampling distributions; without a statistical-inference foundation, use of the OB estimator of the standard error languishes. More recently, focus has been on sequential procedures, other performance measures, and other foundations for inference, for example Alexopoulos et al. (2014, 2016), Singham and Schruben (2012), Singham (2014). Yeh and Schmeiser (2015a,b) returned to fixed sample sizes, the mean, and batching to argue that OB is a good basis for confidence intervals. Standardized time series (STS) methods (Schruben 1983) also have been employed as part of overlapping batching procedures to enable confidence interval generation via cancellation of the variance estimator (Alexopoulos et al. 2007a,b, Calvin and Nakayama 2013). But again, the lack of an exact sampling distribution for overlapping standardized time series (OSTS) methods prohibits generation of confidence intervals without employing approximation approaches.

Estimation of the variance of the sample mean, along with other statistical functionals by the overlapping batching method in general, (including both OB and OSTS) has been well-developed and extensively studied over the years. However, despite desirable statistical properties, these methods have not gained widespread popularity among simulation practitioners. A major challenge lies in the difficulty in obtaining the correct distribution for the estimators derived from overlapping batches. Many approximations rely on the use of a t-distribution with modified degrees of freedom as a surrogate to account for the dependence between overlapping batches. This paper provides a numerical way to calculate the CDF of OB-t and OSTS-t distributions, from which the quantile values can be easily calculated. Once we have the variance estimate via the overlapping batching method, all that remains is to look up the value from the table (using our associated code) of the OB-t or OSTS-t distribution to construct confidence intervals.

Our contribution here is to define and closely approximate the OB-χ^2, OB-t and OSTS-t distributions, analogous to the classical χ^2 and Student's t distributions, to provide a foundation for OB and OSTS inferences on the standard error of the sample mean, and more generally, of statistical functionals. We present two tables that are analogous to the tables of classical Student's t quantile values found in statistics textbooks and online for the OB-t and OSTS-t distributions. Additionally, we provide code delivering quantiles from a lookup table to allow for generation of confidence intervals. This allows the user to obtain the known benefits of overlapping batches in the well-established batched means domains when using sectioning or STS methods. Our approach closes the gap on this long-standing hurdle to achieving the well-documented numerical benefits of overlapping methods.

We summarize our contributions. The first contribution of this paper is to present a method for calculating the quantiles for the OB-t and OSTS-t distributions. Recent work closely parallels ours, and we highlight the key distinctions. Su et al. (2024) propose a general confidence interval approach for statistical functionals using overlapping batching, but do not provide a computational method for generating the quantiles of the t-distributions. We describe a computational method for computing the t-distributions and provide the resulting quantiles as a lookup table. One option is to prove limiting results for batching estimators of the sample mean under a Functional Central Limit Theorem (FCLT). However, Su et al. (2024) broadened the class of batching estimators to include all statistical functionals, but needed to make a more-restrictive strong invariance assumption. This paper classifies and organizes the literature in these areas, before providing the numerical approach which will allow each of these past theoretical results to be employed in practice.

Su et al. (2024) demonstrate convergence for OB and OSTS estimators to a t-distribution, but omit the proof for OSTS estimators. As a second contribution, we derive the proof for OSTS

t-distribution in this paper using a strong invariance assumption. Our third contribution is to demonstrate the performance of the exact distributions relative to prior approximation methods using coverage functions (Schruben 1980). These coverage functions provide a more comprehensive analysis of overlapping batching performance compared to the usual numerical tables provided for arbitrary confidence coefficients, and we believe they should set the standard for future evaluation of confidence interval procedures.

The organization of this paper is as follows. In Section 2 we review simulation output analysis, focusing on the NB, OB, and OSTS estimators' use in statistical inference for the sample mean with data from a stationary stochastic process. Section 3 extends the estimators for the mean to encompass a large class of statistical functionals for which confidence intervals can be generated. We define the OB-t and OSTS-t distributions in Section 4, which also contains the two lookup tables of quantile values. In Section 5 we present the asymptotic OB and OSTS distributions and provide a proof to fill a last gap in the literature. In Section 6, we discuss the method for computing the quantiles from overlapping distributions, which relies on a covariance calculation between overlapping batches in our eigenvalue-based numerical method. In Section 7 we derive covariances between overlapping batches in both the OB and OSTS case. Section 8 delivers numerical performance results using coverage functions for our exact t-distribution, while Section 9 concludes. The appendix contains proofs and supporting calculations.

2 Overlapping Batch Estimators: A Review

We begin with a review of batching estimators for the sample mean as a familiar framework. Here we review relevant statistical analysis, including a discussion of typical data assumptions, followed by a review of three key batching estimators: NB in Section 2.1, OB in Section 2.2 and OSTS in Section 2.3. In Section 3, we will extend our review from sample mean analysis to general statistical functionals, and show how overlapping batching estimators can be used to construct confidence intervals for these functionals.

A common use of stochastic simulation is to estimate the mean μ and variance σ^2 of a stationary time series $Y_1, Y_2, \ldots$. For sample size n, after possibly deleting some initial data to avoid initialization bias, the usual point estimator of μ is the sample mean

$$\overline{Y} = n^{-1} \sum_{i=1}^{n} Y_i$$

and the usual estimator of σ^2 is the sample variance

$$S^2 = (n-1)^{-1} \sum_{i=1}^{n} (Y_i - \overline{Y})^2.$$

Throughout this paper, we make the following assumption that the data is identically distributed and covariance stationary, which allows related statistical-inference statements.

Data Assumption *Identical.* *The data process $Y \equiv Y_1, Y_2, \ldots$ is composed of identically distributed observations with mean μ and variance σ^2. Asymptotically (as n goes to infinity) $\overline{Y}$ is normally distributed. The asymptotic value of $n\,\mathrm{var}(\overline{Y})$ is $\tau = \gamma_0\sigma^2$, where $\gamma_0 = 1 + 2\sum_{h=1}^{\infty} \rho_h$ is the asymptotic sum of all lag-h autocorrelations. The value of τ is finite.*

Because the value of τ is seldom known, using only the data $(Y_1, Y_2, \ldots, Y_n)$ to estimate τ is at the heart of the analysis. In our context, simulation output analysis seeks statistical-inference statements about the quality of $\overline{Y}$ as an estimator of μ and statements about the quality of $\widehat{\tau}$, the point estimator of τ. Before continuing, we define a second data assumption, which will hold only whenever the phrase "i.i.d. normal" is used.

Data Assumption *i.i.d. normal.* *The data process $Y \equiv Y_1, Y_2, \ldots$ is composed of mutually independent observations that are each identically normally distributed.*

There are various ways that i.i.d. normal data can occur. The most common is by fiat, as is done for the classical χ^2 and Student's t distributions. In the simulation context, where nonnormality and autodependence are commonly encountered, i.i.d. normal data arise by assuming that the observations are differences of Brownian Motion. Typically the Brownian Motion process is assumed to arise via a Functional Central Limit Theorem (FCLT) that assumes asymptotically large amounts of more-primitive data. Data Assumption *i.i.d. normal* is weaker than assuming that data arise from Brownian Motion, because a Brownian Motion cumulative process implies i.i.d. normal observations Y_i, but i.i.d. normal observations imply nothing about a corresponding continuous-time process. The i.i.d. normal assumption is used in Schmeiser (1982) to reach general conclusions about the desired number of NB batches.

Having i.i.d. normal data greatly simplifies the theory and application of inference. All autocorrelations are zero, so $\gamma_0 = 1$ and $\tau = \sigma^2$. Two classical sampling distributions arise: (a) the scaled sample variance, $(n-1)S^2/\sigma^2$, has the χ^2 distribution with $n-1$ degrees of freedom and (b) the sample mean scaled by its estimated standard error, $\overline{Y}/(S/\sqrt{n})$, has the Student's t distribution with $n-1$ degrees of freedom.

When the data are not i.i.d. normal, an approach to statistical inference is to batch the data and then analyze batch means rather than the original data. For $j = 1, 2, \ldots, n-m+1$, define the jth batch of size m to be $Y_j, Y_{j+1}, \ldots, Y_{j+m-1}$ and its batch mean to be the simple average

$$\overline{Y}_j = m^{-1} \sum_{i=j}^{j+m-1} Y_i. \tag{2.1}$$

Batching methods can work because μ is the expected value of the observations Y_i, the batch means $\overline{Y}_j$, and the grand mean $\overline{Y}$.

Batching methods differ primarily in which subsets of the $n-m+1$ batch means are used. OB uses all $n-m+1$ batches to maximize statistical efficiency in estimating τ. Welch (1987) suggested using partially overlapping batched means for computational savings because fewer batches are used. The most common choice is NB, which uses $k = \lfloor n/m \rfloor$ adjacent batches. Fox et al. (1991) investigated spaced batch means, which uses fewer batches by leaving some intermediate observations unused.

Variations exist. Bischak et al. (1993) investigates using unequal weights for the observations within each batch. Alexopoulos et al. (2007b,a) discuss properties of more-general overlapping-batch estimators (including OB). Standardized time series methods can also be applied to each batch by employing cancellation in variance estimation. Our interest here, however, is to focus on three key approaches which we review in the next sections: NB using contiguous batches, OB using all batches (fully overlapping) and OSTS methods using fully overlapping batches. The order of the computing efficiency is $O(n)$ for NB, OB, and OSTS.

2.1 The NB Estimator of τ

The NB estimator of τ partitions the data into k adjacent batches of size m, where $k = \lfloor n/m \rfloor$. The n observations Y_i are used (only) to compute the batch means and the grand mean. NB analyses assume that the batch size m is chosen to be large enough that the batch means are i.i.d. normal and that the effect of the few observations lying in no batch can be ignored. Then, $\mathrm{var}(\overline{Y}) = (m/n)\mathrm{var}(\overline{Y}_j)$. Therefore, $\tau \equiv n\,\mathrm{var}(\overline{Y}) = m\,\mathrm{var}(\overline{Y}_j)$. Given at least two batch means, $\mathrm{var}(\overline{Y}_j)$ is estimated by its sample variance to yield the NB estimator of τ

$$\widehat{\tau}_{\mathrm{nb}} \equiv d_{\mathrm{nb}}^{-1} \sum_{j=1}^{k} (\overline{Y}_{(j-1)m+1} - \overline{Y})^2, \tag{2.2}$$

where $d_{\mathrm{nb}} \equiv (k-1)/m$. Statistical inference on τ is based on

$$\widehat{V}_{\mathrm{nb}} \equiv \frac{(k-1)\widehat{\tau}_{\mathrm{nb}}}{n\,\mathrm{var}(\overline{Y})} \tag{2.3}$$

and statistical inference on μ is based on

$$T_{\mathrm{nb}} \equiv (\overline{Y} - \mu)/\sqrt{\widehat{\tau}_{\mathrm{nb}}/n}. \tag{2.4}$$

With the assumption of i.i.d. normal batch means, properties of these random variables are well known. Estimation of τ is unbiased; that is, $\mathrm{E}(\widehat{\tau}_{\mathrm{nb}}/n) = \mathrm{var}(\overline{Y})$. $\widehat{V}_{\mathrm{nb}}$ has the classical χ^2 distribution with $k-1$ degrees of freedom and T_{nb} has the Student's t distribution with $k-1$ degrees of freedom. Denote $v_{\mathrm{nb}} = \lfloor n/m \rfloor - 1 = k-1$ as the degrees of freedom for the NB estimator. Tables of the χ^2 quantile values and Student's t quantile values are ubiquitous in statistics textbooks and online, as well as easily computed using statistical software.

2.2 The OB Estimator of τ

The OB estimator of τ is analogous to NB. For $m = 1$, both estimators are identical. Both assume that m is large enough so that $\tau \equiv n\,\mathrm{var}(\overline{Y}) = m\,\mathrm{var}(\overline{Y}_j)$. OB, however, uses all $n - m + 1$ batch means to estimate $\mathrm{var}(\overline{Y}_j)$; that is, the OB estimator of τ is

$$\widehat{\tau}_{\mathrm{ob}} \equiv d_{\mathrm{ob}}^{-1} \sum_{j=1}^{n-m+1} (\overline{Y}_j - \overline{Y})^2, \tag{2.5}$$

where $d_{\mathrm{ob}} \equiv (n-m)(n-m+1)/(nm)$. (Meketon and Schmeiser (1984) used a slightly different denominator. With the current denominator, when n/m is an integer the means of $\widehat{\tau}_{\mathrm{ob}}$ and $\widehat{\tau}_{\mathrm{nb}}$ are equal; the asymptotic variance of $\widehat{\tau}_{\mathrm{ob}}$ is two-thirds that of $\widehat{\tau}_{\mathrm{nb}}$, an implied degrees of freedom increase of fifty percent.) Note that (2.5) is referred to as the *sectioning* estimator which uses the sample mean of the entire data set ($\overline{Y}$) as the point estimate. The *batching* estimator averages the sample means of each batch as the point estimate. Because sectioning methods are generally believed to produce better coverage due to lower bias in the point estimator (for example, see Nakayama (2014)), we focus on sectioning results in this paper. Additionally, we always assume fully overlapping batches using the maximum number of batches $(n-m+1)$.

We need an analogy to the classical distributions' degrees of freedom. In the NB context the degrees of freedom is $\nu_{\mathrm{nb}} = k-1$, where $k = \lfloor n/m \rfloor$ is the number of NB batches of size m. To create an analogy for the OB context, we interpret degrees of freedom to be *the ratio of the number of observations outside any particular batch, $n-m$, to the number within any one batch, m*; in particular,

$$\nu_{\mathrm{ob}} = \frac{n-m}{m}.$$

This degrees of freedom definition allows for a common definition for both NB and OB contexts, since for $m=1$ NB and OB are identical. When n/m is an integer, the degrees of freedom for NB and OB are the same. OB degrees of freedom, however, can be fractional, including being less than one. Analogous to NB, define

$$\widehat{V}_{\mathrm{ob}} \equiv \frac{(\frac{n-m}{m})\widehat{\tau}_{\mathrm{ob}}}{n\,\mathrm{var}(\overline{Y})} \tag{2.6}$$

and

$$T_{\mathrm{ob}} \equiv (\overline{Y}-\mu)/\sqrt{\widehat{\tau}_{\mathrm{ob}}/n}. \tag{2.7}$$

Even for i.i.d. normal data, for $m>1$ the batch means are dependent because observations lie in multiple batches. Therefore $\widehat{V}_{\mathrm{ob}}$ does not have a classical chi-squared distribution and T_{ob} does not have a Student's t distribution.

Properties of $\widehat{V}_{\text{ob}}$ are evaluated in Yeh and Schmeiser (2015a) using Monte Carlo experiments. For estimating τ, OB performs better than NB for every statistical measure considered. Several of the improved measures arise because compared to the NB's classical chi-squared distribution, the sampling distribution of $\widehat{V}_{\text{ob}}$ is closer to the normal distribution (Yeh and Schmeiser 2015a).

Despite OB's efficiency advantage, the sampling distributions of $\widehat{V}_{\text{ob}}$ and T_{ob} have not been developed. Rather, approximations have been used, usually reflecting the observation that the asymptotic ratio of the OB and NB variance estimators is 2/3, so increasing the Student's t degrees of freedom by 50% is a quick and reasonable approximation. Alexopoulos et al. (2007c) rescued, via a personal-communication reference, the degrees of freedom formula $\nu_{\text{ob}} = (3/2)\nu_{\text{nb}}[1 + \nu_{\text{nb}}^{-0.5-0.6(\nu_{\text{nb}}+1)}]$, where $\nu_{\text{nb}} = (n-m)/m$, which was obtained by Schmeiser in 1986 to match the observed (via Monte Carlo) variance of $\widehat{V}_{\text{ob}}$ with the Student's t variance of $\nu_{\text{nb}}/(\nu_{\text{nb}} - 2)$.

2.3 The Overlapping Standardized Time Series Estimator of τ

Standardized time series were introduced in Schruben (1983) as a cancellation method for constructing confidence intervals. While the STS estimator can be applied to non-overlapping batches, here we focus on the overlapping case, OSTS. Construct for each batch j of size m (consisting of samples $Y_j, \ldots, Y_{j+m-1}$ with mean $\overline{Y}_j$) a standardized time series:

$$T_j(t) = \frac{\lfloor mt \rfloor}{\sqrt{m}} \left(\overline{Y}_j - \frac{1}{\lfloor mt \rfloor} \sum_{k=j}^{j+\lfloor mt \rfloor - 1} Y_k \right)$$

for $t \in [0,1]$. A weighted area estimator can be computed using a weight function $w : [0,1] \to \mathbb{R}$. For simplicity in future derivations, we assume a constant weight function $w(t) = \sqrt{12}$ which meets the required properties described in Schruben (1983), though more complex weight functions could also be employed. For each batch, we compute the weighted area estimator as

$$A_j = \left[\frac{1}{m} \sum_{k=1}^{m} w\left(\frac{k}{m}\right) T_j\left(\frac{k}{m}\right) \right]^2.$$

The final overlapping batched area estimator is

$$\hat{\tau}_{\text{osts}} \equiv \frac{1}{n-m+1} \sum_{j=1}^{n-m+1} A_j.$$

Define $\beta := m/n$ as the proportion of the batch size to the total sample size. Also let $W(u)$ be standard Brownian motion over $0 \le u \le 1$. Then, let $\mathcal{B}_u(t)$ be a Brownian bridge over $u \le t \le \beta$ as $\mathcal{B}_u(t) = W(u+t) - W(u) - \frac{t}{\beta}[W(u+\beta) - W(u)]$. From Alexopoulos et al. (2007b) we have convergence in distribution of the variance estimator as $m, n \to \infty$:

$$\hat{\tau}_{\text{osts}} \Rightarrow \frac{\tau}{\beta^3(1-\beta)} \int_0^{1-\beta} \left[\int_0^{\beta} w(t)\mathcal{B}_u(t)dt \right]^2 du, \tag{2.8}$$

where $\Rightarrow$ denotes weak convergence. We define the degrees of freedom $\nu_{\text{osts}} = n/m$ which is $\nu_{\text{ob}} + 1$. For nonoverlapping STS estimators, the degrees of freedom is $n/m = \nu_{\text{nb}} + 1$, so we define the degrees of freedom for overlapping standardized time series analagously, as $\nu_{\text{osts}} = \nu_{\text{ob}} + 1$. We can again define the inference parameters:

$$\widehat{V}_{\text{osts}} \equiv \frac{(\frac{n}{m})\widehat{\tau}_{\text{osts}}}{n\,\text{var}(\overline{Y})}$$

and

$$T_{\text{osts}} \equiv (\overline{Y} - \mu)/\sqrt{\widehat{\tau}_{\text{osts}}/n}.$$

As with OB, OSTS estimators exhibit dependence between batches due to overlapping.

3 Confidence Intervals for Statistical Functionals

While the above discussion focuses on the classical example of using batching for mean estimation, we highlight the details of a few key papers that further the development of overlapping batching estimators for values other than sample means. Calvin and Nakayama (2013) develop confidence intervals for quantiles using standardized time series, including employing estimators using overlapping batches. Pasupathy et al. (2022) develop confidence regions for multidimensional quantile vectors finding coverage benefits from overlapping batches, and this work is extended to infinite-dimensional confidence regions for the quantile field in Pasupathy et al. (2024). Critically, Su et al.

(2024) provides a comprehensive study of the limiting behavior of overlapping batch estimators when deriving confidence intervals for general statistical functionals, and we will rely on a similar limiting analysis here. Their paper considers a large batch setting where the batch sizes m_n change with the sample size and $\lim m_n/n > 0$, and also a small batch setting where $\lim m_n/n = 0$. In this paper, we consider the large batch setting in our limiting results, as Su et al. (2024) already show the elegance of the small batch setting resulting in consistent estimation of the variance and convergence to normal random variables. Su et al. (2024) omits the proof for the large batch setting when using standardized time series estimators, and we include it here for completeness.

We next provide a brief summary of statistical functionals and their associated confidence intervals to establish the conditions for when the OB-t and OSTS-t distributions exist. A functional is, in essence, a function whose domain is a space of functions rather than real numbers. Specifically, it takes a function as input and returns a corresponding value. In nonparametric estimation, the quantity of interest θ is often expressed as a functional, denoted by $\theta(P)$, where P is a probability measure specifying the law that generates the underlying stationary data process. A statistical functional is a mapping from a collection of probability measures to real numbers: given a probability measure, it returns a numerical value. One advantage of statistical functionals is that they provide a straightforward way to estimate the quantity of interest by substituting P_n for P in their expression, where P_n denotes the empirical measure constructed from the output sequence $\{Y_1, \ldots, Y_n\}$. That is, $\theta(P_n)$ serves as the plug-in estimator $\widehat{\theta}_n$ for $\theta(P)$. For example, if $\theta(P)$ is the η-quantile of Y_i, $F^{-1}(\eta) := \min\{y : F(y) \geq \eta\}$, then the natural point estimator $\widehat{\theta}_n := \widehat{\theta}(\{Y_i, 1 \leq i \leq n\})$ is the η-th empirical quantile $\min\{y : F_n(y) \geq \eta\}$, where F is the CDF of Y_i and $F_n(y) = \sum_{i=1}^n \mathbb{I}_{\{Y_i \leq y\}}, y \in \mathbb{R}$. Many statistical functionals are applicable to confidence intervals estimation via the batching method, e.g. expectations, quantiles, conditional value at risk; see Su et al. (2024) for more.

The idea of batching can be extended to a broader range of statistical functionals beyond just the mean. Let $\widehat{\theta}_{i,m} := \widehat{\theta}(\{Y_j, j = i, \ldots, i+m-1\}), i = 1, \ldots, n-m+1$ be the estimator constructed from the i-th batch of the observations $Y_i, \ldots, Y_{i+m-1}$. The OB estimator for statistical functionals is defined in a manner analogous to 2.5, with $\overline{Y}_j$ and $\overline{Y}$ replaced by $\widehat{\theta}_{i,m}$ and $\widehat{\theta}_n$, respectively

$$\widehat{\tau}_{\text{ob}} \equiv d_{\text{ob}}^{-1} \sum_{i=1}^{n-m+1} (\widehat{\theta}_{i,m} - \widehat{\theta}_n)^2, \tag{3.1}$$

where $d_{\text{ob}} \equiv (n-m)(n-m+1)/(nm)$. The standardized time series for a statistical functional can be defined similarly,

$$T_{i,m}(t) := \frac{\lfloor mt \rfloor \left(\hat{\theta}_{i,\lfloor mt \rfloor} - \hat{\theta}_{i,m}\right)}{\sqrt{m}}, \tag{3.2}$$

where $t \in [0,1]$.

Section 5 will explicitly define the Functional Central Limit Theorem (FCLT) assumption and the Strong Invariance (SI) assumption, but we motivate their uses here. The strong invariance assumption is a more stringent condition than the FCLT which ensures a Wiener process approximation holds for a sequence $\{\widehat{\theta}_n, n \geq 1\}$ of batching estimators. The SI assumption specifies how large the sampling error is and how it behaves as $n \to \infty$. Traditional asymptotic results, e.g. the Central Limit Theorem, state that $\sqrt{n}(\widehat{\theta}_n - \theta)$ follows a normal distribution. This only describes "convergence in distribution" and does not provide precise sample-wise error bounds. The SI assumption goes further by constructing an almost sure approximation. Instead of saying that the sampling error approximately follows a normal distribution, one can explicitly approximate the sampling error by a Wiener process plus a small remainder, where the remainder often grows more slowly than the Wiener process. This provides an explicit error bound on how closely a sequence of sampling errors of batching estimators tracks a Wiener process. Although these statistical functionals require more stringent regularity conditions, the implication for confidence interval estimation via batching is substantial. Our study shows that given a batching type, e.g., overlapping batches, the distributions of the pivotal statistics corresponding to different statistical functionals, used to construct confidence intervals via batching, converge to a known distribution.

Next, we provide a brief overview of confidence interval calculation. One example of a statistical functional is using the sample mean $\overline{Y}$ as an estimator for μ. Then, having (somehow) an estimate, $\widehat{\tau}$, of τ, one has an estimate of the standard error of $\overline{Y}$; in particular, $\widehat{\text{ste}}(\overline{Y}) = \sqrt{\widehat{\tau}/n}$. This value can be used in its own right to report the precision of an observed $\overline{Y}$. Song and Schmeiser (1994, 2009) discuss using the estimated standard error to determine the number of digits to report. Song and

Schmeiser (2011) suggest format displays for point estimates and their standard-error estimates.

For a general statistical functional where $\hat{\theta}_n$ is the estimator, we have $\widehat{\text{ste}}(\hat{\theta}_n) = \sqrt{\widehat{\tau}/n}$ where $\widehat{\tau}$ is the estimate of the variance constant of the statistical functional. The more-traditional approach to statistical inference is to calculate a confidence interval for (or, equivalently, test a hypothesis on) the performance measure of interest. The usual two-sided confidence interval for θ is

$$\hat{\theta}_n \pm s_{\nu;\alpha}\sqrt{\widehat{\tau}/n}, \tag{3.3}$$

where the interval procedure is designed to cover θ with nominal probability $1-\alpha$. Define $s_{\nu;\alpha}$ as the $1-\alpha$ quantile of the appropriate random variable with ν degrees of freedom. For estimating the mean μ using i.i.d. normal data, high statistical efficiency (that is, actual coverage matches nominal coverage and intervals are narrow) occurs with $m = 1$ (when NB and OB are identical) and $s_{\nu,\alpha} = t_{n-1,\alpha}$, the Student's t quantile with $n-1$ degrees of freedom. For other data, NB uses k large batches so that the batch means are approximately i.i.d. normal and therefore $s_{\nu,\alpha} = t_{k-1,\alpha}$ is appropriate. OB procedures, in contrast, commonly use the approximation of Student's t quantiles with increased degrees of freedom as discussed in Section 2.2.

Statistical inference can also make statements about the quality of the estimate of the variance constant τ when estimating the mean μ using $\overline{Y}$. The usual two-sided confidence interval for τ is

$$(\nu\widehat{\tau}/c_{\psi,1-\alpha/2},\, \nu\widehat{\tau}/c_{\psi,\alpha/2}), \tag{3.4}$$

where the interval is designed to cover τ with nominal probability $1-\alpha$ and ψ degrees of freedom. The discussion is analogous to that for the mean, μ. For i.i.d. normal data, high statistical efficiency occurs with $m = 1$ and $c_{\psi,\alpha/2} = \chi^2_{n-1,\alpha/2}$ and $c_{\psi,1-\alpha/2} = \chi^2_{n-1,1-\alpha/2}$, the classical chi-squared quantiles. For other data, NB uses k large batches so that the batch means are approximately i.i.d. normal and therefore chi-squared quantiles with $k-1$ degrees of freedom are appropriate. OB procedures, in contrast, commonly use the approximation of chi-squared quantiles with increased degrees of freedom.

4 Overlapping t Distributions

We present the resulting tables for the t-distribution for OB in Section 4.1, and OSTS in Section 4.2. The method for calculating these tables is presented in the remainder of the paper.

4.1 The OB-χ^2 and OB-t Distributions

We now define and present the sampling distributions of the OB estimators $\widehat{V}_{\text{ob}}$ (which we call OB-χ^2) and T_{ob} (which we denote OB-t). The OB-χ^2 and OB-t distributions are analogous to the classical chi-squared (for $\widehat{V}_{\text{nb}}$) and Student's t (for T_{nb}) distributions. Quantiles from these OB distributions are appropriate choices for the c and s confidence-interval coefficients in Section 3.

There could be many OB-χ^2 and OB-t distributions. In particular, the distribution of $\widehat{\tau}_{\text{ob}}$ depends upon the data process. In keeping with the foundation of the classical distributions, in this section we invoke Data Assumption *i.i.d. normal*. In addition to the data process, $\widehat{\tau}_{\text{ob}}$ depends functionally upon both the run length n and the batch size m. These values are known to the practitioner, so the specific sampling distribution could be computed for that particular (n, m) value using a Monte Carlo approach.

Our goal, however, is to construct the OB analogies to the two classical families of distributions. Therefore, we define the OB distributions to correspond to the limit as m goes to infinity. In the definitions, the ν_{ob} degrees of freedom is a fixed distribution parameter. Therefore m going to infinity implies that the run length $n = m(\nu_{\text{ob}} + 1)$ goes to infinity. Also, the number of batches $n - m + 1 = m\nu_{\text{ob}} + 1$ goes to infinity. We next define the OB-χ^2 and OB-t distributions.

Definition OB-χ^2. *Under Data Assumption* i.i.d. normal*, the OB-χ^2 distribution with ν_{ob} degrees of freedom is the sampling distribution of $\widehat{V}_{\text{ob}}$ as $m \to \infty$ with $\nu_{\text{ob}} = (n - m)/m$ constant.*

Definition OB-t. *Under Data Assumption* i.i.d. normal*, the OB-t distribution with ν_{ob} degrees of freedom is the sampling distribution of T_{ob} as $m \to \infty$ with $\nu_{\text{ob}} = (n - m)/m$ constant.*

Table 1, analogous to the commonly encountered Student's t table, contains the OB-t distribution's quantiles. Each row corresponds to a particular value of ν_{ob}; each column corresponds to a quantile. We can construct a similar table for the quantiles of the OB-χ^2 distribution, but since these quantiles are not often used we omit it for brevity.

Both Student's t and the OB-t distributions have ranges on the real-number line and density

Table 1: Numerically computed quantile values of the OB-t distribution.

	0.75	0.9	0.95	0.975	0.98	0.99	0.995	0.9975	0.999	0.9995
df 1	0.753	1.572	2.175	2.784	2.982	3.604	4.233	4.869	5.719	6.367
2	0.739	1.513	2.062	2.603	2.777	3.317	3.859	4.404	5.128	5.678
3	0.723	1.456	1.960	2.450	2.606	3.089	3.572	4.055	4.696	5.184
4	0.713	1.419	1.893	2.347	2.491	2.933	3.371	3.809	4.388	4.828
5	0.706	1.394	1.848	2.277	2.412	2.825	3.232	3.635	4.168	4.570
6	0.701	1.376	1.816	2.227	2.356	2.747	3.130	3.509	4.006	4.380
7	0.698	1.363	1.792	2.190	2.314	2.689	3.055	3.414	3.883	4.236
8	0.695	1.353	1.774	2.162	2.282	2.645	2.997	3.341	3.788	4.123
9	0.693	1.345	1.760	2.140	2.257	2.610	2.950	3.282	3.712	4.033
10	0.691	1.339	1.748	2.122	2.237	2.582	2.913	3.235	3.651	3.960
11	0.689	1.334	1.739	2.107	2.220	2.559	2.883	3.196	3.600	3.900
12	0.688	1.329	1.731	2.095	2.206	2.539	2.857	3.164	3.558	3.849
13	0.687	1.326	1.725	2.084	2.194	2.523	2.835	3.136	3.521	3.806
14	0.686	1.322	1.719	2.075	2.184	2.508	2.816	3.112	3.490	3.769
15	0.685	1.320	1.714	2.068	2.175	2.496	2.800	3.092	3.463	3.737
16	0.685	1.317	1.709	2.061	2.168	2.485	2.786	3.074	3.440	3.708
17	0.684	1.315	1.706	2.055	2.161	2.476	2.773	3.058	3.419	3.684
18	0.684	1.313	1.702	2.050	2.155	2.468	2.762	3.044	3.400	3.661
19	0.683	1.312	1.699	2.045	2.150	2.460	2.752	3.031	3.384	3.642
20	0.683	1.310	1.696	2.040	2.145	2.453	2.743	3.020	3.369	3.624
21	0.682	1.309	1.694	2.037	2.140	2.447	2.735	3.010	3.356	3.608
22	0.682	1.307	1.692	2.033	2.136	2.442	2.728	3.000	3.343	3.593
23	0.682	1.306	1.690	2.030	2.133	2.436	2.721	2.992	3.332	3.580
24	0.681	1.305	1.688	2.027	2.129	2.432	2.715	2.984	3.322	3.567
25	0.681	1.304	1.686	2.024	2.126	2.428	2.709	2.977	3.312	3.556
26	0.681	1.303	1.684	2.022	2.124	2.424	2.704	2.970	3.304	3.546
27	0.681	1.303	1.683	2.019	2.121	2.420	2.699	2.964	3.296	3.536
28	0.680	1.302	1.682	2.017	2.118	2.417	2.695	2.958	3.288	3.527
29	0.680	1.301	1.680	2.015	2.116	2.413	2.691	2.953	3.281	3.519
30	0.680	1.301	1.679	2.013	2.114	2.411	2.687	2.948	3.275	3.511
40	0.679	1.296	1.670	2.000	2.099	2.389	2.659	2.912	3.228	3.455
50	0.678	1.293	1.665	1.992	2.090	2.377	2.642	2.891	3.200	3.422
60	0.677	1.291	1.662	1.987	2.084	2.368	2.631	2.877	3.182	3.400
70	0.677	1.290	1.659	1.983	2.079	2.362	2.623	2.867	3.169	3.384
80	0.677	1.289	1.658	1.980	2.076	2.358	2.617	2.859	3.159	3.372
90	0.676	1.288	1.656	1.978	2.074	2.354	2.612	2.854	3.151	3.363
100	0.676	1.287	1.655	1.976	2.072	2.351	2.609	2.849	3.145	3.356
200	0.675	1.284	1.650	1.968	2.063	2.339	2.592	2.828	3.117	3.323
1000	0.675	1.282	1.646	1.962	2.056	2.329	2.579	2.811	3.096	3.297
∞	0.674	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.090	3.291

functions that are symmetric about zero. Both have identical last rows, containing standard-normal quantiles. The Student's t and OB-t distributions differ in that OB-t has a smaller variance and kurtosis. When degrees of freedom are large, the OB-t quantiles for ν_{ob} degrees of freedom are similar to the Student's t quantiles for $\nu_{\text{nb}} = 3\nu_{\text{ob}}/2$ degrees of freedom; for example, $t^{\text{ob}}_{20,0.9995} = 3.624$ and $t^{\text{nb}}_{30,0.9995} = 3.646$. As with the Student's t distribution, when ν_{ob} is large OB-t quantiles can be approximated with the normal distribution. Substituting $\nu_{\text{nb}} = 3\nu_{\text{ob}}/2$ into the Student's t variance $\nu_{\text{nb}}/(\nu_{\text{nb}}-2)$ yields OB-t variance $3\nu_{\text{ob}}/(3\nu_{\text{ob}}-4)$. Therefore, $t^{\text{ob}}_{\nu_{\text{ob}},\alpha} \approx z_\alpha \sqrt{3\nu_{\text{ob}}/(3\nu_{\text{ob}}-4)}$.

4.2 The OSTS-t Distributions

We can similarly calculate the OSTS-χ^2 estimator using $\widehat{V}_{\text{osts}}$ and the OSTS-t estimator using T_{osts}. Table 2 displays our calculation of the quantiles of the OSTS-t distribution. As in the OB-t calculation, we assume m and the corresponding value of $n = m\nu_{\text{osts}}$ go to infinity, and the underlying data is i.i.d. normal. Again, we see that the limit as ν_{osts} approaches infinity is the standard normal quantile as expected. The OSTS-t distribution appears to have a slightly smaller variance than the OB-t distribution by way of smaller quantile values.

Table 2: Numerically computed quantile values of the OSTS-t distribution.

	0.75	0.9	0.95	0.975	0.98	0.99	0.995	0.9975	0.999	0.9995
df 2	0.730	1.481	2.005	2.516	2.680	3.186	3.693	4.202	4.876	5.390
3	0.709	1.402	1.861	2.295	2.431	2.847	3.255	3.659	4.187	4.585
4	0.699	1.367	1.799	2.199	2.324	2.700	3.065	3.423	3.887	4.234
5	0.693	1.348	1.764	2.146	2.263	2.617	2.957	3.288	3.714	4.031
6	0.690	1.336	1.742	2.111	2.225	2.564	2.888	3.201	3.602	3.899
7	0.688	1.327	1.727	2.088	2.198	2.527	2.839	3.140	3.524	3.806
8	0.686	1.321	1.716	2.071	2.179	2.500	2.804	3.095	3.466	3.738
9	0.684	1.316	1.707	2.057	2.164	2.479	2.777	3.061	3.422	3.685
10	0.683	1.313	1.701	2.047	2.152	2.463	2.756	3.034	3.387	3.644
11	0.683	1.310	1.695	2.039	2.143	2.450	2.738	3.013	3.359	3.610
12	0.682	1.307	1.691	2.032	2.135	2.439	2.724	2.995	3.335	3.582
13	0.681	1.305	1.687	2.026	2.128	2.430	2.712	2.980	3.316	3.559
14	0.681	1.303	1.684	2.021	2.123	2.422	2.702	2.967	3.299	3.539
15	0.680	1.302	1.681	2.017	2.118	2.415	2.693	2.956	3.284	3.522
16	0.680	1.300	1.679	2.013	2.114	2.410	2.686	2.946	3.272	3.507
17	0.680	1.299	1.677	2.010	2.110	2.405	2.679	2.938	3.261	3.494
18	0.679	1.298	1.675	2.007	2.107	2.400	2.673	2.930	3.251	3.482
19	0.679	1.297	1.673	2.004	2.104	2.396	2.668	2.923	3.242	3.472
20	0.679	1.297	1.672	2.002	2.101	2.392	2.663	2.917	3.234	3.462
21	0.679	1.296	1.671	2.000	2.099	2.389	2.659	2.912	3.227	3.454
22	0.678	1.295	1.669	1.998	2.097	2.386	2.655	2.907	3.221	3.446
23	0.678	1.295	1.668	1.996	2.095	2.384	2.651	2.903	3.215	3.439
24	0.678	1.294	1.667	1.995	2.093	2.381	2.648	2.898	3.210	3.433
25	0.678	1.293	1.666	1.993	2.091	2.379	2.645	2.895	3.205	3.427
26	0.678	1.293	1.665	1.992	2.090	2.377	2.642	2.891	3.200	3.422
27	0.678	1.293	1.665	1.991	2.089	2.375	2.640	2.888	3.196	3.417
28	0.678	1.292	1.664	1.990	2.087	2.373	2.637	2.885	3.192	3.412
29	0.677	1.292	1.663	1.989	2.086	2.371	2.635	2.882	3.189	3.408
30	0.677	1.291	1.663	1.988	2.085	2.370	2.633	2.880	3.185	3.404
40	0.677	1.289	1.658	1.981	2.077	2.359	2.619	2.861	3.161	3.375
50	0.676	1.287	1.655	1.976	2.072	2.352	2.610	2.850	3.147	3.358
60	0.676	1.286	1.654	1.974	2.069	2.348	2.604	2.843	3.137	3.347
70	0.676	1.286	1.652	1.972	2.067	2.345	2.600	2.838	3.131	3.339
80	0.676	1.285	1.651	1.970	2.065	2.342	2.597	2.834	3.125	3.333
90	0.675	1.285	1.651	1.969	2.064	2.341	2.595	2.831	3.122	3.328
100	0.675	1.284	1.650	1.968	2.063	2.339	2.593	2.829	3.118	3.324
200	0.675	1.283	1.647	1.964	2.058	2.333	2.584	2.818	3.104	3.307
1000	0.675	1.282	1.645	1.961	2.055	2.328	2.578	2.809	3.093	3.294
∞	0.674	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.090	3.291

5 The Existence of the OB and OSTS Distributions

The OB-t and OSTS-t tables in Section 4 are intended to be accurate representations of the asymptotic (as batch size m goes to infinity) distributions of T_{ob} and T_{osts} for general data processes. Other than numerical computing error, there are two error sources to explain why the OB tables might not be accurate. First, the batch size m might not be large enough for the sequences of random variables T_{ob} and T_{osts} to approach their limits (assuming that those limits exist). Second, the sample size n might not be large enough for the cumulative data process to be well represented by Brownian Motion (with implied i.i.d. normal observations).

Previous work demonstrates that the limits exist for these overlapping variance estimators and that therefore the OB-t and OSTS-t distributions exist. The original proofs assume an FCLT that allows for estimation of the mean, while recent work makes an SI assumption yielding estimation of any statistical functional. Additionally, there are separate results for OB and OSTS methods. Table 3 summarizes the key papers. Aktaran-Kalayci (2006) and Alexopoulos et al. (2007b) establish the limits of overlapping variance estimators for OB and OSTS respectively, under the FCLT case. More recently, Su et al. (2024) establishes the both OB and OSTS limits under the SI case, extending the mean estimators to include statistical functionals.

Table 3: Summary of convergence proofs for overlapping batch estimators.

Assumption	Estimator	OB	OSTS
FCLT	mean only	Aktaran-Kalayci (2006) Thm 10	Alexopoulos et al. (2007b) Thm 3
SI	statistical functionals	Su et al. (2024) Thm 5.1	Su et al. (2024) Thm 7.1 (claimed) This paper, Thm 5.4 (proof)

In this paper, we provide a unified computational method for computing the quantiles of the t-distributions associated with the above estimators for confidence interval calculations. This allows for direct implementation of the limits derived in past work without relying on approximation methods. We briefly summarize the limiting results in this section. For the OSTS case under the SI assumption, we provide a proof of Theorem 5.4 which was omitted in Su et al. (2024).

We assume the data process satisfies Data Assumption *Identical* but not Data Assumption *i.i.d. normal*. Notationally, let $D[0,1]$ denote the space of functions on $[0,1]$ that are right-continuous with left limits, $\Rightarrow_{n\to\infty}$ denote weak convergence, and $W(\cdot)$ denote standard Brownian Motion on

$[0, 1]$. First, we formally define the FCLT assumption, then present the key theorems using our notation.

Assumption FCLT The sequence of random functions

$$X_n(t) = \frac{\lfloor nt \rfloor \left(\sum_{j=1}^{\lfloor nt \rfloor} Y_j / \lfloor nt \rfloor - \mu \right)}{\sqrt{\tau n}}$$

for $t \in [0, 1]$ and $n = 1, 2, \ldots$ satisfies $X_n(\cdot) \Rightarrow_{n \to \infty} W(\cdot)$.

Theorem 5.1. (FCLT for OB - Adapted from Aktaran-Kalayci (2006)) For $\theta_n = \overline{Y}$, $n = (\nu_{\text{ob}} + 1)m = m/\beta$, and a fixed degree of freedom $\nu_{\text{ob}} \in \mathbb{Z}^+$, Assumption FCLT implies

$$\widehat{\tau}_{\text{ob}} \Rightarrow_{n \to \infty} \frac{\tau}{\beta(1-\beta)^2} \int_0^{1-\beta} [W(u+\beta) - W(u) - \beta W(1)]^2 du, \tag{5.1}$$

where $\beta = \lim_{m,n \to \infty} m/n \in (0, 1)$.

Theorem 5.2. (FCLT for OSTS - Adapted from Alexopoulos et al. (2007b)) For $\theta_n = \overline{Y}$, $n = \nu_{\text{osts}}\, m = m/\beta$, and a fixed degree of freedom $\nu_{\text{osts}} \in \mathbb{Z}^+ \setminus \{1\}$, Assumption FCLT implies

$$\widehat{\tau}_{\text{osts}} \Rightarrow_{n \to \infty} \frac{12\tau}{\beta^3 (1-\beta)} \int_0^{1-\beta} \left\{ \int_0^{\beta} W(u+v) - W(u) - \frac{v}{\beta} \left[W(u+\beta) - W(u) \right] dv \right\}^2 du, \tag{5.2}$$

where $\beta = \lim_{m,n \to \infty} m/n \in (0, 1)$.

Next, we formally define the SI assumption and present the two theorems for OB and OSTS.

Assumption of Strong Invariance The sequence $\{\hat{\theta}_n, n \geq 1\}$ of batching estimators satisfies the following strong invariance principle. There exists a standard Wiener process $\{W(t), t \geq 0\}$ and a stationary stochastic process $\{\tilde{Y}_n, n \geq 1\} \stackrel{\text{d}}{=} \{Y_n, n \geq 1\}$ defined on a common probability space such that as $n \to \infty$,

$$\left| \tau^{-1/2} \left(\hat{\theta}_{\lfloor n \rfloor} - \theta(P) \right) - n^{-1} W(n) \right| \leq \Gamma\, n^{-1/2-\delta} \sqrt{\log n} \quad \text{a.s.}, \tag{5.3}$$

where the constant $\delta > 0$ and the real-valued random variable Γ satisfies $\mathbb{E}[\Gamma] < \infty$.

This assumption is used in Su et al. (2024) to prove overlapping batch estimator results for statistical functions (under sectioning, batching, and standardized time series) and to construct the limiting t-distributions $\widetilde{T}_{\text{ob}}$ and $\widetilde{T}_{\text{osts}}$. We highlight the key theorem for the large batch regime for sectioning. Note that the authors define b_∞ as the limiting number of batches, which can be infinite or finite. The version using a finite number of batches will be useful our numerical approach.

Theorem 5.3. (SI for OB - Adapted from Su et al. (2024), Theorem 5.1) For statistical functional θ_n, $n = (\nu_{\text{ob}} + 1)m = m/\beta$, and a fixed degree of freedom $\nu_{\text{ob}} \in \mathbb{Z}^+$, Assumption SI implies

$$\hat{\tau}_{\text{ob}} \Rightarrow_{n\to\infty} \begin{cases} \frac{\tau}{\beta(1-\beta)^2} \int_0^{1-\beta} \left(W(u+\beta) - W(u) - \beta W(1)\right)^2 du, & \text{for } b_\infty = \infty; \\ \frac{\tau}{\beta(1-\beta)b_\infty} \sum_{j=1}^{b_\infty} \left(W(c_j+\beta) - W(c_j) - \beta W(1)\right)^2, & \text{for } b_\infty \in \mathbb{N} \setminus \{1\}, \end{cases} \tag{5.4}$$

where $c_j := (j-1)\frac{1-\beta}{b_\infty - 1}$ which converges to $(j-1)/n$ when $b_\infty \to n-m+1$ in the fully overlapping case. Furthermore,

$$\widetilde{T}_{\text{ob}} \Rightarrow_{n\to\infty} \frac{W(1)}{\sqrt{\hat{\tau}_{\text{ob}}/\tau}}. \tag{5.5}$$

We next present the statistical functional limiting result for the OSTS case, and provide the proof in the appendix.

Theorem 5.4. (SI for OSTS - Adapted from Su et al. (2024), Theorem 7.1) For statistical functional θ_n, $n = \nu_{\text{osts}} m = m/\beta$, and a fixed degree of freedom $\nu_{\text{osts}} \in \mathbb{Z}^+ \setminus \{1\}$, Assumption SI implies

$$\hat{\tau}_{\text{osts}} \Rightarrow_{n\to\infty} \begin{cases} \frac{12\tau}{\beta^3(1-\beta)} \int_0^{1-\beta} \left(\int_0^\beta (W(u+v) - W(u) - \frac{v}{\beta}(W(u+\beta) - W(u))dv\right)^2 du, & \text{for } b_\infty = \infty; \\ \frac{12\tau}{\beta^3 b_\infty} \sum_{j=1}^{b_\infty} \left(\int_0^\beta W(c_j+v) - W(c_j) - \frac{v}{\beta}(W(u+\beta) - W(u))dv\right)^2 du, & \text{for } b_\infty \in \mathbb{N} \setminus \{1\}, \end{cases} \tag{5.6}$$

where $c_j := (j-1)\frac{1-\beta}{b_\infty - 1}$ which converges to $(j-1)/n$ when $b_\infty \to n-m+1$ in the fully overlapping case. Furthermore,

$$\widetilde{T}_{\text{osts}} \Rightarrow_{n\to\infty} \frac{W(1)}{\sqrt{\hat{\tau}_{\text{osts}}/\tau}}. \tag{5.7}$$

Proof: See Appendix A.

6 Computing the OB and OSTS Tables

We now discuss the computation of the OB-t and OSTS-t quantiles the tables in Section 4. First, we compute the inverse of the OB-χ^2 cdf $F_{\text{ob}}^{\chi^2}(s;\nu_{\text{ob}}) \equiv \mathrm{P}(\widetilde{V}_{\text{ob}} \le s)$ which allows us to compute the inverse of the the OB-t cdf $F_{\text{ob}}^{t}(s;\nu_{\text{ob}}) \equiv \mathrm{P}(\widetilde{T}_{\text{ob}} \le s)$. Similar calculations will be performed for deriving the OSTS tables in Section 7.2.

The values are computed numerically, as discussed in Section 6.1. Then, in Section 6.2 we discuss approximating the tables with Monte Carlo experiments. In Section 6.3, we explain our choice of batch sizes m large enough to approximate (either numerically or with Monte Carlo) the quantiles.

We will need an additional result. The Continuous Mapping Theorem yields Corollary 6.1, showing that a limiting distribution analogous to the χ^2 exists.

Corollary 6.1. For positive batch size m and $n = (\nu_{\text{ob}} + 1)m$, Assumption FCLT implies

$$\begin{aligned}\widehat{V}_{\text{ob}} &= \frac{(\frac{n-m}{m})\widehat{\tau}_{\text{ob}}}{n\,\mathrm{var}(\overline{Y})} \\ &\Rightarrow_{m\to\infty} \frac{\nu_{\text{ob}}\hat{\tau}_{\text{ob}}}{\tau} \equiv \widetilde{V}_{\text{ob}}.\end{aligned} \tag{6.1}$$

Then, another application of the CMT implies that the distribution of OB-t exists, and the OB-t random variable can be also written as

$$\widetilde{T}_{\text{ob}} \equiv \frac{Z}{\sqrt{\widetilde{V}_{\text{ob}}/\nu_{\text{ob}}}}. \tag{6.2}$$

6.1 Our Numerical Method

We pursue a deterministic numerical method to compute the inverse of the OB-χ^2 cdf $F_{\text{ob}}^{\chi^2}(s;\nu_{\text{ob}})$ and the inverse of the OB-t cdf $F_{\text{ob}}^{t}(s;\nu_{\text{ob}})$. We describe our numerical method using five building blocks: (1) express $\widetilde{V}_{\text{ob}}$ as a linear combination of $n-m+1$ squared standard-normal random variables, (2) for the batch size m, compute $n-m+1$ eigenvalues as the weights of the linear combination, (3) for any quantile value x, use Laplace transforms to invert the linear combination to obtain $\mathrm{P}(\widetilde{V}_{\text{ob}} \ge x)$, (4) for the same quantile x, numerically integrate (across standard-normal realizations z) the product of conditional (on $Z = z$) OB-t probabilities and standard-normal probabilities to obtain $\mathrm{P}(\widetilde{T}_{\mathrm{ob}} \geq x)$, and (5) if a quantile is desired, use numerical root finding to search across x values to find the desired quantile.

6.1.1 Building Block 1: Linear Combination of Independent $\chi^2(1)$

The first building block is to write the OB variance estimator $\widetilde{V}_{\mathrm{ob}}$ as a linear combination of $n-m+1$ independent chi-squared random variables, each with one degree of freedom. This approach arose from ideas in Mathai and Provosts (1992). The decomposition into the linear combination is as follows. Define a discrete difference map $Q^{m,n} : W \in D[0,1] \mapsto Q_W^{m,n} \in D[0,(n-m)/n]$

$$Q_W^{m,n}(u) := W\left(\frac{\lceil un \rceil - 1 + m}{n}\right) - W\left(\frac{\lceil un \rceil - 1}{n}\right) - \frac{W(1)}{n/m}$$

for $u \in [0,(n-m)/n]$, and a continuous difference map $Q : W \in D[0,1] \mapsto Q_W \in D[0,1-\beta]$ where

$$Q_W(u) := W(u+\beta) - W(u) - \beta W(1), \quad u \in [0,1-\beta].$$

Recall from Corollary 6.1 that

$$\widetilde{V}_{\mathrm{ob}} \equiv \frac{\nu_{\mathrm{ob}}\,\hat{\tau}_{\mathrm{ob}}}{\tau}.$$

Let D denote the $n-m+1$ vector of Standard Brownian Motion differences with components

$$D_j = Q_W^{m,n}\left(\frac{j}{n}\right).$$

Then, from Theorem 5.3 $\widetilde{V}_{\mathrm{ob}}$ has a discrete approximation $cD^T D$, where $c = 1/\left[\beta(1-\beta)(n-m+1)\right]$. Let Σ denote the $(n-m+1) \times (n-m+1)$ matrix containing the covariances between D_i and D_j which will be calculated in Section 7.1. The vector D has a multivariate normal distribution with mean zero and covariance matrix Σ. Define $U = \Sigma^{-1/2} D$. Then,

$$\widetilde{V}_{\mathrm{ob}} = c\, U^T \Sigma U.$$

Decompose $\Sigma = P^T \Lambda P$ where $P^T P = P P^T = I$ and Λ is the diagonal matrix whose elements are the eigenvalues of Σ. Then

$$\widetilde{V}_{\mathrm{ob}} = c\, U^T P^T \Lambda P U.$$

Defining $Z = PU$ yields

$$\widetilde{V}_{\text{ob}} = c\, Z^T \Lambda Z = c \sum_{i=1}^{n-m+1} \lambda_i Z_i^2,$$

where the eigenvalues of Σ, λ_i, are nonnegative because the covariance matrix Σ is positive semi-definite. Because the data are centered on the sample mean, one eigenvalue is zero. The Z_i variables are standard normal, so each Z_i^2 is chi-squared with one degree of freedom; also the Z_i variables are independent since $\text{cov}(Z) = \text{cov}(P\Sigma^{-1/2}D) = P\Sigma^{-1/2}\text{cov}(D)\Sigma^{1/2}P^T = I$.

6.1.2 Building Block 2: Eigenvalues

The second building block is to compute the eigenvalues λ_i for $i = 1, 2, \ldots, n-m+1$. This block is straightforward; see, for example, Smith et al. (2013). It does, however, require order $O((n-m+1)^3)$ computing time and order $O((n-m+1)^2)$ storage. Fortunately, the eigenvalues are computed only once for every value of ν_{ob} to create the t-tables.

6.1.3 Building Block 3: Laplace Transforms

The third building block is to invert the convolution of $\lambda_i Z_i$ with a Laplace transform to obtain $\text{P}(\widetilde{V}_{\text{ob}} \geq x)$ for any constant $x \geq 0$. See, for example, Abate and Whitt (1995). The Laplace transform can be inverted relatively quickly to obtain the OB-χ^2 cdf $F_{\text{ob}}^{\chi^2}(x; \nu_{\text{ob}})$ for any specified set of x values.

6.1.4 Building Block 4: Numerical Quadrature

The fourth building block is to write the OB-t cdf in a tractable way for numerical integration across the standard-normal values. Let ϕ denote the standard-normal density function. For $t \geq 0$ the cdf, from Corollary 6.1, is

$$F_{\text{ob}}^t(t) = \text{P}\Big(\frac{Z}{\sqrt{\widetilde{V}_{\text{ob}}/\nu_{\text{ob}}}} \leq t\Big).$$

Condition on Z to obtain

$$F_{\text{ob}}^t(t; \nu_{\text{ob}}) = \int_{-\infty}^{\infty} \text{P}\Big(\frac{Z}{\sqrt{\widetilde{V}_{\text{ob}}/\nu_{\text{ob}}}} \leq t | Z = z\Big)\phi(z)dz.$$

Because the normal density function is symmetric around zero, we need to integrate for only $z > 0$, yielding

$$F^t_{\mathrm{ob}}(t;\nu_{\mathrm{ob}}) = 1/2 + \int_0^\infty \mathrm{P}\big(z \le t\sqrt{\widetilde{V}_{\mathrm{ob}}/\nu_{\mathrm{ob}}}\big)\phi(z)dz.$$

Rewrite the integrand to obtain

$$F^t_{\mathrm{ob}}(t;\nu_{\mathrm{ob}}) = 1/2 + \int_0^\infty \mathrm{P}\big(\widetilde{V}_{\mathrm{ob}} \ge \frac{\nu_{\mathrm{ob}} z^2}{t^2}\big)\phi(z)dz.$$

Substituting the OB-χ^2 cdf, which we know how to compute from Building Block 3, yields

$$F^t_{\mathrm{ob}}(t;\nu_{\mathrm{ob}}) = 1/2 + \int_0^\infty [1 - F^{\chi^2}_{\mathrm{ob}}(x;\nu_{\mathrm{ob}})\big(\frac{\nu_{\mathrm{ob}} z^2}{t^2};\nu_{\mathrm{ob}}\big)]\phi(z)dz.$$

This last integral is evaluated using numerical quadrature. See, for example, Press (2007).

6.1.5 Building Block 5: Root Finding

If quantiles are needed, as they are for the OB tables, the fifth building block is to use one-dimensional root finding to invert the cdfs. That is, find the pth OB-χ^2 quantile by solving for x in $F^{\chi^2}_{\mathrm{ob}}(x;\nu_{\mathrm{ob}}) = p$ and/or find the pth OB-t quantile solving for t in $F^t_{\mathrm{ob}}(t;\nu_{\mathrm{ob}}) = p$. The resulting quantile values in the lookup table can be called using the supporting code.

We have described the approach for calculating the OB-t tables. The same procedure can be used to calculate the OSTS-t tables by simply changing Σ to represent the covariance matrix of batches D_i formed from the standardized time series batching method, and adjusting ν_{osts} accordingly.

6.2 Monte Carlo Verification

Because Monte Carlo requires one-hundred times more computing to estimate to one more digit of precision, the precision required to create the four-digit t-tables in Section 4 would have required massive computing resources. Also, we wanted software that could quickly and accurately compute any t-quantile, not just those in the t-tables. However, t-quantile values can be estimated with a Monte Carlo experiment. For each degree of freedom ν_{ob}, the experiment generates many replications of $\widehat{\tau}_{\mathrm{ob}}$, and from those many replications of $\widehat{V}_{\mathrm{ob}}$ and T_{ob}. These values are placed into two fine-grid histograms, for example, one for OB-χ^2 and one for OB-t. When the sampling is complete, the histogram is counted left to right to estimate the requested quantiles. The logic for each Monte Carlo observation of $\widehat{\tau}_{\mathrm{ob}}$ is to generate n standard normal i.i.d. observations Y_j and compute $\widehat{\tau}_{\mathrm{ob}}$ using batches of size m.

The Monte Carlo software was verified by using $m = 1$ to estimate quantiles in the classical tables. In turn the Monte Carlo results were used to verify the more-precise and quicker-to-compute results from the numerical method in Section 6.1. Development and implementation of our approach in Section 6.1 was more involved than the Monte Carlo approach, but the numerical results are deterministic and accurate.

6.3 Choosing Batch and Sample Sizes

There has been some research on how to choose the batch size m, and this continues to be an ongoing topic of research. Song (1988) and Song and Schmeiser (1995) discuss the mean-squared error (mse) optimal batch size m. For several batching estimators, the quadratic-form graphs in Song and Schmeiser (1993) illustrate the bias-variance tradeoff as m changes. Yeh and Schmeiser (2004) showed that OB batch-size choice is more robust (as measured by the mse) than NB (and other estimators of τ). Pedrosa (1994) develops the 121-OBM algorithm for estimating mse-optimal batch size from a single vector of observations Y. Yeh (2002) discusses dynamic batch sizing with limited storage as more observations become available. Schmeiser (2004) contains additional discussion and references in that research thread.

Both Monte Carlo and our numerical method require that m and $n = m(\nu_{\mathrm{ob}} + 1)$ be finite integers. Because (for the OB tables) we are interested in the limit $m \to \infty$, for each desired degrees of freedom $\nu_{\mathrm{ob}} \in (0, \infty)$ the value of m needs to be chosen large enough to approximate asymptotic behavior and small enough to be computationally reasonable. After testing sensitivity for the four-digit precision of our OB tables, we chose to use $m = \lfloor 1000/\sqrt{\nu_{\mathrm{ob}}} \rfloor + 1$. Thus, in our table computations the largest batch size, $m = 1001$, was used for $\nu_{\mathrm{ob}} = 1$; similarly, the smallest batch size, $m = 32$, was used for $\nu_{\mathrm{ob}} = 1000$.

When ν_{ob} is very large, both Monte Carlo and our numerical method are computationally slow. Fortunately, the normal approximations from Section 4 can be used. A fast Matlab function for returning the OB-t and OSTS-t values from the tables in Section 4 available at `https://faculty.nps.edu/dsingham/OB_Code.zip`. To calculate these values, we employed parallel computing over

64 cores to distribute the calculation of Σ for large m across multiple processors, since this matrix can become quite large. Parallelization was also used to distribute calculations across multiple values of ν_{ob} and p. We conduct the same computations for OSTS, however numerical integration is needed in this case to calculate the covariance matrix, and we employ additional parallelization here.

7 Covariances of Overlapping Batch Means

The fundamental difference between NB and OB is that overlapping batches have observations in common, which causes positive dependence between batch means. The OB correlation structure is needed to analyze the behavior of $\widehat{\tau}_{\text{ob}}$, and the OSTS correlation structure is needed for $\widehat{\tau}_{\text{osts}}$. We will first perform the calculation of Σ for the OB estimator in Section 7.1, and then for the OSTS estimator in Section 7.2.

7.1 OB batch covariances

Here we derive covariances Σ of D_i for finite batch size m and $n = (\nu_{\text{ob}} + 1)m$. Batch indices i and j range from 1 to $n - m + 1$. Recall that we use $\beta = m/n$ to denote the batch-size proportion relative to the entire data length, so when rescaling the data to a standardized time series on $[0, 1]$, the batch length relative to that unit interval is β. For fixed m, n, we have

$$\begin{aligned}\operatorname{cov}(D_i, D_j) &= \operatorname{cov}\big(Q_W^{m,n}(i/n), Q_W^{m,n}(j/n)\big)\\ &= \operatorname{cov}\left(W\left(\frac{i-1}{n} + \beta\right) - W\left(\frac{i-1}{n}\right) - \beta W(1), W\left(\frac{j-1}{n} + \beta\right) - W\left(\frac{j-1}{n}\right) - \beta W(1)\right).\end{aligned} \tag{7.1}$$

As $m, n \to \infty$, the covariances can be estimated using the fact that the cumulative data process converges to Standard Brownian Motion, for which $\operatorname{cov}\big(W(s_1), W(s_2)\big) = \min\{s_1, s_2\}$, while $0 \leq s_1, s_2 \leq (n - m)/n$. For example, considering the first cross-term in (7.1) with $s_1 = (i - 1)/n, s_2 =$

$(j-1)/n$, and assuming $s_1 \leq s_2$, we have:

$$\begin{aligned}
&\operatorname{cov}\big(W(s_1+\beta)-W(s_1), W(s_2+\beta)-W(s_2)\big) \\
&\quad = \operatorname{cov}\big(W(s_1+\beta), W(s_2+\beta)\big) - \operatorname{cov}\big(W(s_1+\beta), W(s_2)\big) \\
&\quad\ - \operatorname{cov}\big(W(s_1), W(s_2+\beta)\big) + \operatorname{cov}\big(W(s_1), W(s_2)\big) \\
&\quad = \min\{s_1+\beta, s_2+\beta\} - \min\{s_1+\beta, s_2\} - \min\{s_1, s_2+\beta\} + \min\{s_1, s_2\} \qquad (7.2)\\
&\quad = \beta + 2\min\{s_1, s_2\} - \min\{s_1+\beta, s_2\} - \min\{s_1, s_2+\beta\} \\
&\quad = \beta + s_1 - \min\{s_1+\beta, s_2\} \\
&\quad = \max\{0, \beta - s_2 - s_1\}.
\end{aligned}$$

Completing similar calculations for the other cross-terms in (7.1) yields, as $m, n \to \infty$,

$$\operatorname{cov}(D_i, D_j) \to \beta \left(\max\left\{0, 1 - \frac{|i-j|}{m}\right\} - \beta \right). \tag{7.3}$$

These covariances are used in Building Block 1 of our numerical algorithm, as discussed in Section 6.1.1. As discussed in Section 2, a cumulative Brownian Motion data process implies i.i.d. normal observations; other than for scaling, Brownian Motion covariances hold true under Data Assumption *i.i.d. normal*, as used in the definitions of the OB sampling distributions.

7.2 OSTS batch covariances

Next, we calculate Σ for the OSTS case. Let $\mathcal{B}_s(t)$ denote the Brownian bridge starting from s and ending at $s+\beta$ over $s \leq t \leq s+\beta$, constructed by the standard Brownian motion on $[0,1]$ and $\beta = \lim_{m,n\to\infty} m/n$. We can write a discrete approximation of the OSTS variance estimator from Theorem 5.4 using the constant weight function as

$$\hat{\tau}_{\text{osts}} = \frac{12\tau}{\beta^3(1-\beta)} \int_0^{1-\beta} \left[\int_0^\beta \mathcal{B}_s(t)dt \right]^2 ds \tag{7.4}$$

$$\approx \frac{12\tau}{\beta^3(n-m+1)} \sum_{j=1}^{n-m+1} \left[\int_0^\beta \mathcal{B}_{j/n}(t)dt \right]^2. \tag{7.5}$$

Again, we approximate $\widetilde{V}_{\text{osts}} \equiv \nu_{\text{osts}} \hat{\tau}_{\text{osts}}/\tau$ by cD^TD where $c = 12/\left[\beta^3(n-m+1)\right]$ and $D_j =$

$[\int_0^\beta \mathcal{B}_{j/n}(t)dt]$, so it remains to derive the matrix Σ using covariances between batches $[\int_0^\beta \mathcal{B}_{s_1}(t)dt]$ and $[\int_0^\beta \mathcal{B}_{s_2}(t)dt]$ where $s_1 = (i-1)/n$ and $s_2 = (j-1)/n$ for all $i, j \in 1, \ldots, n-m+1$. Consider the following two Brownian bridges starting at s_1 and s_2,

$$\mathcal{B}_{s_1}(t_1) = W(s_1 + t_1) - W(s_1) - \frac{t_1}{\beta}[W(s_1+\beta) - W(s_1)]$$
$$\mathcal{B}_{s_2}(t_2) = W(s_2 + t_2) - W(s_2) - \frac{t_2}{\beta}[W(s_2+\beta) - W(s_2)],$$

where $0 \le s_1, s_2 \le 1-\beta$ and $0 \le t_1, t_2 \le \beta$. Therefore,

$$\mathrm{Cov}\left[\int_0^\beta \mathcal{B}_{s_1}(t_1)\,dt_1, \int_0^\beta \mathcal{B}_{s_2}(t_2)\,dt_2\right] = \int_0^\beta \int_0^\beta \mathrm{Cov}\left[\mathcal{B}_{s_1}(t_1), \mathcal{B}_{s_2}(t_2)\right] dt_1 dt_2 \tag{7.6}$$

by Fubini's theorem. Appendix B shows how Fubini's theorem can be be applied, and derives some closed-form representations of (7.6).

8 Numerical Results

To show the quality in performance gained by using the exact overlapping batch distributions, we employ coverage functions (Schruben 1980) to compare observed coverage to nominal coverage. Coverage functions deliver vastly more information than tables which arbitrarily choose a value of $1-\alpha$, for example 90% or 95% intervals, and report observed coverage at that value. This is because nominal coverage should be achieved for all $\alpha \in (0,1)$. Coverage functions are calculated by collecting, for each simulated replication, the observed coverage (the confidence coefficient needed to exactly cover the true value), and taking the empirical cumulative distribution function of these coverage values. For a confidence interval procedure to be valid, the coverage function should match the uniform distribution, so that the exact desired coverage level is achieved for any desired confidence coefficient. For our experiments, most coverage functions are generated using 10,000 independent replications to ensure sampling uncertainty has been smoothed away (note the effect of coverages estimate variability from low macro-replications is not always visible in standard coverage tables for singular α values). All of our experiments will construct confidence intervals for the mean to compare classical procedures using OB-t and OSTS-t distributions.

Next, we define "approximate" overlapping methods which exist in the past literature. Because overlapping t-distributions were not previously available, researchers obtained the benefits of overlapping batches by using an adjusted degrees of freedom applied to the Student's t distribution. First, we take two approximate methods tested in Yeh and Schmeiser (2015a) for the OB method. The first one is taken from Meketon and Schmeiser (1984) and sets the degrees of freedom

$$a_1 = 1.5\left(\frac{n}{m} - 1\right).$$

This corresponds to a 50% increase in degrees of freedom derived from belief that the variance of the overlapping method is approximately two-thirds that of the non-overlapping estimator. However, Yeh and Schmeiser (2015a) adjust this further to obtain a second approximate method with improved numerical performance using

$$a_2 = -0.4a_1^2 + 1.3a_1 + 2$$

when $0 < a_1 \leq 2$, otherwise

$$a_2 = a_1 + 4/a_1^2.$$

We refer to these approximate methods as OB-a_1 and OB-a_2. Next, we describe an approximate OSTS method from Alexopoulos et al. (2007a) which similarly adjusts degrees of freedom by carefully observing the limit of the overlapping batch means estimators using OSTS methods. For $b = 1/\beta = n/m$ and weight function $w(t) = \sqrt{12}$, we obtain an approximate method OSTS-a which employs degrees of freedom

$$a = \text{round}\left(\frac{2}{(24b-31)/(35(b-1)^2)}\right). \tag{8.1}$$

Note that for large β the value of a may be negative meaning the approximation cannot be used. We are now ready to present coverage functions comparing performance of these approximate methods to the derived OB and OSTS t-quantiles. First, we show coverage functions for the OB approach for *i.i.d. normal* data with mean 0 and variance 1 in Figure 1. We use a time series of $n = 1000$ data points with varying batch sizes to test the effect of overlapping. For $m = 100$, we see that all three OB methods appear to deliver nominal coverage, which is unsurprising given the nature of the data and lower levels of overlap in the batches. However, as m increases to 500 and 900, we see that using the exact OB-t distribution continues to deliver nominal coverage, while the approximate methods OB-a_1 and OB-a_2 exhibit overcoverage, with OB-a_1 showing extreme overcoverage.

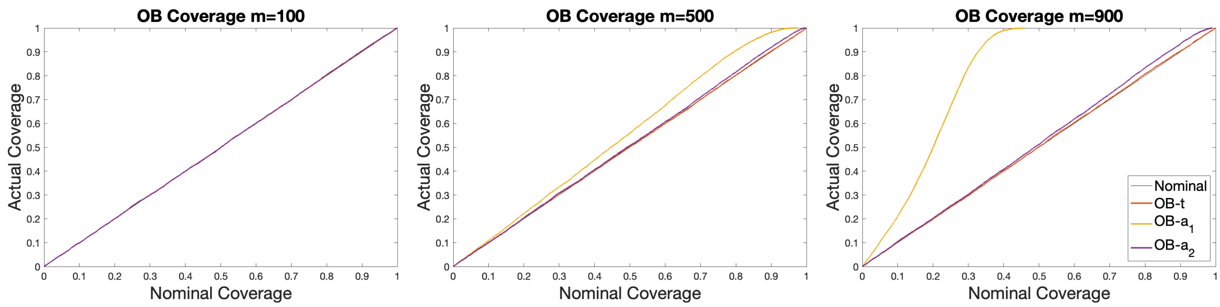

Figure 1: Coverage functions with actual coverage plotted against nominal coverage. Method: OB, Data: $\mathcal{N}(0,1)$, $n = 1000$, Replications: 10000.

Next, we plot the results comparing OSTS-t to the approximate degrees of freedom method, OSTS-a. Figure 2 displays results for *i.i.d. normal* data. We see that as m increases, coverage generally improves, with OSTS-t achieving better coverage than the approximate method for $m = 750$. This implies larger batches may perform better when using OSTS methods. Furthomore, the approximate method does not work for larger m (relative to n) because the degrees of freedom calculation in (8.1) becomes negative. Thus, we cannot employ OSTS-a when $m = 900$.

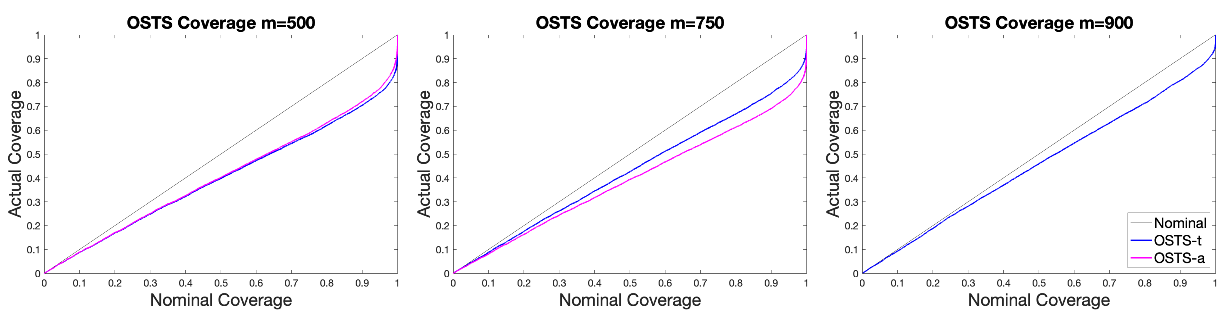

Figure 2: Coverage functions with actual coverage plotted against nominal coverage. Method: OSTS, Data: $\mathcal{N}(0,1)$, $n = 1000$, Replications: 10000.

Next we explore highly dependent data by simulated the waiting times in an M/M/1 queue with utilization $\rho = 0.9$. Figure 3 displays the results using an overall run length of $n = 50000$, with varying large overlapping batches of $m = 10000, 20000, 30000$. Smaller batch sizes than 10000 led to greater levels of undercoverage, and all three methods have almost identical coverage for values of m smaller than 10000. As the batch sizes increase, we see that the naive method OB-a_1 again exhibits overcoverage, while the OB-t method approaches nominal coverage from below. Interestingly, OB-a_2 gives the best results with coverage closest to nominal. We separately ran similar experiments with $n = 10000$ and found coverage behaved similarly, except coverage was lower than that in Figure 3 which used $n = 50000$. In the separate experiment with $n = 10000$, when using extremely large batch sizes ($\beta = 0.9$) coverage started to decrease again.

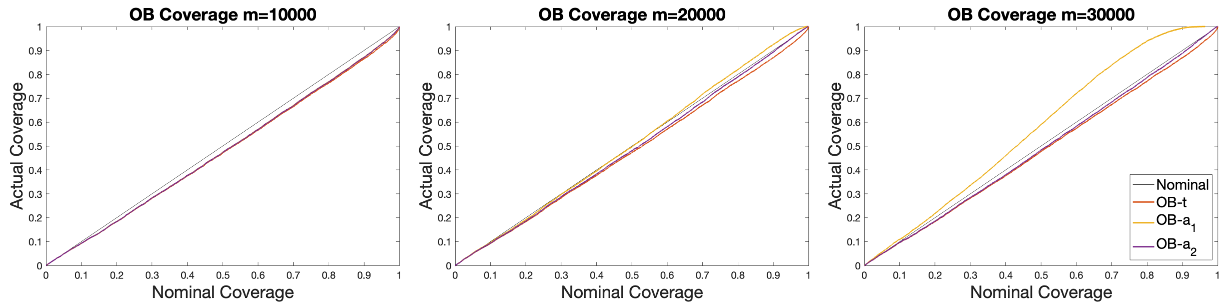

Figure 3: Coverage functions with actual coverage plotted against nominal coverage. Method: OB, Data: M/M/1, $\rho = 0.9$, $n = 50000$, Replications: 10000.

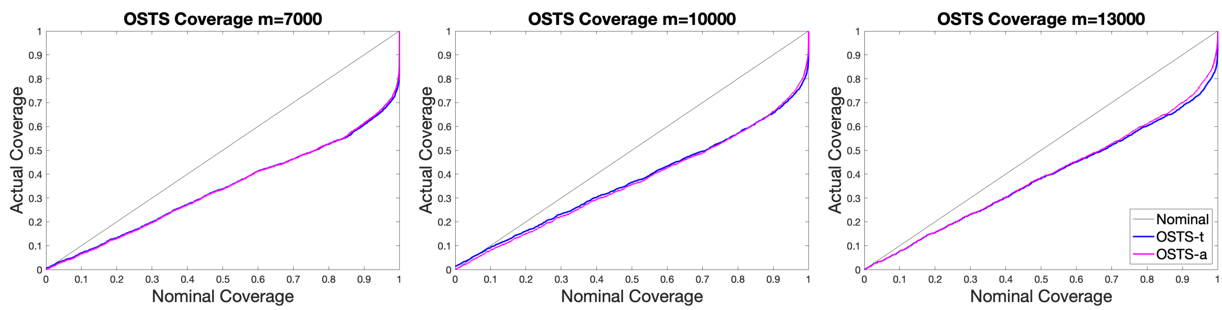

Figure 4: Coverage functions with actual coverage plotted against nominal coverage. Method: OSTS, Data: M/M/1, $\rho = 0.9$, $n = 20000$, Replications: 2000.

Finally, we present the results for M/M/1 data with utilization $\rho = 0.9$ using the OSTS distribution in Figure 4. Evaluating performance for the OSTS method requires higher computing resources than OB for few reasons. The variance estimators themselves require OSTS computations for each batch. Additionally, our calculation of Σ relies on numerical integration for each element, which becomes time consuming for large matrices (large m, n). Thus, we employ $n = 20000$ and only 2000 replications for each plot. Each plot takes 60 hours to generate on a supercomputer employing 64 cores for parallelizing both for the calculation of Σ and for the macro-replications. Despite these limitations compared to the other figures, we feel the coverage functions still present a much more complete picture than a simple coverage table for a single confidence coefficient. Figure 4 demonstrates coverage improvement due to increased overlapping in batching, but perhaps no real benefit of using OSTS-t versus OSTS-a. We again note that when m/n is too large, OSTS-a cannot be used because the proposed degrees of freedom is negative.

The "true" distributions of OB-t and OSTS-t are fundamentally characterized by two parameters: n and m, along with the underlying data process. However, it is difficult to provide a single table for simulation practitioners to use. In contrast, their asymptotic distributions depend only on the ratio m/n rather than the individual values of n and m. A finding from the coverage functions is that the distribution of OB-t converges significantly faster to its asymptotic distribution than OSTS-t. The interpretation is that both OB-t and OSTS-t require the underlying data process to converge to a Brownian motion. However, OSTS-t imposes an additional condition: the sum of the standardized time series must converge to the time integral of a Brownian bridge. This extra requirement makes the convergence of OSTS-t slower and more complex than that of OB-t.

9 Conclusions

We define the OB-t and OSTS-t distributions, analogous to the classical Student's t distributions used for NB statistical inference from stationary stochastic process data. We establish that these distributions exist, and closely approximate the distributions' quantile values with a numerical method that uses eigenvalues to reflect the covariances of the overlapping batch means. Access to these distribution's quantiles potentially simplifies and improves statistical inference based on the overlapping batch estimators of not just the sample mean, but general statistical functionals. We present a unified numerical framework for generating quantiles from overlapping t distributions along with code to enable fast confidence interval generation.

We compare the observed performance of the OB-t and OSTS-t distributions to prior approximate methods using coverage functions. We find that our calculated t-distributions generally perform favorably compared to naive inflationary methods for adjusting the degrees of freedom. While one approximation method, OB-a_2, appears to provide very good coverage, there is not a valid justification for its use, whereas OB-t and OSTS-t now are justified.

Armed with the OB-t and OSTS-t distributions, we can now further establish the quality of overlapping methods under new settings. Future work will test varying batch sizes and levels of overlap to determine guidelines for practitioners. Additionally, recent developments in multi-dimensional confidence or infinite-dimensional quantile fields show potential benefits from overlapping batches, and we plan rigorously develop overlapping methods in these areas.

Acknowledgments

This research is supported by Taiwan's Ministry of Science and Technology grant MOST 105-2410-H-008-36 to the second author. We thank two referees, an associate editor, and an area editor for detailed comments and suggestions on an earlier version of this paper. Finally, we acknowledge Bruce Schmeiser for many useful discussions that contributed to this work.

References

Abate J, Whitt W (1995) Numerical inversion of Laplace transforms of probability distributions. *ORSA Journal on Computing* 7(1):36–43.

Aktaran-Kalayci T (2006) *Steady-state analyses: Variance estimation in simulations and dynamic pricing in service systems*. Ph.D. thesis.

Alexopoulos C, Argon NT, Goldsman D, Steiger NM, Tokol G, Wilson JR (2007a) Efficient computation of overlapping variance estimators for simulation. *INFORMS Journal on Computing* 19(3):314–327.

Alexopoulos C, Argon NT, Goldsman D, Tokol G, Wilson JR (2007b) Overlapping variance estimators for simulation. *Operations Research* 55(6):1090–1103.

Alexopoulos C, Goldsman D, Mokashi A, Nie R, Sun Q, Tien KW, Wilson JR (2014) Sequest: A sequential procedure for estimating steady-state quantiles. *Proceedings of the Winter Simulation Conference 2014*, 662–673 (IEEE).

Alexopoulos C, Goldsman D, Serfozo RF (2007c) Stationary processes: Statistical estimation. N Balakrishnan CR, Vidakovic B, eds., *Encyclopedia of Statistical Sciences*, volume 12, 7991–8006 (Hoboken, New Jersey: John Wiley and Sons), 2nd edition.

Alexopoulos C, Goldsman D, Tang P, Wilson JR (2016) SPSTS: A sequential procedure for estimating the steady-state mean using standardized time series. *IIE Transactions* 48(9):864–880.

Bischak DP, Kelton WD, Pollock SM (1993) Weighted batch means for confidence intervals in steady-state simulations. *Management Science* 39(8):1002–1019.

Calvin JM, Nakayama MK (2013) Confidence intervals for quantiles with standardized time series. Pasupathy R, Kim SH, Tolk A, Hill R, Kuhl ME, eds., *Proceedings of the 2013 Winter Simulations Conference*, 601–612 (Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.).

Fox BL, Goldsman D, Swain JJ (1991) Spaced batch means. *Operations Research Letters* 10(5):255–263.

Law AM, Kelton WD (1984) Confidence intervals for steady-state simulations: I. A survey of fixed sample size procedures. *Operations Research* 32(6):1221–1239.

Mathai AM, Provosts SB (1992) *Quadratic forms in random variables* (Marcel Dekker, New York, New York).

Mechanic H (1966a) *Confidence intervals for averages of dependent data in simulations*, volume 1 (Yorktown Heights, NY: Advanced Systems Development Division, IBM Corporation).

Mechanic H (1966b) *Confidence intervals for averages of dependent data in simulations*, volume 2 (Yorktown Heights, NY: Advanced Systems Development Division, IBM Corporation).

Meketon MS, Schmeiser B (1984) Overlapping batch means: Something for nothing? S Sheppard UP, Pegden D, eds., *Proceedings of the 1984 Winter Simulation Conference*, 227–230 (Piscataway, New Jersey).

Nakayama MK (2014) Confidence intervals for quantiles using sectioning when applying variance-reduction techniques. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 24(4):1–21.

Pasupathy R, Singham D, Yeh Y (2022) Overlapping batch confidence regions on the steady-state quantile vector. Feng B, Pedrielli G, Peng Y, Shashaani S, Song E, Corlu C, Lee L, Lendermann P, eds., *Proceedings of the 2022 Winter Simulation Conference* (Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.).

Pasupathy R, Singham D, Yeh Y (2024) Overlapping batch confidence regions on the steady-state quantile field. *Under Review* .

Pedrosa AC (1994) *Automatic batching in simulation output analysis*. Ph.D. thesis, Purdue University.

Press WH (2007) *Numerical recipes 3rd edition: The art of scientific computing* (Cambridge University Press).

Schmeiser B (1982) Batch size effects in the analysis of simulation output. *Operations Research* 30(3):556–568.

Schmeiser B (2004) Simulation output analysis: A tutorial based on one research thread. RG Ingalls JS MD Rossetti, Peters B, eds., *Proceedings of the 2004 Winter Simulation Conference, 2004.*, 162–170 (Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.).

Schruben L (1980) A coverage function for interval estimators of simulation response. *Management Science* 26(1):18–27.

Schruben LW (1983) Confidence interval estimation using standardized time series. *Operations Research* 31(6):1090–1108.

Singham D (2014) Selecting stopping rules for confidence interval procedures. *ACM Transactions on Modeling and Computer Simulation* 24(3):Article 18.

Singham D, Schruben L (2012) Finite-sample performance of absolute precision stopping rules. *INFORMS Journal on Computing* 24(4):624–635.

Smith BT, Boyle JM, Garbow B, Ikebe Y, Klema V, Moler C (2013) *Matrix eigensystem routines-EISPACK guide*, volume 6 (Springer).

Song WT (1988) *On quadratic-form variance estimators of the sample mean in the analysis of simulation output*. Ph.D. thesis, Doctoral dissertation, Purdue University.

Song WT, Schmeiser BW (1993) Variance of the sample mean: Properties and graphs of quadratic-form estimators. *Operations Research* 41(3):501–517.

Song WT, Schmeiser BW (1994) Reporting the precision of simulation experiments. *New Directions in Simulation for Manufacturing and Communications* (Operations Research Society of Japan).

Song WT, Schmeiser BW (1995) Optimal mean-squared-error batch sizes. *Management Science* 41(1):110–123.

Song WT, Schmeiser BW (2009) Omitting meaningless digits in point estimates: The probability guarantee of leading-digit rules. *Operations Research* 57(1):109–117.

Song WT, Schmeiser BW (2011) Displaying statistical point estimates using the leading-digit rule. *IIE Transactions* 43(12):851–862.

Su Z, Pasupathy R, Yeh Y, Glynn P (2024) Overlapping batch confidence intervals on statistical functionals constructed from time series: Application to quantiles, optimization, and estimation. *ACM Transactions on Modeling and Computer Simulation* 34(2):1–43.

Welch PD (1987) On the relationship between batch means, overlapping means and spectral estimation. *Proceedings of the 19th Conference on Winter Simulation*, 320–323.

Yeh Y (2002) *Steady-state simulation output analysis: MSE-optimal dynamic batch means with parsimonious storage*. Ph.D. thesis, Purdue University.

Yeh Y, Schmeiser B (2015a) OBM confidence intervals: Something for nothing? Yilmaz L, Chan WKV, Moon I, Roeder TMK, Macal C, Rossetti MD, eds., *Proceedings of the 2015 Winter Simulation Conference*, 551–561 (Institute of Electrical and Electronics Engineers, Inc.).

Yeh Y, Schmeiser BW (2004) On the MSE robustness of batching estimators. *Operations Research Letters* 32(3):293–298.

Yeh Y, Schmeiser BW (2015b) VAMP1RE: a single criterion for rating and ranking confidence-interval procedures. *IIE Transactions* 47(11):1203–1216.

A Proof of Theorem 5.4

As in Su et al. (2024), we assume that m and n go to ∞ with β being the limit of m/n. Because this paper assumes fully overlapping batches, we set the number of batches is $n-m+1$ to simplify the algebra. But the proof extends to the partially overlapping (arbitrary number of batches) setting as in Su et al. (2024). From the definition of the OSTS estimator,

$$\widehat{\tau}_{\text{osts}} = \frac{1}{n-m+1} \sum_{i=1}^{n-m+1} A_i$$

where, assuming the weight function $w(t) = \sqrt{12}$,

$$A_{i,m} = \left[\frac{\sqrt{12}}{m} \sum_{j=1}^{m} T_{i,m} \left(\frac{j}{m} \right) \right]^2 .$$

The standardized time series for statistical functional estimators of θ is

$$T_{i,m}\left(t\right) = \frac{\lfloor mt \rfloor \left(\widehat{\theta}_{i,\lfloor mt \rfloor} - \widehat{\theta}_{i,m} \right)}{\sqrt{m}},$$

$i = 1, \ldots, n-m+1$ and $0 \le t \le 1$. The OSTS estimator can be rewritten in terms of $\widehat{\theta}_{i,j}$ as

$$\widehat{\tau}_{\text{osts}} = \frac{12}{n-m+1} \sum_{i=1}^{n-m+1} \left[\frac{1}{m} \sum_{j=1}^{m} \frac{j \left(\widehat{\theta}_{i,j} - \widehat{\theta}_{i,m} \right)}{\sqrt{m}} \right]^2 . \tag{A.1}$$

Let $\widetilde{B}_{i,j} = j^{-1} \left[W\left(i-1+j\right) - W\left(i-1\right) \right],\ i = 1, \ldots, n-m+1$, and $j = 1, \ldots, m$. Consider the terms inside the square of the OSTS estimator. Adding and subtracting terms yields

$$\begin{aligned}
&\frac{1}{m}\sum_{j=1}^{m}\left[\frac{j\left(\widehat{\theta}_{i,j}-\theta\left(P\right)+\theta\left(P\right)-\widehat{\theta}_{i,m}\right)}{\sqrt{m}}-\frac{j\sqrt{\tau}}{\sqrt{m}}\left(\widetilde{B}_{i,j}-\widetilde{B}_{i,m}\right)\right]+\frac{1}{m}\sum_{j=1}^{m}\frac{j\sqrt{\tau}}{\sqrt{m}}\left(\widetilde{B}_{i,j}-\widetilde{B}_{i,m}\right)\\
&=\frac{1}{m}\sum_{j=1}^{m}\left[\frac{j\left(\widehat{\theta}_{i,j}-\theta\left(P\right)\right)}{\sqrt{m}}-\frac{j\sqrt{\tau}\widetilde{B}_{i,j}}{\sqrt{m}}\right]+\frac{1}{m}\sum_{j=1}^{m}\left[\frac{j\sqrt{\tau}\widetilde{B}_{i,m}}{\sqrt{m}}-\frac{j\sqrt{\tau}\left(\widehat{\theta}_{i,m}-\theta\left(P\right)\right)}{\sqrt{\tau m}}\right] \qquad \text{(A.2)}\\
&+\frac{1}{m}\sum_{j=1}^{m}\frac{j\sqrt{\tau}}{\sqrt{m}}\left(\widetilde{B}_{i,j}-\widetilde{B}_{i,m}\right).
\end{aligned}$$

We will show that the first and second terms in (A.2) converge to zero almost surely and the third term converges to the area of a Brownian bridge process, i.e., the integral of a Brownian bridge process with respect to time. Except for a set of measure zero in the probability space implied by the stationary invariance assumption, we show that there exists $\Gamma(\omega)$ such that the first term goes to zero uniformly in j . From the SI asumptions, the first term is equivalent in distribution (using independent increments) as:

$$\begin{aligned}
\left|\frac{j\left(\widehat{\theta}_{i,j}-\theta\left(P\right)\right)}{\sqrt{\tau m}}-\frac{W\left(j\right)}{\sqrt{m}}\right| &\leq \Gamma\left(\omega\right)\left(\log j\right)^{1/2}\frac{j^{1/2-\delta}}{m^{1/2}} \leq \Gamma\left(\omega\right)\left(\log m\right)^{1/2}\frac{m^{1/2-\delta}}{m^{1/2}}\\
&= \Gamma\left(\omega\right)\left(\log m\right)^{1/2}\frac{1}{m^{\delta}}.
\end{aligned}$$

Note that when $\delta > 0$,

$$\lim_{x\to\infty}\frac{\left(\ln x\right)^{1/2}}{x^{\delta}}\lim_{x\to\infty}\left(\frac{\ln x}{x^{2\delta}}\right)^{1/2}=\left(\lim_{x\to\infty}\frac{\ln x}{x^{2\delta}}\right)^{1/2}=\left(\lim_{x\to\infty}\frac{1/x}{2\delta x^{2\delta-1}}\right)^{1/2}=\left(\lim_{x\to\infty}\frac{1}{2\delta x^{2\delta}}\right)^{1/2}=0,$$

so we have

$$\left|\frac{1}{m}\sum_{j=1}^{m}\left[\frac{j\left(\widehat{\theta}_{i,j}-\theta\left(P\right)\right)}{\sqrt{m}}-\frac{\sqrt{\tau}W\left(j\right)}{\sqrt{m}}\right]\right| \leq \sqrt{\tau}\Gamma\left(\omega\right)\left(\log m\right)^{1/2}\frac{1}{m^{\delta}}\to 0,$$

as $m\to\infty$. Similarly, we can write the second term in (A.2) using a distributional equivalent and employing the SI assumption to reach the following inequality,

$$\left|j\left(\frac{\widehat{\theta}_{i,m}-\theta\left(P\right)}{\sigma\sqrt{m}}-\frac{W\left(m\right)}{m\sqrt{m}}\right)\right| \leq \Gamma\left(\omega\right)\left(\log m\right)^{1/2} j\, m^{-1-\delta} \leq \Gamma\left(\omega\right)\left(\log m\right)^{1/2}\frac{1}{m^{\delta}}.$$

Then, we have

$$\left|\frac{1}{m}\sum_{j=1}^{m} j\left[\frac{\widehat{\theta}_{i,j}-\theta\left(P\right)}{\sqrt{m}}-\frac{\sqrt{\tau}W\left(m\right)}{m\sqrt{m}}\right]\right| \leq \sqrt{\tau}\Gamma\left(\omega\right)\left(\log m\right)^{1/2}\frac{1}{m^{\delta}} \to 0,$$

as $m \to \infty$. Hence, the first and second terms converge to zero almost surely. Next, define the lattice $\{0, \delta_m, 2\delta_m, \ldots, \lfloor 1/\delta_m \rfloor \delta_m\}$ with resolution $\delta_m := 1/m$ and corresponding projection operator $\lfloor t \rfloor_{\delta_m} := \max\{k\delta_m : t \geq k\delta_m, k \in \mathbb{Z}\}$, where $t \in [0, 1-\delta_m]$. The third term of (A.2) is

$$\begin{aligned}
&\frac{1}{m}\sum_{j=1}^{m}\frac{j\sqrt{\tau}}{\sqrt{m}}\left(\widetilde{B}_{i,j}-\widetilde{B}_{i,m}\right)\\
&=\frac{1}{m}\sum_{j=1}^{m}\frac{\sqrt{\tau}}{\sqrt{m}}\left\{W\left(i-1+j\right)-W(i-1)-\frac{j}{m}\left[W\left(i-1+m\right)-W\left(i-1\right)\right]\right\}\\
&\overset{d}{=}\frac{\sqrt{\tau}}{m}\sum_{j=1}^{m}\left\{W\left(\frac{i-1}{m}+\frac{j}{m}\right)-W\left(\frac{i-1}{m}\right)-\frac{j}{m}\left[W\left(\frac{i-1}{m}+1\right)-W\left(\frac{i-1}{m}\right)\right]\right\}\\
&=\frac{\sqrt{\tau}}{m}\int_{0}^{1-\delta_m}\left\{W\left(\frac{i-1}{m}+\lfloor t\rfloor_{\delta_m}\right)-W\left(\frac{i-1}{m}\right)-\lfloor t\rfloor_{\delta_m}\left[W\left(\frac{i-1}{m}+1\right)-W\left(\frac{i-1}{m}\right)\right]\right\}\frac{1}{\delta_m}\,dt\\
&=\sqrt{\tau}\int_{0}^{1-\delta_m}\left\{W\left(\frac{i-1}{m}+\lfloor t\rfloor_{\delta_m}\right)-W\left(\frac{i-1}{m}\right)-\lfloor t\rfloor_{\delta_m}\left[W\left(\frac{i-1}{m}+1\right)-W\left(\frac{i-1}{m}\right)\right]\right\}dt\\
&\to_{\delta_m\to 0}\sqrt{\tau}\int_{0}^{1}\left\{W\left(\frac{i-1}{m}+t\right)-W\left(\frac{i-1}{m}\right)-t\left[W\left(\frac{i-1}{m}+1\right)-W\left(\frac{i-1}{m}\right)\right]\right\}dt\\
&\overset{d}{=}\sqrt{\tau}\int_{0}^{\beta}\frac{1}{\sqrt{\beta^{3}}}\left\{W\left(\frac{i-1}{n}+v\right)-W\left(\frac{i-1}{n}\right)-\frac{v}{\beta}\left[W\left(\frac{i-1}{n}+\beta\right)-W\left(\frac{i-1}{n}\right)\right]\right\}dv,
\end{aligned}$$

because $\delta_m \to 0$ as $m \to \infty$.

Define another lattice $\{0, \delta_n, 2\delta_n, \ldots, \lfloor 1/\delta_n \rfloor \delta_n\}$ with resolution $\delta_n := 1/n$ and a corresponding projection operator $\lfloor u \rfloor_{\delta_n} := \max\{k\delta_n : u \geq k\delta_n, k \in \mathbb{Z}\}$, where $u \in [0, 1-\beta-\delta_n]$. Then, returning to our original estimator in (A.1), we have

$$
\begin{aligned}
\widehat{\tau}_{\text{osts}} =& \frac{12}{n-m+1} \sum_{i=1}^{n-m+1} \frac{\tau}{\beta^3} \left\{ \int_0^\beta W\left(\frac{i-1}{n}+v\right) - W\left(\frac{i-1}{n}\right) - \frac{v}{\beta}\left[W\left(\frac{i-1}{n}+\beta\right) - W\left(\frac{i-1}{n}\right)\right] dv \right\}^2 \\
&= \frac{12}{n-m+1} \frac{\tau}{\beta^3} \int_0^{1-\beta-\delta_n} \left\{ \int_0^\beta W\left(\lfloor u \rfloor_{\delta_n} + v\right) - W\left(\lfloor u \rfloor_{\delta_n}\right) - \frac{v}{\beta}\left[W\left(\lfloor u \rfloor_{\delta_n} + \beta\right) - W\left(\lfloor u \rfloor_{\delta_n}\right)\right] dv \right\}^2 \frac{1}{\delta_n} du \\
&= \frac{12n}{n-m+1} \frac{\tau}{\beta^3} \int_0^{1-\beta-\delta_n} \left\{ \int_0^\beta W\left(\lfloor u \rfloor_{\delta_n} + v\right) - W\left(\lfloor u \rfloor_{\delta_n}\right) - \frac{v}{\beta}\left[W\left(\lfloor u \rfloor_{\delta_n} + \beta\right) - W\left(\lfloor u \rfloor_{\delta_n}\right)\right] dv \right\}^2 du \\
&\to_{\delta_n \to 0} \frac{12\tau}{(1-\beta)\beta^3} \int_0^{1-\beta} \left\{ \int_0^\beta W(u+v) - W(u) - \frac{v}{\beta}\left[W(u+\beta) - W(u)\right] dv \right\}^2 du
\end{aligned}
$$

because $\delta_n \to 0$ as $m, n \to \infty$.

B Covariance of STS

First, we show Fubini's theorem can be applied in (7.6). We need to show that

$$
E\left[\left|\mathcal{B}_{s_1}(t_1)\,\mathcal{B}_{s_2}(t_2)\right|\right] \le E\left[\mathcal{B}_{s_1}^2(t_1)\right]^{1/2} E\left[\mathcal{B}_{s_2}^2(t_2)\right]^{1/2} = \sqrt{t_1\left(1-\frac{t_1}{\beta}\right) t_2 \left(1-\frac{t_2}{\beta}\right)} \tag{B.1}
$$

is uniformly bounded on $[0,1] \times [0,1]$. The right-hand equality of (B.1) is derived next. First, expand $\mathcal{B}_{s_1}^2(t_1)$:

$$
\begin{aligned}
\mathcal{B}_{s_1}^2(t_1) &= \left[W(s_1+t_1) - W(s_1)\right]^2 + \frac{t_1^2}{\beta^2}\left[W(s_1+\beta) - W(s_1)\right]^2 \\
&\quad - 2\frac{t_1}{\beta}\left[W(s_1+t_1) - W(s_1)\right]\left[W(s_1+\beta) - W(s_1)\right] \\
&= W^2(s_1+t_1) + W^2(s_1) - 2W(s_1+t_1)\,W(s_1) \\
&\quad + \frac{t_1^2}{\beta^2}\left[W^2(s_1+\beta) + W^2(s_1) - 2W(s_1+\beta)\,W(s_1)\right] \\
&\quad - 2\frac{t_1}{\beta}\left[W(s_1+t_1)\,W(s_1+\beta) - W(s_1+t_1)\,W(s_1) - W(s_1)\,W(s_1+\beta) + W^2(s_1)\right].
\end{aligned}
$$

Then, calculate the expectation as

$$
\begin{aligned}
E\left[\mathcal{B}_{s_1}^2(t_1)\right] &= s_1 + t_1 + s_1 - 2s_1 + \frac{t_1^2}{\beta^2}\left[s_1 + \beta + s_1 - 2s_1\right] - 2\frac{t_1}{\beta}\left[s_1 + t_1 - s_1 - s_1 + s_1\right] \\
&= t_1 - \frac{t_1^2}{\beta} = t_1\left(1 - \frac{t_1}{\beta}\right),
\end{aligned}
$$

which completes the derivation of (B.1). Next we return to (7.6) by expanding:

$$
\begin{aligned}
&\operatorname{Cov}\left[\mathcal{B}_{s_1}(t_1), \mathcal{B}_{s_2}(t_2)\right] \\
&= \operatorname{Cov}\left[W(s_1+t_1) - W(s_1) - \frac{t_1}{\beta}\left[W(s_1+\beta) - W(s_1)\right], W(s_2+t_2) - W(s_2) - \frac{t_2}{\beta}\left[W(s_2+\beta) - W(s_2)\right]\right].
\end{aligned}
$$

Expanding the cross-terms in the covariance yields $\mathcal{B}_{s_1}(t_1) \times \mathcal{B}_{s_2}(t_2)$

$$
\begin{aligned}
&= [W(s_1+t_1) - W(s_1)][W(s_2+t_2) - W(s_2)] - \frac{t_2}{\beta}[W(s_1+t_1) - W(s_1)][W(s_2+\beta) - W(s_2)] \\
&- \frac{t_1}{\beta}[W(s_1+\beta) - W(s_1)][W(s_2+t_2) - W(s_2)] + \frac{t_1 t_2}{\beta^2}[W(s_1+\beta) - W(s_1)][W(s_2+\beta) - W(s_2)] \\
&= W(s_1+t_1)W(s_2+t_2) - W(s_1+t_1)W(s_2) - W(s_1)W(s_2+t_2) + W(s_1)W(s_2) \\
&- \frac{t_2}{\beta}\left[W(s_1+t_1)W(s_2+\beta) - W(s_1+t_1)W(s_2) - W(s_1)W(s_2+\beta) + W(s_1)W(s_2)\right] \\
&- \frac{t_1}{\beta}\left[W(s_1+\beta)W(s_2+t_2) - W(s_1+\beta)W(s_2) - W(s_1)W(s_2+t_2) + W(s_1)W(s_2)\right] \\
&+ \frac{t_1 t_2}{\beta^2}\left[W(s_1+\beta)W(s_2+\beta) - W(s_1+\beta)W(s_2) - W(s_1)W(s_2+\beta) + W(s_1)W(s_2)\right].
\end{aligned}
$$

Using the fact that the expectation of a Brownian bridge is zero yields

$$
\begin{aligned}
&\operatorname{Cov}\left[\mathcal{B}_{s_1}(t_1), \mathcal{B}_{s_2}(t_2)\right] = \min\{s_1+t_1, s_2+t_2\} - \min\{s_1+t_1, s_2\} - \min\{s_1, s_2+t_2\} + \min\{s_1, s_2\} \\
&- \frac{t_2}{\beta}\left[\min\{s_1+t_1, s_2+\beta\} - \min\{s_1+t_1, s_2\} - \min\{s_1, s_2+\beta\} + \min\{s_1, s_2\}\right] \\
&- \frac{t_1}{\beta}\left[\min\{s_1+\beta, s_2+t_2\} - \min\{s_1+\beta, s_2\} - \min\{s_1, s_2+t_2\} + \min\{s_1, s_2\}\right] \\
&+ \frac{t_1 t_2}{\beta^2}\left[\min\{s_1+\beta, s_2+\beta\} - \min\{s_1+\beta, s_2\} - \min\{s_1, s_2+\beta\} + \min\{s_1, s_2\}\right].
\end{aligned}
$$

This expression is used to construct Σ for generating the OSTS-t distribution.