Monte Carlo sampling bias in the microwave uncertainty framework*

Michael Frey¹, Benjamin F Jamroz², Amanda Koepke¹, Jacob D Rezac² and Dylan Williams²

¹ Statistical Engineering Division, National Institute of Standards and Technology, Boulder, CO 80305, United States of America
² Radio Frequency Division, National Institute of Standards and Technology, Boulder, CO 80305, United States of America

E-mail: michael.frey@nist.gov

Received 14 November 2018, revised 20 June 2019
Accepted for publication 24 June 2019
Published 6 August 2019

Abstract
Uncertainty propagation software can have unknown, inadvertent biases introduced by various means. This work treats bias identification and reduction in one such software package, the microwave uncertainty framework (MUF). The MUF provides automated multivariate statistical uncertainty propagation and analysis on a Monte Carlo (MC) basis. Combine is a key module in the MUF, responsible for merging data, raw or transformed, to accurately reflect the variability in the data and in its central tendency. In this work the performance of Combine’s MC replicates is analytically compared against its stated design goals. An alternative construction is proposed for Combine’s MC replicates and its performance is compared, too, against Combine’s design goals. These comparisons reveal that Combine’s MC uncertainty results with the current construction method are biased except under restrictive conditions. The bias with the proposed alternative construction, by contrast, is, without restriction, asymptotically zero (in the large MC sample size limit), and this construction is recommended.

Keywords: Monte Carlo sampling, sampling bias, uncertainty propagation, systematic error, statistical software

(Some figures may appear in colour only in the online journal)
system elements. Model modules are useful, for example, for building calibration models, and they can be cascaded to represent increasingly complex systems. Other processing modules, termed Transform modules, are available to perform oscilloscope and receiver calibrations, Fourier transforms, and other user-defined custom analytical transformations. The MUF’s user-selectable modules are designed to be used at any point in extended analyses, even allowing intermediate results to be used by different users for different purposes. This modular flexibility is a powerful feature of the MUF.

Combine is a key module in the MUF, responsible for merging measurement data, raw or transformed, to accurately reflect the data’s central tendency and corresponding variability. Combine is shown schematically with its inputs and outputs in figure 1. Each of Combine’s J inputs includes a vector nominal value $\vec{N}_j^{(t)}$ of a (possibly transformed or otherwise processed) measurement. Also at each input is an associated sample of vector MC replicates $\vec{M}_{jq}$, $q = 1, \ldots, Q$, describing the measurement’s distribution. Combine merges these inputs, producing a summary nominal value $\vec{N}^{(c)}$ with attendant MC replicates $\vec{M}^{(c)}$ as follows:

$$\vec{N}^{(c)} = \vec{N}^{(t)}$$  \hspace{1cm} (1)

and

$$\vec{M}^{(c)} = \vec{M}^{(t)} + \frac{1}{J} \vec{U}_c \sqrt{\vec{D}_c} \vec{Z}_q,$$  \hspace{1cm} (2)

where the $\vec{Z}_q$ are independent standard normal variates, $\vec{Z}_q \sim N(\vec{0}, \vec{I})$, $q = 1, \ldots, Q$. Further, the $\vec{Z}_q$ are independent of the matrices $\vec{M}^{(t)}_j$ and $\vec{U}_c \sqrt{\vec{D}_c}$, where $\vec{U}_c$ and $\vec{D}_c$ are the unitary and diagonal members, respectively, of the eigendecomposition $\vec{\Sigma}^{(n)} = \vec{U}_c \vec{D}_c \vec{U}_c^T$ of the sample covariance matrix

$$\vec{\Sigma}^{(n)} = \frac{1}{J-1} \sum_{j=1}^{J} \left( \vec{N}^{(n)}_j - \vec{N}^{(t)} \right) \left( \vec{N}^{(n)}_j - \vec{N}^{(t)} \right)^T$$  \hspace{1cm} (3)

associated with the J nominal vectors $\vec{N}^{(n)}_j$ at Combine’s input. Here and throughout we use boldface symbols for matrices, overarrows for vectors, and the overbar-dot convention for averaging (see appendix). The vectors $\vec{Z}_q$ in (2) model standard normal variation along the principal axes of (3). These standard normal variates are scaled by the standard deviations in $\sqrt{\vec{D}_c}$ and then rotated by $\vec{U}_c$ onto the coordinate axes of (3) to reflect the variability associated with the averaging performed by Combine. The components of the $\vec{Z}_q$ are chosen to be normally distributed based on the assumption that the number $J$ of Combine inputs and their independence support a Central Limit Theorem approximation.

This paper presents an analysis of the Combine module in the MUF. Because of the MUF’s distributed, multi-user, multi-purpose nature, Combine can be executed at different stages of uncertainty propagation analysis. We study a generic use of Combine described by the two-stage scenario diagrammed in figure 1. In the first stage of this scenario, multivariate data are joined with shared systematic error in a bank of Transform modules and then, in the second stage, the transformed data are combined within the Combine module. Transform’s outputs take the form of nominal values of a user-selected mathematical transformation with associated uncertainties. Transform’s outputs in this form allow us in our two-stage scenario to study and assess Combine’s operation in which its inputs with their associated uncertainties are used to produce a summary mean output with an associated uncertainty. Our analysis is based on a mathematical statement of the goals for Combine in this two-stage scenario and focuses specifically on the bias in the mean and covariance of Combine’s MC replicates $\vec{M}^{(c)}_j$. This analysis reveals that Combine’s construction of MC replicates is fundamentally biased, and we propose an alternative construction that effectively eliminates this bias.

The remainder of the paper is organized as follows. In section 2 we fully detail the generic two-stage scenario diagrammed in figure 1 and mathematically state Combine’s design goals for that scenario. We analyze Combine’s performance with respect to these design goals, showing that the sample mean of Combine’s MC replicates has zero bias and giving an analytical expression for the bias in the MC replicates’ covariance. This covariance bias is studied for specific systematic error mechanisms in the Transform stage in figure 1. In section 3 we propose an alternative construction for Combine’s MC replicates and show that the sample mean of Combine’s MC replicates has zero bias. We put tight bounds on the corresponding covariance bias and show that this bias is asymptotically zero; in this latter regard the proposed construction is better than the current method. Estimation bias is the primary concern in MC sampling, but MC estimation variability is also an issue. In section 4 we continue our comparison of the current and alternative MC replicate constructions, comparing the variabilities in their sample means and sample covariances. We conclude in section 5 with summary remarks supporting adoption of the proposed alternative MC replicate construction method in place of Combine’s current method.
2. Bias in the Combine module

We suppose in the two-stage scenario in figure 1 that the \( J > 1 \) data vectors \( \tilde{Y}_j \) (of length \( K \)) are identically distributed and mutually independent and write \( \tilde{Y}_j \sim (\tilde{\mu}, \Sigma) \) to identify the mean \( \tilde{\mu} \) and covariance matrix \( \Sigma \) of \( \tilde{Y}_j \). We also suppose that the MC-generated, length-\( K \) errors \( \tilde{S}_q \), \( q = 1, \ldots, Q \) are identically distributed, mutually independent, and independent of the sample of data vectors \( \tilde{Y}_j \). The mean and covariance of \( \tilde{S}_q \) are \( \tilde{\mu}_q \sim (\bar{\mu}, \Upsilon) \). The data covariance \( \Sigma \) represents random error in the measurement of \( \tilde{\mu} \) while the errors \( \tilde{S}_q \) are systematic post-measurement errors introduced among the \( \tilde{Y}_j \) due, for example, to calibration adjustments.

Each data vector \( \tilde{Y}_j \) in figure 1 is operated on individually by Transform, producing for \( \tilde{Y}_j \) a nominal value

\[
\tilde{N}_j^{(n)} = F(\tilde{Y}_j, \tilde{\mu}) \tag{4}
\]

for \( j = 1, \ldots, J \) and a sample of vector Monte Carlo (MC) replicates

\[
\tilde{M}_q^{(c)} = F(\tilde{Y}_j, \tilde{S}_q) \tag{5}
\]

for \( q = 1, \ldots, Q \). The superscripts in (4) and (5) signify that these are Transform outputs in the first stage in our scenario. The MC replicates in (5) vary for a given \( \tilde{Y}_j \) only according to random replicates from the distribution of \( \tilde{S}_q \). The same \( Q \) random replicates \( \tilde{S}_q \) are used to create each \( \tilde{Y}_j \)’s sample of MC replicates. This models systematic errors shared among the \( \tilde{Y}_j \). The Transform function \( F \) in (4) and (5) has the general form

\[
F(\vec{y}, \vec{s}) = \begin{pmatrix}
    f(y_1, s_1) \\
    \vdots \\
    f(y_K, s_K)
\end{pmatrix}
\tag{6}
\]

where \( y_k \) and \( s_k \) are the \( k \)-th components of the vectors \( \vec{y} \) and \( \vec{s} \), respectively. The scalar-valued function \( f \) is a user-specified parameter in Transform. Some choices of \( f \) are \( f(y, s) = y + s \), \( f(y, s) = ys \), \( f(y, s) = \sin(y + s) \), and \( f(y, s) = y^2 \), representing additive, multiplicative, phase, and exponential error, respectively.

Combine in figure 1’s two-stage scenario produces the nominal summary value \( \tilde{N}^{(c)} \) in (1) with the attendant MC replicates in (2). For the Transform outputs given by (4) and (5), these Combine outputs become

\[
\tilde{N}^{(c)} = F(\tilde{Y}_*, \tilde{\mu}) \tag{7}
\]

and

\[
\tilde{M}_q^{(c)} = M_q^{(n)} + \frac{1}{\sqrt{J}} \mathbf{U}_c \sqrt{\mathbf{D}_c} \tilde{Z}_q, \tag{8}
\]

where \( M_q^{(n)} = F(\tilde{Y}_*, \tilde{S}_q) \). The nominal value \( \tilde{N}^{(c)} \) in (7) produced by Combine is a natural, intuitive summary of the center of the transformed data. The purpose of the MC replicates \( \tilde{M}_q^{(c)} \) is to capture the sampling distribution of the MC estimator \( \tilde{M}_q^{(c)} \) of the center. Formally, Combine is required to produce a sample of MC replicates whose mean \( \tilde{M}_q^{(c)} \) is an unbiased estimator of the vector \( E[F(\tilde{Y}_*, \tilde{S}_q)] \) and whose covariance

\[
\Sigma_{\tilde{M}_q^{(c)}} = \frac{1}{Q - 1} \sum_{q=1}^{Q} (\tilde{M}_q^{(c)} - \tilde{M}_q^{(c)}) (\tilde{M}_q^{(c)} - \tilde{M}_q^{(c)})^T \tag{9}
\]

is an unbiased estimator of the covariance \( V[F(\tilde{Y}_*, \tilde{S}_q)] \) of the vector \( F(\tilde{Y}_*, \tilde{S}_q) \). In other words, the MC replicates in (8) should satisfy

\[
E[\tilde{M}_q^{(c)}] = E[F(\tilde{Y}_*, \tilde{S}_q)] \tag{10}
\]

and

\[
E[\Sigma_{\tilde{M}_q^{(c)}}] = V[F(\tilde{Y}_*, \tilde{S}_q)] \tag{11}
\]

We note for later use that under the conditions of our two-stage scenario the estimands in (10) and (11) can be expressed as

\[
E[F(\tilde{Y}_*, \tilde{S}_q)] = E[F(\tilde{Y}_*, \tilde{S}_q)] \tag{12}
\]

and

\[
V[F(\tilde{Y}_*, \tilde{S}_q)] = \frac{1}{J} V[F(\tilde{Y}_*, \tilde{S}_q)] + \frac{J - 1}{J} \text{Cov}[F(\tilde{Y}_*, \tilde{S}_q), F(\tilde{Y}_*, \tilde{S}_q)] \tag{13}
\]

with \( j \neq j' \), where the matrix \( \text{Cov}[	ilde{X}_j, \tilde{X}_j'] = E[(\tilde{X}_j - E[\tilde{X}_j])(\tilde{X}_j - E[\tilde{X}_j])^T] \) is the cross-covariance of vectors \( \tilde{X}_j \) and \( \tilde{X}_j' \).

Combine’s second design goal (11) specifies that Combine’s MC replicates \( \tilde{M}_q^{(c)} \) provide an unbiased estimate of the \( K \times K \) covariance \( V[F(\tilde{Y}_*, \tilde{S}_q)] \). This requirement is essential for state-of-the-art microwave measurement, which relies on high-speed instrumentation including vector network analyzers (VNAs) operating in the frequency domain, temporal sampling oscilloscopes, and an array of other instruments, often used simultaneously in the same experiment. The refined measurements made possible by these arrangements allow investigators to, for example, identify the multiple reflections created by small imperfections in microwave systems, capture distortions due to the systems’ frequency-limited electronics, and study the role of noise. Statistical analysis of the data from this mix of instrumentation, including the conduct of uncertainty analyses, often involves shifts between the time and frequency domains. These shifts require that microwave uncertainty analyses account, particularly, for statistical correlations among measurements [4, 5]. To see this, consider that imperfections in microwave systems are often the source of unwanted reflections and attendant power losses. These temporal effects Fourier-map into the frequency domain as ripples with a characteristic period related to the inverse of the reflections’ time spacing. The VNA is currently the most accurate instrument for measuring these multiple reflections, and the errors made by this frequency sampling instrument manifest themselves as correlated time domain errors in the magnitudes, shapes, and positions of the multiple reflections. Statistical uncertainties in VNA measurements cannot be transformed correctly into the time domain.
without an unbiased accounting of correlations created by the domain transformation [6].

Combine’s construction of its MC replicates is described in (8) in terms of the two-stage scenario in figure 1. We can analytically test these MC replicates against Combine’s design goals (10) and (11). Our analysis, summarized in proposition 1 below, of the MC construction in (8) shows that Combine meets its goals (10) and (11) only under certain conditions, and that without these conditions Combine exhibits bias.

**Proposition 1.** Suppose that, in the two-stage scenario in figure 1, we have \( J > 1 \) independent, identically distributed data vectors \( \vec{Y}_j \sim (\mu, \Sigma) \). Also suppose we have \( Q \) independent, identically distributed errors \( \vec{S}_q \sim (\nu, \Upsilon) \). Assume the sets of \( \vec{Y}_j \) and \( \vec{S}_q \) are independent. Suppose further that the Transform outputs \( \vec{N}_q^{(n)} = F(\vec{Y}_j, \nu) \) and \( \vec{M}_q^{(n)} \) are given by (4) and (5) with \( F \) as in (6), and the Combine outputs \( \vec{N}^{(c)}_q \) and \( \vec{M}^{(c)}_q \) are given by (7) and (8). Then

\[
E[\vec{M}^{(c)}_q] = E[\vec{F}(\vec{Y}_j, \vec{S}_q)]
\]

and

\[
E[\vec{\Sigma}^{(c)}_q] = V[\vec{F}(\vec{Y}_j, \vec{S}_q)] + \frac{1}{J} \Psi
\]

where \( \Psi = V[\vec{F}(\vec{Y}_j, \nu)] - V[E[\vec{F}(\vec{Y}_j, \vec{S}_q)]|\vec{Y}_j] \).

Proposition 1 establishes that design goal (10) is generally met by the MC replicates in (8), but that design goal (11) is not. The covariance in the sample of MC replicates is biased by an amount \( \Psi/J \). We will see in the next section that this bias can be positive or negative. We first prove the two parts (14) and (15) of the proposition.

**Proof of (14).** We first note that \( E[\vec{M}^{(c)}_q] = E[\vec{M}^{(c)}_q] \) since the \( \vec{M}_q^{(c)} \) are identically distributed. Then, conditioning on the factor \( U_c \sqrt{D_c} \) in (8) and using that \( \vec{Z}_q \) and \( U_c \sqrt{D_c} \) are independent, we have

\[
E[\vec{M}^{(c)}_q] = E[E[\vec{M}^{(c)}_q] + \frac{1}{\sqrt{J}} U_c \sqrt{D_c} \vec{Z}_q | U_c \sqrt{D_c}]
\]

\[
= E[E[\vec{M}^{(c)}_q] | U_c \sqrt{D_c}] + \frac{1}{\sqrt{J}} E[U_c \sqrt{D_c} | E[\vec{Z}_q]].
\]

Since \( E[\vec{Z}_q] = 0 \), this yields \( E[\vec{M}^{(c)}_q] = E[\vec{M}^{(c)}_q] \), which proves (14).

To prove (15) in the proposition, we need four lemmas, which we state here. Their proofs are given in the appendix. Lemma 1 concerns the sample covariance of cross-correlated vectors. Lemmas 2 and 3 are elementary conditioning argument-based results for auto- and cross-covariances. Lemma 4 is used here and in the proofs of subsequent propositions.

**Lemma 1.** Let \( \vec{X}_j \sim (\vec{\mu}, \Sigma) \) for \( j = 1, \ldots, J > 1 \) with sample covariance

\[
\Sigma = \frac{1}{J-1} \sum_{j=1}^{J} (\vec{X}_j - \vec{X})(\vec{X}_j - \vec{X})^T.
\]

Suppose the \( \vec{X}_j \) are cross-correlated with \( \text{Cov}[\vec{X}_j, \vec{X}_k] = \Sigma' \) for all \( j \neq k \). Then \( E[\Sigma] = \Sigma - \Sigma' \).

**Lemma 2.** Let \( \vec{Z} \sim (\vec{0}, I) \) be independent of the vector-matrix pair \((\vec{A}, \vec{B})\). Then \( V[\vec{A} + \vec{B}\vec{Z}] = V[\vec{A}] + E[\vec{B}\vec{B}^T] \).

**Lemma 3.** Let \( \vec{Z}_1, \vec{Z}_2 \sim (\vec{0}, I) \), and suppose \( \vec{Z}_1, \vec{Z}_2 \) and \((\vec{A}_1, \vec{A}_2, \vec{B})\) are mutually independent. Then \( \text{Cov}[\vec{A}_1 + \vec{B}\vec{Z}_1, \vec{A}_2 + \vec{B}\vec{Z}_2] = \text{Cov}[\vec{A}_1, \vec{A}_2] \).

**Lemma 4.** Let \( \vec{S}, \vec{S}' \) be independent, identically distributed random vectors independent of the random vector \( \vec{Y} \). Let \( F(\vec{Y}, \vec{S}) \) be a vector function of \( \vec{Y} \) and \( \vec{S} \). Then \( \text{Cov}[F(\vec{Y}, \vec{S}), F(\vec{Y}, \vec{S}')] = V[E[F(\vec{Y}, \vec{S})|\vec{Y}]] \).

**Proof of (15).** The MC replicate vectors \( \vec{M}_q^{(c)} \) created by Combine are correlated with common cross-covariance \( \text{Cov}[\vec{M}_q^{(c)}, \vec{M}_q^{(c)}] \). Therefore, according to lemma 1,

\[
E[\vec{\Sigma}^{(c)}_q] = V[\vec{M}_q^{(c)}] - \text{Cov}[\vec{M}_q^{(c)}, \vec{M}_q^{(c)}]
\]

with \( q \neq q' \). Using definition (8) for \( \vec{M}_q^{(c)} \), lemma 2, definition (5) for \( \vec{M}_q^{(c)} \), and the eigendecomposition \( \vec{\Sigma}^{(c)}_q \) = \( U_q D_q U_q^T \), we have

\[
V[\vec{M}_q^{(c)}] = V \left[ \vec{M}_q^{(c)} + \frac{1}{\sqrt{J}} U_q \sqrt{D_q} \vec{Z}_q \right]
\]

\[
= V[\vec{M}_q^{(c)}] + \frac{1}{\sqrt{J}} E[U_q \sqrt{D_q} \vec{Z}_q]
\]

\[
= V[\vec{F}(\vec{Y}, \vec{S})] + \frac{1}{\sqrt{J}} E[\vec{\Sigma}^{(c)}_q].
\]

The Transform nominal values \( \vec{N}_q^{(n)} \) in (4) are independent and identically distributed so

\[
E[\vec{\Sigma}^{(n)}_q] = V[\vec{N}_q^{(n)}] = V[\vec{F}(\vec{Y}, \nu)]
\]

and (18) becomes

\[
V[\vec{M}_q^{(c)}] = V[\vec{F}(\vec{Y}, \vec{S})_q] + \frac{1}{\sqrt{J}} V[\vec{F}(\vec{Y}, \nu)].
\]

Now consider the cross-covariance \( \text{Cov}[\vec{M}_q^{(c)}, \vec{M}_q^{(c)}] \) in (17). Using definition (8) for \( \vec{M}_q^{(c)} \), lemma 3, and definition (5) for \( \vec{M}_q^{(c)} \), we have
\[ \text{Cov}[\ddot{M}_n, \ddot{M}_n] = \text{Cov}[M_n] + \frac{1}{\sqrt{N}} \text{Cov}[\ddot{D} \ddot{Z}, M_n] + \frac{1}{\sqrt{N}} \text{Cov}[\ddot{D} \ddot{Z}, \ddot{M}_n] \]
\[ = \text{Cov}[M_n] \]
\[ = \frac{1}{T} \text{Cov} \left[ \sum_{j=1}^{T} F(\ddot{Y}_j, \ddot{S}_q), \sum_{j=1}^{T} F(\ddot{Y}_j, \ddot{S}_q) \right] \]
\[ = \frac{1}{T} \text{Cov}[F(\ddot{Y}_j, \ddot{S}_q), F(\ddot{Y}_j', \ddot{S}_q)] , \quad (21) \]

the last equality holding because the data vectors \( \ddot{Y}_j \) are independent. Applying lemma 4 to the covariance in (21) and substituting the result along with (20) back into (17) proves (15).

2.1. Example error models

Proposition 1’s point is that the MC replicates produced by Combine in our two-stage scenario have a covariance bias \( \Psi / J \). In the remainder of this section we evaluate \( \Psi \) for various error models, showing that \( \Psi \) can be positive, negative, or zero. Where \( \Psi \) is non-zero, we show in the univariate case \( K = 1 \) that the relative bias

\[ \text{relbias}[\Sigma_{\ddot{M}_n}] = \frac{E[\Sigma_{\ddot{M}_n}] - V[F(\ddot{Y}_j, \ddot{S}_q)]}{V[F(\ddot{Y}_j, \ddot{S}_q)]} \quad (22) \]

approaches \( \pm 20\% \) in one example and even \( 200\% \) in another.

**Additive error:** The function \( f \) in (6) is \( f(a, b) = a + b \) for additive error. In this case \( F(\ddot{y}, \ddot{x}) = \ddot{y} + \ddot{x} \) and

\[ E[F(\ddot{Y}_j, \ddot{S}_q)] = E[\ddot{Y}_j + \ddot{S}_q][\ddot{Y}_j] \]
\[ = \ddot{Y}_j + E[\ddot{S}_q] \]
\[ = F(\ddot{Y}_j, \ddot{X}) \]

so \( \Psi \) in (16) is identically zero. Thus for additive shared systematic error Combine’s MC replicates have both zero mean bias and zero covariance bias.

**Multiplicative error:** The function \( f \) in (6) is \( f(y, s) = ys \) for multiplicative error and the kth component of \( F(\ddot{Y}_j, \ddot{S}_q) \) is \( f(Y_{jk}, S_{qk}) = Y_{jk}S_{qk} \) where \( Y_{jk} \) and \( S_{qk} \) are the kth components of \( \ddot{Y}_j \) and \( \ddot{S}_q \). We have

\[ E[f(Y_{jk}, S_{qk})] = E[Y_{jk}S_{qk}] = Y_{jk}E[S_{qk}] = k(y) \]

Therefore \( E[F(\ddot{Y}_j, \ddot{S}_q)] = F(\ddot{Y}_j, \ddot{X}) \) and \( \Psi = 0 \). This shows that for multiplicative shared systematic error Combine’s MC replicates have both zero mean bias and zero covariance bias.

**Phase error:** The function \( f \) in (6) is \( f(y, s) = \sin(y + s) \) for phase error. In this case the covariance in Combine’s MC replicates can be biased. We focus on the univariate case \( K = 1 \) in which we have scalars, \( Y_j \) and \( S_p \), and the phase error is uniformly distributed, \( S_p \sim \text{Unif}(\delta, \delta) \), \( \delta > 0 \), with mean \( \nu = 0 \) and range \( 2\delta \). We note first that

\[ E[\sin(Y_j + S_p) | Y_j = y] = \frac{\sin y \sin \delta}{\delta} . \]

Therefore

\[ V[E[\sin(Y_j + S_p) | Y_j]] = V \left[ \frac{\sin y \sin \delta}{\delta} \right] = \frac{\sin^2 \delta}{\delta^2} V[\sin y] \]

and, using \( \nu = E[S_p] = 0 \),

\[ V[\sin(Y_j + 0)] = \frac{\sin^2 \delta}{\delta^2} V[\sin y] = \left( 1 - \frac{\sin^2 \delta}{\delta^2} \right) V[\sin y] > 0. \quad (23) \]

To assess the relative size of the bias associated with \( \Psi > 0 \) in (23), we consider the extremal case where \( \delta = \pi \) and where \( Y_j \) is \( \pm \pi / 2 \) with equal probabilities. Then \( \Psi = 1 \), \( V[\sin(Y_j + S_p)|Y_j] = 1/2 \), and \( \text{Cov}[\sin(Y_j + S_p)|Y_j, \sin(Y_j + S_p)|Y_j] = 0 \) for \( j \neq j' \). Using (13), we find that the relative bias (22) associated with the MC sample variance is

\[ \text{relbias}[\Sigma_{\ddot{M}_n}] = \frac{\frac{1}{2} \Psi}{V[F(\ddot{Y}_j, \ddot{S}_q)]} = \frac{\frac{1}{2} \Psi}{V[\sin y]} = 2 \].

Here the relative bias is 200% for any sample size \( J \). This albeit extreme example demonstrates that very large relative biases are possible with Combine’s current method of MC replicate construction.

**Exponential error:** The function \( f \) in (6) is \( f(y, s) = y^s \) for exponential error. In this case the covariance in Combine’s MC replicates can be positively or negatively biased. We focus on the case \( K = 1 \) of uniformly distributed scalars, \( Y_j \) and \( S_p \). For this case we find that \( \Psi \) is broadly, but not always, negative.

Let \( Y_j \sim \text{Unif}[a, b] \) and \( S_p \sim \text{Unif}[1 - \alpha, 1 + \alpha] \), \( 0 \leq \alpha \leq 1 \). We have \( \nu = E[S_p] = 1 \) so \( \Psi \) in (16) is \( V[Y_j] - V[k(Y_j)] \) where

\[ k(y) = E[y^s] = \frac{y \sinh(\alpha \ln y)}{\alpha \ln y} . \]

Using \( V[Y_j] = \frac{1}{12} (b - a)^2 \) and evaluating \( V[k(Y_j)] \) numerically, we find that \( \Psi \) is slightly positive for small \( b \), as shown in figure 2. Otherwise, in the region \( 0 \leq a \leq b \leq 8 \), \( \Psi \) is negative, increasingly so for larger ranges \( 2a \) and \( b - a \).

In the cases presented in figure 2 for exponential error, the covariance

\[ \text{Cov}[F(\ddot{Y}_j, \ddot{S}_q), F(\ddot{Y}_j, \ddot{S}_q)] = \text{Cov}[Y_j^s, Y_j^s] \]

is positive. According to (13), then, the relative bias (22) associated with \( \Psi \) is strongest at the smallest sample size \( J = 2 \), in which case

\[ \text{relbias}[\Sigma_{\ddot{M}_n}] = \frac{\Psi}{V[Y_j^s] + \text{Cov}[Y_j^s, Y_j^s]} . \]
Numerical evaluation of this expression yields the relative biases shown in figure 2. At its strongest the relative bias approaches $\pm 20\%$ for $\alpha = 0.95$.

3. An alternative MC construction

The previous section shows that Combine’s MC replicates $\tilde{M}_q^{(C)}$ in (8) generated for the two-stage scenario in figure 1 fail to fully meet Combine’s design goals (10) and (11). We propose in this section an alternative construction $\tilde{M}_q^{(A)}$ for Combine’s MC replicates. Like the $\tilde{M}_q^{(C)}$ replicates in (8), the proposed $\tilde{M}_q^{(A)}$ replicates meet goal (10). In contrast to the $\tilde{M}_q^{(C)}$ replicates, the $\tilde{M}_q^{(A)}$ replicates essentially meet goal (11), doing so arbitrarily closely for sufficiently large MC replicate sample size $Q$.

Let

$$\tilde{M}_q^{(A)} = M_{\star q}^{(T)} + \frac{1}{\sqrt{J}} U_A \sqrt{D_A} \tilde{Z}_q$$

(24)

where $M_{\star q}^{(T)} = F(\hat{Y}_q, \hat{S}_q)$ and $\tilde{Z}_q \sim N(0, I), q = 1, \ldots, Q$. Further, the $\tilde{Z}_q$ are independent of both the $M_{\star q}^{(T)}$ and $U_A \sqrt{D_A}$. In this alternative construction the matrices $U_A$ and $D_A$ are now the unitary and diagonal members, respectively, of the eigendecomposition $\hat{\Sigma}_{\tilde{M}_q^{(T)}} = U_A \Lambda_A U_A^T$ of the sample covariance

$$\hat{\Sigma}_{\tilde{M}_q^{(T)}} = \frac{1}{J-1} \sum_{j=1}^J (M_j^{(T)} - \bar{M}^{(T)})(M_j^{(T)} - \bar{M}^{(T)})^T$$

(25)

associated with the means $\bar{M}_q^{(T)}$ of the MC samples at Combine’s input. Proposition 2 below shows that basing the sample...
variability of the stage-two Combine MC replicates on the stage-one MC means \( M_q^n \) instead of on the stage-one nominal values \( \bar{N}_q^0 \) essentially removes the bias \( \Psi / J \) identified in proposition 1. This reduced bias is explained in some part by the greater information retained by using the MC means instead of the nominal values: the \( M_q^0 = F(\bar{Y}_q, \bar{S}_q) \) reflect nonlinearities in \( F \) across the full distribution of \( \bar{S}_q \), while the nominal values \( \bar{N}_q^0 = F(\bar{Y}_q, \bar{U}_q) \) are only exposed to \( F \) at the mean \( \bar{Y}_q \) of the \( S_q \) distribution.

**Proposition 2.** Let the set-up be the same as in proposition 1 except that the Combine-stage MC replicates in figure 1 are given by \( \bar{Y}_q \) in (24). Then

\[
E[M_q^n] = E[F(\bar{Y}_q, \bar{S}_q)]
\]  
(26)

and

\[
E[\vec{\Sigma}_{q:j}^{0:n}] = V[F(\bar{Y}_q, \bar{S}_q)] + \frac{1}{Q} \Phi
\]  
(27)

where \( \Phi \) is the difference of two \( K \times K \) covariances

\[
\Phi = E[V[F(\bar{Y}_q, \bar{S}_q)|\bar{S}_q]] - V[E[F(\bar{Y}_q, \bar{S}_q)|\bar{Y}_q]].
\]  
(28)

**Proof.** The proof of (26) is the same as that of (14) because \( \bar{Z}_q \) and \( U_q \sqrt{D_q} \) in (24) are again independent. To prove (27), we first note that the arguments based on Lemmas 1–3 early in the proof of (15) apply also here, giving

\[
E[\vec{\Sigma}_{q:j}^{0:n}] = V[\bar{M}_q^n] - \text{Cov}[\bar{M}_q^n, \bar{M}_q^n]
\]  
(29)

with

\[
V[\bar{M}_q^n] = V[F(\bar{Y}_q, \bar{S}_q)] + \frac{1}{J} E[\vec{\Sigma}_{q:j}^{0:n}]
\]  
(30)

and

\[
\text{Cov}[\bar{M}_q^n, \bar{M}_q^n] = \frac{1}{J} V[E[F(\bar{Y}_q, \bar{S}_q)|\bar{Y}_q]],
\]  
(31)

in which case

\[
E[\vec{\Sigma}_{q:j}^{0:n}] = V[F(\bar{Y}_q, \bar{S}_q)] + \frac{1}{J} \left( E[\vec{\Sigma}_{q:j}^{0:n}] - V[E[F(\bar{Y}_q, \bar{S}_q)|\bar{Y}_q]] \right).
\]  
(32)

Using lemma 1, we write \( E[\vec{\Sigma}_{q:j}^{0:n}] \) in (30) as

\[
E[\vec{\Sigma}_{q:j}^{0:n}] = V[M_q^n] - \text{Cov}[M_q^n, M_q^n].
\]  
(33)

Next,

\[
V[M_q^n] = V \left[ \frac{1}{Q} \sum_{q=1}^{Q} F(\bar{Y}_q, \bar{S}_q) \right]
\]  
(34)

\[
= \frac{1}{Q} V[F(\bar{Y}_q, \bar{S}_q)] + \frac{Q-1}{Q} \text{Cov}[F(\bar{Y}_q, \bar{S}_q), F(\bar{Y}_q, \bar{S}_q)].
\]  
(35)

Similarly, the covariance \( \text{Cov}[M_q^n, M_q^n] \) in (33) is

\[
\text{Cov}[M_q^n, M_q^n] = \text{Cov} \left[ \frac{1}{Q} \sum_{q=1}^{Q} F(\bar{Y}_q, \bar{S}_q), \frac{1}{Q} \sum_{q'=1}^{Q} F(\bar{Y}_{q'}, \bar{S}_{q'}) \right]
\]  
(36)

where the last step is an application of lemma 4. Substituting (35) and (36) back into (33) yields

\[
E[\vec{\Sigma}_{q:j}^{0:n}] = \frac{1}{Q} V[F(\bar{Y}_q, \bar{S}_q)] + \frac{Q-1}{Q} V[E[F(\bar{Y}_q, \bar{S}_q)|\bar{Y}_q]]
\]  
(37)

Finally, substituting this back into (32) proves (27).

For the scalar case \( K = 1 \) the relative bias

\[
\text{relbias}[\vec{\Sigma}_{q:j}^{0:n}] = \frac{1}{JQ} V[F(\bar{Y}_q, \bar{S}_q)]
\]  
(38)

associated with the MC sample variance in proposition 2 has simple bounds, given in the following proposition. Following the proof of this proposition, we show by simple examples that these bounds are tight.

**Proposition 3.** The relative bias in (38) for the scalar case \( K = 1 \) satisfies

\[
0 \leq \text{relbias}[\vec{\Sigma}_{j}^{0:n}] \leq \frac{1}{Q}
\]  
(39)

**Proof.** We prove first that the relative bias in (39) is non-negative. The variance of a random variable \( X \) can be expressed by

\[
V[X] = \frac{1}{2} E[(X - X')^2]
\]  
(40)

where \( X, X' \) are independent and identically distributed. Using conditional versions of (40), we have

\[
\Phi = E[V[F(Y_j, S_q)|S_q]] - V[E[F(Y_j, S_q)|Y_q]]
\]  
(41)

\[
= E \left[ \frac{1}{2} E[F(Y_j, S_q) - F(Y'_j, S_q)^2|S_q] \right]
\]  
(42)

\[
- \frac{1}{2} E \left[ E[F(Y_j, S_q)|Y_q] - E[F(Y'_j, S_q)|Y'_q] \right]^2.
\]  
(43)
Define \( \Delta(y, y', s) = F(y, s) - F(y', s) \). Then

\[
\Phi = \frac{1}{2} E[\Delta^2(Y, Y', S)] - \frac{1}{2} E[\Delta(Y, Y', S)[Y, Y']]
\]

Then the relative bias in (38) is

\[
\text{relbias}[\hat{\Sigma}^{MC}] = \frac{1}{Q} \frac{1}{2} E[V[F(Y, S)]|S] - V[E[F(Y, S)]|Y]
\]

establishing the upper bound in (39).

In the remainder of this section we look at the examples of additive and multiplicative error to see that the bounds in proposition 3 on the relative bias of the MC sample variance \( \hat{\Sigma}^{MC} \) are tight.

**Additive error.** In this case \( \hat{\Phi} = \bar{\Phi} + \bar{s} \) and \( \Phi \) in (28) is

\[
\Phi = E[V[\bar{Y} + \bar{S} | S]] - V[E[\bar{Y} + \bar{S} | S]] = E[V[\bar{Y}]] - V[\bar{Y}] + E[\bar{S}]
\]

This shows that proposition 3’s lower bound is tight for the additive shared systematic error that Combine MC replicates constructed according to (24) have both zero mean bias and zero covariance bias.

**Multiplicative error.** We focus on the univariate case \( K = 1 \) in which we have scalars, \( Y, S \). Then

\[
\Phi = E[V[Y, S] | S] - V[E[Y, S] | Y] = E[S^2] V[Y] - E^2[S] V[Y]
\]

The corresponding relative bias for our case \( K = 1 \) is

\[
\text{relbias}[\hat{\Sigma}_1^{MC}] = \frac{1}{Q} \frac{1}{2} \Phi = \frac{1}{Q} \frac{1}{2} \frac{1}{Q} V[F(Y, S)]
\]

where in the last step we used the product rule for variance [8]. Expression (44) is less than or equal to \( 1/Q \), achieving \( 1/Q \) for \( E[Y] = E[S] = 0 \), showing that the upper bound in proposition 3 is tight. Thus, for multiplicative error with the alternative MC sample variance, the relative covariance bias can be as great as 100% in the extreme case \( Q = 1 \); but for typical user choices of \( Q \) it is no greater than a small fraction of a percent.

4. **Relative variability**

Sections 2 and 3 compared Combine’s current and proposed alternative constructions of MC replicates from the standpoint of bias in the MC sample means and covariances. Specifically, propositions 1 and 2 established that the MC sample means \( M_{MC} \) and \( M_{MC} \) are each unbiased. These propositions also established that MC sample covariance \( \hat{\Sigma}^{MC} \) is biased, while the bias of \( \hat{\Sigma}^{MC} \) is asymptotically zero as \( Q \to \infty \). In this section we complete our comparison of the two constructions by considering the differences in the variabilities of their MC sample means and covariances.

4.1. **Relative variability in MC sample means**

The MC sample means with the current and alternative MC constructions (8) and (24) are, respectively,

\[
M_{MC} = M_{MC} + \frac{1}{\sqrt{J}} U_c \sqrt{D} Z,
\]

\[
\tilde{M}_{MC} = M_{MC} + \frac{1}{\sqrt{J}} U_c \sqrt{D} Z.
\]

The difference in their degrees of variability (their covariances) is given by the following proposition.

**Proposition 4.** Consider the two-stage scenario in figure 1 with the same set-up as in propositions 1 and 2. Then

\[
V[M_{MC}] - V[M_{MC}] = \frac{1}{Q} \Phi - \frac{1}{Q} \Phi^2
\]

with the matrices \( \Psi \) and \( \Phi \) as defined in (16) and (28).

**Proof.** Lemma 2 yields for the MC sample means in (45) that
\[ V[\hat{M}^{(q)}_q] = V[M^{(q)}_q] + \frac{1}{JQ} E[\Sigma^{(q)}_R], \]
\[ V[\hat{M}^{(q)}_q] = V[M^{(q)}_q] + \frac{1}{JQ} E[\Sigma^{(q)}_S], \]

in which case
\[ V[\hat{M}^{(q)}_q] - V[\hat{M}^{(q)}_q] = \frac{1}{JQ} E[\Sigma^{(q)}_R] - \frac{1}{JQ} E[\Sigma^{(q)}_S]. \]

Using the definitions of \( \Psi \) and \( \Phi \) and the results in (19) and (37) for \( E[\Sigma^{(q)}_R] \) and \( E[\Sigma^{(q)}_S] \), we find
\[
V[\hat{M}^{(q)}_q] - V[\hat{M}^{(q)}_q] = \frac{1}{JQ} V[F(y, \bar{y}) - \frac{1}{Q} V[F(y, S)], \]
\[ \frac{1}{Q} V[E[F(y, S)]_S] + \frac{Q}{Q} V[E[F(y, S)]_P] \]
\[ = \frac{1}{JQ} \Psi - \frac{1}{JQ} \left( \frac{1}{Q} V[F(y, S)] \right) \]
\[ = \frac{1}{Q} V[E[F(y, S)]_S] - \frac{1}{Q} V[E[F(y, S)]_P] \]
\[ = \frac{1}{JQ} \Psi - \frac{1}{JQ} \left( \frac{1}{Q} V[F(y, S)] \right) \]
\[ = \frac{1}{JQ} \Psi - \frac{1}{JQ} \left( \frac{1}{Q} V[F(y, S)] \right). \]

Recalling definition (28) for \( \Phi \), this proves the proposition. \( \square \)

For sufficiently large finite \( Q \), the sign of the difference in the covariances \( V[M^{(q)}_q] \) and \( V[\hat{M}^{(q)}_q] \) is determined by \( \Psi \) in (46). We saw in section 2 that \( \Psi \) can be positive, negative, or zero, so depending on the form of the Transform error model \( F(y, S) \) either of the two estimators \( \hat{M}^{(q)}_q \) or \( M^{(q)}_q \) can exhibit less variability. According to (46), when \( \Psi \) is positive, the alternatively constructed MC replicates \( \hat{M}^{(q)}_q \) are better than the current \( M^{(q)}_q \) both because their sample mean \( \hat{M}^{(q)}_q \) exhibits less variability and because their sample variance \( \Sigma^{(q)}_S \) is asymptotically unbiased. When \( \Psi \) is negative, the picture is mixed: \( \hat{M}^{(q)}_q \) exhibits greater variability than \( M^{(q)}_q \), but \( \hat{M}^{(q)}_q \) is biased. All these considerations of the relative variabilities of \( \hat{M}^{(q)}_q \) and \( M^{(q)}_q \) are, of course, dominated by Proposition 4’s main import that the difference in their variabilities is asymptotically zero,
\[ V[\hat{M}^{(q)}_q] - V[M^{(q)}_q] = 0 \quad \text{for} \quad Q \to \infty. \]

4.2. Relative variability in MC sample covariances

We consider in this subsection the difference in the variabilities of the MC sample covariances with the current and proposed alternative constructions, limiting our considerations to the univariate (\( K = 1 \)) case. For \( K = 1 \)
\[ M^{(q)}_q = \hat{M}^{(q)}_q + \frac{Z_q}{\sqrt{J}} \sqrt{S^{(q)}_R}, \] (47)
where \( S^{(q)}_R = F(Y, \nu) \) and
\[ S^{(q)}_S = \frac{1}{J-1} \sum_{j=1}^J \left( N^{(q)}_j - N^{(q)}_1 \right)^2. \]

Even with the restriction \( K = 1 \), assessing the variance of the sample variance of the \( M^{(q)}_q \) in (47) is difficult, necessarily involving fourth moments. Because the \( Z_q \) in (47) are normal, the following lemma (proved in [7]) is useful.

**Lemma 5.** Let \( X_n = \mu_n + \sigma Z_n \) for \( n = 1, ..., N \) where \( \sigma \) and the \( \mu_n \) are constants and the \( Z_n \) are mutually independent and univariate standard normal-distributed. Let
\[ u^2 = \frac{1}{N-1} \sum_{n=1}^N (\mu_n - \bar{\mu})^2, \quad S_X^2 = \frac{1}{N-1} \sum_{n=1}^N (X_n - \bar{X})^2. \]

Then \( E[S_X^2] = u^2 + \sigma^2 \) and
\[ V[S_X^2] = \frac{2}{N-1} \sigma^4 + \frac{4}{N-1} \sigma^2 u^2. \] (48)

Let \( \hat{M}^{(q)}_q \) be the set of \( Q \) MC sample means \( M^{(q)}_q \) in (47). Conditioned on \( \hat{M}^{(q)}_q \) and \( S^{(q)}_R \), combine’s MC replicates \( M^{(q)}_q \) in (47) are independent and normally distributed with means \( \hat{M}^{(q)}_q \) and variance \( \frac{1}{J} S^{(q)}_R \). Then according to lemma 5,
\[ E[S^{(q)}_R | \hat{M}^{(q)}_q, S^{(q)}_R] = \frac{1}{J} S^{(q)}_R + S^{(q)}_{M^{(q)}_q}. \]

and
\[ V[S^{(q)}_R | \hat{M}^{(q)}_q, S^{(q)}_R] = \frac{2}{J-1} \frac{1}{J} S^{(q)}_R + \frac{4}{J-1} \frac{1}{J} S^{(q)}_R + S^{(q)}_{S^{(q)}_R}, \]
where
\[ S^{(q)}_{S^{(q)}_R} = \frac{1}{J-1} \sum_{q=1}^Q (M^{(q)}_q - \hat{M}^{(q)}_q)^2. \]

Then
\[ V[S^{(q)}_R] = V[E[S^{(q)}_R | \hat{M}^{(q)}_q, S^{(q)}_R]] + E[V[S^{(q)}_R | \hat{M}^{(q)}_q, S^{(q)}_R]] \]
\[ = \frac{1}{J} S^{(q)}_R + S^{(q)}_{M^{(q)}_q} \]
\[ + \frac{2}{J-1} \frac{1}{J} E \left[ S^{(q)}_R \right] \]
\[ + \frac{4}{J-1} \frac{1}{J} E \left[ S^{(q)}_R \right]. \] (49)

A parallel calculation for the alternatively constructed MC replicates
\[ M^{(q)}_q = \hat{M}^{(q)}_q + \frac{Z_q}{\sqrt{J}} \sqrt{S^{(q)}_{S^{(q)}_R}}, \]
yields
\[ V[S_{M^*}^2] = V\left[ \frac{1}{J} S_{M^*} + S_{Y^*}^2 \right] \]
\[ + \frac{2}{Q-1} \frac{1}{E} \left[ \bar{S}^4_{M^*} \right] + \frac{4}{Q-1} E \left[ \bar{S}^2_{M^*} S_{Y^*}^2 \right]. \]

We have, therefore, from (49) and (50) that
\[ V[S_{M^*}^2] - V[S_{Y^*}^2] = V\left[ \frac{1}{J} S_{M^*} + S_{Y^*}^2 \right] \]
\[ - \frac{2}{Q-1} \frac{1}{E} \left[ \bar{S}^4_{M^*} \right] + O(1/Q) \]
\[ \approx \frac{1}{J} \left( V[S_{M^*}^2] - V[S_{Y^*}^2] \right) + O(1/Q). \] (51)

We now pursue expressions for the difference in the two variances in approximation (51) in the cases of additive and multiplicative error.

**Additive error model:** For \( f(y,s) = y + s \) we have \( M^*_s = Y^* + S^* \) and \( N^*_s = Y^* + \nu \) where \( \nu = E[S^*] \). Then \( S_{M^*}^2 = S_{Y^*}^2 \) and \( S_{N^*}^2 = S_{Y^*}^2 + \nu^2 S_{Y^*}^2 \). The factors \( S_{M^*}^2 \) and \( S_{N^*}^2 \) are independent, so using the product rule for variance [8], we find
\[ V[S_{M^*}^2] - V[S_{N^*}^2] = V[S_{Y^*}^2 S_{Y^*}] - V[\nu^2 S_{Y^*}] \]
\[ = V[S_{Y^*}^2 E[S_{Y^*}^2]] + E[S_{Y^*}^4 V[S_{Y^*}]] - \nu^4 V[S_{Y^*}^2]. \] (52)

**Multiplicative error model:** For \( f(y,s) = ys \) we have \( M^*_s = Y^*S^* \) and \( N^*_s = Y^*\nu \). Then \( S_{M^*}^2 = S_{Y^*}^2 S_{Y^*}^2 \) and \( S_{N^*}^2 = \nu^2 S_{Y^*}^2 \). Let \( \tau, \omega, \) and \( \psi \) be the second, third, and fourth central moments of \( S_{\nu} \) and let \( \phi \) be the fourth central moment of \( Y^* \). Moment results for the sample mean and sample variance in \[9, 10\] then yield
\[ E[S_{Y^*}^2] = \sigma^2, \]
\[ V[S_{Y^*}^2] = \frac{\phi^2 - \sigma^4}{J(J-1)}, \]
\[ E[S_{Y^*}^4] = \nu^4 + \frac{6\nu^2 \tau^2}{Q} + \frac{3\tau^4 + 4\nu \omega^3}{Q^2} + \frac{\psi^4 - 3\tau^2}{Q^3}. \]
\[ V[S_{Y^*}^2] = \frac{4\nu^2 \tau^2}{Q} + \frac{2\tau^4 + 4\nu \omega^3}{Q^2} + \frac{\psi^4 - 3\tau^2}{Q^3}. \]

It then follows from (53) that
\[ V[S_{M^*}^2] - V[S_{N^*}^2] = \sigma^4 \left( \frac{4\nu^2 \tau^2}{Q} + \frac{2\tau^4 + 4\nu \omega^3}{Q^2} + \frac{\psi^4 - 3\tau^2}{Q^3} \right) \]
\[ + \left( \frac{6\nu^2 \tau^2}{Q} + \frac{3\tau^4 + 4\nu \omega^3}{Q^2} + \frac{\psi^4 - 3\tau^2}{Q^3} \right) \]
\[ \left( \phi^4 - \sigma^4 + \frac{2\sigma^4}{J(J-1)} \right). \] (54)

This suggests generally for multiplicative error that, for \( Q \to \infty \),
\[ V[S_{M^*}^2] - V[S_{Y^*}^2] \to 0. \] (55)

Computer experiments were done to check the approximation in (51) for the difference in the variances of the MC sample variances. According to results (52) and (55), the difference in the variances should be zero and asymptotically zero for additive and multiplicative noise, respectively. These results were confirmed in computer experiments with different distributions for \( Y^* \) and \( S_{\nu} \) and different data sample sizes \( J \). The results in figure 3 are typical. Figure 3 show the estimated relative difference
\[ \text{reldiff} = \frac{V[S_{M^*}^2] - V[S_{Y^*}^2]}{V[S_{M^*}^2] + V[S_{Y^*}^2]} \] (56)
for a case of additive noise and a case of multiplicative noise. For these two curves in figure 3, the data and noise are standard normal-distributed and \( J = 4 \). The plotted relative differences defined by (56) have a potential range of \( \pm 100\% \). Each plotted point in figure 3 is estimated from an independent set of 100000 computer trials.

Our aim with results (52) and (55) was to discover whether either method of constructing MC replicates dominates the other with respect to variance of the MC sample variance. These results and the experiments confirming them, indicate that neither MC replicate construction method dominates the other when the error is additive or multiplicative. For error models beyond the additive and multiplicative cases, little analytical headway seems possible, so we turn directly to computer experiments.

Also presented in figure 3 are corresponding experiment results obtained for phase and exponential error. The plotted phase error data show the relative difference (56) of variances for phase error for the extremal case discussed in section 2 in which the error \( S_{\nu} \) is distributed \( \text{Unif}[-\pi, \pi] \), the data \( Y^* \) are \(-\pi/2 \) and \( \pi/2 \) with equal probabilities, and \( J = 4 \). The last two curves in figure 3 show two cases of exponential error, both in which the data and exponential error are uniformly distributed, \( Y^* \sim \text{Unif}[0, b] \) and \( S_{\nu} \sim \text{Unif}[1 - \alpha, 1 + \alpha] \), with \( b = 1.8 \) and \( \alpha = .95 \). These two cases of exponential error are the two cases in figure 2 where the relative bias of Combine’s MC sample variance is most extreme—approaching \( \pm 20\% \).

The results in figure 3 show that neither MC construction method dominates the other by having consistently smaller variance in its sample variance. For the case of exponential error with \( b = 1 \), the relative difference (56) in variances is negative, indicating the sample variance with the current method has less variability. For exponential error with \( b = 8 \), though, the alternative method’s sample variance shows less variability. Also, the five plots in figure 3 taken together illustrate that the relative difference in variances can exhibit different degrees and types of transient behavior for small \( Q \). The relative difference for exponential error with \( b = 8 \) shows little transient change,
while for exponential error with \( b = 8 \) the change is significant. The example with phase error shows that the relative difference in the variances can even change sign as \( Q \) increases.

5. Summary remarks

The MUF is a powerful tool for uncertainty modeling and analysis relating to data obtained in high-precision microwave experiments, and the Combine module is a key component of the MUF. We compared the MC replicates currently constructed by Combine with those based on an alternative construction, using bias and variance of the MC sample mean and sample covariance as performance measures. We showed first that, with the current method of Combine MC replicate construction, the MC sample covariance is biased. Examples showed that this bias can be unacceptably large—200% in one extreme example and approaching ±20% in others—and cannot be reduced by choosing the MC sample size \( Q \) sufficiently large. The MC sample covariance using the alternative construction for MC replicates is also biased, but this bias is asymptotically zero with \( Q \).

Bias is the primary concern in MC sampling and the distinction, the current method being biased and the alternative being asymptotically unbiased, is the two construction methods’ most important difference. Looking beyond bias to the difference in the variabilities of the MC sample means with the two methods, we showed that this difference is asymptotically zero, with neither method dominating the other for small \( Q \). Comparing the variabilities of the MC sample variances was even more nuanced and non-determinative: in the cases of additive and multiplicative error, the difference in the variances of the sample variances is zero or asymptotically zero, while for phase and exponential error, neither method consistently out-performs the other.

We showed in this case study of bias in uncertainty propagation software that our proposed alternative MC replicate construction method has an important advantage with regard to MC sample covariance bias, while lacking any clear statistics disadvantage relative to the current method. Consequently, the current method of constructing MC replicates in the MUF Combine module is set to be replaced with our proposed alternative. This study shows both that unknown, inadvertent biases are potentially present in even closely scrutinized statistical software and that computer experiments can successfully identify these biases. We urge that statistical performance tests be standard for modern software uncertainty propagation tools, and we anticipate that the statistical approach used here will be useful to future analyses of MUF performance and of the performance of other similar statistical software for uncertainty propagation.

Acknowledgment

This investigation benefited from early-stage discussions with Chih-Ming Wang and Sarah Streett, members of the Statistical Engineering Division of the National Institute of Standards and Technology.

Appendix

We set out some standard notation and prove the four lemmas used in the proof of proposition 1.

Notation. Overbars indicate sample averaging. The conventional dot notation [11] indicates the indices over which averaging is done; for example, for scalars \( X_{ij} \), \( i = 1, \ldots, I \), \( j = 1, \ldots, J \).
\[ X_{\bullet} = \frac{1}{J} \sum_{j=1}^{J} X_{ij}, \quad \hat{X}_{\bullet\bullet} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij}. \]

This averaging notation suppresses a vector’s overbar; for vectors \( \vec{X}_{ij}, i = 1, \ldots, I; j = 1, \ldots, J, \)
\[ \hat{X}_{\bullet} = \frac{1}{J} \sum_{j=1}^{J} \vec{X}_{ij}, \quad \hat{X}_{\bullet\bullet} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} \vec{X}_{ij}. \]

Probability expectations, including variances and covariances, always involve the joint distribution of all random variables present within the expectation; for example, for two random variables \( X \) and \( Y \) and a deterministic scalar function \( g, E[g(X, Y)] \) is calculated with respect to the joint distribution of \( X \) and \( Y \). Restricted probability averages are expressed by conditional expectations: for example \( E[g(X, Y)|X] \) is the expectation of \( g(X, Y) \) with respect to the distribution of \( (X, Y) \) conditioned on \( X \).

**Lemma A.1.** Let \( \vec{X}_{ij} \sim (\vec{m}, \Sigma) \) for \( j = 1, \ldots, J, J > 1 \) with common cross-covariance \( \text{Cov}[\vec{X}_{ij}, \vec{X}_{i\prime j}] = \Sigma' \) for all \( j \neq k \). Then \( E[\Sigma] = \Sigma - \Sigma' \) where \( \Sigma \) is the sample covariance matrix.

\[ \hat{\Sigma} = \frac{1}{J-1} \sum_{j=1}^{J} (\vec{X}_{ij} - \vec{X})(\vec{X}_{ij} - \vec{X})^{T}. \quad (A.1) \]

**Proof.**
\[
E[\Sigma] = E \left[ \frac{1}{J-1} \sum_{j=1}^{J} \vec{X}_{ij} \vec{X}_{ij}^{T} - \frac{1}{J-1} \vec{X}_{\bullet\bullet} \vec{X}_{\bullet\bullet}^{T} \right] \\
= \frac{1}{J-1} E[\vec{X}_{ij} \vec{X}_{ij}^{T}] - \frac{1}{J-1} \left( \sum_{j=1}^{J} E[\vec{X}_{ij} \vec{X}_{ij}^{T}] + \sum_{j' \neq j} E[\vec{X}_{i j'} \vec{X}_{i j'}^{T}] \right) \\
= \frac{1}{J-1} (\Sigma + \vec{m} \vec{m}^{T}) - \frac{1}{J-1} (\Sigma' + \vec{m} \vec{m}^{T}) \\
= \Sigma - \Sigma'.
\]

**Lemma A.2.** Let \( \vec{Z} \sim (\vec{0}, I) \) be independent of the vector-matrix pair \( (\vec{A}, \vec{B}) \). Then \( \text{V}[\vec{A} + \vec{BZ}] = \text{V}[\vec{A}] + \text{E}[\text{BB}^{T}] \).

**Proof.**
\[
\text{V}[\vec{A} + \vec{BZ}] = \text{V}[E[\vec{A} + \vec{BZ}]] + \text{E}[\text{V}[\vec{A} + \vec{BZ}]] \\
= \text{V}[E[\vec{A}]] + \text{E}[\text{V}[\vec{A}]] + \text{E}[\text{V}[\vec{B}]] + \text{V}[\text{V}][\vec{Z}][\vec{Z}]^{T} \\
= \text{V}[E[\vec{A}]] + \text{E}[\text{V}[\vec{B}]] + \text{E}[\text{BB}^{T}] \\
= \text{V}[\vec{A}] + \text{E}[\text{BB}^{T}].
\]

**Lemma A.3.** Suppose \( \vec{Z}_{1}, \vec{Z}_{2} \) and \( (\vec{A}_{1}, \vec{A}_{2}, \vec{B}) \) are mutually independent with \( \vec{Z}_{1}, \vec{Z}_{2} \sim (\vec{0}, I) \). Then \( \text{Cov}[\vec{A}_{1} + \vec{BZ}_{1}, \vec{A}_{2} + \vec{BZ}_{2}] = \text{Cov}[\vec{A}_{1}, \vec{A}_{2}] \).

**Proof.**
\[
\text{Cov}[\vec{A}_{1} + \vec{BZ}_{1}, \vec{A}_{2} + \vec{BZ}_{2}] \\
= \text{Cov}[E[\vec{A}_{1} + \vec{BZ}_{1}], E[\vec{A}_{2} + \vec{BZ}_{2}][\vec{B}]] \\
+ \text{E}[\text{Cov}[\vec{A}_{1} + \vec{BZ}_{1}, \vec{A}_{2} + \vec{BZ}_{2}][\vec{B}]] \\
= \text{Cov}[\vec{A}_{1}, \vec{A}_{2}] \\
= \text{Cov}[\vec{A}_{1}, \vec{A}_{2}].
\]

**Lemma A.4.** Let \( \vec{S}, \vec{S}^{'}, \vec{Y} \) be independent, identically distributed random vectors independent of the random vector \( \vec{Y} \). Let \( F(\vec{Y}, \vec{S}) \) be a vector function of \( \vec{Y} \) and \( \vec{S} \). Then \( \text{Cov}[F(\vec{Y}, \vec{S}), F(\vec{Y}, \vec{S}^{'})] = \text{V}[E[F(\vec{Y}, \vec{S})|\vec{Y}]] \).

**Proof.**
\[
\text{Cov}[F(\vec{Y}, \vec{S}), F(\vec{Y}, \vec{S}^{'})] = \text{Cov}[E[F(\vec{Y}, \vec{S})]\vec{Y}, E[F(\vec{Y}, \vec{S}^{'})]\vec{Y}] \\
+ \text{E}[\text{Cov}[F(\vec{Y}, \vec{S}), F(\vec{Y}, \vec{S}^{'})]|\vec{Y}]. \quad (A.2)
\]

The two functions \( E[F(\vec{Y}, \vec{S})]\vec{Y} \) and \( E[F(\vec{Y}, \vec{S}^{'})]\vec{Y} \) are identical so the first covariance on the right in \( (A.2) \) is \( \text{V}[E[F(\vec{Y}, \vec{S})]|\vec{Y}] \). Also, \( F(\vec{Y}, \vec{S}) \) and \( F(\vec{Y}, \vec{S}^{'}) \) are conditionally independent given \( \vec{Y} \), so their conditional covariance on the right in \( (A.2) \) is zero.

**References**

1. Joint Committee for Guides in Metrology 2008 Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement (Sèvres: International Bureau of Weights and Measures)
2. Joint Committee for Guides in Metrology 2008 Evaluation of Measurement Data—Supplement 2 to the Guide to the Expression of Uncertainty in Measurement—Models with Any Number of Output Quantities (Sèvres: International Bureau of Weights and Measures)
3. Avolio G, Williams D F, Street S, Frey M, Schreurs D, Ferrero A and Dieudonn M 2017 Software tools for uncertainty evaluation in VNA measurements: a comparative study 89th ARFTG Microwave Measurement Conf. (IEEE) pp 1–7
4. Lewandowski A, Williams D F, Hale P D, Wang J C M and Dienstfrey A 2010 Covariance-based vector-network-analyzer uncertainty analysis for time- and frequency-domain measurements IEEE Trans. Microw. Theory Tech. 58 1877–86
5. Jamroz B F, Williams D F, Remley K A and Horansky R D 2018 Importance of preserving correlations in error-vector-magnitude uncertainty 91st ARFTG Microwave Measurement Conf. (IEEE) pp 1–4
6. Hale P D, Williams D F and Dienstfrey A 2018 Waveform metrology: signal measurements in a modulated world Metrologia 55 S135
7. Knautz H and Trenkler G 1995 Some bounds for bias and variance of \( \hat{S} \) under dependence Scandinavian J. Stat. 22 121–8
[8] Goodman L A 1960 On the exact variance of products J. Am. Stat. Assoc. 55 708–13
[9] Angelova J A 2012 On moments of sample mean and variance Int. J. Pure Appl. Math. 79 67–85
[10] Rose C and Smith M D 2002 Mathematical Statistics with Mathematica (New York: Springer)
[11] Montgomery D C 2017 Design and Analysis of Experiments 9th edn (Hoboken, NJ: Wiley)