Green Simulation: Reusing the Output of Repeated Experiments

Mingbin Feng, Jeremy Staum
Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, 60208
benfeng@u.northwestern.edu, j-staum@northwestern.edu

We introduce a new paradigm in simulation experiment design and analysis, called "green simulation," for the setting in which experiments are performed repeatedly with the same simulation model. Green simulation means reusing outputs from previous experiments to answer the question currently being asked of the simulation model. As a first method for green simulation, we propose estimators that reuse outputs from previous experiments by weighting them with likelihood ratios, when parameters of distributions in the simulation model differ across experiments. We analyze convergence of these green simulation estimators as more experiments are repeated, while a stochastic process changes the parameters used in each experiment. We find that green simulation reduces variance by more than an order of magnitude in examples involving catastrophe bond pricing and credit risk evaluation.

Key words: likelihood ratio method; multiple importance sampling; score function method; simulation metamodeling

1. Introduction

Consider a setting in which simulation experiments are performed repeatedly, using the same simulation model with different values of its inputs. As we discuss in detail below, such settings occur when a simulation model is used routinely to support a business process, and over the lifecycle of a simulation model as it goes from development to application in repeated simulation studies. In these settings, the standard practice is that each new simulation experiment is designed to answer a particular question without using the output of previous simulation experiments. We advocate a paradigm of *green simulation* for repeated experiments, meaning that one should reuse output from previous experiments to answer new questions. The benefit of green simulation is greater computational efficiency. In this article, we show that when old simulation output is reused well, it provides greater accuracy when combined with a new simulation experiment than would be achieved by the same new simulation experiment alone.

Green simulation entails a new perspective on management of simulation experiments. The standard practice is to discard or ignore the output of a simulation experiment after it has delivered the desired answer. When a new question arises, a new simulation experiment is designed to answer it without using the output of previous simulation experiments. From this perspective, running

the simulation model is a computational cost or expense. From the green simulation perspective, running the simulation model is a computational investment that provides future benefits, because simulation output is a valuable resource to be used in answering questions that will be asked of the simulation model in the future.

One setting of repeated simulation experiments occurs when simulation supports a business process. For example, in finance and insurance, simulation models support pricing and risk management decisions that are made periodically. At each period, current information, such as prices and forecasts, is used to update the inputs to the model, and a simulation experiment is performed to answer a question about price or risk. Similarly, in manufacturing, service, and logistics systems, simulation models can be used routinely to provide information about expected completion times and to support decisions about such matters as dispatching and staffing. A simulation experiment is run whenever information is required, using inputs that describe the current state of the system.

Another setting of repeated simulation experiments occurs in the lifecycle of a simulation model as it goes from development to application in simulation studies. First, experiments are performed for purposes of verification and validation of the model. They may also be performed for model calibration: to choose realistic values of unknown inputs. For these purposes, the simulation model is run many times with different values of its inputs, to see how its outputs change with its input, and where this behavior is reasonable and realistic. Once model development is complete, experiments are performed for purposes such as making predictions for particular values of the inputs, metamodeling, sensitivity analysis, and optimization. Moreover, some simulation models are used in many simulation studies. Consequently, each model is used in several or many experiments.

Figure 1 illustrates a sequence of two repeated experiments, each with a single run. To clarify our terminology, by "a run," we mean one or more replications of simulation output generated with the inputs to the model held fixed. By a simulation "experiment," we mean a collection of one or more runs of a simulation model, designed for the purpose of answering a specific question. In Figure 1, the first experiment has a single run with input x_1, and the second experiment has a single run with input x_2. Each run has r replications. The purpose of the nth experiment is to estimate $\mu(x_n)$, the mean output of the simulation model when the input is x_n. The standard practice is to estimate $\mu(x_2)$ in the second experiment using only the single run with input x_2. The green question mark in Figure 1 indicates the question answered in this article: "How can we reuse the simulation output from the first experiment to improve our answer in the second experiment?"

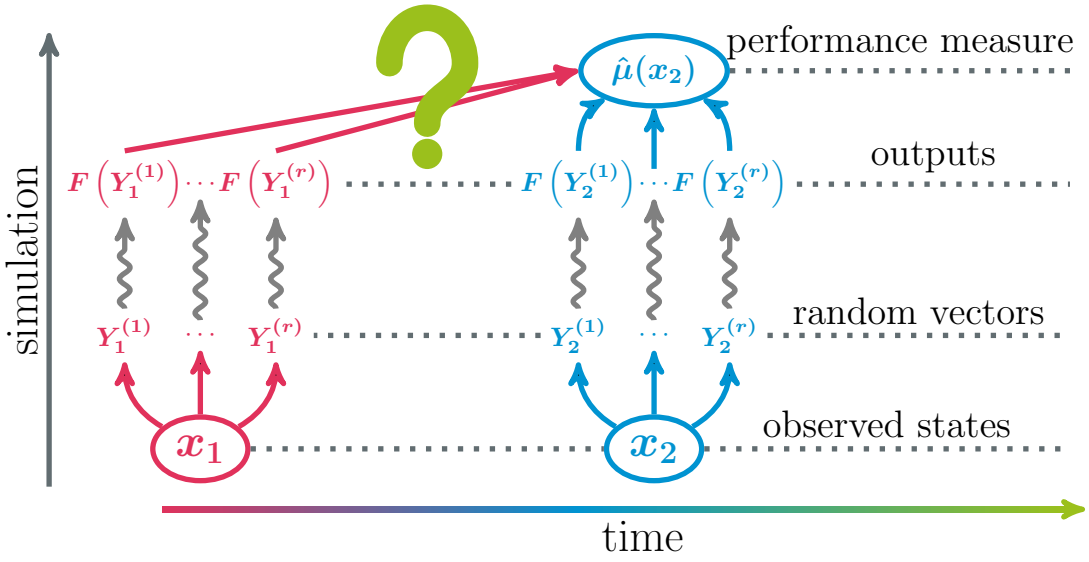

Figure 1 Setting of repeated simulation experiments.

The core idea of green simulation, reusing simulation output, has been applied to isolated experiments that contain multiple runs. Many types of simulation experiments use multiple runs to learn about the model's *response surface*, the function that maps the model's inputs to a performance measure. In stochastic simulation, this performance measure is often the expected output of the simulation model. There are metamodeling and sensitivity analysis experiments that run the model at different input values to learn about how the response surface varies, globally or locally. In nested simulation experiments, an outer-level simulation generates random values of the inputs at which it is desired to learn the value of the response surface of an inner-level simulation. For example, in assessing the impact of uncertainty about a simulation model's input on the conclusions of a simulation study, this model is the inner-level simulation model, and the outer-level simulation samples values of the inputs from an appropriate distribution. In optimization via simulation experiments, the model is run at different input values in a search for optimal input values. If simulation output from runs at some values of the inputs can be reused in estimating the value of the response surface at another value of the inputs, then experiment designs that involve multiple runs can be modified to be cheaper. It is unnecessary to run many replications at every value of the inputs for which it is desired to estimate the value of the response surface if estimates of these values can reuse output from runs at other values of the inputs. For example, the nested simulation methods of Barton et al. (2014) and Xie et al. (2014) use metamodeling to reuse output from a moderate number of runs to estimate the value of the response surface for many values of the inputs, thus reducing the required number of runs in the simulation experiment. The score function method, also known as the likelihood ratio method (Rubinstein and Shapiro 1993, Kleijnen and Rubinstein 1996), has also

been applied to reuse output within isolated experiments for metamodeling, sensitivity analysis, and optimization.

We adopt the likelihood ratio method as the tool for reusing output across experiments. In applications and surveys of the likelihood ratio method that we have seen, such as Beckman and McKay (1987), Rubinstein and Shapiro (1993), Kleijnen and Rubinstein (1996), and Glasserman and Xu (2014), there is an isolated experiment within which each estimator reuses the output from a single run. The exception is Maggiar et al. (2015): they construct an estimator by reusing output from multiple runs in an isolated experiment. We reuse output from multiple runs that come from multiple experiments. Even in the simple setting of a single run per experiment, which we analyze in this article, reusing output from multiple experiments entails reusing output from multiple runs. Estimation using multiple simulation runs, weighted based on likelihood ratios, was given the name *multiple importance sampling* by Veach (1997). On multiple importance sampling, see, for example, Hesterberg (1988, 1995), Owen and Zhou (2000), Veach and Guibas (1995), and Veach (1997). Drawing on this literature on multiple importance sampling, Maggiar et al. (2015) and we use the same two estimators, which we call the individual likelihood ratio (ILR) estimator (Section 2.1) and the mixture likelihood ratio (MLR) estimator (Section 2.2).

The work of Maggiar et al. (2015) and our work differ in setting and findings. The key differences are that their work focuses on an isolated experiment with a deterministic simulation model, whereas ours focuses on repeated experiments with a stochastic simulation model. Their goal is optimization of the response surface after smoothing by convolution with a Gaussian kernel, to reduce the influence of numerical noise in the deterministic simulation. They apply the likelihood ratio method to the corresponding Gaussian random variable. Their convergence theorem describes convergence to an optimal solution within an isolated optimization experiment with an increasing number of iterations. Our green simulation paradigm applies broadly to stochastic simulation, and we emphasize the setting of repeated experiments. Our convergence theorems describe convergence of estimators of values of the response surface at some or all points as the number of repeated experiments increases. Another difference between the work of Maggiar et al. (2015) and our work is in the findings about the ILR and MLR estimators. They report that the the choice between ILR and MLR estimators makes "only a small difference on the performance" of their optimization procedure when applied to a testbed of optimization problems. In Section 4, we find that the difference between ILR and MLR estimators could be large, depending on the simulation model and sequence of repeated experiments. We recommend the MLR estimator, which is theoretically superior, because we find that it can work well in practice even in cases where the ILR estimator yields poor results.

The main contribution of this article is to introduce the paradigm of green simulation for repeated experiments and to demonstrate theoretically and experimentally that it yields significant benefits in computational efficiency. Specifically, we investigate the case in which a simulation experiment is run routinely (e.g., to support a business process) with updated values of inputs that serve as parameters of distributions of random variables generated in the simulation. For this case, we propose (Section 2), analyze (Section 3), and test (Section 4) green simulation estimators based on the likelihood ratio method. In particular, we prove novel theorems about how these green simulation estimators converge as the number of repeated experiments increases, while the number of simulation replications per experiment remains constant. In this setting, the estimator based on standard practice does not converge at all, but our green simulation estimators converge at the canonical rate for Monte Carlo: variance inversely proportional to total computational budget. They converge at this rate even though the total computational budget includes all previous experiments, which are not obviously relevant to answering the question currently being asked of the simulation model. A secondary contribution of this article is to import methods from multiple importance sampling (where importance sampling is used for variance reduction) into the likelihood ratio method (which reuses simulation output), and to provide evidence that there can be great practical value to using the weighting scheme in the MLR estimator. The methods we develop in this article are based on likelihood ratios, so they are applicable only when simulation experiments are repeated with changes to parameters of distributions. Green simulation has broader applicability than the specific methods proposed in this article. Some future research directions that will extend the scope of applicability of green simulation are discussed in Section 5.

2. Green Simulation via the Likelihood Ratio Method

We develop green simulation procedures in a setting of repeated experiments with the same simulation model and some changing parameters that affect the likelihood of simulated random variables. The setting of repeated experiments, and the notation developed in this section, are illustrated in Figure 1. Let X_n represent the parameters in the nth experiment, for example, prices observed in the market on day n or forecasted arrival rates for period n. We treat $\{X_n : n = 1, 2, \cdots\}$ as a discrete-time stochastic process taking values in a general state space $\mathcal{X}$. We use "state" to refer to an element of $\mathcal{X}$ or a random variable taking values in $\mathcal{X}$, and "current" to refer to quantities associated with the nth experiment. Thus, X_n is the current state. We suppose that the current state is observable at the current time step n, when the current experiment needs to be run, but was not observable earlier.

Given the current state X_n, the current simulation experiment samples a random vector Y_n according to the conditional likelihood $h(\cdot; X_n)$. For example, $h(y; x)$ could be the conditional

probability density for a stock's price to be y in one year given that the stock's current price is x. As another example, $h(y;x)$ could be the conditional probability density for the vector y of interarrival times and service times given the arrival rate x, regarding service rate as fixed and not included in x. We will write expectations in the form of integrals, which implicitly assumes that Y_n has a conditional probability density $h(\cdot;X_n)$, but this is not essential; the setting allows for continuous, discrete, and mixed conditional distributions. We assume that $h(\cdot;x)$ has the same support $\mathcal{Y}$ for all $x \in \mathcal{X}$. We also assume that $h(y;x)$ can be evaluated; it is not enough merely to be able to sample according to this likelihood.

The simulation output or simulated performance of the stochastic system is $F(Y_n)$, where the function $F(\cdot) : \mathcal{Y} \mapsto \mathbb{R}$ represents the logic of the simulation model. For example, $F(y)$ could be the discounted payoff of a stock option if the stock's price in one year is y, or $F(y)$ could be the average customer waiting time in a queue given the vector y of interarrival times and service times.

We have assumed that the stochastic process $\{X_n : n = 1, 2, \cdots\}$ affects the simulation only through the conditional likelihood $h(\cdot;X_n)$ of the random vector Y_n. This assumption is necessary to support the green simulation procedures, based on likelihood ratios, that are proposed in this article. However, it is not necessary for green simulation in general.

In the current experiment, we wish to estimate the conditional expected performance $\mu(X_n)$ of the stochastic system given the current state X_n; $\mu(X_n)$ can also be described as the current expected performance. The expected performance for state x is

$$\mu(x) = \mathbb{E}[F(Y_n) | X_n = x] = \int_{\mathcal{Y}} F(y)\, h(y;x)\, dy, \tag{1}$$

which is the same for all n. We assume that $\mu(x)$ is finite for all $x \in \mathcal{X}$.

Standard practice in the setting of repeated experiments is to estimate $\mu(X_n)$ by the *Standard Monte Carlo (SMC)* estimator

$$\hat{\mu}_r^{SMC}(X_n) = \frac{1}{r} \sum_{j=1}^{r} F\left(Y_n^{(j)}\right), \tag{2}$$

based on running r replications of the simulation model with the parameters set according to the current state X_n. For simplicity in notation, we have assumed that the number r of replications is fixed, but this is not essential. We refer to $\left\{F\left(Y_n^{(j)}\right) : j = 1, \cdots, r\right\}$ as the output of the current experiment.

The purpose of the current experiment is to estimate the current expected performance $\mu(X_n)$. This is a random variable because the current state X_n was not observable at time step 0. Figures of merit for an estimator, such as bias and variance, should be evaluated conditional on the current

state X_n. Clearly, $\hat{\mu}_r^{SMC}(X_n)$ is conditionally unbiased for $\mu(X_n)$, given X_n. The variance for state x is

$$\sigma^2(x) = \mathbb{V}\mathrm{ar}\left[F\left(Y_n\right)|X_n = x\right] = \int_{\mathcal{Y}} \left(F\left(y\right) - \mu(x)\right)^2 h\left(y;x\right) dy, \tag{3}$$

which is the same for all n. We assume that $\sigma^2(x)$ is finite for all $x \in \mathcal{X}$. The conditional variance of the SMC estimator, given the current state X_n, is $\sigma^2(X_n)/r$. Thus, to reduce the conditional variance of the SMC estimator, one must increase the number r of replications in the current experiment.

Next, we propose two green simulation estimators that reuse the output of the $n-1$ previous experiments and combine it with the current (nth) experiment's output. These green simulation estimators are also conditionally unbiased. In Section 3, we show that, under some conditions, the conditional variance of these green simulation estimators goes to zero as the number n of experiments goes to infinity, even if the number r of replications per experiment is fixed.

2.1. Individual Likelihood Ratio Estimator

In the current experiment, the *target distribution* appears in the conditional expectation $\mu(X_n) = \mathbb{E}\left[F\left(Y_n\right)|X_n\right]$ that we are estimating. The target distribution has the likelihood $h\left(\cdot;X_n\right)$. In a previous experiment at time step $k < n$, the *sampling distribution* had likelihood $h\left(\cdot;X_k\right)$. To use the output of a previous experiment in estimating $\mu(X_n)$ in a way that is conditionally unbiased given the state history $X_1, \ldots, X_n$, we can adjust the old output by using the likelihood ratio between the target distribution and the sampling distribution:

$$\mathbb{E}\left[\frac{h\left(Y_k;X_n\right)}{h\left(Y_k;X_k\right)} F\left(Y_k\right)|X_1,\ldots,X_n\right] = \int_{\mathcal{Y}} \frac{h\left(y;X_n\right)}{h\left(y;X_k\right)} h\left(y;X_k\right) F\left(y\right) dy = \mathbb{E}\left[F\left(Y_n\right)|X_n\right].$$

Using the likelihood ratio method (Rubinstein and Shapiro 1993), we can combine the output from all previous experiments and the current experiment into the following estimator of $\mu(x)$:

$$\hat{\mu}_{n,r}^{ILR}(x) = \sum_{k=1}^{n} \sum_{j=1}^{r} \frac{1}{nr} \frac{h\left(Y_k^{(j)};x\right)}{h\left(Y_k^{(j)};X_k\right)} F\left(Y_k^{(j)}\right) \tag{4a}$$

$$= \sum_{k=1}^{n} \frac{1}{n} \left[\frac{1}{r} \sum_{j=1}^{r} \frac{h\left(Y_k^{(j)};x\right)}{h\left(Y_k^{(j)};X_k\right)} F\left(Y_k^{(j)}\right)\right]. \tag{4b}$$

In particular, $\hat{\mu}_{n,r}^{ILR}(X_n)$ is our estimator of the current expected performance $\mu(X_n)$. However, the use of the likelihood ratio enables us to estimate expected performance given a state x that we never used in a sampling distribution. We refer to the green simulation estimator in (4) as the *Individual Likelihood Ratio (ILR)* estimator. It contains likelihood ratios that each involve one individual sampling distribution; this distinguishes the ILR estimator from the next green simulation estimator we will propose.

The ILR estimator in (4b) can be seen as the average of n individual importance-sampling estimators; the nth is the SMC estimator if state $x = X_n$. It is worth emphasizing the difference between the present application of likelihood ratios and the typical application of importance sampling. In importance sampling, the designer of the simulation experiment chooses the sampling distribution with the aim of reducing variance. In this article, we do not address the choice of sampling distribution. We assume that the sampling distribution in the current experiment is determined by the current state X_n, as is standard practice in the setting of repeated experiments. We use likelihood ratios not to reduce variance, but to enable reuse of simulation output based on different sampling distributions.

Of course, the effect of likelihood ratios on variance needs to be considered. For any states $x, x' \in \mathcal{X}$, define the target-x-sample-x' variance as

$$\sigma_x^2(x') = \int_{\mathcal{Y}} \left(F(y) \frac{h(y;x)}{h(y;x')} - \mu(x) \right)^2 h(y;x')\, dy. \tag{5}$$

It could be more or less than the target-x-sample-x variance $\sigma_x^2(x) = \sigma^2(x)$ associated with standard Monte Carlo; it could even be infinite. Based on (4b), given the state history $X_1, \ldots, X_n$, the conditional variance of the estimator of current expected performance, $\hat{\mu}_{n,r}^{ILR}(X_n)$, is $\sum_{k=1}^n \sigma_{X_n}^2(X_k)/(n^2 r)$. If none of $\sigma_{X_n}^2(X_1), \ldots, \sigma_{X_n}^2(X_n)$ is too large, then the ILR estimator has lower conditional variance than the SMC estimator. However, the ILR estimator could have higher or infinite conditional variance. Considering that $\sigma_{X_n}^2(X_1), \ldots, \sigma_{X_n}^2(X_n)$ may be unequal, one might try to construct a lower-variance estimator by using unequal weights instead of the equal weights $1/n$ in (4b). Even better results are possible if we replace the equal weights $1/nr$ in (4a) with unequal weights that depend on the random vector $Y_k^{(j)}$. This topic is addressed in the research literature on multiple importance sampling, which leads to the next green simulation estimator we propose.

2.2. Mixture Likelihood Ratio Estimator

In our setting of repeated experiments, we have r replications sampled from each of n distributions. The collection of nr observations can be viewed as a stratified sample from an equally-weighted mixture of these n distributions. The likelihood of the mixture is denoted by $\bar{h}(\cdot; X_1, \ldots, X_n)$, where

$$\bar{h}(y; x_1, \ldots, x_n) = \frac{1}{n} \sum_{k=1}^n h(y; x_k). \tag{6}$$

Veach and Guibas (1995) advocated replacing the equal weights $1/nr$ in (4a) with "balance heuristic" weights which, in our setting, are $h\left(Y_k^{(j)}; X_k\right) / nr \bar{h}\left(Y_k^{(j)}; X_1, \ldots, X_n\right)$. This leads to the *mixture likelihood ratio (MLR)* estimator

$$\hat{\mu}_{n,r}^{MLR}(x) = \sum_{k=1}^n \sum_{j=1}^r \frac{1}{nr} \frac{h\left(Y_k^{(j)}; x\right)}{\bar{h}\left(Y_k^{(j)}; X_1, \ldots, X_n\right)} F\left(Y_k^{(j)}\right). \tag{7}$$

The MLR estimator is the likelihood ratio estimator that arises when we consider the pooled outputs of all simulation experiments performed so far, $\{Y_k^{(j)} : j = 1, \ldots, r, k = 1, \ldots, n\}$, as a stratified sample from a mixture distribution (Hesterberg 1995). It follows from this interpretation, or immediately from the results of Veach and Guibas (1995, Section 3.2), that the MLR estimator is conditionally unbiased for $\mu(x)$ given $X_1, \ldots, X_n$.

The MLR estimator is superior to the ILR estimator in the sense that

$$\mathbb{V}\mathrm{ar}\left[\hat{\mu}_{n,r}^{MLR}(x) \middle| X_1, \ldots, X_n\right] \leq \mathbb{V}\mathrm{ar}\left[\hat{\mu}_{n,r}^{ILR}(x) \middle| X_1, \ldots, X_n\right] \tag{8}$$

for any n, r, and x. This inequality follows immediately from Theorem A.2 of Martino et al. (2014). Despite the inferiority of the ILR estimator to the MLR estimator in terms of variance, we continue to consider both ILR and MLR estimators. One reason is that the ILR estimator is easier to analyze theoretically (Section 3). Another is that it has lower computation cost (Section 2.3). This makes it worth investigating, in Section 4, whether the MLR estimator has much less variance than the LIR estimator in practical examples; we find that it does.

Finally, we consider an additional merit of the MLR estimator when used to estimate the current expected performance $\mu(X_n)$. To obtain robustness against a poor choice of sampling distribution in importance sampling, Hesterberg (1995) proposed to use a defensive mixture distribution, which is a mixture of a proposed sampling distribution with the target distribution. When we estimate $\mu(X_n)$, the target distribution is the nth distribution in the mixture, so the mixture is defensive. In this case, the likelihood ratio $h\left(Y_k^{(j)}; X_n\right) / \bar{h}\left(Y_k^{(j)}; X_1, \ldots, X_n\right) \leq n$ (Hesterberg 1995). It follows that the conditional variance of $\hat{\mu}_{n,r}^{MLR}(X_n)$ given the state history $X_1, \ldots, X_n$ is at most $n\sigma^2(X_n)$.

2.3. Green Algorithms for Green Simulation Estimators

This section proposes and analyzes algorithms for the green simulation estimators in the setting of repeated experiments. The algorithms are also green, in the sense that they store and reuse likelihood evaluations as well as simulation output. Suppose that one simulation replication has computational cost C_F, one evaluation of a likelihood has computational cost C_h, and the computational cost of basic arithmetic operations such as addition, multiplication, and division is negligible in comparison to these. We envision a situation in which C_F is large, C_h is smaller but need not be negligible, storage space is abundant, and memory access is fast. We consider a sequence of experiments, indexed $n = 1, 2, \cdots$, of r replications each. For each experiment in the sequence, the SMC, ILR, or MLR estimators of the current expected performance $\mu(X_n)$ are computed, and the green simulation procedures store some information to be used in the next experiment. We analyze the storage requirement and computation cost of the nth experiment.

For benchmarking purposes, we first consider the SMC estimator. It has zero storage requirement in the sense that no information is stored from one experiment to the next. Its computation cost is rC_F.

Consider the ILR estimator $\hat{\mu}_{n,r}^{ILR}(X_n)$ in (4a). The likelihood $h\left(Y_k^{(j)}; X_k\right)$ in the denominator does not change as n increases. Therefore we propose to store and reuse likelihoods from one experiment to the next in Algorithm 1. The storage requirements and non-negligible computation costs for the nth experiment are shown on the right in Algorithm 1. The algorithm has storage requirement $2nr$ and computation cost $rC_F + nrC_h$ for the nth experiment. The linear growth rate in n is reassuring; it suggests that it is affordable to reuse the outputs of many experiments in the ILR estimator.

Algorithm 1 Green implementation of ILR estimator

for $n = 1, 2, \cdots$, **do**
  Observe X_n
  Set $\hat{\mu}_{n,r}^{ILR}(X_n) \leftarrow 0$
  for $j = 1, \cdots, r$ **do**
    Sample $Y_n^{(j)}$ and evaluate $F\left(Y_n^{(j)}\right)$ ▷ rC_F computation
    Append to output storage $F\left(Y_n^{(j)}\right)$ ▷ nr storage
    Set $\hat{\mu}_{n,r}^{ILR}(X_n) \leftarrow \hat{\mu}_{n,r}^{ILR}(X_n) + F\left(Y_n^{(j)}\right)$
    Calculate likelihood $h\left(Y_n^{(j)}; X_n\right)$ ▷ rC_h computation
    Append to likelihood storage $h\left(Y_k^{(n)}; X_n\right)$ ▷ nr storage
  end for
  for $k = 1, \cdots, n-1$ **do**
    for $j = 1, \cdots, r$ **do**
      Retrieve $F\left(Y_k^{(j)}\right)$ and $h\left(Y_k^{(j)}; X_k\right)$ from storage
      Calculate likelihood $h\left(Y_k^{(j)}; X_n\right)$ ▷ $(n-1)rC_h$ computation
      Set $\hat{\mu}_{n,r}^{ILR}(X_n) \leftarrow \hat{\mu}_{n,r}^{ILR}(X_n) + F\left(Y_k^{(j)}\right) h\left(Y_k^{(j)}; X_n\right) / h\left(Y_k^{(j)}; X_k\right)$
    end for
  end for
  Set $\hat{\mu}_{n,r}^{ILR}(X_n) \leftarrow \hat{\mu}_{n,r}^{ILR}(X_n)/nr$ and output
end for

For the MLR estimator $\hat{\mu}_{n,r}^{MLR}(X_n)$, a green algorithm is especially valuable. Inspection of (6) and (7) suggests that the MLR estimator requires $n^2 r$ likelihood evaluations: $h\left(Y_k^{(j)}; x_\ell\right)$, for all $j = 1, \ldots, r$ and $k, \ell = 1, \cdots, n$. A quadratic growth rate of computation cost in n could be an

obstacle for using the MLR estimator when reusing the output of many experiments. By storing and reusing likelihoods from one experiment to the next in Algorithm 2, we avoid this quadratic growth and achieve linear growth of the computation cost in n, as was the case for the ILR estimator. Algorithm 2 has storage requirement $2nr$ and computation cost $rC_F + (2n-1)rC_h$ for the nth experiment. This result for MLR is similar to the result for ILR, but MLR requires almost twice as many likelihood evaluations.

Algorithm 2 Green implementation of MLR estimator

for $n = 1, 2, \cdots,$ **do**
  Observe X_n
  Set $\hat{\mu}_{n,r}^{MLR}(X_n) \leftarrow 0$
  for $j = 1, \cdots, r$ **do**
    for $k = 1, \ldots, n$ **do**
      if $k < n$ **then**
        Retrieve $F\left(Y_k^{(j)}\right)$ and $\bar{h}(Y_k^{(j)})$ from storage
      else
        Sample $Y_n^{(j)}$ and evaluate $F\left(Y_n^{(j)}\right)$ ▷ rC_F computation
        Append to output storage $F\left(Y_n^{(j)}\right)$ ▷ nr storage
        Set $n\bar{h}(Y_n^{(j)}) \leftarrow 0$
        for $\ell = 1, \cdots, n-1$ **do**
          Calculate likelihood $h\left(Y_n^{(j)}; X_\ell\right)$ ▷ $(n-1)rC_h$ computation
          Set $n\bar{h}(Y_n^{(j)}) \leftarrow n\bar{h}(Y_n^{(j)}) + h\left(Y_n^{(j)}; X_\ell\right)$
        end for
      end if
      Calculate likelihood $h\left(Y_k^{(j)}; X_n\right)$ ▷ nrC_h computation
      Set $n\bar{h}(Y_k^{(j)}) \leftarrow n\bar{h}(Y_k^{(j)}) + h\left(Y_k^{(j)}; X_n\right)$
      Append to likelihood storage $n\bar{h}(Y_n^{(j)})$ ▷ nr storage
      Set $\hat{\mu}_{n,r}^{MLR}(X_n) \leftarrow \hat{\mu}_{n,r}^{MLR}(X_n) + F\left(Y_k^{(j)}\right) h\left(Y_k^{(j)}; X_n\right) / n\bar{h}(Y_k^{(j)})$
    end for
  end for
  Set $\hat{\mu}_{n,r}^{MLR}(X_n) \leftarrow \hat{\mu}_{n,r}^{MLR}(X_n)/r$ and output
end for

3. Convergence of Green Simulation Estimators

In this section, we analyze the convergence of the green simulation estimators as the number n of experiments grows while the number r of replications per experiment is fixed. To this end, we introduce further assumptions on the stochastic process $\{X_n : n = 1, 2, \ldots\}$ that determines the sampling distributions and on its relationship to the likelihood ratios between target distribution and sampling distribution. Although not all of the assumptions in the theorems are transparent, we show in the Appendix that they can be verified in a realistic example.

Our first assumption is that $\{X_n : n = 1, 2, \ldots\}$ is ergodic. The intuition for making this assumption is as follows. For any target state $x \in \mathcal{X}$, we envision that it has a neighborhood such that, for every x' in this neighborhood, the sampling distribution associated with x' is a good sampling distribution for the target distribution associated with x. A good sampling distribution would be one for which the target-x-sample-x' variance $\sigma_x^2(x')$ defined in (5) is sufficiently small. An ergodic process returns to this neighborhood infinitely often. The consequence is that, as $n \to \infty$, the number of good samples to be used in estimating $\mu(x)$ also grows without bound.

For simplicity of exposition, we consider ergodic Markov chains with general state space. The following summary is based on Meyn and Tweedie (2009) and Nummelin (2004), particularly on Theorems 13.3.3 and 17.1.7 of Meyn and Tweedie (2009) and Proposition 6.3 of Nummelin (2004).

DEFINITION 1. A Markov chain $\{X_n : n = 1, 2, \cdots\}$ is called *ergodic* if it is positive Harris recurrent and aperiodic.

THEOREM 1. *If $\{X_n : n = 1, 2, \cdots\}$ is an ergodic Markov chain that takes values in a general state space $\mathcal{X}$ and has initial probability measure ν and transition kernel P, then there exists a probability measure π on the Borel sigma-algebra $\mathcal{B}(\mathcal{X})$ of $\mathcal{X}$ such that*

$$\lim_{n\to\infty} \left\| \int_{\mathcal{X}} \nu(dx) P^n(x, \cdot) - \pi \right\| = 0. \tag{9}$$

The ergodicity of a Markov chain and the existence and uniqueness of its stationary probability measure π are equivalent. In the following theorem, the notation $f \in L_1(\mathcal{X}, \mathcal{B}(\mathcal{X}), \pi)$ means that the random variable $f(X)$ has a finite expectation under π.

THEOREM 2. *If $\{X_n : n = 1, 2, \cdots\}$ is an ergodic Markov chain with stationary probability measure π and $f \in L_1(\mathcal{X}, \mathcal{B}(\mathcal{X}), \pi)$, then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_k) = \int_{\mathcal{X}} f(x) d\pi(x) \tag{10}$$

Based on Theorem 2, Theorem 3 shows that the conditional variance of the green simulation estimators evaluated at a fixed target state x, given the state history $X_1, \ldots, X_n$, goes to zero at

the rate $\mathcal{O}(n^{-1})$. Because both green simulation estimators are unbiased, it follows that they are consistent as $n \to \infty$. Theorem 3 also provides a result about unconditional variance, which may be easier to interpret.

THEOREM 3. *Assume that $\{X_n : n = 1, 2, \cdots\}$ is an ergodic Markov chain with stationary probability measure π. For any target state $x \in \mathcal{X}$, if σ_x^2 defined in (5) is in $L_1(\mathcal{X}, \mathcal{B}(\mathcal{X}), \pi)$, then*

$$\begin{aligned}\lim_{n\to\infty} nr\mathbb{V}ar\left[\hat{\mu}_{n,r}^{MLR}(x)\middle| X_1, \ldots, X_n\right] &\le \lim_{n\to\infty} nr\mathbb{V}ar\left[\hat{\mu}_{n,r}^{ILR}(x)\middle| X_1, \ldots, X_n\right] \\ &= \int_{\mathcal{X}} \sigma_x^2(x')d\pi(x'), \quad a.s. \end{aligned} \tag{11}$$

If, furthermore, the target-x-sample-X_n variance process $\{\sigma_x^2(X_n), n = 1, 2\cdots\}$ is uniformly integrable, then

$$\lim_{n\to\infty} nr\mathbb{V}ar\left[\hat{\mu}_{n,r}^{MLR}(x)\right] \le \lim_{n\to\infty} nr\mathbb{V}ar\left[\hat{\mu}_{n,r}^{ILR}(x)\right] = \int_{\mathcal{X}} \sigma_x^2(x')d\pi(x'). \tag{12}$$

Proof. We derive the results for the ILR estimator; the results for the MLR estimator follow due to (8). By the conditional independence of the random vectors $\{Y_k^{(j)} : j = 1, \ldots, r, k = 1, \cdots, n\}$ given $X_1, \ldots, X_n$, we have

$$\mathbb{V}\text{ar}\left[\hat{\mu}_{n,r}^{ILR}(x)\middle| X_1, \ldots, X_n\right] = \frac{1}{n^2 r^2}\sum_{k=1}^{n}\sum_{j=1}^{r}\mathbb{V}\text{ar}\left[F\left(Y_k^{(j)}\right)\frac{h\left(Y_k^{(j)}; x\right)}{h\left(Y_k^{(j)}; X_k\right)}\right] = \frac{1}{n^2 r}\sum_{k=1}^{n}\sigma_x^2(X_k). \tag{13}$$

Therefore, by Theorem 2, we obtain (11):

$$\lim_{n\to\infty} nr\mathbb{V}\text{ar}\left[\hat{\mu}_{n,r}^{ILR}(x)\middle| X_1, \ldots, X_n\right] = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\sigma_x^2(X_k) = \int_{\mathcal{X}}\sigma_x^2(x')d\pi(x'), \quad \text{a.s.}$$

To establish (12), consider that

$$\begin{aligned}\mathbb{V}\text{ar}\left[\hat{\mu}_{n,r}^{ILR}(x)\right] &= \mathbb{E}\left[\mathbb{V}\text{ar}\left[\hat{\mu}_{n,r}^{ILR}(x)\middle| X_1, \ldots, X_n\right]\right] + \mathbb{V}\text{ar}\left[\mathbb{E}\left[\hat{\mu}_{n,r}^{ILR}(x)\middle| X_1, \ldots, X_n\right]\right] \\ &= \mathbb{E}\left[\mathbb{V}\text{ar}\left[\hat{\mu}_{n,r}^{ILR}(x)\middle| X_1, \ldots, X_n\right]\right] + \mathbb{V}\text{ar}\left[\mu(x)\right] \\ &= \mathbb{E}\left[\frac{1}{n^2 r}\sum_{k=1}^{n}\sigma_x^2(X_k)\right],\end{aligned}$$

using (13). Therefore

$$\lim_{nr\to\infty} nr\mathbb{V}\text{ar}\left[\hat{\mu}_{n,r}^{ILR}(x)\right] = \lim_{n\to\infty}\mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}\sigma_x^2(X_k)\right] = \mathbb{E}\left[\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\sigma_x^2(X_k)\right] = \int_{\mathcal{X}}\sigma_x^2(x')d\pi(x'),$$

where the exchange of limit and expectation holds by uniform integrability, and Theorem 2 justifies the last step.

Theorem 3 shows the asymptotic superiority of the green simulation estimators to standard Monte Carlo (SMC), as the number n of repeated experiments increases. Recall, from the discussion of (3), that the conditional variance of the SMC estimator for the current expected performance $\mu(X_n)$, given the state history, is $\sigma^2(X_n)/r$. This does not converge to zero as $n \to \infty$, assuming that sampling variances are positive. Yet the green simulation estimators converge to the true value $\mu(x)$ as $n \to \infty$. This means that we can obtain arbitarily high accuracy without increasing the budget r per experiment, merely by reusing output from experiments that are repeated at each time step with budget r. Under the assumptions of Theorem 3, reusing old simulation output is highly effective in the sense that the variance of the green simulation estimators converges as $\mathcal{O}((nr)^{-1})$, which is the standard rate of convergence for Monte Carlo in terms of the computational budget nr expended on all experiments that were ever run in the past.

Theorem 3 is about convergence of green simulation estimators for any fixed target state x that satisfies certain conditions. Theorem 4 provides the conclusion that the convergence of the ILR estimator is uniform across all states $x \in \mathcal{X}$. That is, after a sufficiently large number of experiments, the function $\hat{\mu}^{ILR}$ has small error as an estimator of the function μ that maps a state x to the expected performance $\mu(x)$. As a corollary, it follows that the ILR estimator of the current expected performance, $\mu(X_n)$, eventually has small error. Theorem 4 is similar to Theorem 1 of Staum et al. (2015), but it simplifies one of the assumptions of the latter theorem by relying on the ergodicity of $\{X_n : n = 1, 2, \cdots\}$.

THEOREM 4. *Assume that $\{X_n : n = 1, 2, \cdots\}$ is an ergodic Markov chain with stationary probability measure π, and the following assumptions hold:*

A1. The state space $\mathcal{X}$ is a subset of a Euclidean space. The function $\mu : \mathcal{X} \mapsto \mathbb{R}$ is bounded.

A2. The support $\mathcal{Y}$ is a subset of a Euclidean space. For any $x \in \mathcal{X}$, the function $F(\cdot)\, h(\cdot; x)$ is measurable. For any $y \in \mathcal{Y}$, the function $h(y; \cdot)$ is bounded on $\mathcal{X}$.

A3. There exist non-negative constants a and b such that, for all $x, x' \in \mathcal{X}$ and all $y \in \mathcal{Y}$,

$$|F(y)| \frac{h(y; x)}{h(y; x')} \leq a \exp(b\|y\|_2).$$

A4. Let $M(x,t) = \int \exp(t\|y\|_2) h(y; x)\, dy$. There exists $t > b$, where b is as in A3, such that $M(\cdot, t) \in L_1(\mathcal{X}, \mathcal{B}(\mathcal{X}), \pi)$ and that the sequence of random variables $\{M(X_n, t) : n = 1, 2, \cdots\}$ is uniformly integrable.

Then $\lim_{n\to\infty} \sup_{x\in\mathcal{X}} |\hat{\mu}^{ILR}_{n,r}(x) - \mu(x)| = 0$, almost surely.

Proof. The conclusion of this theorem is the same as that of Theorem 1 of Staum et al. (2015), so it suffices to verify the five assumptions of that theorem. Assumptions 1, 3, and 4 of Staum et al. (2015) hold as direct consequences of A1, A2 and A3. Their Assumption 2 holds because we

have assumed that the sampling distribution $h(\cdot;x)$ has a common support $\mathcal{Y}$ for all states $x \in \mathcal{X}$. It remains to verify Assumption 5 of Staum et al. (2015) using A4 and ergodicity.

By ergodicity, the distribution of $M(X_n,t)$ converges to a limiting distribution as $n \to \infty$. Because $M(\cdot,t) \in L_1(\mathcal{X},\mathcal{B}(\mathcal{X}),\pi)$, this limiting distribution of $M(X_n,t)$ has a finite mean; call it m. Uniform integrability of $\{M(X_n,t) \ : n = 1,2,\cdots\}$ implies that $\lim_{n\to\infty} \mathbb{E}[M(X_n,t)] = \mathbb{E}\left[\lim_{n\to\infty} M(X_n,t)\right] = m$. Therefore there exists a finite c such that $\mathbb{E}[M(X_n,t)] \le c$ for all $n = 1,2,\cdots$. Let $p = t/b > 1$. We now have

$$\sum_{i=1}^{\infty} i^{-p}\mathbb{E}[M(X_n,pb)] \le c\sum_{i=1}^{\infty} i^{-p} < \infty,$$

which verifies Assumption 5 of Staum et al. (2015).

COROLLARY 1. *Under the assumptions of Theorem 4,* $\lim_{n\to\infty} |\hat{\mu}^{ILR}_{n,r}(X_n) - \mu(X_n)| = 0$ *almost surely.*

Proof. From Theorem 4, it follows that for every $\epsilon > 0$, there exists a random variable N_ϵ such that, almost surely, for all $x \in \mathcal{X}$ and $n > N_\epsilon$, we have $|\hat{\mu}^{ILR}_{n,r}(x) - \mu(x)| < \epsilon$. Therefore, for all $n > N_\epsilon$, we have $|\hat{\mu}^{ILR}_{n,r}(X_n) - \mu(X_n)| < \epsilon$.

4. Numerical Examples

In this section, we use two numerical examples to illustrate green simulation, demonstrate its value, and compare our two green simulation estimators with each other and with standard Monte Carlo. First is a reinsurance example of pricing catastrophe bonds. For this example, we verify the conditions of Theorem 3 and Theorem 4 in the Appendix, which shows that the conditions are applicable to a realistic example. The experiment results conform to the theoretical predictions and show that the MLR estimator is superior to the ILR estimator, which is superior to the SMC estimator. The second example involves measuring the credit risk of a loan portfolio. In this example, the conditions of the theorems do not hold, but the experiment results show that green simulation can still deliver valuable results in such a situation. Although the ILR estimator is not reliable in this example, the MLR estimator is superior to the SMC estimator.

4.1. Catastrophe Bond Pricing with Compound Losses

A catastrophe bond ("CAT bond") is an important reinsurance contract that helps insurance companies to hedge against losses from catastrophic events (Munich Re Geo Risks Research 2015). This example is relevant beyond insurance; for example, senior tranches of structured financial instruments are essentially economic catastrophe bonds, because they suffer credit losses only in the event of an economic catastrophe (Coval et al. 2009). Simulation of CAT bonds can be computationally intensive, because it involves fairly rare events in a complex geophysical models. Specifically, the example is of a hurricane CAT bond.

In practice, reinsurance contracts are subject to periodic renewals. This is the source of the repeated experiments: the same CAT bond is priced every period, using the same hurricane simulation model, with parameters X_n updated to reflect the current climatological forecast. In this example, the period is semi-annual. The state process $\{X_n : n = 1, 2, \cdots\}$ is modeled by an ergodic Markov chain. Given the state X_n, we can simulate hurricanes that take place during the lifetime of the CAT bond and the resulting total insured loss Y_n underlying the CAT bond. Finally, we compute the payoff of the CAT bond per dollar invested,

$$F(Y_n) = \mathbf{1}_{\{Y_n \le K\}} + p\mathbf{1}_{\{Y_n > K\}},$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function, K is the trigger level, and $p \in [0, 1)$ is the fraction of face value that is received if insured losses exceed the trigger level. The fair price of this CAT bond can be obtained in terms of the expected payoff $\mu(X_n) = \mathbb{E}[F(Y_n)|X_n]$. We consider a hurricane CAT bond with lifetime 10 years, trigger level $K = 25$ million dollars, and recovery fraction $p = 0.5$.

In this example, we use a simplified version of the model of Dassios and Jang (2003). The insured loss is modeled as a compound random variable: $Y_n = \sum_{i=1}^{M_n} Z_n^i$, where M_n denotes the number of claims and Z_n^i denotes the ith claim size. In this model, $Z_n^{(i)}$, $i = 1, 2 \cdots$ are i.i.d. and independent of M_n. This is a popular loss model due to its flexibility and suitability for many practical applications (Klugman et al. 2012), yet it can provide mathematical tractability. Let the probability mass function of M_n be $p(m; \lambda_n)$ and the probability density of Z_n^i be $f(z; \theta_n)$, where λ_n and θ_n are parameters determined by the state X_n. Specifically, we take M_n to be Poisson with mean λ_n and Z_n^i to be exponential with mean θ_n. In this model, the expected payoff

$$\mu(X_n) = \mathbb{E}[F(Y_n)|X_n] = \mathbb{E}\left[p + (1-p)\mathbf{1}_{\{Y_n \le K\}}\right] = p + \sum_{m=1}^{\infty} p(m; \lambda_n) F(K; \theta_n, m) \tag{14}$$

where $p(m; \lambda)$ is the Poisson probability mass function with mean λ and $F(K; \theta, m)$ is the Gamma cumulative distribution function with scale parameter θ and shape parameter m. From 1981-2010, the average number of major hurricanes was 2.7 per decade and the average cost per hurricane was about \$5,000 million (Blake and Gibney 2011). Therefore, we set up a stochastic model of the states $\{X_n : n = 1, 2, \cdots\}$ and a transformation $(\lambda, \theta) = \psi(x)$ so that λ_n is usually around 2.7 and θ_n is usually around 5 (measured in thousands of millions of dollars).

In this example, the ergodic Markov chain driving the parameters of the loss model is a stationary AR(1) process with state space $\mathcal{X} = \mathbb{R}^2$, given by

$$X_n = \mu_\infty + \varphi X_{n-1} + \varepsilon_n,$$

where $\{\varepsilon_n : n = 1, 2, \cdots\}$ is an i.i.d. sequence of bivariate normal random vectors with mean zero and variance $\mathrm{diag}(\sigma_\varepsilon^2)$, and the parameters are

$$\mu_\infty = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \varphi = \begin{bmatrix} 0.6 \\ 0.5 \end{bmatrix}, \quad \text{and} \quad \sigma_\varepsilon^2 = \begin{bmatrix} 0.8^2 \\ 0.5^2 \end{bmatrix}.$$

The state space $\mathcal{X} = \mathbb{R}^2$ is inappropriate for parameters that must be non-negative because they represent an expected number of hurricanes and an expected loss per hurricane. We introduce a transformation $\psi : \mathbb{R}^2 \mapsto (\underline{\lambda}, \bar{\lambda}) \times (\underline{\theta}, \bar{\theta})$ so that the parameters $(\lambda, \theta) = \psi(x)$ lie between plausible lower and upper bounds. The transformation is sigmoidal and maps $x = [x^1, x^2]$ to

$$(\lambda, \theta) = \psi(x) = \left[\underline{\lambda} + \tfrac{\bar{\lambda} - \underline{\lambda}}{1 + e^{-x^1}}, \, \underline{\theta} + \tfrac{\bar{\theta} - \underline{\theta}}{1 + e^{-x^2}} \right].$$

In particular, we took $(\underline{\lambda}, \bar{\lambda}) = (2, 4)$ as the range for expected number of hurricanes per decade and $(\underline{\theta}, \bar{\theta}) = (4, 6)$ as the range for expected loss per hurricane. To give a picture of the variability of the parameters for repeated experiments in this example, Figure 2 shows histograms of the parameters λ and θ resulting from sampling from the stationary distribution of the AR(1) process.

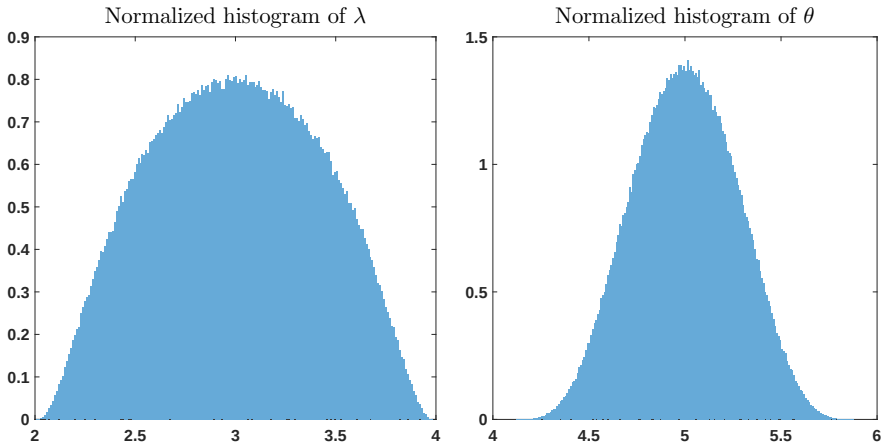

Figure 2 **Histograms of parameters sampled based on the stationary distribution of the AR(1) process.**

To clarify the model, Algorithm 3 shows how the standard Monte Carlo simulation works.

To investigate the effectiveness of green simulation, we performed a sequence of 100 repeated simulation experiments (i.e., $n = 1, 2, \cdots, 100$) with $r = 100$ replications each. Using the same sample path $\{X_n : n = 1, 2, \cdots, 100\}$ and the same simulation output $\{Y_n^{(j)} : n = 1, 2, \cdots, 100, j = 1, 2, \cdots, 100\}$, we evaluated the SMC, ILR, and MLR estimators at each period $n = 1, 2, \cdots, 100$, in each of three states: the current state X_n, the central state $x_{mi} = (0, 0)$ corresponding to the moderate parameters $\lambda = 3$ and $\theta = 5$, and an extreme state $x_{hi} = (\infty, \infty)$ corresponding to extreme parameters $\lambda = 4$ and $\theta = 6$. (Strictly speaking, the extreme state x_{hi} does not belong to the

Algorithm 3 Standard Monte Carlo Simulation for Catastrophe Bond Pricing

Sample state X_0 from the stationary distribution of the AR(1) process
for $n = 1, 2, \cdots,$ **do**
  Sample state X_n from AR(1) process, conditional on X_{n-1}
  Set parameters $(\lambda_n, \theta_n) \leftarrow \psi(x_n)$
  Set $\hat{\mu}_{n,r}^{SMC}(X_n) \leftarrow 0$
  for $j = 1, \cdots, r$ **do**
    Sample number of hurricanes $M_n^{(j)} \sim p(\cdot; \lambda_n)$
    for $i = 1, \cdots, M_n^{(j)}$ **do**
      Sample loss for ith hurricane $Z_n^{(i,j)} \sim f(\cdot; \theta_n)$
    end for
    Set $Y_n^{(j)} \leftarrow \sum_{i=1}^{M_n} Z_n^{i,j}$
    if $Y_n^{(j)} \leq K$ **then**
      Set $F\left(Y_n^{(j)}\right) \leftarrow 1$
    else
      Set $F\left(Y_n^{(j)}\right) \leftarrow p$
    end if
    Set $\hat{\mu}_{x_n}^{SMC}(X_n) \leftarrow \hat{\mu}_{x_n}^{SMC}(X_n) + F\left(Y_n^{(j)}\right)$
  end for
  Set $\hat{\mu}_{n,r}^{SMC}(X_n) \leftarrow \hat{\mu}_{n,r}^{SMC}(X_n)/r$ and output
end for

state space $\mathcal{X} = \mathbb{R}^2$, but the parameter vector $(\lambda, \theta) = (4, 6)$ is a limit point of the range of the transformation ψ that maps states to parameters.)

For the purpose of accurately estimating the unconditional variances of the all these estimators, we performed such a sequence of experiments 10,000 times. These 10,000 macro-replications of the sequence of experiments have independent sample paths and simulation output. The estimated variance of a fixed-state estimator $\hat{\mu}(x)$ of $\mu(x)$ was $\sum_{k=1}^{10,000} (\hat{\mu}^{(k)}(x) - \mu(x))^2/10,000$, where $\hat{\mu}^{(k)}(x)$ is the value of the estimator on the kth macro-replication. Likewise, the estimated variance of a current-state estimator $\hat{\mu}(X_n)$ of $\mu(X_n)$ was $\sum_{k=1}^{10,000} (\hat{\mu}^{(k)}(X_n^{(k)}) - \mu(X_n^{(k)}))^2/10,000$. Due to using 10,000 macro-replications, the standard errors of these estimated variances are less than 1% of the corresponding estimated variance.

Figure 3 is a log-log plot of the variances of the SMC, ILR, and MLR estimators for two fixed states, x_{mi} and x_{hi}, for each experiment $n = 1, 2, \cdots, 100$. The horizontal solid blue line is the SMC variance, the dotted red line is the ILR variance, and the dashed purple line is the MLR

variance. The fixed-state SMC estimators $\hat{\mu}_{n,r}^{SMC}(x_{mi})$ and $\hat{\mu}_{n,r}^{SMC}(x_{hi})$ were generated by sampling according to $h(\cdot; x_{mi})$ and $h(\cdot; x_{mi})$, respectively, in experiments distinct from the experiments described in the preceding paragraph. The SMC variance forms a horizontal line because, for each experiment n, $\hat{\mu}_{n,r}^{SMC}(x_{mi})$ and $\hat{\mu}_{n,r}^{SMC}(x_{hi})$ use a fixed number r of replications drawn from the sampling distribution associated with the fixed state x_{mi} or x_{hi}. In addition, a black solid line with slope -1 and intercept equal to the SMC variance is plotted for reference. This line shows the variance of an SMC estimator with nr replications, which is the cumulative number of replications simulated up through the nth experiment. We compare the lines for green simulation estimators against the solid lines. When they go below the horizontal line, they have lower variance than a SMC simulation with r replications, the budget for a single experiment. When they go near the line with slope -1, they have variance nearly as low as a SMC simulation with nr replications, the cumulative budget for all experiments so far. That is a major success for green simulation, because it shows that reusing old simulation output is nearly as effective as generating new simulation output in the current experiment.

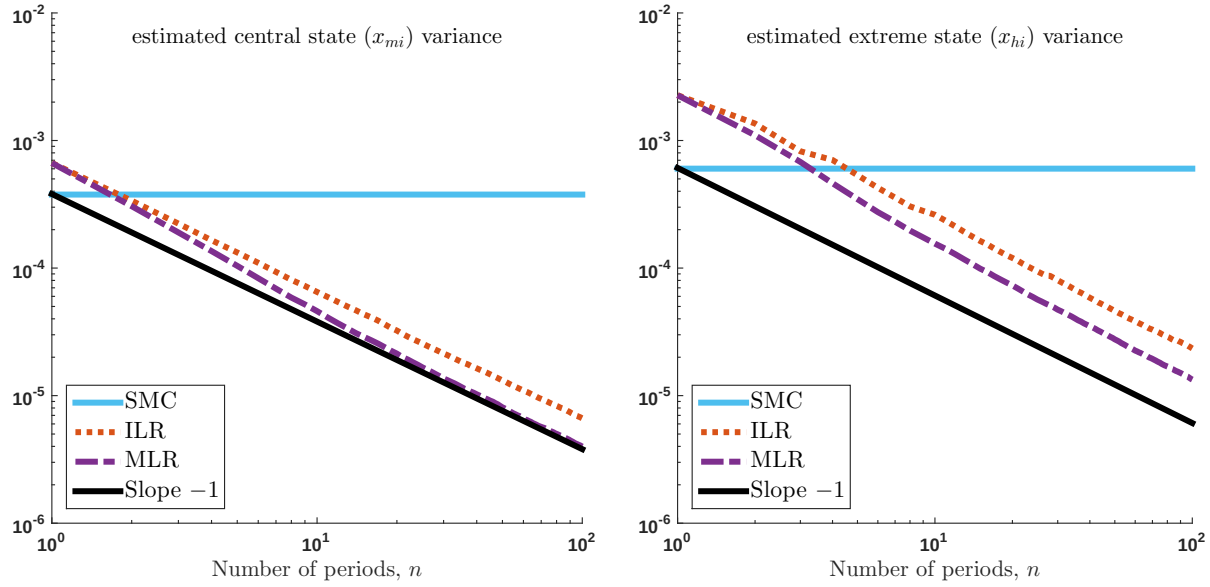

Figure 3 Log-log plots of estimated variances of fixed-state estimators.

We see from Figure 3 that the MLR estimator has lower variance than the ILR estimator, which is consistent with (8). In this example, the gap between them grows to be substantial as the number n of repeated experiments increases: for $n = 100$, the ratio between ILR and MLR variances is about 1.7 for both x_{mi} and x_{hi}. Initially, for very small n, the green simulation estimators have higher variances than an SMC estimator based on a simulation in the fixed state x_{mi} or x_{hi}. The cause is the probability that none of the states visited so far, $X_1, X_2, \cdots, X_n$, were near x_{mi} or x_{hi}.

This event is more likely for the extreme state x_{hi} than for the center state x_{mi}, which is why the higher variances persist longer for the extreme state (for the first 5 experiments) than for the center state (only for the first experiment). The variances of the green simulation soon become lower than those of SMC and continue to decrease as the outputs of more experiments are reused. After 100 experiments, the MLR estimator's variance is over 45 times smaller than the SMC estimator's variance for x_{hi} and over 95 times smaller for x_{mi}. In other words, by using 10,000 total replications simulated in 100 simulation experiments, with simulation based on parameters corresponding to $X_1, X_2, \cdots, X_{100}$ and not x_{hi}, the MLR estimator achieves higher accuracy in estimating $\mu(x_{hi})$ than standard Monte Carlo with $9,500$ replications simulated based on parameters corresponding to x_{hi}. Comparing to the black solid reference line, we see that the green simulation estimators' variances eventually decrease approximately at a rate of n^{-1}, as discussed in connection with Theorem 3.

Next we consider the variance of the SMC, ILR, and MLR estimators for the current-state expected performance $\mu(X_n)$. Figure 4 shows their variances. For the first experiment ($n = 1$), there is no stored simulation output from a previous experiment, so all three estimators are the same. In this example, the green simulation estimators' variances decrease from the beginning and are always less than the SMC variance. Otherwise, the performance of the green simulation estimators is similar to what was seen when the state was fixed. After 100 experiments, the MLR estimator's variance is over 61 times smaller than the SMC estimator's variance.

4.2. Credit Risk Management

In simulation for financial risk management, experiments using the same simulation model are performed periodically, as often as daily. Information observed in the markets is used to update parameters that affect risk, and the simulation model is run again with new parameters. In this example, the risk management simulation measures the credit risk exposure of a portfolio containing loans to companies with listed equity. The asset values of these debtor companies are observable and serve as parameters in the risk model: the lower the asset value of a debtor company, the more likely it is to default in the future. Thus, in our setting of repeated experiments, the current state X_n contains the current asset values of all debtor companies.

In this example, we work with a structural model of default based on the influential work of Merton (1974); for an exposition, see McNeil et al. (2005), for example. At any period n, the asset value of a company equals the sum of its equity and debt values. The equity value can be observed in the stock market and the debt value can be observed from public records, so the asset value can be computed. For simplicity of exposition, we assume that debt remains constant. In this model, the asset value follows geometric Brownian motion, and a company defaults when its asset value falls below its debt. Because geometric Brownian motion is not an ergodic process, Theorems 3

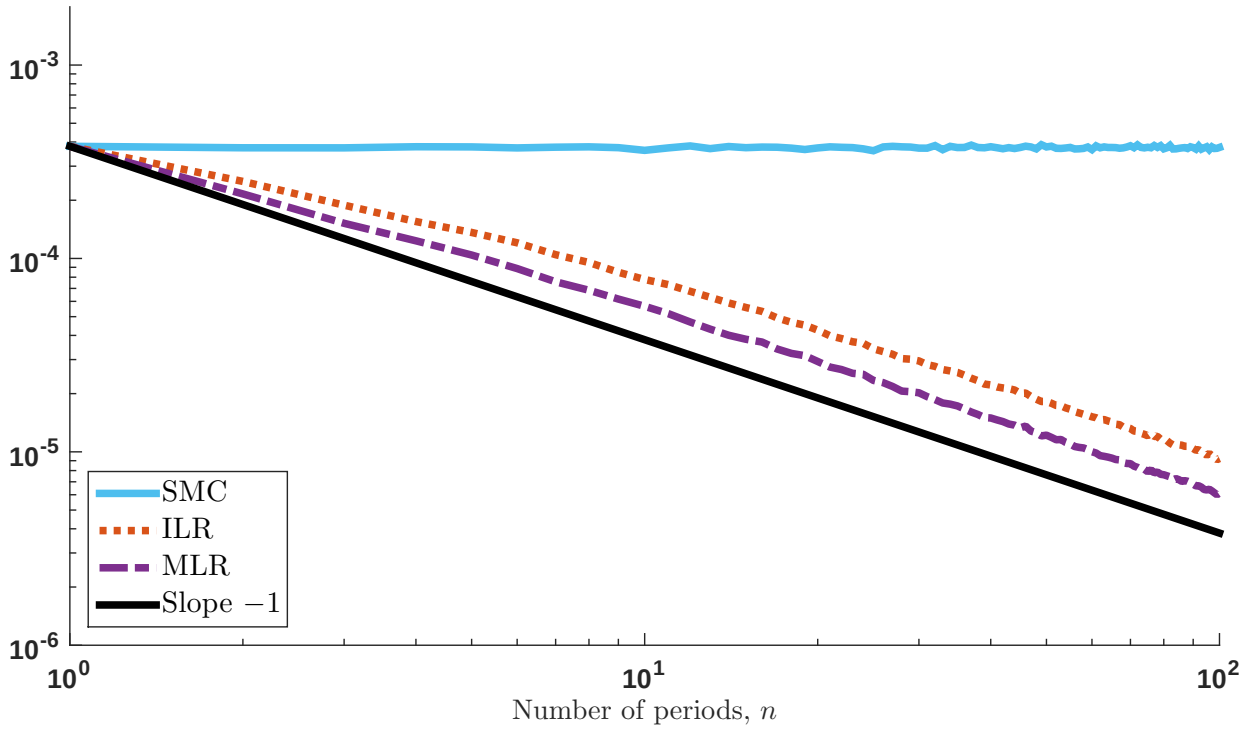

Figure 4 **Log-log plot of estimated variances for current-state estimators.**

and 4 do not apply. As n increases, the current state X_n tends to drift further away from an earlier state, such as X_1. Therefore, intuition suggests that the benefit of reusing old simulation output would diminish over time. We consider this example to show that green simulation nonetheless delivers valuable results.

We consider a loan portfolio whose composition remains constant over time. Many loan portfolios are dynamic: as outstanding loans are being repaid, new loans are being initiated. Despite the dynamic nature of such portfolios, there are lending businesses in which portfolios retain a similar composition over time. For example, in the business of accounts receivable, customers may place regular, periodic orders of the same size, each resulting in payment due in 90 days. An investment fund may target a loan portfolio with fixed characteristics such as maturity and portfolio weights on different types of loans.

In our simplified example, we consider risk management of the value at time horizon $t = 0.5$ years of a portfolio in which there are two loans, both having maturity $T = 5$ years. Simulation experiments are repeated weekly, i.e., with a period of $\Delta t = 1/52$ years. In the nth experiment, the quantity $\mu(X_n)$ being estimated is the conditional probability, given the current asset values X_n, that the cost of defaults and anticipated defaults after t years will exceed a threshold $\kappa = 6$. The random vector Y_n that we simulate in the nth experiment is the asset values $S_t = [S_{t,1}, S_{t,2}]$ at time t, given the current asset values $S_0 = [S_{0,1}, S_{0,2}] = X_n$. For each company $i = 1, 2$, the marginal

distribution of the asset return $S_{t,i}/S_{0,i}$ is lognormal, determined by the drift η_i and volatility ς_i of the geometric Brownian motion for asset value. Specifically, $\eta_1 = 15\%$, $\eta_2 = 10\%$, $\varsigma_1 = 30\%$, and $\varsigma_2 = 20\%$. The joint distribution of the asset returns $S_{t,1}/S_{0,1}$ and $S_{t,2}/S_{0,2}$ is specified by a Student t copula with 3 degrees of freedom and correlation 0.5. The initial asset values of the two companies are $X_{0,1} = 100$ and $X_{0,2} = 90$, and their debt is $D_1 = D_2 = 85$. The loss if company i defaults is denoted a_i, and $a_1 = 5$ and $a_2 = 4$. The discount rate is $r = 5\%$. The cost of a default or anticipated default by company i, as of time t, is

$$L_i = \ell_i(S_{t,i}) = \begin{cases} a_i \text{ if } S_{t,i} < D_i \\ a_i e^{-r(T-t)} \Phi\left(\frac{\ln(S_{t,i}/D_i)+(r-\varsigma_i^2/2)(T-t)}{\varsigma_i\sqrt{T-t}}\right) \text{ if } S_{t,i} \geq D_i \end{cases}.$$

The first line of the formula represents the loss if company i defaults at time t. The second line of the formula represents a risk-neutral conditional expectation, as of time t, of the discounted loss at time T if company i defaults then. The portfolio's loss is $L_1 + L_2$, the sum of losses over the debtor companies. The simulation output $F(Y_n)$ is 1 if $L_1 + L_2 > \kappa$, and 0 otherwise. To clarify the model, Algorithm 4 shows how the standard Monte Carlo simulation works.

Algorithm 4 Standard Monte Carlo Simulation for Credit Risk Management

for $n = 1, 2, \cdots,$ **do**
  Sample state X_n from bivariate lognormal distribution with Student t copula, based on time increment Δt, conditional on X_{n-1}
  Set $\hat{\mu}_{n,r}^{SMC}(X_n) \leftarrow 0$
  for $j = 1, \cdots, r$ **do**
    Sample state $Y_n^{(j)}$ from bivariate lognormal distribution with Student t copula, based on time increment t, conditional on X_n
    Set $L_n^{(j)} \leftarrow \ell_i(Y_{n,1}^{(j)}) + \ell_i(Y_{n,2}^{(j)})$
    if $L_n^{(j)} \leq \kappa$ **then**
      Set $F\left(Y_n^{(j)}\right) \leftarrow 1$
    else
      Set $F\left(Y_n^{(j)}\right) \leftarrow 0$
    end if
    Set $\hat{\mu}_{x_n}^{SMC}(X_n) \leftarrow \hat{\mu}_{x_n}^{SMC}(X_n) + F\left(Y_n^{(j)}\right)$
  end for
  Set $\hat{\mu}_{n,r}^{SMC}(X_n) \leftarrow \hat{\mu}_{n,r}^{SMC}(X_n)/r$ and output
end for

To investigate the effectiveness of green simulation, we performed a sequence of 52 simulation experiments repeated weekly (i.e., $n = 1, 2, \cdots, 52$) with $r = 1,000$ replications each. Using the same

sample path $\{X_n : n = 1, 2, \cdots, 100\}$ and the same simulation output $\{Y_n^{(j)} : n = 1, 2, \cdots, 100, j = 1, 2, \cdots, 100\}$, we evaluated the SMC, ILR, and MLR estimators at each period $n = 1, 2, \cdots, 100$, in the current state X_n. For the purpose of accurately estimating the unconditional variances of the all these estimators, we performed such a sequence of experiments 10,000 times. These 10,000 macro-replications of the sequence of experiments have independent sample paths and simulation output. The estimated variance of a current-state estimator $\hat{\mu}(X_n)$ of $\mu(X_n)$ was $\sum_{k=1}^{10,000} (\hat{\mu}^{(k)}(X_n^{(k)}) - \mu(X_n^{(k)}))^2/10,000$. These estimated variances appear in a log-log plot in Figure 5, along with vertical error bars representing their 95% approximate-normal confidence intervals.

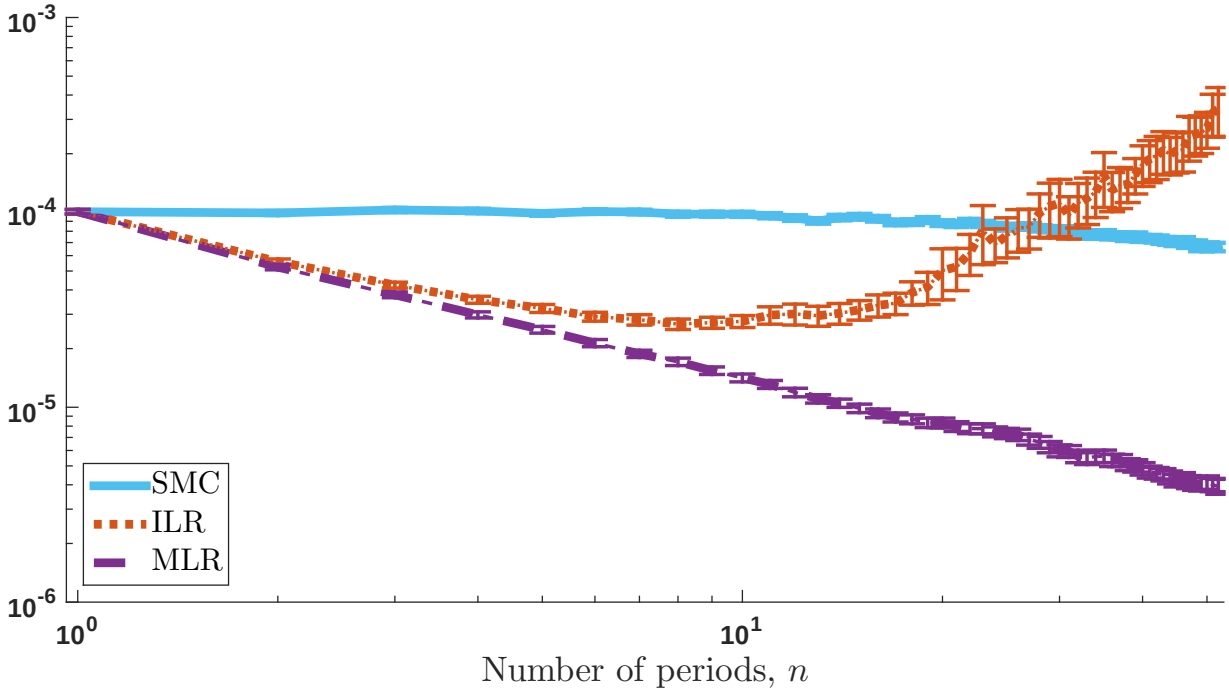

Figure 5 **Log-log plot with error bars for estimated variances for current state estimators.**

Figure 5 shows behavior for the MLR estimator similar to what was seen for the previous example in Figure 4. The MLR estimator's variance is less than the SMC estimator's variance, and it decreases as the number n of repeated experiments increases. By $n = 52$, it is over 17 times smaller than the SMC estimator's variance. However, in Figure 5, we see effects of the non-ergodic nature of the state process $\{X_n : n = 1, 2, \cdots\}$. Due to the positive drifts of the asset prices, debtor companies tend to become less likely to default, so the SMC variance decreases slightly as the number of periods n increases, instead of forming a straight line as in Figure 4. The dramatic effect is on the behavior of the ILR estimator. Its variance decreases over the first 6 experiments, but after about 20 experiments, its variance increases again. Eventually, its variance exceeds the SMC estimator's variance. Apparently, the difference between sampling distributions for state X_1 and

X_{52} is likely to become so large that the use of likelihood ratios in the ILR estimator (4) inflates variance. The failure of the ILR estimator and the success of the MLR estimator in this example demonstrate the practical importance of the MLR estimator. The evident disutility of some of the old simulation output in this example also raises a future research question in green simulation: how to determine which old simulation output is worthwhile to reuse in estimating the expected performance in the current state.

5. Conclusions and Future Research

In Section 3, we established theorems about the convergence of green simulation estimators as the number of repeated experiments increases. We tested their practical performance for small and moderate numbers of repeated experiments in two examples in Section 4. In the example of Section 4.1, the conditions of the theorems held. Our ILR and MLR estimators were both successful in significantly reducing variance compared to standard Monte Carlo, but MLR was somewhat better. In the example of Section 4.2, the conditions of the theorems did not hold. The MLR estimator was successful in significantly reducing variance, whereas the ILR estimator had problems. Consequently, we recommend the MLR estimator for doing green simulation in the setting analyzed in this article: where simulation experiments are repeated with changes to parameters of distributions of random variables generated in the simulation. The variance reduction achieved by green simulation depends on several aspects of the particular example: the stochastic process that describes changing parameters, the particular distributions whose parameters change, and the number of repeated experiments. Our theorems show that, when their conditions hold, green simulation eventually becomes greatly superior to standard Monte Carlo. Our experiment results suggest that green simulation is extremely promising: in the only two examples that we investigated, the MLR estimator achieved variance lower than standard Monte Carlo by factors of 17 and 61 after a moderate number of repeated experiments.

Because green simulation is a new paradigm, there are several good directions for future research. Here we call attention to a few that are most relevant to this article.

Some future research topics are relevant to the specific methods proposed in this article. We found the MLR estimator to be satisfactory for our purposes. However, further enhancements have been considered in the literature on importance sampling. For example, Hesterberg (1995) investigated different ways to normalize weights, and Owen and Zhou (2000) proposed to use likelihoods that appear in the MLR estimator as control variates. At the end of Section 4.2, the experiment results raised the question of which old simulation output to reuse in green simulation. This question is worthy of investigation in connection with the methods proposed in this article and also with other methods. In general, there are two possible drawbacks to using all of the old simulation output.

One is that some of the old simulation output makes an estimator worse if it is reused than if it is not reused (as seen in Section 4.2). The other is that if the amount of old simulation output is extremely large, reusing more of it generates diminishing returns in terms of improved estimator quality compared to the computational cost of reusing it. The theoretical and practical benefit of various green simulation methods could be enhanced by good rules for selecting the subset of old simulation output to reuse.

Various green simulation methods are possible. Besides those proposed in this article, others should be investigated because they would be more broadly applicable or more effective in some cases. In this article, we used the likelihood ratio method for green simulation. The likelihood ratio method is directly applicable only when simulation experiments are repeated with changes to parameters of distributions of random variables generated in the simulation, not when a changing parameter affects something other than a distribution, e.g., numbers of servers or sizes of buffers in a simulation of a queuing system. Even if these methods are applicable, they would not be highly effective if it is unlikely to visit a state that is sufficiently similar to a previously visited state, where similarity is measured according to (5). Rubinstein and Shapiro (1993) address the likelihood ratio method's effectiveness and extensions of its applicability. More broadly applicable green simulation methods can be designed based on metamodeling. We have obtained encouraging initial results based on stochastic kriging (Ankenman et al. 2010). Green simulation based on Database Monte Carlo, used as in Rosenbaum and Staum (2015), is also more broadly applicable.

We focused on showing that when old simulation output is reused well, it provides greater accuracy when combined with a new simulation experiment than would be achieved by the same new simulation alone. We analyzed how the accuracy of an answer to the current question improves as the number of repeated experiments increases. However, it might be possible to answer the current question sufficiently accurately with no further experiment. If a new simulation experiment was indeed required, one could design it in light of the current question and the currently available information. This leads to future research in experiment design not from a blank slate. Also, take consideration of possible future questions when designing the current one, in light of knowledge of the state process.

Appendix. Verifying the Conditions of Theorem 3 and Theorem 4 for the Catastrophe Bond Pricing Example

In this appendix, we verify the conditions of Theorem 3 and Theorem 4 for the catastrophe bond (CAT bond) pricing example. In this example, the underlying state process $\{X_n : n = 1, 2, \cdots\}$ is AR(1), so it is ergodic. Recall that the loss in the CAT bond pricing example is given by $Y_n = \sum_{i=1}^{M_n} Z_n^i$. Conditional on $X_n = x$, and denoting $\phi(x) = (\lambda, \theta)$, the number M_n of claims is Poisson distributed with mean λ and independent of

$\{Z_n^i : i = 1, \cdots, M_n\}$, which are independent random variables, exponentially distributed with mean θ. The conditional distribution of Y_n given X_n places probability mass $h(0;x) = e^{-\lambda}$ on $y = 0$ and has probability density

$$h(y;x) = \sum_{m=1}^{\infty} \frac{\lambda^m}{m!} e^{-\lambda} \frac{y^{m-1} e^{-y/\theta}}{\Gamma(m)\theta^m} = \sqrt{\frac{\lambda}{\theta y}} e^{-\lambda - y/\theta} I_1\left(2\sqrt{\frac{\lambda y}{\theta}}\right)$$

for $y > 0$, where Γ is the Gamma function and I_1 is the modified Bessel function of the first kind of order 1.

We will first establish two lemmas that are useful for verifying the conditions of Theorems 3 and 4 for the CAT bond pricing example. Define the domain $\mathcal{K} = (0,\infty) \times (0,\infty) \times \mathbb{R}$ and the functions $A : \mathcal{K} \mapsto \mathbb{R}$ and $a : \mathcal{K} \times (0,\infty) \mapsto \mathbb{R}$ by

$$A(k) = \int_0^\infty a(k,y)dy, \quad \text{where} \quad a(k,y) = \frac{e^{-k_1 y + k_3\sqrt{y}}}{\sqrt{y}} I_1(k_2\sqrt{y}) \geq 0. \tag{15}$$

LEMMA 1. *The function A is continuous on $\mathcal{K}$.*

Proof. For any $k, k' \in \mathcal{K}$, we have

$$|A(k) - A(k')| \leq \textstyle\int_N^\infty a(k,y)dy + \int_N^\infty a(k',y)dy + \int_0^N |a(k,y) - a(k',y)|\, dy.$$

For any $k \in \mathcal{K}$ and $\epsilon > 0$, we will show that there exist $N > 0$ and $\delta > 0$ such that the right side is bounded above by ϵ if $\|k - k'\|_2 < \delta$. First, we will show that, for any $k \in \mathcal{K}$ and $\epsilon > 0$, there exists $N_1 > 0$ such that $\int_{N_1}^\infty a(k,y)dy \leq \epsilon/3$. Applying the same argument to $k' \in \mathcal{K}$, there exists N_2 such that $\int_{N_2}^\infty a(k',y)dy \leq \epsilon/3$. We then let $N = \max\{N_1, N_2\}$. Finally, we show that for any $k \in \mathcal{K}$, $\epsilon > 0$, and $N > 0$, there exists some $\delta > 0$ such that $\int_0^N |a(k,y) - a(k',y)|\, dy \leq \epsilon/3$ if $\|k - k'\|_2 \leq \delta$.

First, it is proved by Luke (1972) that $\Gamma(\nu+1)(2/y)^\nu I_\nu(y) < \cosh(y)$ for $y > 0$ and $\nu > -1/2$. Taking $\nu = 1$ in this inequality, and observing that $\cosh(y) < e^y$, we have $I_1(y) < (y/2)e^y$. Let $\tilde{k}_1 = k_1/2 > 0$. Then

$$a(k,y) = \frac{e^{-k_1 y + k_3\sqrt{y}}}{\sqrt{y}} I_1(k_2\sqrt{y}) < \frac{e^{-k_1 y + k_3\sqrt{y}}}{\sqrt{y}} \left(\frac{k_2\sqrt{y}}{2} e^{k_2\sqrt{y}}\right) \leq \frac{k_2 e^{(k_2+k_3)^2/(4\tilde{k}_1)}}{2} e^{-\tilde{k}_1 y} =: Ce^{-\tilde{k}_1 y},$$

where the second inequality holds because $(k_2 + k_3)\sqrt{y} - \tilde{k}_1 y \leq (k_2 + k_3)^2/(4\tilde{k}_1)$, and the constant $C > 0$ is defined for ease of notation. Therefore $\int_N^\infty a(K,y)dy \leq C\int_N^\infty e^{-\tilde{k}_1 y}dy = C(e^{-N\tilde{k}_1}/\tilde{k}_1)$. Take

$$N_1 = -\frac{\ln[\epsilon\tilde{k}_1/3C]}{\tilde{k}_1} = -\frac{2\ln[(\epsilon k_1)/(6C)]}{k_1}.$$

Then $\int_{N_1}^\infty a(K,y)dy \leq \epsilon/3$.

The function I_1 is a solution of Bessel's differential equation, so it is continuous on $(0,\infty)$. Consequently, the function $\tilde{a} : \mathcal{K} \times (0,\infty) \mapsto \mathbb{R}$ defined as $\tilde{a}(k,y) := e^{-k_1 y + k_3\sqrt{y}} I_1(k_2\sqrt{y}) = a(k,y)\sqrt{y}$ is continuous on $\mathcal{K} \times [0,\infty)$. Choose any $\delta_0 > 0$ and define the compact neighborhood $\mathcal{N}_k(\delta_0) := \{k' \in \mathcal{K} : \|k - k'\|_2 \leq \delta_0\}$. The function $\tilde{a}$ is continuous on the compact set $\mathcal{N}_k(\delta_0) \times [0,N]$. Therefore, it is uniformly continuous on $\mathcal{N}_k(\delta_0) \times [0,N]$. Consequently, there exists $\delta \in (0,\delta_0]$ such that, for all $(k,y),(k',y') \in \mathcal{N}_k(\delta_0) \times [0,N]$ that satisfy $\|(k,y) - (k',y')\|_2 \leq \delta$, we have $|\tilde{a}(k,y) - \tilde{a}(k',y')| \leq \epsilon/(6\sqrt{N})$. Therefore, for any $k' \in \mathcal{K}$ such that $\|k - k'\|_2 \leq \delta$, we have

$$\int_0^N |a(k,y) - a(k',y)|\, dy = \int_0^N \frac{1}{\sqrt{y}} |\tilde{a}(k,y) - \tilde{a}(k',y)|\, dy \leq \frac{\epsilon}{6\sqrt{N}} \int_0^N \frac{1}{\sqrt{y}} dy = \frac{\epsilon}{3}.$$

Define the domain $\mathcal{K}^B = (0,\infty)\times(0,\infty)\times(0,\infty)$ and the functions $B:\mathcal{K}^B \mapsto \mathbb{R}$ and $b:\mathcal{K}^B\times(0,\infty)\mapsto\mathbb{R}$ by

$$B(k) = \int_0^\infty b(k,y)dy, \quad \text{where} \quad b(k,y) = \frac{e^{-k_1 y}}{\sqrt{y}} \frac{[I_1(k_2\sqrt{y})]^2}{I_1(k_3\sqrt{y})} \geq 0.$$

LEMMA 2. *If $\bar{\mathcal{K}} \subset \mathcal{K}$ is compact, then* $\sup\{A(K)|K\in\bar{\mathcal{K}}\} < \infty$. *If $\bar{\mathcal{K}}^B \subset \mathcal{K}^B$ is compact, then* $\sup\{B(K)|K\in\bar{\mathcal{K}}^B\} < \infty$.

Proof. Because A is continuous in $\mathcal{K}$, by Lemma 1, and $\bar{\mathcal{K}}\subset\mathcal{K}$ is compact, it follows that $\sup\{A(K)|K\in\bar{\mathcal{K}}\} = \max\{A(K)|K\in\bar{\mathcal{K}}\} < \infty$.

For any $k\in\mathcal{K}^B$, let $k^* = \max\{k_2,k_3\}$ and $k_* = \min\{k_2,k_3\}$. Then it follows from Theorem 2.1 of Laforgia (1991) that

$$b(k_1,k_2,k_3,y) < \frac{e^{-k_1 y}}{\sqrt{y}}\left[e^{2(k^*-k_*)\sqrt{y}}\frac{k^*}{k_*}\right] I_1(k_2\sqrt{y}) = a(k_1,k_2,2(k^*-k_*),y).$$

Therefore $B(k_1,k_2,k_3) \leq A(k_1,k_2,2(k^*-k_*))$ for any $k\in\mathcal{K}^B$. Moreover, the compactness of $\bar{\mathcal{K}}^B$ implies the compactness of the set

$$\mathcal{K}^* := \{(k_1,k_2,2(k^*-k_*))|(k_1,k_2,k_3)\in\bar{\mathcal{K}}^B, k^* = \max\{k_2,k_3\}, k_* = \min\{k_2,k_3\}\},$$

which is a subset of $\mathcal{K}$. Therefore $\sup\{B(K)|K\in\bar{\mathcal{K}}^B\} \leq \sup\{A(K)|K\in\mathcal{K}^*\} < \infty$.

PROPOSITION 1. *In the catastrophe bond example, if $\bar{\lambda} \geq \underline{\lambda} > 0$ and $2\underline{\theta} > \bar{\theta} \geq \underline{\theta} > 0$, then for any target state $x\in\mathbb{R}^2$, $\int_{\mathbb{R}^2}\sigma_x^2(x')\,d\pi(x') < \infty$ and the sequence $\{\sigma_x^2(X_n), n = 1,2\cdots\}$ is uniformly integrable.*

Proof. Consider any target state $x\in\mathbb{R}^2$ and any sampling state $x'\in\mathbb{R}^2$. The likelihood ratio is $\ell_x(y;x') = h(y;x)/h(y;x')$. Because the simulation output $F(Y_n)$ is between 0 and 1, for all n, the target-x-sample-x' variance $\sigma_x^2(x')$ defined in Equation (5) satisfies

$$0 \leq \sigma_x^2(x') \leq \mathbb{E}\left[(F(Y_n)\,\ell_x(Y_n;x'))^2|X_n = x'\right] \leq \mathbb{E}\left[(\ell_x(Y_n;x'))^2|X_n = x'\right].$$

To establish the desired conclusions, it suffices to show that this conditional second moment has a finite upper bound over $x'\in\mathbb{R}^2$, for any fixed $x\in\mathbb{R}^2$. Denote $(\lambda,\theta) = \varphi(x) > 0$ and $(\lambda',\theta') = \varphi(x')$. We have

$$\mathbb{E}\left[(\ell_x(Y_n;x'))^2|X_n = x'\right] = \left(\frac{e^{-\lambda}}{e^{-\lambda'}}\right)^2 e^{-\lambda'} + \int_0^\infty \left(\frac{h(y;x)}{h(y;x')}\right)^2 h(y;x')dy.$$

The first term is bounded above by $e^{\bar{\lambda}-2\underline{\lambda}}$. For the second term, we have

$$\begin{aligned}
\int_0^\infty \left(\tfrac{h(y;x)}{h(y;x')}\right)^2 h(y;x')dy &= \int_0^\infty \frac{\left[\sqrt{\frac{\lambda}{\theta y}}e^{-\lambda-\frac{y}{\theta}} I_1\left(2\sqrt{\frac{\lambda y}{\theta}}\right)\right]^2}{\sqrt{\frac{\lambda'}{\theta' y}}e^{-\lambda'-\frac{y}{\theta'}} I_1\left(2\sqrt{\frac{\lambda' y}{\theta'}}\right)}dy \\
&= \sqrt{\tfrac{\lambda^2\theta'}{\lambda'\theta^2}}e^{\lambda'-2\lambda}\int_0^\infty \tfrac{1}{\sqrt{y}} e^{-\left(\frac{2\theta'-\theta}{\theta\theta'}\right)y} \frac{\left[I_1\left(2\sqrt{\frac{\lambda y}{\theta}}\right)\right]^2}{I_1\left(2\sqrt{\frac{\lambda' y}{\theta'}}\right)}dy \qquad (16) \\
&= \sqrt{\tfrac{\lambda^2\theta'}{\lambda'\theta^2}}e^{\lambda'-2\lambda} B\left(\tfrac{2\theta'-\theta}{\theta\theta'}, 2\sqrt{\tfrac{\lambda}{\theta}}, 2\sqrt{\tfrac{\lambda'}{\theta'}}\right).
\end{aligned}$$

For all $x'\in\mathbb{R}^2$, we have that $(\lambda',\theta') = \varphi(x')$ is in a compact set $[\underline{\lambda},\bar{\lambda}]\times[\underline{\theta},\bar{\theta}]$ that does not contain zero. On this set, the arguments of B are all bounded so $\left\{\left(\frac{2\theta'-\theta}{\theta\theta'}, 2\sqrt{\frac{\lambda}{\theta}}, 2\sqrt{\frac{\lambda'}{\theta'}}\right)|(\lambda',\theta')\in[\underline{\lambda},\bar{\lambda}]\times[\underline{\theta},\bar{\theta}]\right\} = \bar{\mathcal{K}}^B$ is compact. Thus, it follows from the second claim of Lemma 2 that Equation (16) has a finite upper bound over $x'\in\mathbb{R}^2$.

PROPOSITION 2. *In the catastrophe bond pricing example, if* $\bar{\lambda} \geq \underline{\lambda} > 0$ *and* $2\underline{\theta} > \bar{\theta} \geq \underline{\theta} > 0$*, then Assumptions A1, A2, A3, and A4 in Theorem 4 are satisfied.*

Proof. Since $F(y) = \mathbf{1}_{\{y \leq K\}} + p\mathbf{1}_{\{y > K\}} \in [p,1]$ so $\mu_x = \mathbb{E}[F(Y_x)] \in [p,1]$ for all $x \in \mathcal{X}$ and therefore A1 holds.

Clearly, for $y = 0$, $h(y;\lambda,\theta)$ is bounded by $e^{-\underline{\lambda}}$ and $e^{-\bar{\lambda}}$. For $y \in (0,\infty)$, the modified Bessel function of the first kind $I_1(\sqrt{\lambda s/\theta})$ is bounded over $(\lambda,\theta) \in [\underline{\lambda},\bar{\lambda}] \times [\underline{\theta},\bar{\theta}]$. Therefore A2 holds.

Define $k^* = \max\{2\sqrt{\lambda/\theta} | (\lambda,\theta) \in \Psi\}$, $k_* = \max\{2\sqrt{\lambda/\theta} | (\lambda,\theta) \in \Psi\}$. It follows from Theorem 2.1 in Laforgia (1991) we have, for all $\lambda, \lambda' \in [\underline{\lambda},\bar{\lambda}]$, $\theta, \theta' \in [\underline{\theta},\bar{\theta}]$ and all $y \in (0,\infty)$,

$$\omega(y) = \frac{\sqrt{\frac{\lambda}{\theta y}} e^{-\lambda - \frac{y}{\theta}} I_1\left(2\sqrt{\frac{\lambda y}{\theta}}\right)}{\sqrt{\frac{\lambda'}{\theta' y}} e^{-\lambda' - \frac{y}{\theta'}} I_1\left(2\sqrt{\frac{\lambda' y}{\theta'}}\right)} < \sqrt{\frac{\lambda\theta'}{\lambda'\theta}} e^{\lambda' - \lambda} e^{\left(\frac{1}{\theta'} - \frac{1}{\theta}\right) y} \left(\frac{k^*}{k_*} e^{(k^* - k_*)\sqrt{y}}\right) \leq \tilde{a} e^{\tilde{b} y + (k^* - k_*)\sqrt{y}}$$

where $\tilde{a} = \frac{k^*}{k_*}\sqrt{\frac{\bar{\lambda}\bar{\theta}}{\underline{\lambda}\underline{\theta}}} e^{\bar{\lambda} - \underline{\lambda}}$ and $\tilde{b} = \frac{1}{\underline{\theta}} - \frac{1}{\bar{\theta}}$. Let $\epsilon = \frac{1}{\bar{\theta}} - \frac{1}{2\underline{\theta}} > 0$ and $b = \tilde{b} + \epsilon = \frac{1}{2\underline{\theta}}$. Then we have $\omega(y) < a e^{by}$ for $a = \tilde{a} e^{\frac{(k^* - k_*)^2}{4\epsilon}}$ because $(k^* - k_*)\sqrt{s} - \epsilon s \leq \frac{(k^* - k_*)^2}{4\epsilon}$. Finally, $\omega(0) = e^{\lambda' - \lambda} < a$. Therefore A3 holds.

For A4, consider any $t < 1/\bar{\theta}$ and $x \in \mathcal{X}$, the moment generating function for the compound CAT bond loss is given by

$$M(x,t) = e^{-\lambda} + \int_0^\infty e^{ty} \sqrt{\frac{\lambda}{\theta y}} e^{-\lambda - \frac{y}{\theta}} I_1\left(2\sqrt{\frac{\lambda y}{\theta}}\right) dy = e^{-\lambda}\left(1 + \sqrt{\frac{\lambda}{\theta}} A\left(\frac{1}{\theta} - t, 2\sqrt{\frac{\lambda}{\theta}}, 0\right)\right). \tag{17}$$

For all $x \in \mathbb{R}^2$, we have that $(\lambda,\theta) = \varphi(x)$ is in a compact set $[\underline{\lambda},\bar{\lambda}] \times [\underline{\theta},\bar{\theta}]$ that does not contain zero. On this set, the arguments of A are all bounded, so $\left\{\left(\frac{1}{\theta} - t, 2\sqrt{\frac{\lambda}{\theta}}, 0\right) | (\lambda,\theta) \in [\underline{\lambda},\bar{\lambda}] \times [\underline{\theta},\bar{\theta}]\right\} = \bar{\mathcal{K}}$ is compact. Thus, it follows from the first claim of Lemma 2 that Equation (17) has a finite upper bound over $x \in \mathbb{R}^2$.

Acknowledgments

Portions of this article were published in the *Proceedings of the Winter Simulation Conference* (Feng and Staum 2015). The authors thank Shane Henderson and Barry Nelson for providing references. They are grateful for financial support from the Intel Parallel Computing Center at Northwestern University.

References

Ankenman, Bruce, Barry L. Nelson, Jeremy Staum. 2010. Stochastic kriging for simulation metamodeling. *Operations Research* **58**(2) 371–382.

Barton, Russell R., Barry L. Nelson, Wei Xie. 2014. Quantifying input uncertainty via simulation confidence intervals. *INFORMS Journal on Computing* **26**(1) 74–87.

Beckman, Richard J., Michael D. McKay. 1987. Monte Carlo estimation under different distributions using the same simulation. *Technometrics* **29**(2) 153–160.

Blake, Eric S., Ethan J. Gibney. 2011. *The Deadliest, Costliest, and Most Intense United States Tropical Cyclones from 1851 to 2010 (and Other Frequently Requested Hurricane Facts).* NOAA Technical Memorandum NWS NHC-6, National Weather Service, National Hurricane Center, Miami, Florida.

Coval, Joshua D., Jakub W. Jurek, Erik Stafford. 2009. Economic catastrophe bonds. *The American Economic Review* **99**(3) 628–666.

Dassios, Angelos, Ji-Wook Jang. 2003. Pricing of catastrophe reinsurance and derivatives using the Cox process with shot noise intensity. *Finance and Stochastics* **7**(1) 73–95.

Feng, Mingbin, Jeremy Staum. 2015. Green simulation designs for repeated experiments. L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, M. D. Rossetti, eds., *Proceedings of the 2015 Winter Simulation Conference*. Institute of Electrical and Electronics Engineers, Inc., Piscataway, New Jersey, 403–413.

Glasserman, Paul, Xingbo Xu. 2014. Robust risk measurement and model risk. *Quantitative Finance* **14**(1) 29–58.

Hesterberg, Tim. 1995. Weighted average importance sampling and defensive mixture distributions. *Technometrics* **37**(2) 185–194.

Hesterberg, Timothy Classen. 1988. Advances in importance sampling. Ph.D. thesis, Stanford University.

Kleijnen, Jack P.C., Reuven Y. Rubinstein. 1996. Optimization and sensitivity analysis of computer simulation models by the score function method. *European Journal of Operational Research* **88** 413–427.

Klugman, Stuart A., Harry H. Panjer, Gordon E. Willmot. 2012. *Loss Models: From Data to Decisions*. 4th ed. John Wiley & Sons.

Laforgia, Andrea. 1991. Bounds for modified Bessel functions. *Journal of Computational and Applied Mathematics* **34**(3) 263–267.

Luke, Yudell L. 1972. Inequalities for generalized hypergeometric functions. *Journal of Approximation Theory* **5**(1) 41–65.

Maggiar, Alvaro, Andreas Wächter, Irina S. Dolinskaya, Jeremy Staum. 2015. A derivative-free algorithm for the optimization of functions smoothed via Gaussian convolution using multiple importance sampling. *Optimization Online* Last accessed on Nov 11, 2015.

Martino, Luca, Victor Elvira, David Luengo, Jukka Corander. 2014. An adaptive population importance sampler: Learning from uncertainty. *IEEE Transactions on Signal Processing* **63**(16) 4422 – 4437.

McNeil, Alexander J., Rüdiger Frey, Paul Embrechts. 2005. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, Princeton, NJ, USA.

Merton, Robert C. 1974. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance* **29**(2) 449–470.

Meyn, Sean, Richard L. Tweedie. 2009. *Markov Chains and Stochastic Stability*. 2nd ed. Cambridge University Press, New York, NY, USA.

Munich Re Geo Risks Research. 2015. Loss events worldwide 1980 - 2014, 10 costliest events ordered by overall losses. Tech. rep., Munich Re.

Nummelin, Esa. 2004. *General Irreducible Markov Chains and Non-Negative Operators*, *Cambridge Tracts in Mathematics*, vol. 83. Cambridge University Press.

Owen, Art B., Yi Zhou. 2000. Safe and effective importance sampling. *Journal of the American Statistical Association* **95**(449) 135–143.

Rosenbaum, Imry, Jeremy Staum. 2015. Database Monte Carlo for simulation on demand. L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, M. D. Rossetti, eds., *Proceedings of the 2015 Winter Simulation Conference*. Institute of Electrical and Electronics Engineers, Inc., Piscataway, New Jersey, 679–688.

Rubinstein, Reuven Y., Alexander Shapiro. 1993. *Discrete event systems: sensitivity analysis and stochastic optimization by the score function method*. Wiley.

Staum, Jeremy, Andreas Wächter, Alvaro Maggiar, Mingbin Feng. 2015. Uniform convergence of sample average approximation with adaptive importance sampling. *Optimization Online* Last accessed on Nov 11, 2015.

Veach, Eric. 1997. Robust Monte Carlo methods for light transport simulation. Ph.D. thesis, Stanford University.

Veach, Eric, Leonidas J Guibas. 1995. Optimally combining sampling techniques for Monte Carlo rendering. *SIGGRAPH 95 Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques*. ACM, New York, NY, USA, 419–428.

Xie, Wei, Barry L. Nelson, Russell R. Barton. 2014. A Bayesian framework for quantifying uncertainty in stochastic simulation. *Operations Research* **62**(6) 1439–1452.