Proceedings of the 2020 Winter Simulation Conference
K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, eds.

REUSING SIMULATION OUTPUTS OF REPEATED EXPERIMENTS VIA LIKELIHOOD RATIO REGRESSION

Ben Feng

Department of Statistics and Actuarial Science
University of Waterloo
200 University Ave W
Waterloo, ON N2L 3G1, CANADA

Guangxin Jiang

School of Management
Harbin Institute of Technology
2 Yikuang Street
Harbin, Heilongjiang, 150001, CHINA

ABSTRACT

Simulation experiments are sometimes conducted periodically, with updated parameters of the stochastic system being modeled. Storing and reusing the past simulation experiment data may be helpful for the current simulation experiment. In this paper, we consider reusing simulation data in repeated experiments to develop high-quality metamodels. Specifically, we propose a generalized least square regression metamodel whose input data include simulation outputs from the current and the past experiments. Moreover, the past simulation outputs are reused via the likelihood ratio method. Asymptotic variance analysis is provided to show the benefits of reusing past simulation data in prediction accuracy, and the numerical results show the effectiveness of the proposed method.

1 INTRODUCTION

Simulation is a powerful tool to study different properties and performances for stochastic systems. For example, investors can use dynamic asset simulation models for pricing complex financial derivatives and estimating various risk measures for financial portfolios. In the current COVID-19 pandemic, we see that simulation models can help health experts to predict the spread of the virus and guide major social and economic decisions. In practice, simulation experiments are often ran periodically as the stochastic system evolves over time. For example, in many financial institutions simulations are ran monthly or even daily to reflect new market information and to constantly monitor the changing risk positions. We also see health experts around the world frequently update their simulation models and provide new projections on the numbers of infected patients and deaths due to COVID-19.

Typically, when the stochastic system of interest evolves, the simulation model is updated accordingly. Then, new simulation experiments are conducted from blank slate, i.e., simulation outputs from past simulation experiments are usually, if not always, discarded. This is a wasteful use of simulation outputs, especially when model updates are minimal so outputs from past simulation experiments can still be useful for the current simulation model. Also, simulation experiments are sometimes computationally expensive so the outputs are scarce resources. This scarcity further justifies the reusage of existing simulation outputs from past experiments. Lastly, past experiments can be viewed as some offline simulations that aid the online decision-making in the current experiment. Specifically, simulation outputs from past experiments can be used to calibrate predictive models (offline simulation stage), which are then used to make predictions about the current simulation problem (online application stage).

This study is based on two main inspirations: *Green simulation* (Feng and Staum 2015; Feng and Staum 2017) and offline simulation online application (OSOA) (Hong and Jiang 2019; Jiang et al. 2019) . Green simulation is a novel experiment design paradigm that views simulation outputs as valuable resources that should be recycled and reused to improve the efficiency of repeated experiments. Different implementations

of green simulation designs, such as likelihood ratio based methods (Feng and Staum 2017), control variate based methods (Feng and Staum 2016), and metamodeling based methods (Dong et al. 2018) have been studied in the literature. The likelihood ratio method (Beckman and McKay 1987; L'Ecuyer 1990), which is mathematically similar to the well-known importance sampling variance reduction technique (Owen 2013; Glasserman 2013). The proposed method in this study differs from (Feng and Staum 2017) in that the likelihood ratio estimates are not used directed, but rather used as data points for a regression metamodel, which is different from the likelihood ratio metamodel in (Dong et al. 2018). Lastly, unlike (Feng and Staum 2016), control variate is not considered in this study. Green simulation via likelihood ratio method adapts the likelihood ratio method in the context of repeated experiments and study the temporal convergence properties as the experiments are repeated indefinitely. Besides estimation problems, green simulation via likelihood ratio has also been applied in optimization (Maggiar et al. 2018; Feng et al. 2018) and input uncertainty quantification (Zhou and Liu 2018; Feng and Song 2019).

The OSOA framework proposes to run experiments at different values of the covariates and build predictive models (offline simulation stage) so that the performance of the models may be predicted once the values of the covariates are observed. When the real-time problems arises (online application stage), the predictive models are used to assist the decision-making process. In a financial engineering application, Jiang et al. (2019) proposed a logistic regression approach to monitor online financial risk measures, and derived several enhanced methods to improve the prediction accuracy; Jiang et al. (2020) adopted the OSOA with stochastic cokriging in financial derivatives pricing and risk management. The OSOA has also been applied in simulation optimization, see Shen et al. (2018) and Gao et al. (2019).

The contribution of this paper is as follows. We combine green simulation and OSOA to improve the efficiency of simulation experiments. Specifically, we develop accurate functional approximation for stochastic systems via regression by reusing simulation outputs that were generated in the past. The stochastic systems of interest evolve over time in a certain way so that past simulation outputs can be appropriately re-weighted via likelihood ratios to produce unbiased estimates of the current stochastic system's performances. These likelihood-ratio weighted estimates, combined with the current simulation outputs, are used to calibrate a regression metamodel. We show that such augmentation of data set always improves the regression's prediction accuracy, compared to only using the current simulation outputs. The remainder of this article is organized as follow: Section 2 articulates the mathematical settings of our study and reviews two main building blocks of our proposal. Section 3 provides the methodological details and variance analysis of the likelihood ratio regression method. Section 4 presents two numerical examples to illustrates the advantages of the likelihood ratio regression method. Section 5 concludes

2 SETTINGS AND PROBLEM FORMULATION

We are interested in approximating a parameterized function $f : (\mathscr{X}, \Theta) \mapsto \mathbb{R}$ given by

$$f_\theta(x) = \int_\Xi F(x,\xi) h_\theta(\xi) d\xi = \mathrm{E}[F(x,\xi_\theta)], \tag{1}$$

where $F(x,\cdot)$ is measurable for all x and $h_\theta(\xi)$ is a probability density or mass function for the random variable ξ_θ; some or all of x, θ, and ξ can be vectors. The simulation model $F(x,\xi)$ is a semi-black-box model where the users have access to the embedded random variable ξ_θ but the input-output relationship F is unknown or complex. The design variable x encodes the deterministic components of the simulation model that does not affect the model's embedded random variables. In contrast, θ encodes the distributional parameters that govern the random variable ξ_θ in the stochastic simulation.

We are interested in repeated experiments where the distributional parameter θ changes in each experiment. For example, the experiment may be repeated periodically with updated parameter θ based on a new real-world information collected between two experiments. In each experiment, the expected simulation performance $f_\theta(x)$ for multiple values of x or even the whole response surface for all $x \in \mathscr{X}$ is of interest to perform what-if analysis, sensitivity analysis, and/or optimization. For example, suppose

an investor rebalances her portfolio weekly. Every week, she runs simulation experiments to estimate the sufficient amount of cash that she needs to set up the portfolio in the next week. The simulation model projects many plausible asset values in a week (i.e., the x's) then estimates the corresponding portfolio value (i.e., the $f_\theta(x)$'s) using updated market volatility parameter (i.e., the θ).

For simplicity, in this study we consider one past simulation experiment that was ran using parameter θ_0 and one current experiment whose parameter has been updated as θ_1; the proposed methodology can be extended to more than two repeated experiments. In a typical workflow for repeated experiments, new simulation experiments are ran using the updated parameter θ_1; past simulation outputs associated with parameter θ_0 are discarded. In this section, we develop a novel simulation methodology that reuses these past simulation outputs to enhance the accuracy and efficiency of the current experiment.

2.1 Regression using Simulation Outputs

Consider a fixed parameter $\theta \in \Theta$, the function $f_\theta(x)$ can be approximated via regression using the simulation outputs $F(x,\xi_\theta)$. For ease of exposition, consider a linear model

$$Y_\theta(x) := f_\theta(x) = [\boldsymbol{\phi}(x)]^\top \boldsymbol{\beta}_\theta, \tag{2}$$

where $\boldsymbol{\phi}(\cdot) = (\phi_1(\cdot),\ldots,\phi_k(\cdot))^\top$ is a given set of basis functions and $\boldsymbol{\beta}_\theta = (\beta_{\theta,1},\ldots,\beta_{\theta,k})^\top$ are the corresponding regression coefficients. This is a linear model in $\boldsymbol{\phi}$, but the basis functions $\boldsymbol{\phi}(\cdot)$ can be nonlinear functions of the covariate x. Asymptotic properties for regression methods have long been studied (White 1984), some of which are summarized in Section 3. Simulation outputs satisfy the assumptions required of the data by regression models, sometimes trivially. Regression using simulation outputs has been studied and used in financial applications such as American option pricing (Longstaff and Schwartz 2001; Tsitsiklis and Van Roy 2001; Stentoft 2004) and risk measurement (Broadie et al. 2015; Jiang et al. 2019).

Let $x_1,\ldots,x_m \in \mathscr{X}$ be a given set of covariates that are of interest in the current experiment. Suppose for every covariate $x_1,\ldots,x_m$ one can obtain point estimates $\widehat{Y}_\theta(x_1),\ldots,\widehat{Y}_\theta(x_m)$ via simulation, then the simulation outputs can be organized as a data set $\mathscr{Y}_\theta := \{(\widehat{Y}_\theta(x_j),x_j) : j = 1,\ldots,m\}$. This data set can then be used to calibrate the regression coefficients $\widehat{\boldsymbol{\beta}}_\theta = \widehat{\boldsymbol{\beta}}(\mathscr{Y}_\theta)$ via the least-square regression, the maximum likelihood estimation (MLE), etc. For example, for any design point $x \in \mathscr{X}$, a standard Monte Carlo estimator for $Y_\theta(x)$ is

$$\widehat{Y}^{MC}_\theta(x) = \frac{1}{n}\sum_{i=1}^{n} F(x,\xi^{(i)}_\theta), \quad \xi^{(i)}_\theta \overset{i.i.d.}{\sim} h_\theta,\ i = 1,\ldots,n. \tag{3}$$

We assume that $\mathrm{Var}[F(x,\xi_\theta)] < \infty$ so the Monte Carlo estimator (3) has a finite variance for all $x \in \mathscr{X}$ and all $\theta \in \Theta$. Then, after running simulation experiments at x_j, $j = 1,\ldots,m$, the simulation outputs $\mathscr{Y}^{MC}_\theta := \{(\widehat{Y}^{MC}_\theta(x_j),x_j) : j = 1,\ldots,m\}$ form a data set for regression. The Monte Carlo estimator (3) is unbiased, so the sample average can be written as

$$\widehat{Y}^{MC}_\theta(x_j) = Y_\theta(x_j) + \varepsilon(x_j), \quad \mathrm{E}[\varepsilon(x_j)] = 0,\ \forall j = 1,\ldots,m.$$

Setting $\theta = \theta_1$ for the current experiments, the regression coefficient can be estimated by using these simulation outputs, i.e., $\widehat{\boldsymbol{\beta}}^{MC}_{\theta_1} := \widehat{\boldsymbol{\beta}}(\mathscr{Y}^{MC}_{\theta_1})$.

It is natural to use only the current simulation outputs $\mathscr{Y}^{MC}_{\theta_1}$ to calibrate the regression model $Y_{\theta_1}(x)$. However, in the settings of repeated experiments, past simulation outputs associated with a different parameter θ_0 can also be useful for approximating $f_{\theta_1}(x)$. Firstly, the simulation model F is the same in both the past and the current experiments, so the past simulation outputs contain useful information about the simulation model. Also, in the regression model (2), the basis functions are the same in both the past and the current experiments; only the regression parameters $\boldsymbol{\beta}_\theta$ vary with θ. Lastly, past simulation outputs can be conveniently stored in a database and be rapidly retrievable without running any simulation experiments. So, rather than discarding these relevant and cheaply available past simulation outputs, we propose to reuse them.

2.2 Reusing Simulation Outputs via Likelihood Ratio Estimators

Consider any fixed design point $x \in \mathscr{X}$. Suppose the past simulation experiment was ran at x using parameter θ_0 and the simulation outputs $\{(F(x,\xi_{\theta_0}^{(i)}),x) : i = 1,\ldots,n\}$ are stored and readily accessible. The parameter for the current experiment is θ_1.

Recall the probability density function (pdf) $h_\theta(\xi)$ in (1). We assume throughout this study that h_θ is a well-defined pdf for all $\theta \in \Theta$ and that all these pdfs have a common support Ξ. Then, for any $x \in \mathscr{X}$, the following identity holds:

$$f_{\theta_1}(x) = \int_\Xi F(x,\xi)h_{\theta_1}(\xi)d\xi = \int_\Xi F(x,\xi)\frac{h_{\theta_1}(\xi)}{h_{\theta_0}(\xi)}h_{\theta_0}(\xi)d\xi = \mathrm{E}\left[F(x,\xi_{\theta_0})\frac{h_{\theta_1}(\xi_{\theta_0})}{h_{\theta_0}(\xi_{\theta_0})}\right]. \tag{4}$$

Based on (4), a likelihood ratio (LR) estimator for $f_{\theta_1}(x)$ is given by

$$\widehat{Y}^{LR}_{\theta_1,\theta_0}(x) = \frac{1}{n}\sum_{i=1}^{n} F(x,\xi_{\theta_0}^{(i)})\frac{h_{\theta_1}(\xi_{\theta_0}^{(i)})}{h_{\theta_0}(\xi_{\theta_0}^{(i)})}, \quad \xi_{\theta_0}^{(i)} \overset{i.i.d.}{\sim} h_{\theta_0},\ i = 1,\ldots,n. \tag{5}$$

Mathematically, the LR estimator (5) is identical to a standard importance sampling estimator, but they are used in different situations and for different purposes: Importance sampling selects a good sampling parameter θ_0 so that the resulting estimator for $f_{\theta_1}(x)$ has a small variance. Users of LR estimator could not select the sampling parameter θ_0 in the past experiment, but rather aim to reuse the past simulation outputs to estimate $f_{\theta_1}(x)$ for the current parameter θ_1.

Clearly, by identity (4), the LR estimator (5) is unbiased. So, suppose the past experiment was ran using parameter θ_0 at different covariates $x_1^0,\ldots,x_{m_0}^0$, and we can write

$$\widehat{Y}^{LR}_{\theta_1,\theta_0}(x_j^0) = Y_{\theta_1}(x_j^0) + \varepsilon(x_j^0), \quad \mathrm{E}[\varepsilon(x_j^0)] = 0,\ j = 1,\ldots,m_0.$$

Moreover, if the LR-weighted simulation outputs $F(x_j^0,\xi_{\theta_0})\frac{h_{\theta_1}(\xi_{\theta_0})}{h_{\theta_0}(\xi_{\theta_0})}$ all have finite variances, the error terms $\varepsilon(x_j^0)$, $j = 1,\ldots,m$ are asymptotically normally distributed as $n \to \infty$ by the central limit theorem. Lastly, assuming the simulation experiments at different points x_j^0 were ran independently, the error terms are mutually independent. We denote the LR-weighted simulation outputs by $\mathscr{Y}^{LR}_{\theta_1,\theta_0} := \left\{\left(\widehat{Y}^{LR}_{\theta_1,\theta_0}(x_j^0),x_j^0\right) : j = 1,\ldots,m_0\right\}$. The covariates in the past experiment can be different from those in the current experiment; we assume they are all distinct for simplicity.

We propose to combine the LR-weighted simulation outputs $\mathscr{Y}^{LR}_{\theta_1,\theta_0}$ with the current simulation outputs $\mathscr{Y}^{MC}_{\theta_1}$ to calibrate the regression model $f_{\theta_1}(x)$. Detailed discussions and analysis are provided in in Section 3. Before proceeding, we note some limitations of the regression method and the likelihood ratio method, which reflect the importance of our approach:

- One limitation of the LR estimator (5) is that the simulation output $F(x,\xi_{\theta_0})$ can only be reused to estimate $Y_{\theta_1}(x)$ for the same covariate x; this limitation is rooted in the identity (4) and cannot be lifted easily. In the presence of covariate x that does not affect the distributions of the random variates in a stochastic simulation, (5) cannot be used to approximate the $f_\theta(x')$ for any $x' \neq x$. If only the distributional parameter θ changes between the past and the current experiments, the unbiased LR-based metamodel in Dong et al. (2018) would be applicable.
- One can argue, correctly and understandably, that the distributional parameter θ can be amalgamated into the covariate x in the regression model. Then the resulting regression parameter $\boldsymbol{\beta}$ would not depend on the parameter θ. This argument is mathematically correct, but the suggested implementation can be difficult or undesirable in the setting of repeated experiments.

- The distributional parameter θ is only updated once in each experiment, so one needs to wait for many repetitions of experiments to get simulation outputs associated with a reasonable number of different values of θ. It may not be straightforward to determine the right number of waiting periods. Also, during the waiting period, simulation outputs would likely be wasted.
- Even granted a sufficient amount repeated experiments to justify regression on θ, the current experiment's parameter can be close to or even outside the boundary of the parameters seen in the past. This leads to extrapolating the regression model, which often results in inaccurate predictions. In contrast, users often have some controls over the design points to avoid or to alleviates extrapolation.
- Lastly, viewing θ as merely a covariate in the regression model does not take full advantage of the known the density functions h_θ. In other words, the unbiased LR estimators are overlooked. Discarding these unbiased estimators and instead run an unnecessarily high-dimension regression (with both x and θ as the covariates) appears to be counter-intuitive and undesirable.

3 LIKELIHOOD RATIO REGRESSION

We propose a novel methodology that reuses existing simulation outputs $\mathscr{Y}_{\theta_0}$ to enhance the approximation accuracy of the function $f_{\theta_1}(x)$ for the current parameter θ_1. As alluded in Section 2, we propose to combine the past simulation outputs with the current ones, i.e., $\mathscr{Y}_{\theta_1^+} = \mathscr{Y}^{LR}_{\theta_1,\theta_0} \bigcup \mathscr{Y}^{MC}_{\theta_1}$, then approximate $f_{\theta_1}(x)$ via regression. Intuitively, by reusing the past simulation outputs, the regression parameter estimated from the combined data set, i.e., $\widehat{\boldsymbol{\beta}}_{\theta_1^+} := \widehat{\boldsymbol{\beta}}(\mathscr{Y}_{\theta_1^+})$, produces a more accurate regression model than the one based on only the current data set, i.e., $\widehat{\boldsymbol{\beta}}_{\theta_1} = \widehat{\boldsymbol{\beta}}(\mathscr{Y}^{MC}_{\theta_1})$. This section elaborates the methodological details and provides asymptotic analysis of the proposed method.

3.1 Reusing Simulation Outputs via Likelihood Ratio Regression

Consider the following regression model at given covariates $x_1,\ldots,x_m$,

$$\widehat{Y}_\theta(x_j) := f_\theta(x_j) = [\boldsymbol{\phi}(x_j)]^\top \boldsymbol{\beta}_\theta + \varepsilon(x_j), \tag{6}$$

where the error terms $\varepsilon(x_j)$, $j = 1,\ldots,m$, satisfy $\mathrm{E}[\varepsilon(x_j)] = 0$, $\mathrm{Var}[\varepsilon(x_j)] = \sigma_j^2$, and are mutually independent. Here we use m, with appropriate subscripts, to denote the number of design points in a simulation experiment. For instance, the past and the current experiments have m_0 and m_1 design points, respectively. Consequently the combined data set $\mathscr{Y}_{\theta_1^+}$ has $m_0 + m_1$ covariates in total.

In matrix notation, the linear model (6) is written as $\widehat{\boldsymbol{Y}}_\theta = \boldsymbol{\Phi}\boldsymbol{\beta}_\theta + \boldsymbol{\varepsilon}$, where $\widehat{\boldsymbol{Y}}_\theta = [\widehat{Y}_\theta(x_1),\ldots,\widehat{Y}_\theta(x_m)]^\top$ is an $m \times 1$ vector, $\boldsymbol{\Phi}$ is an $m \times k$ matrix, $\boldsymbol{\beta}_\theta$ a $k \times 1$ vector, and $\boldsymbol{\varepsilon}$ an $m \times 1$ vector. For notational convenience, denote the $m \times m$ covariance matrix of error vector by $\boldsymbol{\Omega} = \mathrm{Var}[\boldsymbol{\varepsilon}]$. Given a data set $\mathscr{Y}_\theta := \{(\widehat{Y}_\theta(x_j), x_j) : j = 1,\ldots,m\}$, one can organize the response variables into a vector $\widehat{\boldsymbol{Y}}_\theta$ and calculate the matrix $\boldsymbol{\Phi}$ using the covariates $x_1,\ldots,x_m$. Suppose the error covariance matrix $\boldsymbol{\Omega}$ is known, then one can show (White 1984, Chapter 1) that the generalized least square (GLS) estimator for $\boldsymbol{\beta}_\theta$ is

$$\widehat{\boldsymbol{\beta}}_\theta := \widehat{\boldsymbol{\beta}}(\mathscr{Y}_\theta) = (\boldsymbol{\Phi}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^\top \boldsymbol{\Omega} \widehat{\boldsymbol{Y}}_\theta. \tag{7}$$

Asymptotic properties for $\widehat{\boldsymbol{\beta}}_\theta$ are summarized in Lemma 1.

Lemma 1 [Paraphrase of Theorem 1.3 in (White 1984)] Suppose that

(i) The data $\mathscr{Y}_\theta$ is generated by $\widehat{\boldsymbol{Y}}_\theta = \boldsymbol{\Phi}\boldsymbol{\beta}_\theta + \boldsymbol{\varepsilon}$ for some $\boldsymbol{\beta}_\theta < \infty$.
(ii) The design matrix $\boldsymbol{\Phi}$ is a finite nonstochastic $m \times k$ matrix.
(iii) The covariance matrix $\boldsymbol{\Omega}$ is finite and positive definite and $\boldsymbol{\Phi}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{\Phi}$ is nonsingular for all $m \geq k$.
(iv) $\mathrm{E}[\boldsymbol{\varepsilon}] = \mathbf{0}$, i.e., every error term has a zero mean.

(v) $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$, i.e., $\boldsymbol{\varepsilon}$ follows a multi-variate normal distribution with zero mean and covariance matrix $\boldsymbol{\Omega}$.

Then

(a) Given (i)-(iii), $\widehat{\boldsymbol{\beta}}_\theta$ exists for all $m \geq k$ and is unique.
(b) Given (i)-(iv), $\mathrm{E}[\widehat{\boldsymbol{\beta}}_\theta] = \boldsymbol{\beta}_\theta$.
(c) Given (i)-(v), $\widehat{\boldsymbol{\beta}}_\theta \sim \mathcal{N}\left(\boldsymbol{\beta}_\theta, (\boldsymbol{\Phi}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{\Phi})^{-1}\right)$.

Methodologically, our proposal is rather straightforward: We propose combine the past and the current simulation outputs to calibrate the regression model (6) for the current parameter θ_1, i.e.,

$$\widehat{\boldsymbol{\beta}}_{\theta_1^+} = \widehat{\boldsymbol{\beta}}(\mathscr{Y}_{\theta_1^+}). \tag{8}$$

It is then of our interest to analyze the properties of the combined GLS estimator (8). The variance analysis in the following discussions show that $\widehat{\boldsymbol{\beta}}_{\theta_1^+}$ is asymptotically more accurate than the GLS estimator $\widehat{\boldsymbol{\beta}}_{\theta_1} = \widehat{\boldsymbol{\beta}}(\mathscr{Y}^{MC}_{\theta_1})$ that uses only the current simulation outputs. For notational convenience, in subsequent discussions we add the subscripts 0, 1, and 1^+ to our notations, such as $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Omega}_{1^+}$, to denote quantities associated with the past data set $\mathscr{Y}_{\theta_1}$, the current data set $\mathscr{Y}_{\theta_1}$, and the combined data set $\mathscr{Y}_{\theta_1^+}$, respectively.

3.2 Variance Analysis of Likelihood Ratio Regression

The following assumptions were alluded to previous discussions and are restated for clarity. These assumptions are standard in analyzing likelihood ratio estimators, as they ensure that the LR estimator (5) is unbiased and has a finite variance.

Assumption 1 Assume that the density function $h_\theta(\xi)$ has a common support Ξ for all $\theta \in \Theta$. Also, assume that $\mathrm{Var}[F(x, \xi_{\theta'}) \frac{h_\theta(\xi_{\theta'})}{h_{\theta'}(\xi_{\theta'})}] < \infty$ for all $\theta, \theta' \in \Theta$ and for all $x \in \mathscr{X}$.

Assumption 1 ensures that all our LR estimators are well-defined and that the error covariance matrix $\boldsymbol{\Omega}_{1^+}$ is finite. In subsequent analysis, we assume that the basis functions $\boldsymbol{\phi}(\cdot) = (\phi_1(\cdot), \ldots, \phi_k(\cdot))^\top$ and the corresponding design matrix $\boldsymbol{\Phi}_{1^+}$ satisfy (i)-(iii). So, by (a) in Lemma 1, $\widehat{\boldsymbol{\beta}}_{\theta_1^+}$ exists and is unique. For condition (iv), both the Monte Carlo estimators $\widehat{Y}_\theta(x_j)$, $j = 1, \ldots, m$, and the LR estimators $\widehat{Y}^{LR}_{\theta,\theta'}(x_j^0)$, $j = 1, \ldots, m_0$, are unbiased, so $\mathrm{E}[\boldsymbol{\varepsilon}] = \mathbf{0}$ in the combined data. Therefore, $\mathrm{E}[\widehat{\boldsymbol{\beta}}_{\theta_1^+}] = \boldsymbol{\beta}_{\theta_1}$ by Lemma 1.

Assumption 2 For the current data set $\mathscr{Y}^{MC}_{\theta_1}$, assume that the error terms $\varepsilon(x_j) = Y_{\theta_1}(x_j) - \widehat{Y}^{MC}_{\theta_1}(x_j) \sim \mathcal{N}(0, \sigma^2(x_j))$ for some $\sigma^2(x_j) > 0$, for all $j = 1, \ldots, m$. For the LR-weighted data set $\mathscr{Y}^{LR}_{\theta_1,\theta_0}$, assume that the error terms $\varepsilon(x_j^0) = Y_{\theta_1}(x_j^0) - \widehat{Y}^{LR}_{\theta_1,\theta_0}(x_j^0) \sim \mathcal{N}(0, \sigma^2(x_j^0))$ for some $\sigma^2(x_j^0) > 0$, for all $j = 1, \ldots, m_0$.

We note that Assumption 2 does not hold in general, as the distribution of the simulation outputs $F(x, \xi)$ and that of the LR-weighted outputs are often unknown and complex. Nonetheless, the estimators $\widehat{Y}_\theta(x_j)$ and $\widehat{Y}^{LR}_{\theta,\theta'}(x_j^0)$ are usually sample averages of i.i.d. random variables. Therefore, as the number of replications at each covariate grows, the central limit theorem dictates that the error terms are asymptotically normally distributed. One of our future research objective is to adapt the conclusions of Lemma 1 to such asymptotic normality. If Assumptions 1 and 2 hold, then $\widehat{\boldsymbol{\beta}}_{\theta_1} \sim \mathcal{N}\left(\boldsymbol{\beta}_{\theta_1}, (\boldsymbol{\Phi}_1^\top \boldsymbol{\Omega}_1^{-1} \boldsymbol{\Phi}_1)^{-1}\right)$ and $\widehat{\boldsymbol{\beta}}_{\theta_1^+} \sim \mathcal{N}\left(\boldsymbol{\beta}_{\theta_1}, (\boldsymbol{\Phi}_{1^+}^\top \boldsymbol{\Omega}_{1^+}^{-1} \boldsymbol{\Phi}_{1^+})^{-1}\right)$ by Lemma 1.

We observe that, as the simulation experiments at different covariates are ran independently, all covariance matrices $\boldsymbol{\Omega}_0$, $\boldsymbol{\Omega}_1$, and $\boldsymbol{\Omega}_{1^+}$ are diagonal. Specifically, the diagonal elements are $\boldsymbol{\Omega}_0 = \mathrm{diag}[\sigma^2(x_1^0), \ldots, \sigma^2(x_{m_0}^0)]$, $\boldsymbol{\Omega}_1 = \mathrm{diag}[\sigma^2(x_1), \ldots, \sigma^2(x_{m_1})]$, and $\boldsymbol{\Omega}_{1^+} = \mathrm{diag}[\boldsymbol{\Omega}_0, \boldsymbol{\Omega}_1]$. Also, by construction, $\boldsymbol{\Phi}_{1^+}^\top = [\boldsymbol{\Phi}_0^\top, \boldsymbol{\Phi}_1^\top]$. These observations lead to the variance analysis in Proposition 1.

Proposition 1 Suppose Assumptions 1 and 2 hold and that $\boldsymbol{\Phi}_1^\top \boldsymbol{\Omega}_1^{-1} \boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_0^\top \boldsymbol{\Omega}_0^{-1} \boldsymbol{\Phi}_0$ are non-singular. Then

(i) Every element in the combined GLS estimator $\widehat{\boldsymbol{\beta}}_{\theta_1^+}$ has a smaller variance than the variance its counterpart in $\widehat{\boldsymbol{\beta}}_{\theta_1}$, i.e.,

$$\mathrm{Var}[\widehat{\beta}_{\theta_1^+,\ell}] < \mathrm{Var}[\widehat{\beta}_{\theta_1,\ell}], \quad \forall \ell = 1,\ldots,k. \tag{9}$$

(ii) The prediction variance of the regression model with the combined GLS estimator $\widehat{\boldsymbol{\beta}}_{\theta_1^+}$ is smaller than one with $\widehat{\boldsymbol{\beta}}_{\theta_1}$, i.e., for any covariate x and corresponding explanatory variables $\boldsymbol{\phi} = \boldsymbol{\phi}(x)$,

$$\mathrm{Var}\left[\boldsymbol{\phi}^\top \widehat{\boldsymbol{\beta}}_{\theta_1^+}\right] < \mathrm{Var}\left[\boldsymbol{\phi}^\top \widehat{\boldsymbol{\beta}}_{\theta_1}\right]. \tag{10}$$

Proof. By the Woodbury identity we have

$$\begin{aligned}
(\boldsymbol{\Phi}_{1^+}^\top \boldsymbol{\Omega}_{1^+}^{-1} \boldsymbol{\Phi}_{1^+})^{-1} &= \left(\begin{bmatrix} \boldsymbol{\Phi}_0^\top & \boldsymbol{\Phi}_1^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{\Omega}_0^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_1^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}_0 \\ \boldsymbol{\Phi}_1 \end{bmatrix} \right)^{-1} \\
&= (\boldsymbol{\Phi}_0^\top \boldsymbol{\Omega}_0^{-1} \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1^\top \boldsymbol{\Omega}_1^{-1} \boldsymbol{\Phi}_1)^{-1} \\
&= \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}(\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1} \boldsymbol{A}^{-1}.
\end{aligned}$$

where $\boldsymbol{A} := \boldsymbol{\Phi}_1^\top \boldsymbol{\Omega}_1^{-1} \boldsymbol{\Phi}_1$ and $\boldsymbol{B} := \boldsymbol{\Phi}_0^\top \boldsymbol{\Omega}_0^{-1} \boldsymbol{\Phi}_0$. For convenience, let $\boldsymbol{A}_{\ell,\ell}$, $\boldsymbol{A}_{\ell\cdot}$, and $\boldsymbol{A}_{\cdot\ell}$ be the (ℓ,ℓ)-th element, the ℓ-th row, and the ℓ-th column of any matrix $\boldsymbol{A}$.

Since $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_0$ are diagonal and positive definite, $\boldsymbol{A}$ and $\boldsymbol{B}$ are positive definite. Consequently, $\boldsymbol{A}^{-1}$ and $\boldsymbol{B}^{-1}$ are positive definite. Finally, $\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1}$ and $(\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1}$ are positive definite. Therefore, for any $\ell = 1,\ldots,k$, we have

$$\boldsymbol{A}_{\ell\cdot}^{-1}(\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1} \boldsymbol{A}_{\cdot\ell}^{-1} = [\boldsymbol{A}_{\cdot\ell}^{-1}]^\top (\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1} \boldsymbol{A}_{\cdot\ell}^{-1} > 0,$$

and so

$$\mathrm{Var}[\beta_{\theta_1^+,\ell}] = [(\boldsymbol{\Phi}_{1^+}^\top \boldsymbol{\Omega}_{1^+}^{-1} \boldsymbol{\Phi}_{1^+})^{-1}]_{\ell,\ell} = \boldsymbol{A}_{\ell,\ell}^{-1} - \boldsymbol{A}_{\ell\cdot}^{-1}(\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1} \boldsymbol{A}_{\cdot\ell}^{-1} < [(\boldsymbol{\Phi}_1^\top \boldsymbol{\Omega}_1^{-1} \boldsymbol{\Phi}_1)^{-1}]_{\ell,\ell} = \mathrm{Var}[\beta_{\theta_1,\ell}].$$

For any covariate x and corresponding basis function values $\boldsymbol{\phi}(x)$,

$$\begin{aligned}
\mathrm{Var}\left[\boldsymbol{\phi}^\top \widehat{\boldsymbol{\beta}}_{\theta_1^+}\right] &= \boldsymbol{\phi}^\top \mathrm{Var}\left[\widehat{\boldsymbol{\beta}}_{\theta_1^+}\right] \boldsymbol{\phi} \\
&= \boldsymbol{\phi}^\top (\boldsymbol{\Phi}_{1^+}^\top \boldsymbol{\Omega}_{1^+}^{-1} \boldsymbol{\Phi}_{1^+})^{-1} \boldsymbol{\phi} \\
&= \boldsymbol{\phi}^\top \left[(\boldsymbol{\Phi}_1^\top \boldsymbol{\Omega}_1^{-1} \boldsymbol{\Phi}_1)^{-1} - \boldsymbol{A}^{-1}(\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1} \boldsymbol{A}^{-1} \right] \boldsymbol{\phi} \\
&= \boldsymbol{\phi}^\top (\boldsymbol{\Phi}_1^\top \boldsymbol{\Omega}_1^{-1} \boldsymbol{\Phi}_1)^{-1} \boldsymbol{\phi} - \boldsymbol{\phi}^\top \left[\boldsymbol{A}^{-1}(\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1} \boldsymbol{A}^{-1} \right] \boldsymbol{\phi} \\
&= \mathrm{Var}\left[\boldsymbol{\phi}^\top \widehat{\boldsymbol{\beta}}_{\theta_1}\right] - [\boldsymbol{A}^{-1}\boldsymbol{\phi}]^\top (\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1} [\boldsymbol{A}^{-1}\boldsymbol{\phi}] < \mathrm{Var}\left[\boldsymbol{\phi}^\top \widehat{\boldsymbol{\beta}}_{\theta_1}\right],
\end{aligned}$$

where the last inequality holds because $(\boldsymbol{A}^{-1} + \boldsymbol{B}^{-1})^{-1}$ is positive definite. □

Proposition 1 shows the benefits of reusing past simulation outputs: (9) indicates a more accurate GLS estimator and (10) indicates more accurate predictions from the regression model. More importantly, the improved accuracy is obtained without running any additional simulation experiments, as we merely reuse the simulation outputs from experiments that were ran in the past.

There are a few caveats in the proposed method. Compared to a regression that uses only the current simulation outputs, reusing the LR-weighted simulation outputs from the past experiment requires:

- Additional storage. When implementing the likelihood ratio regression method, the individual replications of unweighted simulation outputs

$$\{F(x_j^0, \xi_j^{(i)}) : i = 1, \ldots, n, j = 1, \ldots, m_0\}$$

  are stored and reused. In the current era, abundant low-cost storage and quick retrieval of stored data are fair technological assumptions.
- Additional computations for calculating the likelihood ratios and calibrating a larger regression model.
  - As the unweighted simulation outputs $\{F(x_j^0, \xi_j^{(i)}) : i = 1, \ldots, n, j = 1, \ldots, m_0\}$ are stored, the likelihood ratios $\frac{h_{\theta_1}(\xi_{\theta_0}^{(i)})}{h_{\theta_0}(\xi_{\theta_0}^{(i)})}$ for all $i = 1, \ldots, n$ are needed for unbiased LR estimates.
  - The unbiased LR estimates from the past experiment add more covariates into the regression model, which increases the size of the design matrix $\mathbf{\Phi}_{1^+}$, the covariance matrix $\mathbf{\Omega}_{1^+}$, and the responses $\widehat{\boldsymbol{Y}}_\theta$. So it requires more computations to calculate $\widehat{\boldsymbol{\beta}}(\mathscr{Y}_{\theta_1^+})$ than calculating $\widehat{\boldsymbol{\beta}}(\mathscr{Y}_{\theta_1})$.

Nonetheless, we envision that the proposed methodology is valuable for applications where the simulation model $F(x, \xi)$ is computationally expensive. Then the above additional computations are negligible compared to running additional simulation experiments.

4 NUMERICAL EXAMPLES

Two numerical examples are considered in this section to illustrate the advantages of the proposed likelihood ratio regression method. Firstly, a toy example is considered to demonstrate the improved accuracy of the proposed method in a simple setting. The second example considers estimating the risk of a financial option portfolio, which reflects the advantage of the proposed method in a practical application.

4.1 Illustrative Toy Example

Consider a simulation experiment with distributional parameter $\theta = (\mu, \sigma^2)$, random variable $\xi_\theta \sim \mathcal{N}(\mu, \sigma^2)$, and $F(x, \xi_\theta) = \xi_\theta x$. For any distributional parameter $\theta = (\mu, \sigma^2)$, it is of interest to approximate the function

$$f_\theta(x) = \mathrm{E}[F(x, \xi_\theta)] = \mathrm{E}[\xi_\theta x] = \mu x.$$

We will use simulation outputs to calibrate a linear regression model

$$Y_\theta(x) = \beta_0 + \beta_1 x.$$

Note that the estimated intercept can be non-zero.

The parameters in the current experiment are $\theta_1 = (\mu_1, \sigma_1^2) = (-1, 2)$, so the desired response surface is $f_{\theta_1}(x) = -x$. In the current experiment, we first randomly select x_j, $j = 1, \ldots, m_1$ covariates in the range $[0, 5]$, then run $n = 5$ replications at each of the covariates. Also, for an illustration purpose, suppose the past experiment was ran with parameters $\theta_0 = (\mu_0, \sigma_0^2) = (1, 5)$. In the past experiment, the covariates x_j^0, $j = 1, \ldots, m_0$ are also randomly and uniformly selected in the range $[0, 5]$, each had $n = 5$ replications.

Three sets of simulation outputs can be constructed from the past and the current experiments: The current simulation outputs, $\mathscr{Y}_{\theta_1}^{MC} = \{(\widehat{Y}_{\theta_1}^{MC}(x_j), x_j) : j = 1, \ldots, m_1\}$; The unweighted simulation outputs from the past experiment, $\mathscr{Y}_{\theta_0}^{MC} = \{(\widehat{Y}_{\theta_0}^{MC}(x_j^0), x_j^0) : j = 1, \ldots, m_0\}$; The LR-weighted simulation outputs from the past experiment, $\mathscr{Y}_{\theta_1, \theta_0}^{LR} = \{(\widehat{Y}_{\theta_1, \theta_0}^{LR}(x_j^0), x_j^0) : j = 1, \ldots, m_0\}$.

We consider these three different data sets to calibrate the regression model:

(i) The current data set $\mathscr{Y}_{\theta_1}^{MC}$, denoted by P2 in Figure 1.

(ii) Directly combine the past and the current data sets without any LR-weight, i.e., $\mathscr{Y}_{\theta_0}^{MC} \bigcup \mathscr{Y}_{\theta_1}^{MC}$; denoted by CD in Figure 1.

(iii) Combine the LR-weighted outputs from the past experiment with the current simulation outputs (denoted by LR), i.e., $\mathscr{Y}_{\theta_1,\theta_0}^{LR} \bigcup \mathscr{Y}_{\theta_1}^{MC}$; denoted by LR in Figure 1.

For each of the above three data sets, we estimate the GLS estimates of β_0 and β_1, where the variances at each covariate is estimated via the independent replications at that covariate. We compare the prediction accuracies of the three resulting regression models to the true model $f_{\theta_1}(x) = \mu_1 x$ at 100 equally-spaced prediction points $x = 0.05, 0.1, \ldots, 5$. Specifically, we repeated the entire two-periods simulation and calibrate the three regression models 100 times, i.e., macro replications. In each macro replication, the root mean squared errors (RMSE) of the three regression model on these predictions points are calculated. Figure 1 depicts the box plots of the RMSEs for the three different approaches in different settings.

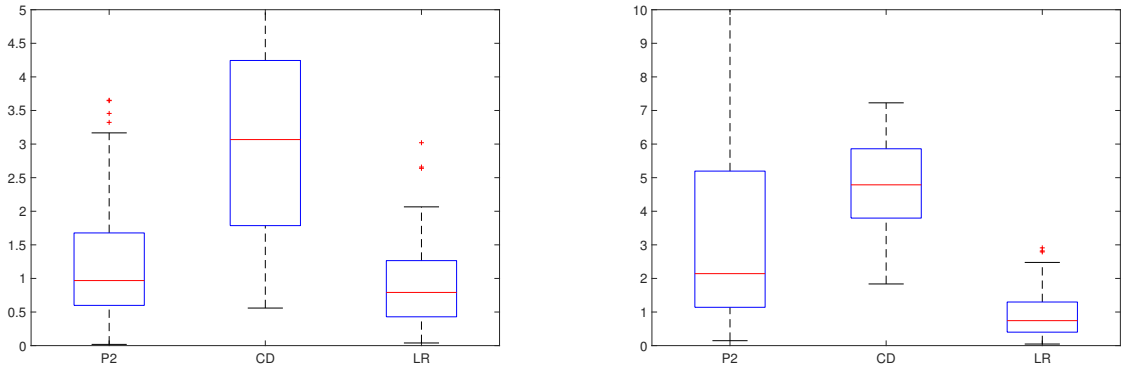

Figure 1: Boxplots of the three approaches for $m_1 = m_2 = 5$ (left panel) and $m_1 = 10, m_2 = 2$ (right panel).

We make the following observations from Figure 1: Firstly, in the left panel, the "CD" box plots show higher average and wider spread of RMSEs than those of "P2". This means that, while reusing existing simulation outputs is an intuitive idea, it requires innovations and cautions to improve, rather than worsening, the computational efficiency. Secondly, the "LR" box plots have lower average and smaller spreads than the "P2" box plots, which indicates that the proposed LR regression approach is better than the regression approach without reusing past simulation outputs. Thirdly, in the right panel, we only simulate at two covariates points in the current experiment. This emulates cases where very limited computation is available in the current period. In this case, the proposed LR regression, which reuses the past simulation outputs, produces much more accurate regression models than regression using only the (very limited) current simulation outputs.

4.2 Option Portfolio

In this example, the portfolio consists of European options, whose mature time $T = 1$ year, based on the same underlying asset driven by a geometric Brownian motion: (i) Long one call option with strike $K_1 = 8$; (ii) Short two call options with strike $K_2 = 10$; (iii) Long one call option with strike $K_3 = 12$; (iv) Long one call option with strike $K_4 = 16$; (v) Short two call options with strike $K_5 = 18$; (vi) Long one call option with strike $K_6 = 20$. Notice that the first three options form a butterfly strategy that designs to earn a profit when the future underlying asset is close to K_2. Similarly the last three options are designed to earn a profit when underlying asset is closed to K_5. So the entire option portfolio will earn a profit when the underlying asset is close to either K_2 or K_5. Using to the Black-Scholes formula for vanilla European options, we can obtain a closed-form price formula, as a function of the current underlying stock price, of the option portfolio.

Now suppose that we want to learn the price formula of the portfolio based on the data generated by simulation experiments (offline simulation) and use the formula in solving real-time problems (online application). Specifically, we learn the price formula every weekend and use the learned formula for the following week. We also assume that the risk-free interest rate and volatility are re-calibrated every weekend, so the price distribution of underlying asset are changed every week due to the updated model parameters. We consider a three-week procedure (three periods) of reusing simulation data, and set $\theta_1 = (r_1, \sigma_1) = (0.05, 0.3)$, $\theta_2 = (r_2, \sigma_2) = (0.04, 0.25)$, and $\theta_3 = (r_3, \sigma_3) = (0.05, 0.28)$. The following basis functions are used:

$$\boldsymbol{\phi}(x) = (x, x^2, x^3, (x-K_1)^+, (x-K_2)^+, \ldots, (x-K_6)^+)),$$

where x is the price of the underlying asset at the time of learning the formula, i.e., the design point.

Suppose that we focus on the price formula of the portfolio on the range $x \in [5, 25]$. Notice that, in this example, if the number of design points in one week is small, $(x-K_i)^+$ for some $i = 1, \ldots, 6$, may just equal to $x - K_i$, so combing x with these basis functions may cause multicollinearity (or near multicollinearity), and the matrix $\boldsymbol{\Phi}^\top \boldsymbol{\Omega} \boldsymbol{\Phi}$ may not be invertible (close to singular). Such numerical difficulty makes the prediction result via the price formula unreliable and inaccurate at least. However, if we reuse the past simulation data in the LR regression, more design points are added and the multicollinearity issue is much less likely. We set $m_1 = m_2 = m_3 = 10$, 15, and 20, and for each case, we set $n = 100$, $n = 1000$, and $n = 10000$. We use *W*3 and *LR* to denote regression using only the current simulation outputs and the combined LR-weighted simulation outputs, respectively. Table 1 shows that our proposed LR regression significantly alleviate the aforementioned multicollinearity issue.

Table 1: The percentage of the non-invertible cases of $\boldsymbol{\Phi}^\top \boldsymbol{\Omega} \boldsymbol{\Phi}$ based on 1000 replications.

	$m_1 = m_2 = m_3 = 10$			$m_1 = m_2 = m_3 = 15$			$m_1 = m_2 = m_3 = 20$		
	$n = 10^2$	$n = 10^3$	$n = 10^4$	$n = 10^2$	$n = 10^3$	$n = 10^4$	$n = 10^2$	$n = 10^3$	$n = 10^4$
W3	58.3%	53.5%	55.2%	15.1%	13.5%	11.1%	1.9%	3.0%	3.7%
LR	0%	0%	0%	0%	0%	0%	0%	0%	0%

Next, we consider different approaches to learn the price formula: (i) Only using the simulation data conducted on the last weekend (denoted by W3), (ii) only using the simulation data conducted on the last weekend but excluding the case of non-invertible matrix (denoted by W3D), (iii) combine the LR-weighted outputs from the past experiment with the current simulation outputs (denoted by LR), and (iv) using simulation data conducted on the last weekend with three times number of design points. The fourth approach can be regarded as the best case in learning the price formula with the same number of total design points. Similar to Example 1, we compare the predictions by each calibrated regression model to the true model at 200 equally-spaced points $S_t = 5.05, 5.10, 5.15, \ldots, 25$ and calculate the RMSE of the different approaches on these test points. We repeated the entire three-period simulation 100 times, each time the RMSE of the four approaches on these predictions points are calculated. Figure 2 depicts the box plots of the RMSEs for the four different approaches in different settings.

We make the following observations from Figure 2: Firstly, the "LR" box plots have a smaller RMSE than those of "W3", which shows the benefits of reusing existing simulation outputs. Even excluding the numerical non-invertible issue ("W3D"), the LR regression approach still performs much better. Secondly, comparing to the best case "W3x3", the LR regression approach almost achieves the same accuracy as if these simulation data were all conducted at the last weekend, and it makes the LR regression approach an appealing method in reusing simulation data to build metamodels in the repeated simulation experiments.

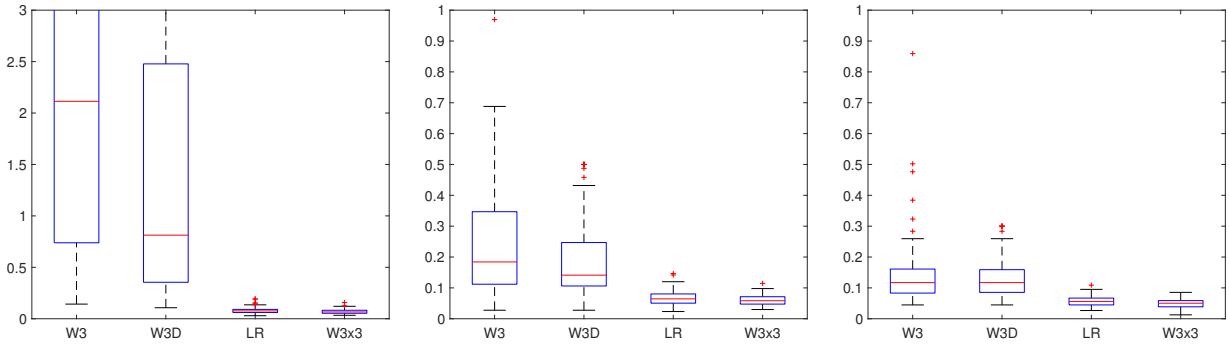

Figure 2: Boxplots of RMSE for different approaches in Example 2. $n = 1000$; $m = 10$ (left panel), $m = 15$ (middle panel), $m = 20$ (right panel).

5 CONCLUSION

This paper proposes a likelihood ratio regression method to reuse simulation data generated by repeated simulation experiments in metamodeling. As time goes, the parameters of simulation models are updated periodically. By multiplying the likelihood ratio, the past simulation outputs seem to generate from the new simulation experiments with the latest model parameters, then a generalized least square method is applied to estimate the coefficients of the preset linear metamodel. Asymptotic variance analysis is conducted to compare the variances of the coefficients estimated by the generalized linear regression with and without reusing past simulation outputs, and to compare the prediction variances of the regression model. A toy example and a realistic example of an option portfolio are provided to show the effectiveness of our method. In the future research, we will derive theoretical results to determine the prediction improvement with periodic model parameters, develop new methods that combine green simulation with offline simulation online application, and consider simulation model misspecification issues.

REFERENCES

Beckman, R. J., and M. D. McKay. 1987. "Monte carlo estimation under different distributions using the same simulation". *Technometrics* 29(2):153–160.

Broadie, M., Y. Du, and C. C. Moallemi. 2015. "Risk estimation via regression". *Operations Research* 63(5):1077–1097.

Dong, J., M. Feng, and B. L. Nelson. 2018. "Unbiased metamodeling via likelihood ratios". In *Proceedings of the 2018 Winter Simulation Conference*, edited by M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, 1778–1789. Piscataway, New Jersey: IEEE.

Feng, M., A. Maggiar, J. Staum, and A. Wächter. 2018. "Uniform convergence of sample average approximation with adaptive multiple importance sampling". In *Proceedings of the 2018 Winter Simulation Conference*, edited by M. Rabe, A. A. Juan, A. S. N. Mustafee, S. Jain, and B. Johansson, 1646–1657. Piscataway, New Jersey: IEEE.

Feng, M., and E. Song. 2019. "Efficient input uncertainty quantification via green simulation using sample path likelihood ratios". In *Proceedings of the 2019 Winter Simulation Conference*, edited by N. Mustafee, K.-H. Bae, S. Lazarova-Molnar, M. Rabe, C. Szab, P. Haas, and Y.-J. Son, 3693–3704. Piscataway, New Jersey: IEEE.

Feng, M., and J. Staum. 2015. "Green simulation designs for repeated experiments". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 403–413. Piscataway, New Jersey: IEEE.

Feng, M., and J. Staum. 2016. "Green simulation with database Monte Carlo". In *Proceedings of the 2016 Winter Simulation Conference (WSC)*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 1108–1118. Piscataway, New Jersey: IEEE.

Feng, M., and J. Staum. 2017. "Green simulation: Reusing the output of repeated experiments". *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 27(4):1–28.

Gao, S., C. Li, and J. Du. 2019. "Rate analysis for offline simulation online application". In *Proceedings of the 2019 Winter Simulation Conference (WSC)*, edited by N. Mustafee, K.-H. Bae, S. Lazarova-Molnar, M. Rabe, C. Szab, P. Haas, and Y.-J. Son, 3468–3479. Piscataway, New Jersey: IEEE.

Glasserman, P. 2013. *Monte Carlo methods in financial engineering*. Springer Science & Business Media.

Hong, L. J., and G. Jiang. 2019. "Offline simulation online application: A new framework of simulation-based decision making". *Asia-Pacific Journal of Operational Research* 36(6):No.1940015.

Jiang, G., L. J. Hong, and B. L. Nelson. 2019. "Online risk monitoring using offline simulation". *INFORMS Journal on Computing*. Forthcoming.

Jiang, G., L. J. Hong, and H. Shen. 2020. "Building closed-form formula for real-time derivative pricing and Greeks calculation using offline simulation". Working paper.

L'Ecuyer, P. 1990. "A Unified View of the IPA, SF, and LR Gradient Estimation Techniques". *Management Science* 36(11):1364–1383.

Longstaff, F. A., and E. S. Schwartz. 2001. "Valuing American options by simulation: A simple least-squares approach". *Review of Financial Studies* 14(1):113–147.

Maggiar, A., A. Wachter, I. S. Dolinskaya, and J. Staum. 2018. "A derivative-free trust-region algorithm for the optimization of functions smoothed via gaussian convolution using adaptive multiple importance sampling". *SIAM Journal on Optimization* 28(2):1478–1507.

Owen, A. B. 2013. *Monte Carlo theory, methods and examples*.

Shen, H., L. J. Hong, and X. Zhang. 2018. "Ranking and selection with covariates for personalized decision making". Working paper.

Stentoft, L. 2004. "Convergence of the least squared Monte Carlo approach to American option valuation". *Management Science* 50(9):1193–1203.

Tsitsiklis, J. N., and B. Van Roy. 2001. "Regression methods for pricing complex American-style options". *IEEE Transactions on Neural Networks* 12(4):694–703.

White, H. 1984. *Asymptotic Theory for Econometricians*. New York: Academic Press.

Zhou, E., and T. Liu. 2018. "Online quantification of input uncertainty for parametric models". In *Proceedings of the 2018 Winter Simulation Conference*, edited by M. Rabe, A. A. Juan, A. S. N. Mustafee, S. Jain, and B. Johansson, 1587–1598. Piscataway, New Jersey: IEEE.

AUTHOR BIOGRAPHIES

BEN MINGBIN FENG is an assistant professor in actuarial science at the University of Waterloo. He earned his Ph.D. in the Department of Industrial Engineering and Management Sciences at Northwestern University. He is an Associate of the Society of Actuaries (ASA). His research interests include stochastic simulation design and analysis, optimization via simulation, nonlinear optimization, and financial and actuarial applications of simulation and optimization methodologies. His e-mail address is ben.feng@uwaterloo.ca. His website is http://www.math.uwaterloo.ca/∼mbfeng/.

GUANGXIN JIANG is a professor in the School of Management at Harbin Institute of Technology. He earned his Ph.D. in Applied Mathematics from Tongji University. His research interests include stochastic models and simulation, machine learning, financial engineering and risk management, FinTech, etc. His email address is gxjiang@hit.edu.cn. His website is http://homepage.hit.edu.cn/jiangguangxin/.