

UNBIASED MLMC STOCHASTIC GRADIENT-BASED OPTIMIZATION OF BAYESIAN EXPERIMENTAL DESIGNS*

TAKASHI GODA[†], TOMOHIKO HIRONAKA[†], WATARU KITADE[†], AND ADAM FOSTER[‡]

Abstract. In this paper we propose an efficient stochastic optimization algorithm to search for Bayesian experimental designs such that the expected information gain is maximized. The gradient of the expected information gain with respect to experimental design parameters is given by a nested expectation, for which the standard Monte Carlo method using a fixed number of inner samples yields a biased estimator. In this paper, applying the idea of randomized multilevel Monte Carlo (MLMC) methods, we introduce an unbiased Monte Carlo estimator for the gradient of the expected information gain with finite expected squared ℓ_2 -norm and finite expected computational cost per sample. Our unbiased estimator can be combined well with stochastic gradient descent algorithms, which results in our proposal of an optimization algorithm to search for an optimal Bayesian experimental design. Numerical experiments confirm that our proposed algorithm works well not only for a simple test problem but also for a more realistic pharmacokinetic problem.

Key words. Bayesian experimental design, expected information gain, multilevel Monte Carlo, nested expectation, stochastic gradient descent

AMS subject classifications. 62K05, 62L20, 65C05, 92C45, 94A17

1. Introduction. In this paper we study optimization of Bayesian experimental designs which aim to maximize the expected amount of information experimental outcomes convey about unobservable, or hidden/latent, random variables of interest by carefully designing an experimental setup. Here we measure the expected amount of information by the Shannon's expected information gain whose definition is given below. Our motivation comes from applications to a number of disciplines, such as mechanical engineering [34], neuroscience [40], bioinformatics [36], psychology [23], and pharmacokinetics [33, 32] among many others.

Let $\theta = (\theta_1, \dots, \theta_s) \in \Theta \subseteq \mathbb{R}^s$ be a vector of continuous unobservable random variables, and we denote the prior probability density of θ by $\pi_0(\theta)$. The information entropy, or the differential entropy, of θ is defined by

$$\mathbb{E}_\theta [-\log \pi_0(\theta)] = \int_{\Theta} -\pi_0(\theta) \log \pi_0(\theta) d\theta.$$

Let us consider a situation where, by conducting some experiments under an experimental design ξ , an observation $Y = (Y_1, \dots, Y_t) \in \mathcal{Y} \subseteq \mathbb{R}^t$ is obtained according to the forward model

$$(1.1) \quad Y = f_\xi(\theta, \epsilon),$$

where $\epsilon = (\epsilon_1, \dots, \epsilon_{s'}) \in \mathcal{E} \subseteq \mathbb{R}^{s'}$, representing the observation noise, is another vector of continuous random variables with its density $\varphi(\epsilon)$, and f_ξ is a deterministic bi-variate function parametrized by the design ξ , possibly with multiple outputs. Here we assume that the experimental design ξ is controllable and can be chosen as an element in an open set $\mathcal{X} \subset \mathbb{R}^d$. Throughout this paper, we assume that the

*Submitted to the editors DATE.

Funding: The work of T.G. is supported by JSPS KAKENHI Grant Number 20K0374. The work of A.F. is kindly supported by EPSRC grant no. EP/N509711/1.

[†]School of Engineering, University of Tokyo, Tokyo, Japan (goda@frcer.t.u-tokyo.ac.jp, hironaka-tomohiko@g.ecc.u-tokyo.ac.jp, kitade-wataru114@g.ecc.u-tokyo.ac.jp).

[‡]Department of Statistics, University of Oxford, Oxford, UK (adam.foster@stats.ox.ac.uk).

domain \mathcal{Y} is independent of ξ , that ϵ is independent both of θ and ξ , and also that the likelihood function $\rho(Y | \theta, \xi)$ is strictly positive and can be computed explicitly with unit cost for any pair of θ, ξ and Y . As is well known, Bayes' theorem states that the posterior probability density of θ given Y , denoted by $\pi^{Y|\xi}$, is given by

$$(1.2) \quad \pi^{Y|\xi}(\theta) = \frac{\rho(Y | \theta, \xi) \pi_0(\theta)}{\rho(Y | \xi)},$$

with $\rho(Y | \xi)$ being the marginal likelihood of Y , i.e.,

$$\rho(Y | \xi) = \mathbb{E}_\theta [\rho(Y | \theta, \xi)] = \int_{\Theta} \rho(Y | \theta, \xi) \pi_0(\theta) d\theta,$$

see for instance [38]. Then, the posterior information entropy of θ after observing Y is given by

$$\mathbb{E}_{\theta|Y,\xi} [-\log \pi^{Y|\xi}(\theta)] = \int_{\Theta} -\pi^{Y|\xi}(\theta) \log \pi^{Y|\xi}(\theta) d\theta,$$

and hence, the expected posterior information entropy of θ by conducting an experiment under an experimental design ξ is given by integrating the posterior information entropy of θ over Y using the marginal likelihood $\rho(Y | \xi)$, i.e.,

$$\mathbb{E}_{Y|\xi} \mathbb{E}_{\theta|Y,\xi} [-\log \pi^{Y|\xi}(\theta)] = \int_{\mathcal{Y}} \int_{\Theta} -\pi^{Y|\xi}(\theta) \log \pi^{Y|\xi}(\theta) d\theta \rho(Y | \xi) dY.$$

Now the difference

$$U(\xi) := \mathbb{E}_\theta [-\log \pi_0(\theta)] - \mathbb{E}_{Y|\xi} \mathbb{E}_{\theta|Y,\xi} [-\log \pi^{Y|\xi}(\theta)]$$

is called the *expected information gain*, the quantity originally introduced in [21] as a measure of experimental designs. By using Bayes' theorem (1.2), we see that $U(\xi)$ is equivalently given by

$$(1.3) \quad \begin{aligned} U(\xi) &= \mathbb{E}_\theta \mathbb{E}_{Y|\theta,\xi} [\log \rho(Y | \theta, \xi)] - \mathbb{E}_{Y|\xi} [\log \rho(Y | \xi)] \\ &= \mathbb{E}_\theta \mathbb{E}_{Y|\theta,\xi} [\log \rho(Y | \theta, \xi)] - \mathbb{E}_{Y|\xi} [\log \mathbb{E}_\theta [\rho(Y | \theta, \xi)]] . \end{aligned}$$

The aim of Bayesian experimental designs is to construct an optimal experimental design $\xi = \xi^*$ which maximizes the expected information gain U [6]. As can be seen from the second term of (1.3), however, estimating $U(\xi)$ is inherently a nested expectation problem with an outer expectation with respect to Y and an inner expectation with respect to θ , which has been considered computationally challenging. The standard, nested Monte Carlo method generates N outer random samples for Y first and then, for each sample of Y , generates M inner random samples for θ . To estimate $U(\xi)$ with root-mean-square accuracy ε ,¹ we typically need $N = O(\varepsilon^{-2})$ and $M = O(\varepsilon^{-1})$, resulting in a total computational complexity of $O(\varepsilon^{-3})$ [34, 2, 27]. Recently there have been some attempts in [16, 3] to reduce this cost to $O(\varepsilon^{-2})$ or $O(\varepsilon^{-2}(\log \varepsilon^{-1})^2)$ by applying a multilevel Monte Carlo (MLMC) method [12, 13] in conjunction with Laplace approximation-based importance sampling [22]. Here the difference between the orders of complexity for the MLMC method is a direct consequence from the basic MLMC theorem, see for instance [13, Theorem 2.1], which

¹Here and in what follows, the difference between the noise ϵ and the accuracy ε should not be confused.

itself depends on the properties of the constructed MLMC estimators. Nevertheless, these results are an intermediate step towards an efficient construction of optimal experimental designs since design optimization has been left behind.

In this paper we deal with this optimization problem more directly. More precisely, under the assumption that the experimental setup, or the set of design parameters, ξ lives in a continuous space such that U is differentiable with respect to ξ , we consider applying stochastic gradient descent optimizations to search for an optimal ξ . As we shall see, the gradient $\nabla_{\xi}U$ is again given by a nested expectation, for which the standard, nested Monte Carlo method using a fixed number of inner samples yields a biased estimator. By applying an unbiased MLMC method from [29], a randomized version of the original MLMC method, we can construct an unbiased estimator of $\nabla_{\xi}U$. This way, in this paper, we arrive at a stochastic gradient-based optimization algorithm in which unbiased random samples to estimate $\nabla_{\xi}U$ are generated at each iteration step.

Here we have to mention that the idea of using stochastic gradient-based methods in Bayesian experimental designs already exists in the literature [18, 9, 10, 5, 20]. In particular, a work by Carlon et al. [5] takes a similar standpoint in that an analytical expression of the gradient $\nabla_{\xi}U$ is derived and then stochastic gradient-based method is applied in conjunction with Monte Carlo estimation of $\nabla_{\xi}U$. However, the expression of $\nabla_{\xi}U$ given in [5, Proposition 1] is proven only for the additive Gaussian noise ϵ , that is, the case where the forward model is given by the form $Y = f_{\xi}(\theta) + \epsilon$ with $\epsilon \sim N(0, \Sigma)$, and the standard (biased) Monte Carlo estimator is used at each iteration step within stochastic gradient-based methods. In this paper we consider a more general form of the forward model as shown in (1.1), which is useful in some applications [33, 32]. Moreover, given that stochastic gradient-based methods are usually established under the assumption that each sample is drawn from the underlying true distribution, using an unbiased estimator of $\nabla_{\xi}U$ should be favorable, and by doing so, we do not need to take care of the bias-variance tradeoff. Although application of MLMC methods to stochastic approximation algorithms have been investigated recently in [11, 7], neither of them considers using a randomized MLMC method to generate unbiased random samples at each iteration step.

The rest of this paper is organized as follows. In Section 2, we provide an analytical expression of the gradient $\nabla_{\xi}U$ and also briefly review some of stochastic gradient-based optimization methods. Although there are a number of stochastic optimization algorithms, one can use any of them in our proposal to optimize Bayesian experimental designs (Algorithm 3.1), and we do not give any recommendation on which method should be used in our algorithm, since it is not the objective of this paper. Again we emphasize that the main contribution of this paper is to provide an unbiased estimator for the gradient $\nabla_{\xi}U$, which is non-trivial but a key assumption in stochastic gradient-based optimization. In Section 3, after introducing a standard, nested Monte Carlo estimator of $\nabla_{\xi}U$, which is obviously biased, we provide an unbiased, multilevel Monte Carlo estimator of $\nabla_{\xi}U$ and prove under some conditions that our estimator has a finite expected squared ℓ_2 -norm with finite computational cost per sample. Our proposal for optimizing Bayesian experimental designs is given in Algorithm 3.1. To demonstrate the effectiveness of our proposed algorithm, we conduct numerical experiments not only for a simple test problem but also for a more realistic pharmacokinetic (PK) problem in Section 4. We conclude this paper with some remarks in Section 5.

2. Stochastic gradient-based optimization.

2.1. Gradient of expected information gains. In what follows, we give an explicit form of the gradient $\nabla_\xi U$. As a preparation, let us rewrite the expected information gain $U(\xi)$ according to (1.1) in the following way. First, by noting that generating Y randomly conditional on θ and ξ is equivalent to computing $f_\xi(\theta, \epsilon)$ for a randomly generated ϵ with both θ and ξ given, the independence between ϵ and the pair (θ, ξ) ensures that the first term of (1.3) is equal to

$$\mathbb{E}_\theta \mathbb{E}_\epsilon [\log \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)] = \mathbb{E}_{\theta, \epsilon} [\log \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)].$$

Similarly, generating Y randomly conditional only on ξ is equivalent to computing $f_\xi(\theta, \epsilon)$ for randomly generated θ and ϵ with a fixed ξ . Therefore, by denoting an i.i.d. copy of θ by θ' , the second term of (1.3) is equal to

$$\mathbb{E}_{Y \mid \xi} [\log \mathbb{E}_{\theta'} [\rho(Y \mid \theta', \xi)]] = \mathbb{E}_{\theta, \epsilon} [\log \mathbb{E}_{\theta'} [\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]].$$

Thus we end up with the following expression of $U(\xi)$:

$$(2.1) \quad U(\xi) = \mathbb{E}_{\theta, \epsilon} [\log \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)] - \mathbb{E}_{\theta, \epsilon} [\log \mathbb{E}_{\theta'} [\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]].$$

As we have stated in the previous section, we assume throughout this paper that the likelihood function can be computed explicitly with unit cost for any pair of inputs. Here we give some examples for which such an explicit computation of the likelihood function is possible.

Example 2.1 (Additive noise). Let us consider a forward model given by

$$f_\xi(\theta, \epsilon) = g_\xi(\theta) + \epsilon,$$

for a uni-variate function $g_\xi : \Theta \rightarrow \mathcal{Y} (= \mathbb{R}^t)$ and $\epsilon \sim N(0, \Sigma)$ with a covariance matrix Σ and $s' = t$. Then, denoting the density of ϵ by φ , we have

$$\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) = \varphi(\epsilon + g_\xi(\theta) - g_\xi(\theta')),$$

with a special case $\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi) = \varphi(\epsilon)$.

Example 2.2 (Multiplicative noise). Let $s' = t = 1$ for simplicity, and consider a forward model given by

$$f_\xi(\theta, \epsilon) = g_\xi(\theta) \times (1 + \epsilon),$$

with $g_\xi : \mathbb{R}^s \rightarrow \mathbb{R}_{>0}$ and $\epsilon \sim N(0, \sigma^2)$. Denoting the density of ϵ by φ , we have

$$\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) = \varphi\left(\frac{g_\xi(\theta)}{g_\xi(\theta')}(1 + \epsilon) - 1\right),$$

with a special case $\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi) = \varphi(\epsilon)$.

Example 2.3 (Mixture of additive and multiplicative noises). Finally, for $t = 1$ and $s' = 2$, i.e., $\epsilon = (\epsilon_1, \epsilon_2) \in \mathbb{R}^2$, let us consider a forward model described by

$$f_\xi(\theta, \epsilon) = g_\xi(\theta) \times (1 + \epsilon_1) + \epsilon_2,$$

with $g_\xi : \mathbb{R}^s \rightarrow \mathbb{R}_{>0}$, $\epsilon_1 \sim N(0, \sigma_1^2)$ and $\epsilon_2 \sim N(0, \sigma_2^2)$. Denoting the density of the standard normal random variable by φ , we have

$$\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) = \frac{1}{\sqrt{|g_\xi(\theta')|^2 \sigma_1^2 + \sigma_2^2}} \varphi\left(\frac{g_\xi(\theta) \times (1 + \epsilon_1) + \epsilon_2 - g_\xi(\theta')}{\sqrt{|g_\xi(\theta')|^2 \sigma_1^2 + \sigma_2^2}}\right),$$

with a special case

$$\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi) = \frac{1}{\sqrt{|g_\xi(\theta)|^2 \sigma_1^2 + \sigma_2^2}} \varphi \left(\frac{\epsilon_1 g_\xi(\theta) + \epsilon_2}{\sqrt{|g_\xi(\theta)|^2 \sigma_1^2 + \sigma_2^2}} \right).$$

Now we are ready to derive the gradient $\nabla_\xi U$. Note that our claim does not assume that the noise ϵ is additive and a Gaussian random variable, as discussed in the last two examples.

PROPOSITION 2.4. *Let θ' be an i.i.d. copy of θ . Assume that the likelihood functions $\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)$ and $\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)$ and their gradients $\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)$ and $\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)$ are all continuous with respect to $\theta, \theta', \epsilon$ and ξ . Then we have*

$$\nabla_\xi U(\xi) = \mathbb{E}_{\theta, \epsilon} \left[\frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)} - \frac{\mathbb{E}_{\theta'} [\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]}{\mathbb{E}_{\theta'} [\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]} \right].$$

Proof. Under the continuity assumption on the likelihood function, the Leibniz integral rule applies and we have

$$\begin{aligned} \nabla_\xi U(\xi) &= \mathbb{E}_{\theta, \epsilon} [\nabla_\xi \log \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)] - \mathbb{E}_{\theta, \epsilon} [\nabla_\xi \log \mathbb{E}_{\theta'} [\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]] \\ &= \mathbb{E}_{\theta, \epsilon} \left[\frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)} \right] - \mathbb{E}_{\theta, \epsilon} \left[\frac{\nabla_\xi \mathbb{E}_{\theta'} [\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]}{\mathbb{E}_{\theta'} [\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]} \right] \\ &= \mathbb{E}_{\theta, \epsilon} \left[\frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)} \right] - \mathbb{E}_{\theta, \epsilon} \left[\frac{\mathbb{E}_{\theta'} [\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]}{\mathbb{E}_{\theta'} [\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]} \right]. \quad \square \end{aligned}$$

As is clear from this proposition, because of the ratio of inner expectations, the gradient $\nabla_\xi U$ is inherently given by a nested expectation with an inner expectation with respect to θ' and an outer expectation with respect to θ and ϵ .

2.2. Basics of stochastic gradient-based optimization. We recall that the aim of Bayesian experimental designs is to find an optimal experimental setup $\xi = \xi^*$ which satisfies

$$\xi^* = \arg \max_{\xi \in \mathcal{X}} U(\xi),$$

where we recall that an open set $\mathcal{X} \subset \mathbb{R}^d$ denotes the feasible domain of ξ . To achieve this goal, one of the reasonable approaches is to use some gradient-based optimization methods in which we set an initial experimental setup $\xi_0 \in \mathcal{X}$ and recursively update itself as

$$\xi_{t+1} = g_t(\xi_t, \nabla_\xi U(\xi_t)) \quad \text{for } t = 0, 1, \dots,$$

until a certain stopping criterion is met. However, computing $\nabla_\xi U$ is already challenging since it is given by a nested expectation. As inferred from the results shown in the next section, it is possible to construct an antithetic MLMC estimator which efficiently estimates $\nabla_\xi U$, but we avoid such a “pointwise” accurate gradient estimation by using *stochastic* gradient-based optimization methods. What we need here is an unbiased estimator of $\nabla_\xi U$ with finite variance and computational cost.

To simplify the presentation, let us define a vector of random variables

$$(2.2) \quad \psi_\xi := \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)} - \frac{\mathbb{E}_{\theta'} [\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]}{\mathbb{E}_{\theta'} [\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]},$$

with $\theta \sim \pi_0$ and $\epsilon \sim \varphi$ being the underlying stochastic variables. It follows from Proposition 2.4 that $\mathbb{E}[\psi_\xi] = \nabla_\xi U(\xi)$. Suppose at this moment that we are able to

generate i.i.d. random samples of ψ_ξ . We emphasize that random sampling of ψ_ξ is far from trivial but we shall show in the next section that this is indeed possible.

In stochastic gradient-based optimization methods, after setting an initial experimental setup $\xi_0 \in \mathcal{X}$, we recursively update itself as

$$\xi_{t+1} = g_t(\xi_t, \psi_{\xi_t}) \quad \text{for } t = 0, 1, \dots,$$

or more generally,

$$\xi_{t+1} = g_t \left(\xi_t, \frac{1}{N} \sum_{n=1}^N \psi_{\xi_t}^{(n)} \right) \quad \text{for } t = 0, 1, \dots,$$

where $\psi_{\xi_t}^{(1)}, \dots, \psi_{\xi_t}^{(N)}$ are i.i.d. realizations of ψ_{ξ_t} for a sample size $N \in \mathbb{Z}_{>0}$. This means that, at each iteration, we only need (rough) unbiased Monte Carlo estimate of $\mathbb{E}[\psi_\xi]$ instead of the true value. There have been many examples for this recursion g_t proposed in the literature.

For instance, one of the most classical methods due to Robbins and Monro [30] is simply given by

$$\xi_{t+1} = \Pi_{\mathcal{X}} \left(\xi_t + a_t \cdot \frac{1}{N} \sum_{n=1}^N \psi_{\xi_t}^{(n)} \right),$$

with a sequence of non-negative reals called *learning rates* a_0, a_1, \dots such that

$$\sum_{t=0}^{\infty} a_t = \infty \quad \text{and} \quad \sum_{t=0}^{\infty} a_t^2 < \infty,$$

where $\Pi_{\mathcal{X}}$ denotes the projection operator which maps the input to a closest point in \mathcal{X} , i.e., $\Pi_{\mathcal{X}}(\xi') = \arg \min_{\xi \in \mathcal{X}} \|\xi - \xi'\|$ with $\|\cdot\|$ being the Euclidean norm of vector.² As described in [37, Chapter 5.9], for instance, if \mathcal{X} is convex, U is strongly concave and differentiable with respect to ξ , and $\mathbb{E}[\|\psi_\xi\|_2^2] < \infty$ for any $\xi \in \mathcal{X}$, then the estimate ξ_t converges to the optimal ξ^* with the mean squared error of $O(1/t)$.

There have been many variants of the classical Robbins-Monro algorithm proposed in the literature, notably such as Polyak-Ruppert averaging [26, 31] and stochastic counterpart of Nesterov's acceleration [24]. More recently, the idea of using not only the first moment of the gradient estimate but also its second moment to set the learning rates for individual design parameters in ξ adaptively has been explored insensitively, especially in the machine learning community, see [8, 39, 19, 28].

3. Monte Carlo gradient estimation. Here we introduce two Monte Carlo estimators of the gradient $\nabla_\xi U(\xi) = \mathbb{E}[\psi_\xi]$. Subsequently we propose an algorithm to efficiently search for optimal Bayesian experimental designs.

²Note that most of the textbooks on stochastic algorithms such as [1, 37] consider minimization problems for which the update rule should be replaced by

$$\xi_{t+1} = \Pi_{\mathcal{X}} \left(\xi_t - a_t \cdot \frac{1}{N} \sum_{n=1}^N \psi_{\xi_t}^{(n)} \right),$$

and the objective function is often assumed to be convex instead of concave.

3.1. Standard Monte Carlo. The standard Monte Carlo method is one of the easiest and the most straightforward methods to approximate ψ_ξ . Let us estimate two expectations with respect to θ' by the Monte Carlo averages using common random samples of θ' , respectively. Namely, for randomly chosen θ and ϵ , let

$$\psi_{\xi,M} := \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)} - \frac{\nabla \varrho_{\xi,M}(\theta, \epsilon)}{\varrho_{\xi,M}(\theta, \epsilon)},$$

with

$$\begin{aligned} \varrho_{\xi,M}(\theta, \epsilon) &= \frac{1}{M} \sum_{m=1}^M \rho(f_\xi(\theta, \epsilon) \mid \theta'^{(m)}, \xi), \\ \nabla \varrho_{\xi,M}(\theta, \epsilon) &= \frac{1}{M} \sum_{m=1}^M \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta'^{(m)}, \xi), \end{aligned}$$

where $\theta'^{(1)}, \dots, \theta'^{(M)}$ are independent samples from the prior distribution π_0 . More generally, for an importance distribution q which may depend on the value of $f_\xi(\theta, \epsilon)$ or the outer random variables θ and ϵ , we can consider

$$(3.1) \quad \psi_{\xi,M,q} := \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)} - \frac{\nabla \varrho_{\xi,M,q}(\theta, \epsilon)}{\varrho_{\xi,M,q}(\theta, \epsilon)},$$

with

$$\begin{aligned} \varrho_{\xi,M,q}(\theta, \epsilon) &= \frac{1}{M} \sum_{m=1}^M \frac{\rho(f_\xi(\theta, \epsilon) \mid \theta'^{(m)}, \xi) \pi_0(\theta'^{(m)})}{q(\theta'^{(m)})}, \\ \nabla \varrho_{\xi,M,q}(\theta, \epsilon) &= \frac{1}{M} \sum_{m=1}^M \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta'^{(m)}, \xi) \pi_0(\theta'^{(m)})}{q(\theta'^{(m)})}, \end{aligned}$$

where $\theta'^{(1)}, \dots, \theta'^{(M)}$ are independent samples from the distribution q .

Although it holds from the linearity of expectation that

$$\begin{aligned} \mathbb{E}[\varrho_{\xi,M,q}(\theta, \epsilon) \mid \theta, \epsilon] &= \mathbb{E}_{\theta'}[\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)], \\ \mathbb{E}[\nabla \varrho_{\xi,M,q}(\theta, \epsilon) \mid \theta, \epsilon] &= \mathbb{E}_{\theta'}[\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)], \end{aligned}$$

for any M , i.e., both the denominator and the numerator themselves are estimated without any bias, respectively, taking the ratio between these two yields

$$\mathbb{E}[\psi_{\xi,M}], \mathbb{E}[\psi_{\xi,M,q}] \neq \mathbb{E}[\psi_\xi] = \nabla_\xi U(\xi)$$

unless $q = \pi^{f_\xi(\theta, \epsilon) \mid \xi}$. This means that neither $\psi_{\xi,M}$ nor $\psi_{\xi,M,q}$ is an unbiased estimator of the gradient $\nabla_\xi U(\xi)$.

3.2. Unbiased multilevel Monte Carlo. Here we introduce an unbiased multilevel Monte Carlo estimator by using the debiasing technique from [29] which itself is an extension of the multilevel Monte Carlo method due to Giles [12, 13]. Let us consider an increasing sequence $0 < M_0 < M_1 < \dots$ such that $M_\ell \rightarrow \infty$ as $\ell \rightarrow \infty$. Then the strong law of large numbers ensures that

$$\mathbb{P} \left[\lim_{\ell \rightarrow \infty} \psi_{\xi, M_\ell, q} = \psi_\xi \right] = 1,$$

see for instance [25, Theorem 9.2], and so the following telescoping sum holds:

$$\nabla_{\xi} U(\xi) = \mathbb{E}[\psi_{\xi}] = \lim_{\ell \rightarrow \infty} \mathbb{E}[\psi_{\xi, M_{\ell}, q}] = \mathbb{E}[\psi_{\xi, M_0, q}] + \sum_{\ell=1}^{\infty} \mathbb{E}[\psi_{\xi, M_{\ell}, q} - \psi_{\xi, M_{\ell-1}, q}].$$

More generally, suppose at this moment that we have a sequence of *correction* random variables $\Delta\psi_{\xi, 0}, \Delta\psi_{\xi, 1}, \dots$ such that $\mathbb{E}[\Delta\psi_{\xi, 0}] = \mathbb{E}[\psi_{\xi, M_0, q}]$ and

$$\mathbb{E}[\Delta\psi_{\xi, \ell}] = \mathbb{E}[\psi_{\xi, M_{\ell}, q} - \psi_{\xi, M_{\ell-1}, q}] \quad \text{for } \ell > 0.$$

Then it holds that

$$(3.2) \quad \nabla_{\xi} U(\xi) = \mathbb{E}[\psi_{\xi}] = \sum_{\ell=0}^{\infty} \mathbb{E}[\Delta\psi_{\xi, \ell}].$$

For any sequence of positive reals w_0, w_1, \dots such that $w_0 + w_1 + \dots = 1$, the expectation of the random variable

$$\frac{\Delta\psi_{\xi, \ell}}{w_{\ell}}$$

with the index $\ell \geq 0$ being selected randomly with probability w_{ℓ} , is equal to the gradient $\nabla_{\xi} U(\xi)$. In fact, it is easy to see that

$$\mathbb{E}\left[\frac{\Delta\psi_{\xi, \ell}}{w_{\ell}}\right] = \sum_{\ell=0}^{\infty} \frac{\mathbb{E}[\Delta\psi_{\xi, \ell}]}{w_{\ell}} w_{\ell} = \sum_{\ell=0}^{\infty} \mathbb{E}[\Delta\psi_{\xi, \ell}] = \nabla_{\xi} U(\xi).$$

Therefore, for any number of outer samples $N \in \mathbb{Z}_{>0}$,

$$\frac{1}{N} \sum_{n=1}^N \frac{\Delta\psi_{\xi, \ell^{(n)}}}{w_{\ell^{(n)}}}$$

with $\ell^{(1)}, \dots, \ell^{(N)}$ being independent and randomly chosen with probability w_{ℓ} is an unbiased Monte Carlo estimator of $\nabla_{\xi} U(\xi)$.

Let C_{ℓ} denote the expected cost of computing $\Delta\psi_{\xi, \ell}$, which is proportional to M_{ℓ} . In order for the random variable $\Delta\psi_{\xi, \ell}/w_{\ell}$ to have finite expected squared ℓ_2 -norm and finite expected computational cost, we must have

$$(3.3) \quad \sum_{\ell=0}^{\infty} \frac{\mathbb{E}[\|\Delta\psi_{\xi, \ell}\|_2^2]}{w_{\ell}} < \infty \quad \text{and} \quad \sum_{\ell=0}^{\infty} C_{\ell} w_{\ell} < \infty.$$

Thus construction of such correction variables $\Delta\psi_{\xi, \ell}$ in conjunction with an associated sequence w_0, w_1, \dots , which has not been discussed yet, becomes a central issue.

3.2.1. Naive construction. Throughout this paper let us consider a geometric progression $M_{\ell} = M_0 2^{\ell}$ for some $M_0 \in \mathbb{Z}_{\geq 0}$. Although it is possible to change the base of the progression to a general integer $b \geq 2$, we restrict ourselves to the case $b = 2$ for simplicity of exposition.

Probably the most straightforward form of the correction variables $\Delta\psi_{\xi, 0}, \Delta\psi_{\xi, 1}, \dots$ is $\Delta\psi_{\xi, 0} = \psi_{\xi, M_0, q}$ and

$$\Delta\psi_{\xi, \ell} = \psi_{\xi, M_0 2^{\ell}, q} - \psi_{\xi, M_0 2^{\ell-1}, q},$$

for $\ell > 0$, where both $\psi_{\xi, M_0 2^{\ell-1}, q}$ and $\psi_{\xi, M_0 2^\ell, q}$ are given as in (3.1) with $M = M_0 2^{\ell-1}$ and $M = M_0 2^\ell$, respectively. Here, instead of using mutually independent $M_0 2^{\ell-1}$ and $M_0 2^\ell$ samples on θ' to compute $\psi_{\xi, M_0 2^{\ell-1}, q}$ and $\psi_{\xi, M_0 2^\ell, q}$, respectively, a subset with size $M_0 2^{\ell-1}$ of the $M_0 2^\ell$ samples on θ' used to compute $\psi_{\xi, M_0 2^\ell, q}$, can be reused to compute $\psi_{\xi, M_0 2^{\ell-1}, q}$ by the linearity of expectation. By doing so, it is expected that $\mathbb{E}[\|\Delta\psi_{\xi, \ell}\|_2^2]$ is much smaller in magnitude than $\mathbb{E}[\|\psi_{\xi, M_0 2^\ell, q}\|_2^2]$ (or $\mathbb{E}[\|\psi_{\xi, M_0 2^{\ell-1}, q}\|_2^2]$).

However, it seems not possible that the order of $\mathbb{E}[\|\Delta\psi_{\xi, \ell}\|_2^2]$ is better than $O(2^{-\ell})$. Recalling that $C_\ell \propto M_\ell \propto 2^\ell$, a faster decay of $\mathbb{E}[\|\Delta\psi_{\xi, \ell}\|_2^2]$ is required to find a sequence of positive reals w_0, w_1, \dots which satisfies the condition (3.3). We conjecture that a lower bound on $\mathbb{E}[\|\Delta\psi_{\xi, \ell}\|_2^2]$ of order $2^{-\ell}$ exists for this naive construction.

3.2.2. Antithetic construction. Motivated by the MLMC literature [15, 4, 14, 16, 17], we address this issue by considering the following *antithetic coupling* in this paper. A key ingredient here is that we can take two disjoint subsets with equal size $M_0 2^{\ell-1}$ of the $M_0 2^\ell$ samples on θ' used to compute $\psi_{\xi, M_0 2^\ell, q}$, which results in two independent realizations of $\psi_{\xi, M_0 2^{\ell-1}, q}$, denoted by $\psi_{\xi, M_0 2^{\ell-1}, q}^{(a)}$ and $\psi_{\xi, M_0 2^{\ell-1}, q}^{(b)}$, respectively. To be more precise, for the independent samples $\theta'^{(1)}, \dots, \theta'^{(M_0 2^\ell)}$ generated from the distribution q , we write

$$\begin{aligned}\psi_{\xi, M_0 2^\ell, q} &= \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)} - \frac{\nabla \varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}, \\ \psi_{\xi, M_0 2^{\ell-1}, q}^{(a)} &= \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)} - \frac{\nabla \varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)}{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)}, \quad \text{and} \\ \psi_{\xi, M_0 2^{\ell-1}, q}^{(b)} &= \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)} - \frac{\nabla \varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)}{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)},\end{aligned}$$

where, for the second term of each, we have defined

$$\begin{aligned}\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon) &= \frac{1}{M_0 2^\ell} \sum_{m=1}^{M_0 2^\ell} \frac{\rho(f_\xi(\theta, \epsilon) \mid \theta'^{(m)}, \xi) \pi_0(\theta'^{(m)})}{q(\theta'^{(m)})}, \\ \nabla \varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon) &= \frac{1}{M_0 2^\ell} \sum_{m=1}^{M_0 2^\ell} \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta'^{(m)}, \xi) \pi_0(\theta'^{(m)})}{q(\theta'^{(m)})}, \\ \varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon) &= \frac{1}{M_0 2^{\ell-1}} \sum_{m=1}^{M_0 2^{\ell-1}} \frac{\rho(f_\xi(\theta, \epsilon) \mid \theta'^{(m)}, \xi) \pi_0(\theta'^{(m)})}{q(\theta'^{(m)})}, \\ \nabla \varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon) &= \frac{1}{M_0 2^{\ell-1}} \sum_{m=1}^{M_0 2^{\ell-1}} \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta'^{(m)}, \xi) \pi_0(\theta'^{(m)})}{q(\theta'^{(m)})}, \\ \varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon) &= \frac{1}{M_0 2^{\ell-1}} \sum_{m=M_0 2^{\ell-1}+1}^{M_0 2^\ell} \frac{\rho(f_\xi(\theta, \epsilon) \mid \theta'^{(m)}, \xi) \pi_0(\theta'^{(m)})}{q(\theta'^{(m)})}, \\ \nabla \varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon) &= \frac{1}{M_0 2^{\ell-1}} \sum_{m=M_0 2^{\ell-1}+1}^{M_0 2^\ell} \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta'^{(m)}, \xi) \pi_0(\theta'^{(m)})}{q(\theta'^{(m)})},\end{aligned}$$

respectively.

Now a sequence of the correction random variables $\Delta\psi_{\xi,0}, \Delta\psi_{\xi,1}, \dots$ is defined by $\Delta\psi_{\xi,0} = \psi_{\xi,M_0,q}$ and

$$(3.4) \quad \begin{aligned} \Delta\psi_{\xi,\ell} &= \psi_{\xi,M_0 2^\ell,q} - \frac{\psi_{\xi,M_0 2^{\ell-1},q}^{(a)} + \psi_{\xi,M_0 2^{\ell-1},q}^{(b)}}{2} \\ &= \frac{1}{2} \left(\frac{\nabla \varrho_{\xi,M_0 2^{\ell-1},q}^{(a)}(\theta, \epsilon)}{\varrho_{\xi,M_0 2^{\ell-1},q}^{(a)}(\theta, \epsilon)} + \frac{\nabla \varrho_{\xi,M_0 2^{\ell-1},q}^{(b)}(\theta, \epsilon)}{\varrho_{\xi,M_0 2^{\ell-1},q}^{(b)}(\theta, \epsilon)} \right) - \frac{\nabla \varrho_{\xi,M_0 2^\ell,q}(\theta, \epsilon)}{\varrho_{\xi,M_0 2^\ell,q}(\theta, \epsilon)}, \end{aligned}$$

for $\ell > 0$. The difference between this antithetic construction and the naive construction is that $\psi_{\xi,M_0 2^{\ell-1},q}$ has been replaced by the mean of $\psi_{\xi,M_0 2^{\ell-1},q}^{(a)}$ and $\psi_{\xi,M_0 2^{\ell-1},q}^{(b)}$. This means that each of the $M_0 2^\ell$ samples is used exactly twice in antithetic construction: once in $\psi_{\xi,M_0 2^\ell,q}$ and once in either $\psi_{\xi,M_0 2^{\ell-1},q}^{(a)}$ or $\psi_{\xi,M_0 2^{\ell-1},q}^{(b)}$. For this novel version of $\Delta\psi_{\xi,\ell}$, the linearity of expectation ensures

$$\begin{aligned} \mathbb{E}[\Delta\psi_{\xi,\ell}] &= \mathbb{E}[\psi_{\xi,M_0 2^\ell,q}] - \frac{1}{2} \left(\mathbb{E}[\psi_{\xi,M_0 2^{\ell-1},q}^{(a)}] + \mathbb{E}[\psi_{\xi,M_0 2^{\ell-1},q}^{(b)}] \right) \\ &= \mathbb{E}[\psi_{\xi,M_0 2^\ell,q}] - \frac{1}{2} \left(\mathbb{E}[\psi_{\xi,M_0 2^{\ell-1},q}] + \mathbb{E}[\psi_{\xi,M_0 2^{\ell-1},q}] \right) \\ &= \mathbb{E}[\psi_{\xi,M_0 2^\ell,q} - \psi_{\xi,M_0 2^{\ell-1},q}], \end{aligned}$$

so that it fits with the telescoping sum representation (3.2) of the gradient $\nabla_\xi U(\xi)$. Despite the distinction being subtle, we will show that the antithetic construction has better properties than the naive construction. Hereafter, $\Delta\psi_{\xi,\ell}$ refers to the antithetic construction given in (3.4).

It is clear that the cost C_ℓ to compute $\Delta\psi_{\xi,\ell}$ is proportional to 2^ℓ , and also that the following *antithetic* properties hold for $\Delta\psi_{\xi,\ell}$:

$$(3.5) \quad \begin{aligned} \varrho_{\xi,M_0 2^\ell,q}(\theta, \epsilon) &= \frac{1}{2} \left(\varrho_{\xi,M_0 2^{\ell-1},q}^{(a)}(\theta, \epsilon) + \varrho_{\xi,M_0 2^{\ell-1},q}^{(b)}(\theta, \epsilon) \right), \quad \text{and} \\ \nabla \varrho_{\xi,M_0 2^\ell,q}(\theta, \epsilon) &= \frac{1}{2} \left(\nabla \varrho_{\xi,M_0 2^{\ell-1},q}^{(a)}(\theta, \epsilon) + \nabla \varrho_{\xi,M_0 2^{\ell-1},q}^{(b)}(\theta, \epsilon) \right), \end{aligned}$$

which play a crucial role in showing that this antithetic construction achieves a faster decay rate of $\mathbb{E}[\|\Delta\psi_{\xi,\ell}\|_2^2]$ than the naive construction, making it possible to find a sequence of positive reals w_0, w_1, \dots which satisfies the condition (3.3). The following claim is the main theoretical result of this paper.

THEOREM 3.1. *Assume that*

$$\sup_{\theta, \theta', \epsilon} \|\nabla_\xi \log \rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)\|_\infty < \infty,$$

and that there exists $u > 2$ such that

$$\mathbb{E}_{\theta \sim \pi_0, \theta' \sim q, \epsilon} \left[\left| \frac{\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) \pi_0(\theta')}{\rho(f_\xi(\theta, \epsilon) \mid \xi) q(\theta')} \right|^u \right] < \infty.$$

Then the following holds true:

1. *For a fixed ℓ , we have*

$$\mathbb{E}[\|\Delta\psi_{\xi,\ell}\|_2^2] = O(2^{-\beta\ell}) \quad \text{with} \quad \beta = \frac{\min(u, 4)}{2}.$$

2. In order to have (3.3), it suffices to choose $w_\ell \propto 2^{-\tau\ell}$ with $1 < \tau < \beta$.

We postpone the proof of the theorem to Appendix A.

Remark 3.2. It follows from the first item of Theorem 3.1 that

$$\mathbb{E}[\|\Delta\psi_{\xi,\ell}\|_2] = O(2^{-\ell}),$$

for a fixed ℓ . Using this property, the bias of the standard Monte Carlo estimator $\psi_{\xi, M_0 2^L, q}$ with $M = M_0 2^L$ inner samples is bounded as

$$\begin{aligned} \|\nabla_\xi U(\xi) - \mathbb{E}[\psi_{\xi, M_0 2^L, q}]\|_2 &= \left\| \sum_{\ell=L+1}^{\infty} \mathbb{E}[\Delta\psi_{\xi,\ell}] \right\|_2 \leq \sum_{\ell=L+1}^{\infty} \mathbb{E}[\|\Delta\psi_{\xi,\ell}\|_2] \\ &= O(2^{-L}) = O(M^{-1}). \end{aligned}$$

This means that, for small M , the standard Monte Carlo estimator may lead to a wrong trajectory of an experimental design in stochastic gradient-biased optimization and the resulting design will not be close to optimal.

3.3. Unbiased MLMC stochastic optimization. Finally we arrive at our proposal of a stochastic algorithm to search for an optimal Bayesian experimental design $\xi^* \in \mathcal{X}$ as summarized in Algorithm 3.1. Here we note that Algorithm 3.1 assumes that the conditions appearing in Theorem 3.1 hold for any $\xi \in \mathcal{X}$ with a common value of u . Given additional assumptions that the domain \mathcal{X} is convex and that U is strongly concave and differentiable with respect to ξ , most of the stochastic gradient-based optimization algorithms have a theoretical guarantee that ξ_t converges to the optimal $\xi^* \in \mathcal{X}$ with some decay rate, typically with the mean square error of $O(1/t)$ as mentioned in Section 2.2.

Algorithm 3.1 Unbiased MLMC stochastic optimization

For a given $1 < \tau < \beta$, set $w_0, w_1, \dots > 0$ such that $w_0 + w_1 + \dots = 1$ and $w_\ell \propto 2^{-\tau\ell}$. For the feasible set \mathcal{X} , initialize $\xi_0 \in \mathcal{X}$ and $t = 0$. For $N \in \mathbb{Z}_{>0}$, do the following:

1. Choose $\ell^{(1)}, \dots, \ell^{(N)} \in \mathbb{Z}_{\geq 0}$ independently and randomly with probability w_ℓ .
2. Compute an unbiased MLMC estimate of the gradient $\nabla_\xi U$ at $\xi = \xi_t$:

$$\frac{1}{N} \sum_{n=1}^N \frac{\Delta\psi_{\xi_t, \ell^{(n)}}}{w_{\ell^{(n)}}}.$$

3. Apply a stochastic gradient-based algorithm to get ξ_{t+1} :

$$\xi_{t+1} = g_t \left(\xi_t, \frac{1}{N} \sum_{n=1}^N \frac{\Delta\psi_{\xi_t, \ell^{(n)}}}{w_{\ell^{(n)}}} \right).$$

4. Check whether a certain stopping criterion is satisfied. If yes, stop the iteration. Otherwise, go to Step 1 with $t \leftarrow t + 1$.
-

As in [2, 16, 5], using Laplace approximation-based importance distribution for q helps not only reduce the expected squared ℓ_2 -norm of the Monte Carlo gradient estimator but also avoid numerical instability coming from concentrated posterior measures of θ' given $f_\xi(\theta, \epsilon)$. We also refer to [35] for some theoretical analyses on the Laplace approximation.

4. Numerical experiments. Here, we conduct numerical experiments on two example problems in Bayesian experimental design. The first example is aimed at verifying our proposed algorithm by using a simple test problem. Then, in order to see practical performance of our algorithm, we consider a PK model used in [33] for our second example. The Python code used in our experiments is available from https://github.com/Goda-Research-Group/MLMC_stochastic_gradient.

4.1. Simple test case. Let $\theta = (\theta_1, \theta_2) \in \mathbb{R}_{>0}^2$ with $\theta_1, \theta_2 \stackrel{\text{iid}}{\sim} \text{lognormal}(\mu, \sigma_0^2)$. For an experimental design $\xi \in \mathbb{R}_{>0}$, let an observation $Y = (Y_1, Y_2) \in \mathbb{R}_{>0}^2$ follow

$$\begin{aligned} Y_1 \mid \theta, \xi &\sim \text{lognormal}(g(\xi) \log \theta_1, \sigma_\epsilon^2), \\ Y_2 \mid \theta, \xi &\sim \text{lognormal}(h(\xi) \log \theta_2, \sigma_\epsilon^2), \end{aligned}$$

for some functions g and h . This is equivalent to consider the following forward model:

$$\begin{aligned} Y_1 &= e^{g(\xi) \log \theta_1 + \sigma_\epsilon \epsilon_1}, \\ Y_2 &= e^{h(\xi) \log \theta_2 + \sigma_\epsilon \epsilon_2}, \end{aligned}$$

for $\epsilon_1, \epsilon_2 \stackrel{\text{iid}}{\sim} N(0, 1)$ independently of θ and ξ , which is obviously a special case of (1.1). The expected information gain for a given ξ is analytically calculated as

$$U(\xi) = \frac{1}{2} \log \left((g(\xi))^2 \frac{\sigma_0^2}{\sigma_\epsilon^2} + 1 \right) \left((h(\xi))^2 \frac{\sigma_0^2}{\sigma_\epsilon^2} + 1 \right).$$

Also, applying Jensen's inequality to (2.1), we see that $U(\xi)$ is bounded above by

$$\begin{aligned} U(\xi) &\leq \tilde{U}(\xi) := \mathbb{E}_{\theta, \epsilon} [\log \rho(f_\xi(\theta, \epsilon) \mid \theta, \xi)] - \mathbb{E}_{\theta, \theta', \epsilon} [\log \rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)] \\ &= (g(\xi))^2 \frac{\sigma_0^2}{\sigma_\epsilon^2} + (h(\xi))^2 \frac{\sigma_0^2}{\sigma_\epsilon^2}. \end{aligned}$$

Here we note that the standard Monte Carlo gradient estimator with $M = 1$ inner sample from the prior distribution, i.e., $\psi_{\xi, 1}$, is nothing but an unbiased estimator of $\nabla_\xi \tilde{U}(\xi)$. Therefore, as long as

$$\xi^* = \arg \max_{\xi \in \mathbb{R}_{>0}} U(\xi) \neq \arg \max_{\xi \in \mathbb{R}_{>0}} \tilde{U}(\xi)$$

holds, stochastic gradient-based optimization based on $\psi_{\xi, 1}$ will not converge to the optimal design ξ^* . In our experiments below, let $\mu = 0$, $\sigma_0 = \sigma_\epsilon = 1$,

$$g(\xi) = e^{-\xi^2/2} \quad \text{and} \quad h(\xi) = \sqrt{\frac{3}{2}} (1 - e^{-\xi^2}).$$

Fig. 1 compares U and \tilde{U} as functions of ξ for this setting. The optimal design which maximizes U is given by $\xi^* = \sqrt{\log 3} \approx 1.048 \dots$ and we can observe that U is concave around ξ^* . On the other hand, its upper bound \tilde{U} is a strictly monotone increasing function and its supremum attains for $\xi \rightarrow \infty$.

Throughout this subsection, we do not use any importance sampling for the unbiased MLMC estimator of $\nabla_\xi U$ and set M_0 , the number of level 0 inner samples, to 1. The left panel of Fig. 2 shows the convergence behavior of the MLMC correction variables $\Delta\psi_{\xi, \ell}$ at $\xi = 1.5$. Here the mean squares (expected squared ℓ_2 -norms) of

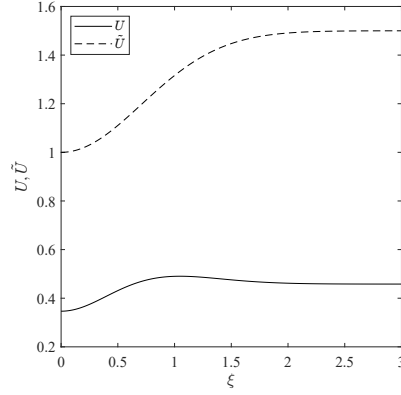


FIG. 1. The expected information gain U and its upper bound \tilde{U} for the test case.

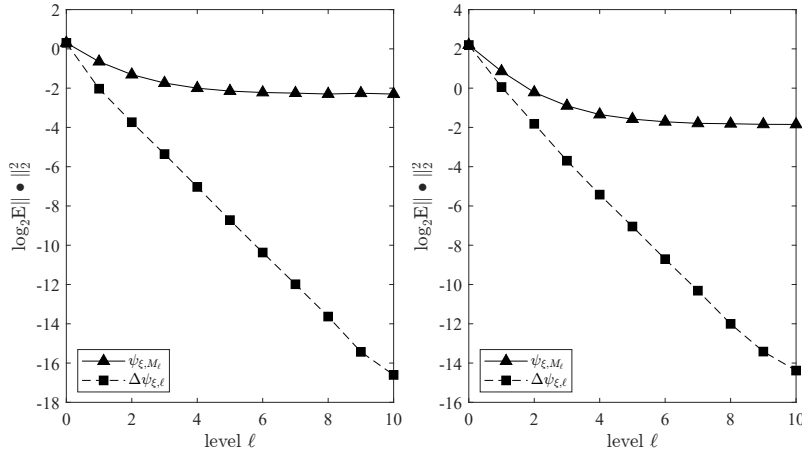


FIG. 2. The mean squares of the variables ψ_{ξ, M_ℓ} and $\Delta\psi_{\xi, \ell}$ for the test case at $\xi = 1.5$ (left) and at $\xi = \xi^* = \sqrt{\log 3}$ (right).

ψ_{ξ, M_ℓ} and $\Delta\psi_{\xi, \ell}$ are plotted on a \log_2 scale as functions of the level ℓ , where the means are estimated empirically by using 10^5 i.i.d. samples at each level. While the mean square of ψ_{ξ, M_ℓ} takes an almost constant value for $\ell > 4$, that of $\Delta\psi_{\xi, \ell}$ decreases geometrically as the level increases. The linear regression of the data for the range $1 \leq \ell \leq 10$ provides an estimation of β as 1.64, which agrees well with the theoretical result in Theorem 3.1. As shown in the right panel of Fig. 2, a similar convergence behavior of the MLMC correction variables $\Delta\psi_{\xi, \ell}$ can be observed at the optimal design $\xi = \xi^* = \sqrt{\log 3}$, where β is estimated as 1.63.

Such a fast geometric decay of the correction variables $\Delta\psi_{\xi, \ell}$ justifies us to apply Algorithm 3.1 to search for the optimal design ξ^* . In order to randomly choose the level ℓ , we set $\tau = 1.5$ and $w_\ell = 2^{-3\ell/2}(1 - 2^{-3/2})$. This implies that the expected number of inner samples used in the MLMC estimator is given by

$$\sum_{\ell=0}^{\infty} 2^\ell w_\ell = (1 - 2^{-3/2}) \sum_{\ell=0}^{\infty} 2^{-\ell/2} = \frac{1 - 2^{-3/2}}{1 - 2^{-1/2}} \approx 2.21.$$

For comparison, we also consider the standard Monte Carlo estimators $\psi_{\xi, M}$ for the

gradient of the expected information gain with various values of $M = 1, 2, 4, \dots, 64$ within stochastic gradient descent. We fix the number of outer samples N to 2000 throughout all the iteration steps for all the estimators. We use the Robbins-Monro algorithm with Polyak-Ruppert averaging and the learning rates $\alpha_t = 5/(t+1)$ as a stochastic descent algorithm, and as the computational cost is proportional to the number of inner samples, we set the maximum iteration steps T to $\lfloor 10^7/M \rfloor$, the largest integer less than or equal to $10^7/M$. Although the number of inner samples is a random variable for the MLMC estimator, we simply set T to $\lfloor 10^7/2.21 \rfloor$. The initial design candidate at $t = 0$ is given by $\xi_0 = 1.5$ and the feasible set \mathcal{X} is set to $\mathbb{R}_{>0}$. Hence the maximum increment of the expected information gain is $U(\sqrt{\log 3}) - U(1.5) \approx 0.0148$. For each gradient estimator, we conduct 10 independent runs and compute the average of the distance $\|\xi_t - \xi^*\|_2^2$ and its standard error for all the iteration steps, which correspond to the line and the shaded area of Fig. 3, respectively.

Fig. 3 shows the convergence behaviors of the estimated experimental design ξ_t for the considered estimators of the gradient $\nabla_\xi U$. Note here that the horizontal axis is given by $M \times t$ as a measure of the total computational cost (here again, we simply let $M = 2.21$ for the MLMC estimator) and both axes use the logarithmic scales. As expected, the standard Monte Carlo estimator with $M = 1$ leads to larger values of ξ_t which make \tilde{U} large, so that the search goes in wrong direction. Even for $M = 2$, the situation is not improved so much and the experimental design ξ_t remains almost the same throughout the iterations. For larger values of M , the standard Monte Carlo estimator works in the early stages, making the distance $\|\xi_t - \xi^*\|_2^2$ small. However, after some iteration steps, the estimate ξ_t converges to some point away from the optimal ξ^* . Although it is natural that such bias can be reduced simply by increasing M , a proper choice of M in practical applications is far from trivial since larger M means a larger computational cost and the bias seems extremely hard to estimate in advance. This is exactly the point where the unbiased MLMC estimator can help. As the black line shows, the distance $\|\xi_t - \xi^*\|_2^2$ decreases consistently from the early stage and overtakes the standard Monte Carlo estimators with fixed M , leading to a better estimate of the optimal experimental design. The linear regression of the data for the whole range $0 < \log_{10}(Mt) \leq 7$ shows that the estimate ξ_t converges to ξ^* with the mean squared error of order $t^{-1.12}$ approximately, which is almost consistent with the standard stochastic optimization theory [37, Chapter 5.9]. A slightly faster decay of the standard Monte Carlo estimators with $M \geq 8$ in the early stages could be because that they estimate the gradients of biased objective functions which are steeper than the gradient of $U(\xi)$ around the initial estimate $\xi_0 = 1.5$ in this case.

4.2. Pharmacokinetic model. Let us consider a PK design problem introduced in [33]. Suppose that a drug with a fixed dose $D = 400$ is administrated to subjects at time $\mathcal{T} = 0$. In order to reduce the uncertainty about a set of PK parameters, which affect the absorption, distribution and the elimination of the drug in the subjects' body, it would be helpful to take blood samples of the subjects at several different times and to measure the concentration of the drug in the samples. Blood samples are assumed to be taken 15 times at $\mathcal{T} = \xi^{(1)}, \dots, \xi^{(15)}$ hours after the drug administration. Given the set of 15 drug concentration measurements, it is expected that the uncertainty of PK parameters of interest θ can be reduced. Our objective here is to optimize sampling times $\xi = (\xi^{(1)}, \dots, \xi^{(15)}) \in \mathbb{R}_{>0}^{15}$ such that the expected information gain brought from blood sampling is maximized.

Let $\theta = (\log k_a, \log k_e, \log V) \in \mathbb{R}^3$ where k_a represents the first-order absorption rate constant, k_e does the first-order elimination rate constant and V does the volume

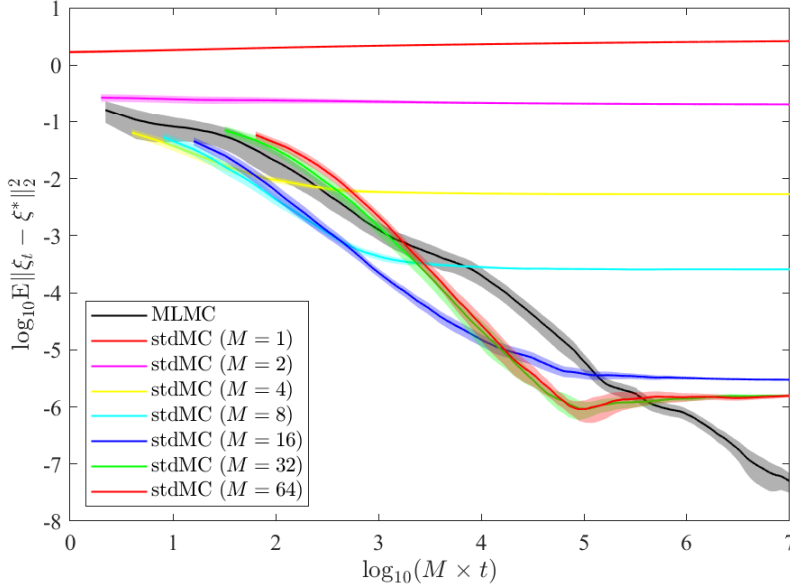


FIG. 3. The convergence of the estimated experimental design ξ_t to the optimal ξ^* for various Monte Carlo estimators of the gradient $\nabla_{\xi} U$. For each estimator, the line and the shaded area represent the average and its standard error estimated from 10 independent runs, respectively.

of distribution. Following [33], assume that the drug concentration of blood sample taken at time $\mathcal{T} \geq 0$ is described as

$$Y_{\mathcal{T}} = \frac{Dk_a}{V(k_a - k_e)} (e^{-k_e \mathcal{T}} - e^{-k_a \mathcal{T}}) (1 + \epsilon_1) + \epsilon_2 =: g_{\mathcal{T}}(\theta, \epsilon),$$

with $\epsilon = (\epsilon_1, \epsilon_2)$, where ϵ_1 and ϵ_2 represent the multiplicative and additive Gaussian noises, respectively. Then our forward model is given by

$$Y = (Y_{\xi(1)}, \dots, Y_{\xi(15)}) = (g_{\xi(1)}(\theta, \epsilon_{\xi(1)}), \dots, g_{\xi(15)}(\theta, \epsilon_{\xi(15)})) \in \mathbb{R}^{15},$$

where $\epsilon_{\xi(1)}, \dots, \epsilon_{\xi(15)}$ are assumed mutually independent and follow the same bi-variate normal distribution

$$\epsilon_{\xi(j)} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.01 & 0 \\ 0 & 0.1 \end{pmatrix} \right).$$

The input random variables in θ are assumed independent and the corresponding probability distributions are given by $\log k_a \sim N(0, 0.05)$, $\log k_e \sim N(\log(0.1), 0.05)$ and $\log V \sim N(\log(20), 0.05)$, respectively. This means that the prior information entropy of θ is equal to $3 \log(\sqrt{2\pi e} \times 0.05) \approx -0.2368$. Moreover, the likelihood function is given by the product of $\rho(g_{\xi(j)}(\theta, \epsilon_{\xi(j)}) | \theta', \xi^{(j)})$ that can be computed explicitly by following Example 2.3.

In this setting the posterior distribution of θ given Y cannot be computed analytically. In order to reduce the expected squared ℓ_2 -norm of the unbiased MLMC estimator of the gradient $\nabla_{\xi} U$, we use Laplace approximation-based importance sampling. Since not only the additive noise but also the multiplicative noise are included in the forward model, we consider a simple modification of the original method in [22]

as follows. Let us write

$$\overline{g_T}(\theta) = \frac{Dk_a}{V(k_a - k_e)} (e^{-k_e T} - e^{-k_a T}) \quad \text{and} \quad \overline{g_\xi}(\theta) = (\overline{g_{\xi^{(1)}}}(\theta), \dots, \overline{g_{\xi^{(15)}}}(\theta)).$$

Then, for the data Y generated conditionally on the known value of $\theta = \theta^*$ from the forward model, we approximate the posterior distribution $\pi^{Y|\xi}(\theta)$ by a Gaussian distribution $N(\hat{\theta}, \hat{\Sigma})$ with

$$\begin{aligned} \hat{\theta} &= \theta^* - (J(\theta^*)^\top \Sigma_\epsilon^{-1} J(\theta^*) + H(\theta^*)^\top \Sigma_\epsilon^{-1} E - \nabla_\theta \nabla_\theta \log \pi_0(\theta^*))^{-1} J(\theta^*)^\top \Sigma_\epsilon^{-1} E, \\ \hat{\Sigma} &= \left(J(\hat{\theta})^\top \Sigma_\epsilon^{-1} J(\hat{\theta}) - \nabla_\theta \nabla_\theta \log \pi_0(\hat{\theta}) \right)^{-1}. \end{aligned}$$

Here J and H denote the Jacobian and Hessian of $-\overline{g_\xi}$, respectively, that is, $J(\theta) = -\nabla_\theta \overline{g_\xi}(\theta)$ and $H(\theta) = -\nabla_\theta \nabla_\theta \overline{g_\xi}(\theta)$. Also we write $E := Y^\top - \overline{g_\xi}(\theta^*)^\top$ and

$$\Sigma_\epsilon = \text{diag} \left(0.01 (\overline{g_{\xi^{(1)}}}(\theta))^2 + 0.1, \dots, 0.01 (\overline{g_{\xi^{(15)}}}(\theta))^2 + 0.1 \right).$$

We use this $N(\hat{\theta}, \hat{\Sigma})$ as an importance distribution q . The only difference from the one in [22] is that the matrix Σ_ϵ depends on the mean response $\overline{g_\xi}(\theta)$ due to the multiplicative noise in our setting. Although a first-order approximation argument similar to [22] might be possible and lead to different forms of $\hat{\theta}$ and $\hat{\Sigma}$, such a detailed analysis on the Laplace approximation is beyond the scope of this paper.

In order to search for optimal design parameters $\xi = (\xi^{(1)}, \dots, \xi^{(15)})$, we do not represent them by a smaller number of parameters as considered in [33], but instead we optimize them directly. We set a design at the initial iteration step $t = 0$ to equi-spaced times $\xi_0 = (1, 2, \dots, 15)$. In Algorithm 3.1, we fix $M_0 = 1$, set $w_0 = 0.9$ and $w_\ell \propto 2^{-3\ell/2}$ for $\ell \geq 1$ such that they are summed up to 1, and set the number of outer samples to $N = 2000$ at each iteration step. This implies that the expected number of inner samples used in the MLMC estimator is given by

$$\sum_{\ell=0}^{\infty} 2^\ell w_\ell = \frac{9}{10} + \frac{2^{3/2} - 1}{10} \sum_{\ell=1}^{\infty} 2^{-\ell/2} \approx 1.34.$$

We use the AMSGrad optimizer with constant learning rate $\alpha_t = 0.004$ and exponential moving average parameters $\beta_1 = 0.9, \beta_2 = 0.999$ as a stochastic descent algorithm, and set the maximum iteration steps T to 10000 as a stopping criterion. The feasible domain \mathcal{X} is restricted to $[0, 24]^{15}$. For comparison, we also consider the standard (biased) Monte Carlo estimator for the gradient $\nabla_\xi U$ with a fixed number of inner samples $M = 1$ and the Laplace approximation-based importance sampling within stochastic gradient descent. As expected from the numerical results shown in [5], the Laplace approximation-based importance sampling helps reduce the bias of the Monte Carlo estimator significantly even for $M = 1$.

Fig. 4 shows the set of design parameters $\xi = (\xi^{(1)}, \dots, \xi^{(15)})$ obtained at the iteration steps $t = 0, 100, 500, 1000, 5000, 10000$ for a single run. The overall convergence behaviors both for the standard Monte Carlo estimator and the MLMC estimator look quite similar to each other. That is, the allocations of 15 sampling times become irregular at the earlier steps compared to the initially equi-spaced design, but then some of sampling times gradually get quite close to each other, ending up with three well-separated clusters. It is interesting to see that stochastic gradient-based optimization

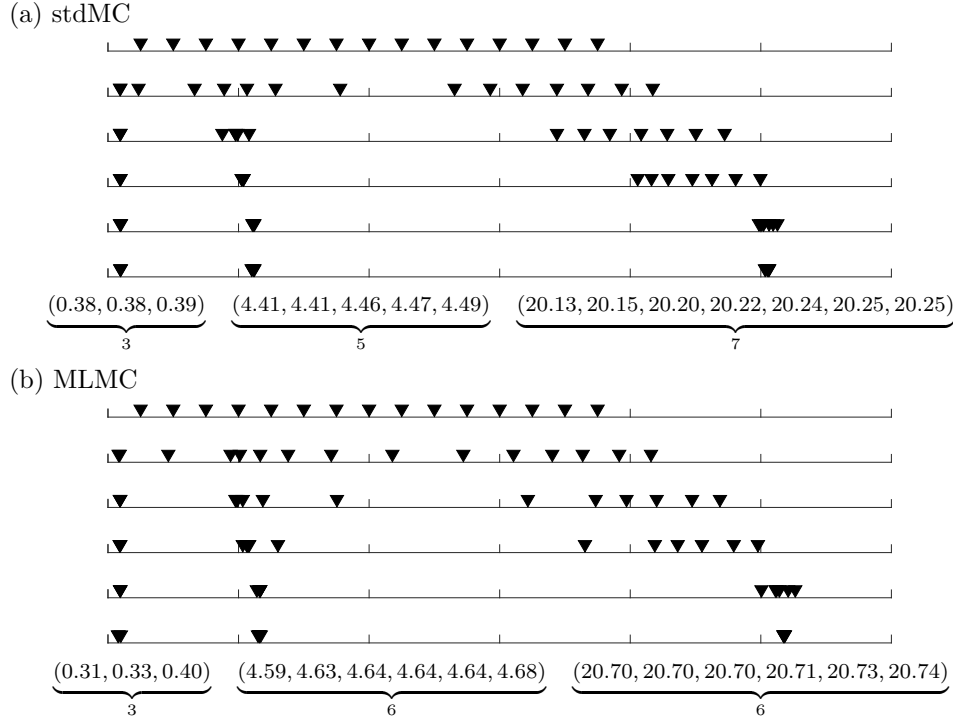


FIG. 4. Design parameters (ξ_1, \dots, ξ_{15}) within the interval $[0, 24]$ during the optimization process for a single run at the iteration steps $t = 0, 100, 500, 1000, 5000, 10000$ (in descending order): (a) the result for stdMC and (b) the result for MLMC. The resulting design is shown in detail respectively at the bottom.

naturally finds such so-called *replicate* design that is often considered in the PK applications [32, 33]. Looking into the details, there is a difference between the resulting designs obtained by the standard Monte Carlo estimator and the MLMC estimator. For the standard Monte Carlo estimator, the number of sampling times allocated to each cluster is 3, 5, 7 (from earlier one to later one), respectively, whereas the corresponding number is 3, 6, 6, respectively, for the MLMC estimator. These allocations of sampling times are consistent among 10 independent runs for both the estimators. The average sampling time (with its standard deviation) within each cluster, estimated from 10 independent runs, is 0.385 (0.003), 4.442 (0.008), 20.202 (0.006) for the standard Monte Carlo estimator, and is 0.367 (0.010), 4.652 (0.018), 20.699 (0.017) for the MLMC estimator. The two-sample Wilcoxon test yields the p-value about 10^{-5} for all of the three clusters, which supports that the differences between the centers of the clusters obtained by the two estimators are statistically significant.

Fig. 5 shows the convergence behaviors of the MLMC correction variables $\Delta\psi_{\xi, \ell}$ at the iteration steps $t = 0, T/2, T$ for a single run. Similarly to Fig. 2, the mean squares (expected squared ℓ_2 -norms) of ψ_{ξ, M_ℓ} and $\Delta\psi_{\xi, \ell}$ are plotted on a \log_2 scale as functions of the level ℓ , where the means are estimated empirically by using 10^5 i.i.d. samples at each level. While the mean square of ψ_{ξ, M_ℓ} takes an almost constant value for $\ell > 4$, that of $\Delta\psi_{\xi, \ell}$ decreases geometrically as the level increases. The linear regression of the data for the range $1 \leq \ell \leq 10$ provides estimations of β as 0.80, 1.36, 1.47, respectively. The result on the case $\beta \leq 1$ is not covered by Theorem 3.1, and in such

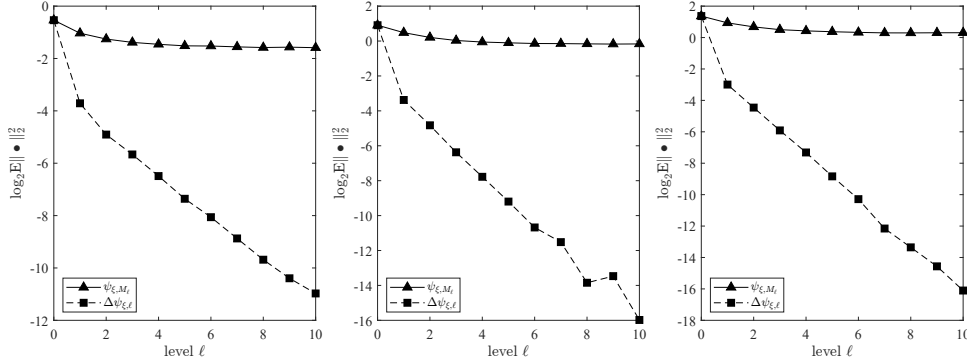


FIG. 5. The mean squares of the variables ψ_{ξ, M_ℓ} and $\Delta\psi_{\xi, \ell}$ for the PK model at the iteration steps $t = 0, T/2, T$

a case, we do not have a right choice of w_ℓ which leads to both finite expected cost and finite expected squared ℓ_2 -norm. Further theoretical investigation is needed to address this issue. On the other hand, the result $\beta > 1$ for the steps $t = T/2, T$ is as expected from our theoretical result. Nonetheless, our choice $w_\ell \propto 2^{-3\ell/2}$ might be a bit aggressive in the sense that the expected squared ℓ_2 -norm of the MLMC estimator possibly does not converge, although we see no evidence of this in our experiments. A practical issue on how to choose w_ℓ properly depending on the problem at hand is also left open for future research.

Finally, Fig. 6 shows the behaviors of the expected information gain U as a function of the number of iteration steps. For this problem, the expected information gain for any design parameter ξ cannot be evaluated exactly, so that we use a randomized variant of the MLMC estimator introduced in [16] with 10^6 outer samples to estimate the expected information gain for every 500 steps. As 10 independent runs are performed, we plot the average of 10 estimated values in mark, while the shaded area represents the linearly interpolated standard error. We can see that the expected information gain increases with some fluctuation as the iteration proceeds, and converges to a constant value. The average converged value for the MLMC estimator is 4.544, which is slightly larger than 4.535 obtained for the standard Monte Carlo estimator. Note that the expected information gain for the initial design is estimated as 3.774, which is well below the maximum values obtained both for the standard Monte Carlo estimator and the MLMC estimator. Just to provide an intuition of this improvement, assume that each individual variable in θ remains independent and follows a normal distribution with an equal variance even after observing Y , which is usually not true. Then it is inferred that the variance of each variable after observing Y with the initial design is reduced on average by the factor $(\exp(3.774/3))^2 \approx 12.379$, whereas that with the resulting design by our proposed optimization algorithm is $(\exp(4.554/3))^2 \approx 20.822$. Although the increment of the maximum expected information gain by using the MLMC estimator seems marginal as compared to the standard Monte Carlo estimator in this example, it is important to emphasize again that the resulting experimental designs are qualitatively different.

5. Conclusion. In this paper we have developed an efficient stochastic algorithm to optimize Bayesian experimental designs such that the expected information gain is maximized. Since the gradient of the expected information gain with respect to design parameters is expressed as a nested expectation, a straightforward use of stochastic

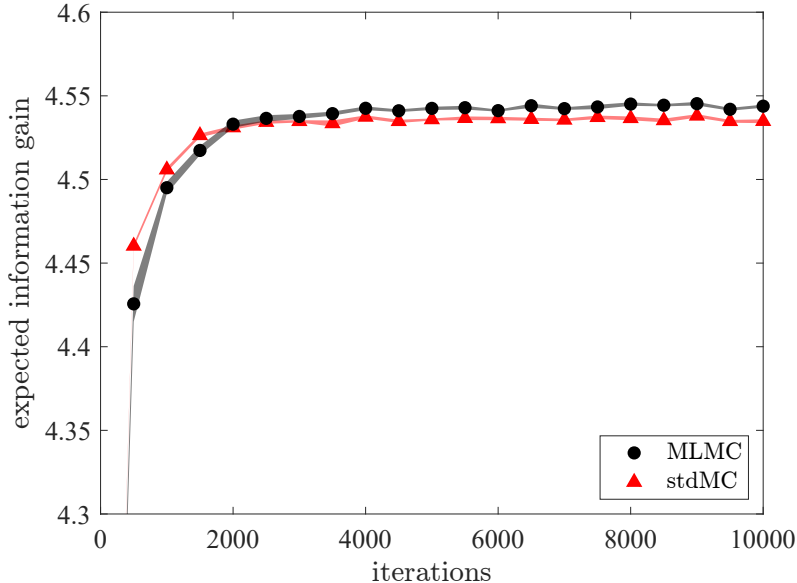


FIG. 6. *The behavior of the expected information gain as a function of the number of iteration steps for the PK model*

gradient-based optimization algorithms in which the number of inner Monte Carlo samples is kept fixed only gives a biased solution of Bayesian experimental design unless i.i.d. sampling from the exact posterior distribution is possible. To overcome this issue, we have introduced an unbiased antithetic multilevel Monte Carlo estimator for the gradient of the expected information gain, and have proven under some conditions that our estimator is unbiased and has finite expected squared ℓ_2 -norm and finite computational cost per one sample. This way, combining our unbiased multilevel estimator with stochastic gradient-based optimization algorithms leads to a novel stochastic algorithm to search for optimal Bayesian experimental designs without suffering from any bias. Numerical experiments for a simple test case show that our proposed algorithm can find the true optimal Bayesian experimental design with the convergence behavior as expected from the standard stochastic optimization theory which is built upon the underlying assumption that an unbiased gradient estimation is possible. In contrast, using the standard Monte Carlo estimator with a fixed number of inner samples fails to reach the optimal design. Moreover, our proposed algorithm performs well for a more realistic pharmacokinetic test problem and gives a higher expected information gain and qualitatively different sampling times compared to designs obtained by the existing standard Monte Carlo estimator.

Acknowledgements. The authors would like to thank the reviewers for their helpful comments and suggestions which lead to a significant improvement over the original manuscript. The authors are also grateful to Takuro Mori (University of Tokyo) for his help on numerical experiments during a revision process. TG would like to thank Professor Mike Giles (University of Oxford) for useful discussions and comments at an early stage of this research.

Appendix A. Proof of Theorem 3.1. The proof for the first assertion follows an argument similar to that of [17, Lemma 3.9] which considers a nested expectation

involving the ratio of two scalar inner conditional expectations. Since the numerator is vector-valued in our setting, however, we give a proof for the sake of completeness.

First let us recall the following result proven, for instance, in [14, Lemma 1].

LEMMA A.1. *Let X be a real-valued random variable with mean zero, and let \bar{X}_N be an average of N i.i.d. samples of X . If $\mathbb{E}[|X|^u] < \infty$ for $u > 2$, there exists a constant $C_u > 0$ depending only on u such that*

$$\mathbb{E}[|\bar{X}_N|^u] \leq C_u \frac{\mathbb{E}[|X|^u]}{N^{u/2}} \quad \text{and} \quad \mathbb{P}[|\bar{X}_N| > c] \leq C_u \frac{\mathbb{E}[|X|^u]}{c^u N^{u/2}},$$

for any $c > 0$.

For any θ, ϵ and ξ , we write $\rho(f_\xi(\theta, \epsilon) \mid \xi) = \mathbb{E}_{\theta'}[\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]$ and also $\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi) = \mathbb{E}_{\theta'}[\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)]$. For randomly chosen θ and ϵ , we define an extreme event A by

$$A := \left\{ \left| \frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right| > \frac{1}{2} \right\} \cup \left\{ \left| \frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right| > \frac{1}{2} \right\}.$$

Then we have

$$(A.1) \quad \mathbb{E}[\|\Delta\psi_{\xi, \ell}\|_2^2] = \mathbb{E}[\|\Delta\psi_{\xi, \ell}\|_2^2 \mathbf{1}_A] + \mathbb{E}[\|\Delta\psi_{\xi, \ell}\|_2^2 \mathbf{1}_{A^c}],$$

where $\mathbf{1}_\bullet$ denotes the indicator function of an event \bullet and A^c denotes the complement of the event A .

Let us look at the first term on the right-hand side of (A.1). Since we use the same i.i.d. samples of $\theta' \sim q$ in the denominator and numerator for the three terms of $\Delta\psi_{\xi, \ell}$, i.e., $\psi_{\xi, M_0 2^\ell, q}, \psi_{\xi, M_0 2^{\ell-1}, q}^{(a)}, \psi_{\xi, M_0 2^{\ell-1}, q}^{(b)}$, it follows from the assumption

$$\sup_{\theta, \theta', \epsilon} \|\nabla_\xi \log \rho(f_\xi(\theta, \epsilon) \mid \theta', \xi)\|_\infty =: \varrho_{\max} < \infty$$

that $\|\psi_{\xi, M_0 2^\ell, q}\|_2^2, \|\psi_{\xi, M_0 2^{\ell-1}, q}^{(a)}\|_2^2, \|\psi_{\xi, M_0 2^{\ell-1}, q}^{(b)}\|_2^2 \leq 2d\varrho_{\max}^2$ where d denotes the cardinality of ξ . Applying Jensen's inequality leads to a bound

$$\begin{aligned} \|\Delta\psi_{\xi, \ell}\|_2^2 &\leq \left(\|\psi_{\xi, M_0 2^\ell, q}\|_2 + \frac{\|\psi_{\xi, M_0 2^{\ell-1}, q}^{(a)}\|_2}{2} + \frac{\|\psi_{\xi, M_0 2^{\ell-1}, q}^{(b)}\|_2}{2} \right)^2 \\ &\leq 2\|\psi_{\xi, M_0 2^\ell, q}\|_2^2 + \|\psi_{\xi, M_0 2^{\ell-1}, q}^{(a)}\|_2^2 + \|\psi_{\xi, M_0 2^{\ell-1}, q}^{(b)}\|_2^2 \leq 8d\varrho_{\max}^2. \end{aligned}$$

Thus we have

$$\mathbb{E}[\|\Delta\psi_{\xi, \ell}\|_2^2 \mathbf{1}_A] \leq 8d\varrho_{\max}^2 \mathbb{E}[\mathbf{1}_A] = 8d\varrho_{\max}^2 \mathbb{P}[A].$$

Noting that both $\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)$ and $\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)$ are unbiased estimates of the target quantity $\rho(f_\xi(\theta, \epsilon) \mid \xi)$ using $M_0 2^{\ell-1}$ random samples of $\theta' \sim q$, it follows from the assumption of the theorem and Lemma A.1 that

$$\mathbb{P}[A] \leq \mathbb{P}\left[\left| \frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right| > \frac{1}{2}\right] + \mathbb{P}\left[\left| \frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right| > \frac{1}{2}\right]$$

$$\begin{aligned}
&\leq \frac{2^{u+1}C_u}{(M_0 2^{\ell-1})^{u/2}} \mathbb{E}_{\theta, \theta', \epsilon} \left[\left| \frac{\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) \pi_0(\theta')}{\rho(f_\xi(\theta, \epsilon) \mid \xi) q(\theta')} - 1 \right|^u \right] \\
&\leq \frac{2^{u+1}C_u}{(M_0 2^{\ell-1})^{u/2}} \left(\mathbb{E}_{\theta, \theta', \epsilon} \left[\left| \frac{\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) \pi_0(\theta')}{\rho(f_\xi(\theta, \epsilon) \mid \xi) q(\theta')} \right|^u \right] + 1 \right).
\end{aligned}$$

This gives a bound on the term $\mathbb{E}[\|\Delta\psi_{\xi, \ell}\|_2^2 \mathbf{1}_A]$ of order $2^{-(u/2)\ell}$.

Next let us look at the second term on the right-hand side of (A.1). By using the antithetic properties (3.5), we have

$$\begin{aligned}
&\Delta\psi_{\xi, \ell} \\
&= \frac{1}{2} \left(\nabla \varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi) \right) \left(\frac{1}{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)} - \frac{1}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right) \\
&\quad + \frac{1}{2} \left(\nabla \varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi) \right) \left(\frac{1}{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)} - \frac{1}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right) \\
&\quad - \left(\nabla \varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi) \right) \left(\frac{1}{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)} - \frac{1}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right) \\
&\quad + \frac{1}{2} \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)} \left(\frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^2 \\
&\quad + \frac{1}{2} \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)} \left(\frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^2 \\
&\quad - \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)} \left(\frac{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^2.
\end{aligned}$$

Noting that

$$\left| \frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right|, \left| \frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right| \leq \frac{1}{2}$$

on A^c and that $\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)$, $\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)$ and $\rho(f_\xi(\theta, \epsilon) \mid \xi)$ are strictly positive by assumption, it holds that

$$\frac{1}{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)}, \frac{1}{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)} \leq \frac{2}{\rho(f_\xi(\theta, \epsilon) \mid \xi)}.$$

The same bound exists also for $\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)$ because of the antithetic property (3.5). By applying Jensen's inequality and then using these bounds, we obtain

$$\begin{aligned}
\|\Delta\psi_{\xi, \ell}\|_2^2 &\leq 2 \left\| \frac{\nabla \varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^2 \\
&\quad + 2 \left\| \frac{\nabla \varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^2 \\
&\quad - \left\| \nabla \varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi) \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^2.
\end{aligned}$$

$$\begin{aligned}
& + 4 \left\| \frac{\nabla \varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^2 \\
& + 2 \left\| \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^4 \\
& + 2 \left\| \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^4 \\
& + 4 \left\| \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^4 \\
& \leq 8 \left\| \frac{\nabla \varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^2 \\
& + 8 \left\| \frac{\nabla \varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^2 \\
& + 16 \left\| \frac{\nabla \varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^2 \\
& + 8 \left\| \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(a)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^4 \\
& + 8 \left\| \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^{\ell-1}, q}^{(b)}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^4 \\
& + 16 \left\| \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^4.
\end{aligned}
\tag{A.2}$$

Let us focus on the third term of (A.2). Applying Hölder's inequality gives

$$\begin{aligned}
& \mathbb{E} \left[\left\| \frac{\nabla \varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^2 \mathbf{1}_{A^c} \right] \\
& \leq \mathbb{E} \left[\left\| \frac{\nabla \varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_2^2 \right. \\
& \quad \times \left. 2^{\max(4-u, 0)} \left| \frac{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right|^{\min(u, 4) - 2} \right] \\
& \leq \left(\mathbb{E} \left[\left\| \frac{\nabla \varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_2^{\min(u, 4)} \right] \right)^{2/\min(u, 4)} \\
& \quad \times 2^{\max(4-u, 0)} \left(\mathbb{E} \left[\left| \frac{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right|^{\min(u, 4)} \right] \right)^{1-2/\min(u, 4)}.
\end{aligned}$$

Using Jensen's inequality and Lemma A.1, the first factor above is bounded by

$$\begin{aligned}
& \mathbb{E} \left[\left\| \frac{\nabla \varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_2^{\min(u, 4)} \right] \\
& \leq d^{\min(u, 4)/2-1} \mathbb{E} \left[\left\| \frac{\nabla \varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon) - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_{\min(u, 4)}^{\min(u, 4)} \right] \\
& \leq \frac{d^{\min(u, 4)/2-1} C_{\min(u, 4)}}{(M_0 2^\ell)^{\min(u, 4)/2}} \\
& \quad \times \mathbb{E}_{\theta, \theta', \epsilon} \left[\left\| \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) \pi_0(\theta')/q(\theta') - \nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_{\min(u, 4)}^{\min(u, 4)} \right] \\
& \leq \frac{2^{\min(u, 4)-1} d^{\min(u, 4)/2-1} C_{\min(u, 4)}}{(M_0 2^\ell)^{\min(u, 4)/2}} \\
& \quad \times \mathbb{E}_{\theta, \theta', \epsilon} \left[\left\| \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) \pi_0(\theta')/q(\theta')}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_{\min(u, 4)}^{\min(u, 4)} + \left\| \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_{\min(u, 4)}^{\min(u, 4)} \right] \\
& \leq \frac{2^{\min(u, 4)-1} d^{\min(u, 4)/2-1} C_{\min(u, 4)}}{(M_0 2^\ell)^{\min(u, 4)/2}} \\
& \quad \times \mathbb{E}_{\theta, \theta', \epsilon} \left[\left\| \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) \pi_0(\theta')/q(\theta')}{\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) \pi_0(\theta')/q(\theta')} \cdot \frac{\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) \pi_0(\theta')/q(\theta')}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_{\min(u, 4)}^{\min(u, 4)} \right. \\
& \quad \left. + \left\| \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_{\min(u, 4)}^{\min(u, 4)} \right] \\
& \leq \frac{2^{\min(u, 4)-1} d^{\min(u, 4)/2} \varrho_{\max}^{\min(u, 4)} C_{\min(u, 4)}}{(M_0 2^\ell)^{\min(u, 4)/2}} \\
& \quad \times \left(\mathbb{E}_{\theta, \theta', \epsilon} \left[\left| \frac{\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) \pi_0(\theta')}{\rho(f_\xi(\theta, \epsilon) \mid \xi) q(\theta')} \right|^{\min(u, 4)} \right] + 1 \right),
\end{aligned}$$

whereas a bound on the second factor directly follows from Lemma A.1, i.e., we have

$$\begin{aligned}
& \mathbb{E} \left[\left| \frac{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right|^{\min(u, 4)} \right] \\
& \leq \frac{C_{\min(u, 4)}}{(M_0 2^\ell)^{\min(u, 4)/2}} \mathbb{E}_{\theta, \theta', \epsilon} \left[\left| \frac{\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) \pi_0(\theta')}{\rho(f_\xi(\theta, \epsilon) \mid \xi) q(\theta')} - 1 \right|^{\min(u, 4)} \right] \\
& \leq \frac{C_{\min(u, 4)}}{(M_0 2^\ell)^{\min(u, 4)/2}} \left(\mathbb{E}_{\theta, \theta', \epsilon} \left[\left| \frac{\rho(f_\xi(\theta, \epsilon) \mid \theta', \xi) \pi_0(\theta')}{\rho(f_\xi(\theta, \epsilon) \mid \xi) q(\theta')} \right|^{\min(u, 4)} \right] + 1 \right).
\end{aligned}$$

Substituting these bounds shows that the third term is of order

$$\left(2^{-\min(u, 4)\ell/2} \right)^{2/\min(u, 4)} \cdot \left(2^{-\min(u, 4)\ell/2} \right)^{1-2/\min(u, 4)} = 2^{-\min(u, 4)\ell/2}$$

for given $u > 2$.

Similarly, the expectation of the sixth term of (A.2) can be bounded above by

$$\mathbb{E} \left[\left\| \frac{\nabla_\xi \rho(f_\xi(\theta, \epsilon) \mid \xi)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} \right\|_2^2 \left(\frac{\varrho_{\xi, M_0 2^\ell, q}(\theta, \epsilon)}{\rho(f_\xi(\theta, \epsilon) \mid \xi)} - 1 \right)^4 \mathbf{1}_{A^c} \right]$$

$$\begin{aligned}
&\leq 2^{\max(4-u,0)} \mathbb{E} \left[\left\| \frac{\nabla_{\xi} \rho(f_{\xi}(\theta, \epsilon) \mid \xi)}{\rho(f_{\xi}(\theta, \epsilon) \mid \xi)} \right\|_2^2 \left| \frac{\varrho_{\xi, M_0 2^{\ell}, q}(\theta, \epsilon)}{\rho(f_{\xi}(\theta, \epsilon) \mid \xi)} - 1 \right|^{\min(u,4)} \right] \\
&\leq 2^{\max(4-u,0)} d \varrho_{\max}^2 \mathbb{E} \left[\left| \frac{\varrho_{\xi, M_0 2^{\ell}, q}(\theta, \epsilon)}{\rho(f_{\xi}(\theta, \epsilon) \mid \xi)} - 1 \right|^{\min(u,4)} \right] \\
&\leq \frac{2^{\max(4-u,0)} d \varrho_{\max}^2 C_{\min(u,4)}}{(M_0 2^{\ell})^{\min(u,4)/2}} \left(\mathbb{E}_{\theta, \theta', \epsilon} \left[\left| \frac{\rho(f_{\xi}(\theta, \epsilon) \mid \theta', \xi) \pi_0(\theta')}{\rho(f_{\xi}(\theta, \epsilon) \mid \xi) q(\theta')} \right|^{\min(u,4)} \right] + 1 \right).
\end{aligned}$$

It is obvious that the other terms of (A.2) can be bounded similarly. This way we obtain a bound on the term $\mathbb{E}[\|\Delta\psi_{\xi, \ell}\|_2^2 \mathbf{1}_{A^c}]$ of order $2^{-\min(u,4)\ell/2}$, which completes the proof of the first assertion of the theorem.

Let us move on to the second assertion. By choosing $w_{\ell} \propto 2^{-\tau\ell}$, it follows from the first assertion that

$$\sum_{\ell=0}^{\infty} \frac{\mathbb{E}[\|\Delta\psi_{\xi, \ell}\|_2^2]}{w_{\ell}} \propto \sum_{\ell=0}^{\infty} 2^{-(\beta-\tau)\ell},$$

and

$$\sum_{\ell=0}^{\infty} C_{\ell} w_{\ell} \propto \sum_{\ell=0}^{\infty} 2^{-(\tau-1)\ell}.$$

Thus, if $1 < \tau < \beta$, these two quantities are obviously bounded. It is important to remark that we have these finite bounds on the expected squared ℓ_2 -norm and the expected computational cost of the random variable $\Delta\psi_{\xi, \ell}/w_{\ell}$, since we assume $u > 2$, which ensures $\beta > 1$.

REFERENCES

- [1] S. ASMUSSEN AND P. W. GLYNN, *Stochastic Simulation*, Springer, New York, 2007.
- [2] J. BECK, B. M. DIA, L. F. R. ESPATH, Q. LONG, AND R. TEMPONE, *Fast Bayesian experimental design: Laplace-based importance sampling for the expected information gain*, Computer Methods in Applied Mechanics and Engineering, 334 (2018), pp. 523–553, <https://doi.org/10.1016/j.cma.2018.01.053>.
- [3] J. BECK, B. M. DIA, L. F. R. ESPATH, AND R. TEMPONE, *Multilevel double loop Monte Carlo and stochastic collocation methods with importance sampling for Bayesian optimal experimental design*, International Journal for Numerical Methods in Engineering, 121 (2020), pp. 3482–3503, <https://doi.org/10.1002/nme.6367>.
- [4] K. BUJOK, B. HAMBLY, AND C. REISINGER, *Multilevel simulation of functionals of Bernoulli random variables with application to basket credit derivatives*, Methodology and Computing in Applied Probability, 17 (2015), pp. 579–604, <https://doi.org/10.1007/s11009-013-9380-5>.
- [5] A. G. CARLON, B. M. DIA, L. F. R. ESPATH, R. H. LOPEZ, AND R. TEMPONE, *Nesterov-aided stochastic gradient methods using Laplace approximation for Bayesian design optimization*, Computer Methods in Applied Mechanics and Engineering, 363 (2020), 112909, <https://doi.org/10.1016/j.cma.2020.112909>.
- [6] K. CHALONER AND I. VERDINELLI, *Bayesian experimental design: a review*, Statistical Science, 10 (1995), pp. 273–304, <https://doi.org/10.1214/ss/1177009939>.
- [7] S. DEREICH AND T. MÜLLER-GRONBACH, *General multilevel adaptations for stochastic approximation algorithms of Robbins-Monro and Polyak-Ruppert type*, Numerische Mathematik, 142 (2019), pp. 279–328, <https://doi.org/10.1007/s00211-019-01024-y>.
- [8] J. DUCHI, E. HAZAN, AND Y. SINGER, *Adaptive subgradient methods for online learning and stochastic optimization*, Journal of Machine Learning Research, 12 (2011), pp. 2121–2159.
- [9] A. FOSTER, M. JANKOWIAK, E. BINGHAM, P. HORSFALL, Y. W. TEH, T. RAINFORTH, AND N. GOODMAN, *Variational Bayesian optimal experimental design*, in 33rd Conference on

- Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada, 2019, <https://arxiv.org/abs/1903.05480>.
- [10] A. FOSTER, M. JANKOWIAK, M. O'MEARA, Y. W. TEH, AND T. RAINFORTH, *A unified stochastic gradient approach to designing Bayesian-optimal experiments*, in 23rd International Conference on Artificial Intelligence and Statistics (AISTATS 2020), Palermo, Italy, 2020, <https://arxiv.org/abs/1911.00294>.
 - [11] N. FRIKHA, *Multilevel stochastic approximation algorithms*, Annals of Applied Probability, 26 (2016), pp. 933–985, <https://doi.org/10.1214/15-AAP1109>.
 - [12] M. B. GILES, *Multilevel Monte Carlo path simulation*, Operations Research, 56 (2008), pp. 607–617, <https://doi.org/10.1287/opre.1070.0496>.
 - [13] M. B. GILES, *Multilevel Monte Carlo methods*, Acta Numerica, 24 (2015), pp. 259–328, <https://doi.org/10.1017/S096249291500001X>.
 - [14] M. B. GILES AND T. GODA, *Decision-making under uncertainty: using MLMC for efficient estimation of EVPPI*, Statistics and Computing, 29 (2019), pp. 739–751, <https://doi.org/10.1007/s11222-018-9835-1>.
 - [15] M. B. GILES AND L. SZPRUCH, *Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation*, Annals of Applied Probability, 24 (2014), pp. 1585–1620, <https://doi.org/10.1214/13-AAP957>.
 - [16] T. GODA, T. HIRONAKA, AND T. IWAMOTO, *Multilevel Monte Carlo estimation of expected information gains*, Stochastic Analysis and Applications, 38 (2020), pp. 581–600, <https://doi.org/10.1080/07362994.2019.1705168>.
 - [17] T. HIRONAKA, M. B. GILES, T. GODA, AND H. THOM, *Multilevel Monte Carlo estimation of the expected value of sample information*, SIAM/ASA Journal on Uncertainty Quantification, 8 (2020), pp. 1236–1259, <https://doi.org/10.1137/19M1284981>.
 - [18] X. HUAN AND Y. M. MARZOUK, *Gradient-based stochastic optimization methods in Bayesian experimental design*, International Journal for Uncertainty Quantification, 4 (2014), pp. 479–510, <https://doi.org/10.1615/Int.J.UncertaintyQuantification.2014006730>.
 - [19] D. P. KINGMA AND J. L. BA, *Adam: A method for stochastic optimization*, Dec. 2014, <https://arxiv.org/abs/1412.6980>.
 - [20] S. KLEINEGESSE AND M. U. GUTMANN, *Bayesian experimental design for implicit models by mutual information neural estimation*, in 37th International Conference on Machine Learning (ICML 2020), 2020, <https://arxiv.org/abs/2002.08129>.
 - [21] D. V. LINDLEY, *On a measure of the information provided by an experiment*, The Annals of Mathematical Statistics, 27 (1956), pp. 986–1005, <https://doi.org/10.1214/aoms/1177728069>.
 - [22] Q. LONG, M. SCAVINO, R. TEMPONE, AND S. WANG, *Fast estimation of expected information gains for Bayesian experimental designs based on Laplace approximations*, Computer Methods in Applied Mechanics and Engineering, 259 (2013), pp. 24–39, <https://doi.org/10.1016/j.cma.2013.02.017>.
 - [23] J. I. MYUNG, D. R. CAVAGNARO, AND M. A. PITT, *A tutorial on adaptive design optimization*, Journal of Mathematical Psychology, 57 (2013), pp. 53–67, <https://doi.org/10.1016/j.jmp.2013.05.005>.
 - [24] Y. E. NESTEROV, *A method of solving a convex programming problem with convergence rate $o(1/k^2)$* , Soviet Mathematics Doklady, 27 (1983), pp. 372–376.
 - [25] A. B. OWEN, *Monte carlo theory, methods and examples*, 2019, <https://statweb.stanford.edu/~owen/mc/>.
 - [26] B. T. POLYAK, *A new method of stochastic approximation type (in Russian)*, Avtomatika i Telemekhanika, 7 (1990), pp. 98–107.
 - [27] T. RAINFORTH, R. CORNISH, H. YANG, A. WARRINGTON, AND F. WOOD, *On nesting Monte Carlo estimators*, in 35th International Conference on Machine Learning, Stockholm, Sweden, 2018, <http://proceedings.mlr.press/v80/rainforth18a.html>.
 - [28] S. J. REDDI, S. KALE, AND S. KUMAR, *On the convergence of Adam and beyond*, Apr. 2019, <https://arxiv.org/abs/1904.09237>.
 - [29] C. H. RHEE AND P. GLYNN, *Unbiased estimation with square root convergence for SDE models*, Operations Research, 63 (2015), pp. 1026–1043, <https://doi.org/10.1287/opre.2015.1404>.
 - [30] H. ROBBINS AND S. MONRO, *A stochastic approximation method*, The Annals of Mathematical Statistics, 22 (1951), pp. 400–407, <https://doi.org/10.1214/aoms/1177729586>.
 - [31] D. RUPPERT, *Stochastic approximation*, in Handbook of Sequential Analysis, B. K. Ghosh and P. K. Sen, eds., Dekker, New York, 1991, pp. 503–529.
 - [32] E. G. RYAN, C. D. DROVANDI, AND A. N. PETTITT, *Fully Bayesian experimental design for pharmacokinetic studies*, Entropy, 17 (2015), pp. 1063–1089, <https://doi.org/10.3390/e17031063>.

- [33] E. G. RYAN, C. D. DROVANDI, M. THOMPSON, AND A. N. PETTITT, *Towards Bayesian experimental design for nonlinear models that require a large number of sampling times*, Computational Statistics and Data Analysis, 70 (2014), pp. 45–60, <https://doi.org/10.1016/j.csda.2013.08.017>.
- [34] K. J. RYAN, *Estimating expected information gains for experimental designs with application to the random fatigue-limit model*, Journal of Computational and Graphical Statistics, 12 (2003), pp. 585–603, <https://doi.org/10.1198/1061860032012>.
- [35] C. SCHILLINGS, B. SPRUNGK, AND P. WACKER, *On the convergence of the Laplace approximation and noise-level-robustness of Laplace-based Monte Carlo methods for Bayesian inverse problems*, Numerische Mathematik, 145 (2020), pp. 915–971, <https://doi.org/10.1007/s00211-020-01131-1>.
- [36] B. SHABABO, B. PAIGE, A. PAKMAN, AND L. PANINSKI, *Bayesian inference and online experimental design for mapping neural microcircuits*, in Advances in Neural Information Processing Systems 26 (NIPS 2013), 2013.
- [37] A. SHAPIRO, D. DENTCHEVA, AND A. RUSZCZYŃSKI, *Lectures on Stochastic Programming*, SIAM, Philadelphia, 2009.
- [38] A. M. STUART, *Inverse problems: A Bayesian perspective*, Acta Numerica, 19 (2010), pp. 451–559, <https://doi.org/10.1017/S0962492910000061>.
- [39] T. TIELEMAN AND G. HINTON, *Lecture 6.5 – RMSProp*, COURSE: Neural networks for machine learning, 4 (2012), pp. 26–31.
- [40] J. VANLIER, C. A. TIEMANN, P. A. J. HILBERS, AND N. A. W. VAN RIEL, *A Bayesian approach to targeted experiment design*, Bioinformatics, 28 (2012), pp. 1136–1142, <https://doi.org/10.1093/bioinformatics/bts092>.