A BANDIT-LEARNING APPROACH TO MULTIFIDELITY APPROXIMATION*

YIMING XU†, VAHID KESHAVARZZADEH‡, ROBERT M. KIRBY§, AND AKIL NARAYAN†

Abstract. Multifidelity approximation is an important technique in scientific computation and simulation. In this paper, we introduce a bandit-learning approach for leveraging data of varying fidelities to achieve precise estimates of the parameters of interest. Under a linear model assumption, we formulate a multifidelity approximation as a modified stochastic bandit and analyze the loss for a class of policies that uniformly explore each model before exploiting it. Utilizing the estimated conditional mean-squared error, we propose a consistent algorithm, adaptive Explore-Then-Commit (AETC), and establish a corresponding trajectory-wise optimality result. These results are then extended to the case of vector-valued responses, where we demonstrate that the algorithm is efficient without the need to worry about estimating high-dimensional parameters. The main advantage of our approach is that we require neither hierarchical model structure nor *a priori* knowledge of statistical information (e.g., correlations) about or between models. Instead, the AETC algorithm requires only knowledge of which model is a trusted high-fidelity model, along with (relative) computational cost estimates of querying each model. Numerical experiments are provided at the end to support our theoretical findings.

Key words. multifidelity, bandit learning, linear regression, Monte Carlo method, consistency

AMS subject classifications. 62-08, 62J05, 65N30, 65C05

1. Introduction.

Computational models are ubiquitous tools in different aspects of science and engineering, from the discovery of new constitutive laws in physics and mechanics to the design of novel systems such as multifunctional materials. Such models are primarily developed to describe the system of interest accurately such that they can replace time-consuming real-life experiments. In addition to accuracy, an important aspect to consider is the computational cost of model evaluations, which typically increases as the accuracy of the model increases. The trade-off between cost and accuracy constitutes a fundamental challenge in computational science that has been the subject of active research. This challenge has given rise to a widely studied class of methods that involves several computational models with various levels of accuracy and cost. In different contexts and formulations, such methods are called multifidelity, multilevel, multiresolution, etc. The common theme among all these methods is the availability of several models that describe the same system of interest but with varying accuracy and computational cost.

1.1. Multifidelity models.

The notion of fidelity may have different meanings in different problems. A typical scenario corresponds to the level of discretization for solving PDEs, where the finest mesh is referred to as the high-fidelity model and coarser meshes are lower fidelity models. Other common multifidelity settings

*Submitted to the editors.

Funding: Y. Xu and A. Narayan are partially supported by NSF DMS-1848508. V. Keshavarzzadeh and A. Narayan are partially supported by AFOSR under award FA9550-20-1-0338. V. Keshavarzzadeh and R.M. Kirby acknowledge that their part of this research was sponsored by ARL under cooperative agreement number W911NF-12-2-0023.

†Department of Mathematics, and Scientific Computing and Imaging Institute, University of Utah (yxu@math.utah.edu, akil@sci.utah.edu).

‡Scientific Computing and Imaging Institute, University of Utah (vkeshava@sci.utah.edu).

§School of Computing, and Scientific Computing and Imaging Institute, University of Utah (kirby@cs.utah.edu.)

include the interpolation and regression models [21, 22], projection-based reduced-order models [53, 49, 31, 10], machine learning models [58, 18, 16] and other reduced-order models [44, 40]. The multifidelity setting considered in this paper is a general abstraction that includes low-fidelity models instantiated as coarse discretizations, via reduced-order models (e.g., reduced-basis methods) or other types of emulators such as Gaussian processes [33, 48].

Combining the low-fidelity models with the high-fidelity model to accelerate computational efficiency is a fruitful idea in engineering science [47]. It can be construed as a model management strategy with three main categories of approaches: 1) Adaptation of the low-fidelity model to the high-fidelity model by the use of high-fidelity information; approaches based on model correction fall into this category [20, 1, 2], 2) fusion of multiple fidelity modes such as control variate approaches [12, 32, 43, 36, 29]; and 3) filtering approaches such as importance sampling where low-fidelity models serve as filters and determine when to use the high-fidelity model [17, 23, 42, 45]. The method developed in this paper belongs to the realm of fusion approaches, where we assume in particular that a subset of low-fidelity models can be combined linearly to effectively predict the value of the high-fidelity model.

1.2. Bandit learning. The idea of bandit learning first appeared in a study [55] of treatment design problems in clinical trials. It was later developed as a powerful tool to study sequential decision-making problems in an uncertain environment. The primary goal of bandit learning is to find an adaptive strategy that produces close-to-optimal decisions by minimizing the loss or regret. The uncertainty in an environment can be modeled using different mechanisms, such as random feedback [37, 6, 19, 50], adversarial manipulation [8, 7, 14], Markov chains [3, 28, 34], etc., giving rise to a wealth of bandit setups that find use in numerous applications in practice [11, 62]. Regret, which plays a similar role as a loss function in statistical learning, can be chosen in various forms, often depending on the specific application of interest. For example, cumulative loss [6, 8] is widely used as regret to measure the overall performance of a strategy, and a discounted version [28] can be used to account for the elapsed time during the strategy. In other situations where only the final decision matters, the loss at a fixed (final) step is adopted [15, 5, 25]. Despite the diversity of choices for regret, the overall goal is quite thematic: find a strategy that performs similarly to an oracle by investing a reasonable amount of computational complexity in a procedure that adaptively learns. The key philosophy behind bandit learning thus lies in efficiently learning a competitive strategy and then utilizing the learned knowledge for future use. A comprehensive treatment of the subject can be found in, for instance, [39, 13].

1.3. Related work. Our approach in this article centers around using non-hierarchical linear regression for constructing a multifidelity estimator. The idea of using linear regression for multifidelity approximation is not completely new. For instance, recent work has introduced a general linear regression framework to study telescoping-type estimators in multifidelity inference [52, 51]. This work treats different fidelities as factor models with heterogeneous noise structures. Assuming that the covariance matrix of the noise is known, the authors proposed a re-weighted least squares estimator which is shown to be the best linear unbiased estimator for the quantity of interest (under a fixed allocation choice). This framework offers a uniform perspective on many well-studied multifidelity and multilevel estimators [26, 27, 29, 46] and provides certain limits on their theoretical performance. However, the construction utilizes oracle information about model correlations, which are often not known

beforehand and must first be learned from the data, placing practical restrictions on the immediate deployment of such procedures. There is limited work on multifidelity methods with non-hierarchical dependence, i.e., methods that do not necessarily assume an ordering of models based on cost/accuracy [30], and algorithms for identifying non-hierarchical relationships are an area of active research.

In [61], the authors proposed to compute a high-fidelity model using a linear combination of low-fidelity surrogates corrected by a deterministic discrepancy function. This approach is easy to implement and allows classical tools to be used in the analysis. Nevertheless, certain drawbacks exist: The assumption on the discrepancy function can be restrictive when noise pollutes data in a random fashion, and realizing the optimal performance of the method requires selecting the 'best' surrogate model in practice, which is a nontrivial task when the computational budget is limited.

The work [35] has a similar title to this paper, but the setup and goals considered there are completely different.

1.4. Contributions of this paper. This paper provides techniques to address some of the challenges identified in Section 1.3 that limit the applicability of current approaches. Under a type of linear model assumption (that is slightly stronger than what is assumed in [52]), we introduce a bandit-learning approach that, under a budget constraint, facilitates computational learning of the relation between different models before an exploitation choice is committed. Our algorithm seeks to identify the best selection of regressors by investing a portion of the budget in an exploration phase. This selection is then utilized in an exploitation phase to construct a surrogate for the high-fidelity model, which is significantly less expensive but exhibits comparable accuracy.

We make three contributions in this paper:

- We formulate multifidelity approximation as a bandit-learning type problem under a linear model assumption. In particular, we introduce a class of policies called uniform exploration policies and define their associated loss (regret). We also derive an asymptotic formula for their loss as the computational budget goes to infinity.
- We propose an adaptive Explore-Then-Commit (AETC) algorithm (Algorithms 5.1) that automatically identifies a trade-off point between exploration and exploitation. We prove that the estimator produced by the algorithm is consistent and asymptotically matches the best regression model with optimal exploration. In particular, this algorithm requires only model cost information and identification of a high-fidelity model. No model correlations or statistics are needed as input.
- We initially consider scalar-valued responses, but subsequently generalize to vector-valued high-fidelity models, and demonstrate that the algorithm is efficient despite some potentially high-dimensional parameter estimation procedures.

Our numerical results strongly support our theoretical findings. Our methodology enjoys great generality as it does not require any particular knowledge of the models, i.e., hierarchical structure or correlation statistics. The procedure requires only the identification of a trusted high-fidelity model, the ability to query the models themselves, and a relative cost estimate of each model relative to the high-fidelity model. It is worth mentioning that our approach can be extended to estimate more complicated statistics such as the cumulative distribution functions (CDF) of QoIs [60], but this requires stronger assumptions.

The rest of the paper is organized as follows. In Section 2, we introduce the abstract setup of multifidelity approximation with which we will be working in the rest of the paper, followed by a brief review of the key ideas behind stochastic bandits. In Section 3, we formulate multifidelity approximation as a modified bandit-learning problem, and in Section 4, we prove the consistency of the estimators used during the exploitation phase and derive a nonasymptotic convergence rate. In Section 5, we propose an adaptive algorithm, AETC, based on the estimated conditional mean-squared errors from the analysis and establish a trajectory-wise optimality result for it. In Section 6, we extend our results developed in the previous sections to the case of vector-valued high-fidelity models and justify the efficiency of the estimation procedures where high-dimensional parameters are involved. In Section 7, we provide a detailed study of the novel AETC algorithm via numerical experiments that verifies our theoretical statements and demonstrates the utility of the AETC algorithm. In Section 8, we conclude by summarizing the main results in the paper.

2. Background.

2.1. Notation. Fix $n \in \mathbb{N}$. We let $Y, X_1, \cdots, X_n$ be random vectors (possibly of different dimensions) representing the high-fidelity model and n surrogate low-fidelity models, respectively. Let c_i ($i \in [n]$) and c_0 be the respective cost of sampling X_i and Y. For example, in parametrized PDEs, X_i is the approximate solution given by the i-th solver under random coefficients of a PDE, and Y is the solution given by the most accurate solver, which is assumed to be the ground truth. The cost c_i corresponds to the computational time. We make no assumptions about the accuracy or costs of X_i relative to those of X_{i+1}. In particular, the index i does not represent an ordering based on cost, accuracy, or hierarchy. For X_i ($i \in [n]$) and Y, let f_i and f be output quantity of interest maps that bring X_i and Y, respectively, to $f_i(X_i) \in \mathbb{R}^{k_i}$ and $f(Y) \in \mathbb{R}^{k_0}$, which are the quantities of practical interest. Understanding the mean of $f(Y)$ is crucial for many simulation problems in engineering.

Computing the expectation of $f(Y)$ may not seem challenging at all: Under mild assumptions on the distribution of Y, a Monte Carlo (MC) procedure exploits the law of large numbers using an ensemble of independent samples to accurately estimate the mean of $f(Y)$. However, in practice, obtaining numerous samples of Y can be computationally expensive if c_0 is large. This motivates the idea of estimating $\mathbb{E}[f(Y)]$ using low-fidelity model outputs $f_i(X_i)$, which often contain some information about $f(Y)$ and are cheaper to sample. For instance, when $f_i(X_i)$ and $f(Y)$ are scalars and the correlation between them is high, one can construct an unbiased estimator for $\mathbb{E}[f(Y)]$ by taking a telescoping sum of the low-fidelity models plus a few samples of the ground truth $f(Y)$. Optimizing over the allocation parameters under a minimum variance criterion leads to the Multi-Fidelity Monte-Carlo (MFMC) estimator. It is shown in [46] that the MFMC method outperforms MC by a substantial margin on several test datasets. However, successful implementation of the method requires both knowledge of statistical information of and between the models, as well as a hierarchical structure, neither of which may be known beforehand. More precisely, in MFMC, a training stage is required to learn correlations between models, and guidance on how to perform this training is heuristic. Our procedure is a principled approach that effects a similar type of training in an automated way and is accompanied by theoretical guarantees. In the rest of this section, we introduce a general framework in terms of learning the high-fidelity model from its surrogates. Our method utilizes ideas from bandit learning and therefore is less reliant on any existing knowledge of the models. We will need a general linear model assumption that is slightly stronger

than the generic correlation assumption and has been widely used in statistics. For simplicity, we assume $k_i = 1$ for $i \in [n]$ and denote

$$X^{(i)} := f_i(X_i) \in \mathbb{R}. \tag{2.1}$$

We will first deal with the case where $f(Y)$ is a scalar, i.e., $k_0 = 1$. The result will be generalized to the vector-valued responses in Section 6.

2.2. Linear regression. Let $S \subseteq [n]$ be a selection of low-fidelity models with $|S| = s > 0$. Suppose $f(Y)$ and $\{X^{(i)}\}_{i\in S}$ satisfy the following *linear model assumption*:

$$f(Y) = X_S^T \beta_S + \varepsilon_S, \tag{2.2}$$

where

$$X_S = (1, (X^{(i)})_{i\in S})^T \in \mathbb{R}^{s+1}$$

is the regressor vector, which includes a constant (intercept) term; $\beta_S \in \mathbb{R}^{s+1}$ is the coefficient vector; and $\varepsilon_S \in \mathbb{R}$ is model noise, which we assume is independent of X_S. For convenience, we assume in the following discussion that ε_S is a centered *sub-Gaussian* random variable with variance σ_S^2.

Note that even though (2.2) considers only linear interactions with X_S, it is less restrictive than it appears. Indeed, a key assumption implied by (2.2) is that the conditional expectation of $f(Y)$ given X_S (i.e., $\mathbb{E}[f(Y)|X_S]$) is a linear function of X_S. By definition, $\mathbb{E}[f(Y)|X_S]$ is a measurable function of X_S, which can be approximated using, e.g., polynomials of X_S on any compact set under mild regularity conditions. When the linearity assumption is violated, one can add higher order terms of X_S as new regressors to fit a larger linear model to mitigate the model misspecification effect. This procedure will not incur any additional computational cost. Nevertheless, identifying an optimal set of appropriately expressive regressors (features) is a much harder problem that goes beyond the scope of this paper.

Our goal is to seek an efficient estimator for $\mathbb{E}[f(Y)]$, so taking the expectation of (2.2) yields

$$\mathbb{E}[f(Y)] = \mathbb{E}[X_S^T]\beta_S. \tag{2.3}$$

The term $\mathbb{E}[X_S]$ on the right-hand side of (2.3) can be estimated by averaging the independent joint samples of X_S, which is the same as the Monte Carlo applied to the conditional expectation $\mathbb{E}[f(Y)|X_S]$, similar to a model-assisted approach in sampling statistics [24].

What is gained from sampling from $f(Y)|X_S$ instead of $f(Y)$? The two major advantages are:

- The N-sample Monte Carlo estimator for $\mathbb{E}[f(Y)]$ based on sampling $f(Y)$ has mean-squared error $\mathbb{V}[f(Y)]/N$, where $\mathbb{V}[\cdot]$ denotes the variance. The same procedure via sampling $f(Y)|X_S$ has mean-squared error $\mathbb{V}[\mathbb{E}[f(Y)|X_S]]/N$. Since conditioning does not increase variance, the numerator of the latter is bounded by the numerator of the former for any fixed N.
- Given a fixed budget, when $\sum_{i\in S} c_i \ll c_0$, the number of affordable samples for $f(Y)|X_S$ is much larger than $f(Y)$, which can significantly reduce the variance of the estimator.

Both observations provide some heuristics that (2.3) may give rise to a better estimator for $\mathbb{E}[f(Y)]$ under certain circumstances.

$X^{(i)}$	the i-th regressor
$f(Y)$	the response variable (vector)
k_i, k_0	the dimension of $X^{(i)}$ and $f(Y)$
m	the number of samples for exploration
n	the number of regressors
B	total budget
S	the model index set (indices of regressors used in regression)
$(c_i)_{i=0}^n / c_{\text{epr}} / c_{\text{ept}}(S)$	cost parameters/total cost/cost of model S
N_S	the affordable samples in model S
β_S	coefficients (matrix) of model S
$X_{\text{epr},\ell}$	the ℓ-th sample in the exploration phase
Z_S	the design matrix in the exploration phase in model S
ε_S / η_S	the noise variable/vector in model S
x_S	the mean of the regressors in model S
Σ_S	the covariance matrix of the regressors in model S
σ_S^2	the variance in model S (single response)
Q	the weight matrix in defining the Q-weighted risk
Γ_S	the covariance matrix of the noise in the case of vector-valued response

TABLE 1
Notation used throughout this article.

In practice, neither the best model index set S (which is referred to as *model* S in the rest of the article) nor the corresponding β_S is known. Thus, we cannot avoid expending some resources to estimate β_S for every model $S \subseteq [n]$ before deciding which model can best predict $\mathbb{E}[f(Y)]$. One possible way to achieve this is via bandit learning.

2.3. Stochastic bandits. This section provides a brief overview of the philosophy behind bandit learning, and in particular, we focus on stochastic bandits, which is the subfield most relevant to our procedure.

Consider a *multiarmed bandit* setup: we are faced with k slot machines or *arms*, where $k \in \mathbb{N}$ is fixed. At each integer time t, a player pulls exactly one arm and this player will receive a reward. Rewards generated from the same arm at different times are assumed to be independent and identically distributed; rewards generated from different arms are independent and their distributions are stationary in time. Given a fixed (time) horizon N, the goal of stochastic bandit learning is to seek a policy with small 'regret'. A *policy* $\pi = (\pi_1, \cdots, \pi_N) \in [k]^N$ is a random sequence of choices adapted to the natural filtration generated by past observations and actions. 'Regret' is a special type of loss function that measures the quality of a policy. In the classical setup, the regret of a policy π is defined as

$$\begin{aligned} R_N(\pi) :&= N\mu_{\max} - \mathbb{E}\left[\sum_{t\in[N]} z_t(\pi)\right] \\ &= \sum_{i\in[k]} T_N(i)\Delta_i \qquad\qquad T_N(i) = \mathbb{E}\left[\#\{t \le N, \pi_t = i\}\right], \end{aligned} \tag{2.4}$$

where $z_t(\pi)$ is the reward received at time t under policy π (i.e., corresponds to reward π_t at time t), μ_i is the expectation of the reward distribution of arm i, $\mu_{\max} := \max_{i\in[k]} \mu_i$, and $\Delta_i := (\mu_{\max} - \mu_i)$ is the *suboptimality* gap of arm i. Intuitively, (2.4) measures the difference between the average total reward under π compared to the best (oracle) mean reward, which is generally not known.

In practice, we assume that the true reward distributions are not available, and

so it is necessary to expend effort estimating the expected reward of each arm. A simple idea to estimate these average rewards is to pull each arm a fixed number of times m, and use the gathered data to estimate the mean rewards. Such a process is called an *exploration* stage in bandit learning. Based on the estimated rewards, we can then make a decision about which arm performs the best and repeatedly select it for the remaining time (the *exploitation* phase) until the horizon is reached. A direct combination of the exploration and exploitation gives the Explore-Then-Commit (ETC) algorithm, shown in Algorithm 2.1.

Algorithm 2.1 Explore-then-Commit (ETC) algorithm for stochastic bandits

Input: m: the number of exploration on each arm
Output: $\pi = (\pi_t)_{t\in[N]}$
if $t \le mk$ **then**
  $\pi_t = \lceil t \bmod k \rceil$
else
  $\pi_t = \arg\max_{i\in[k]} \widehat{\mu}_i(mk)$, where $\widehat{\mu}_i(mk) = \frac{1}{m}\sum_{t=m\cdot(i-1)+1}^{m\cdot i} z_t(\pi)$
end if

The success of the ETC algorithm closely depends on the input parameter m, which dictates how many times each arm is pulled in the exploration phase, and is therefore indicative of the cost of this phase. A small m may result in poor estimation (and thus poor exploration), and in this case, there is a considerable chance that the chosen exploitation strategy will be suboptimal. A large m, on the other hand, leaves little room for exploitation, and regret will then be dominated by the exploration phase. A good choice of m lies in finding the right trade-off point between exploration and exploitation. When $k = 2$ and reward distributions are sub-Gaussian, a near-optimal m can be explicitly computed [39, Chapter 6]. However, such explicit expressions often rely on the knowledge of the suboptimality gaps Δ_i, which are not available in general. More useful algorithms which can be viewed as adaptive generalizations of the ETC include the Upper-Confidence Bound (UCB) algorithm [6] and the Elimination algorithm [9].

One generalization of stochastic bandits pertaining to the multifidelity approximation problem of our interest is budget-limited bandits [56], where each arm has a cost for pulling, and the number of pulls is constrained by a total budget instead of the time horizon. Similar algorithms as well as the regret analysis in the stochastic bandits can be carried out in the budget-limited setup [56, 57].

3. A bandit-learning perspective of the multifidelity problem. In this section, we demonstrate that multifidelity approximation fits into a modified framework of bandit learning, which we exploit to develop an algorithm. Since our multifidelity goal is different from that in classical stochastic bandits, both the action set and the loss function (regret) must be tailored. We begin with the former.

3.1. Action sets and uniform exploration policies. Our goal is to construct a regression-based MC estimator for $\mathbb{E}[f(Y)]$ using the best regression model. However, the relationship between $f(Y)$ and X_S is unknown, requiring us to first estimate the coefficient vector β_S in (2.3) for each S and then decide which model is optimal. As a consequence, the action set contains both the exploration and exploitation options for each $S \subseteq [n]$. Denote by $\mathcal{A}$ the set of actions. Then, we can write $\mathcal{A}$ as

$$\mathcal{A} = \{a_{\mathrm{epr}}(S), a_{\mathrm{ept}}(S)\}_{S\subseteq[n]} \tag{3.1}$$

where

$$a_{\text{epr}}(S) : \text{collect a sample of } (X_S, f(Y))$$
$$a_{\text{ept}}(S) : \text{use the remaining budget to sample } X_S \ .$$

As opposed to stochastic bandits, actions here are highly correlated. For example, exploration will not be allowed whenever an exploitation action is triggered (which exhausts the budget), separating exploration and exploitation into two distinct phases. In addition, action $a_{\text{epr}}(S)$ yields data for every S' satisfying $S' \subseteq S$. For simplicity, we specialize our action set to consider only *uniform exploration policies*, i.e, each exploration yields a sample of all regressors, incurring the cost

$$c_{\text{epr}} := \sum_{i=0}^{n} c_i. \tag{3.2}$$

This procedure is similar to the ETC algorithm (Algorithm 2.1) where arms are explored to the same degree during exploration.

Let $B > 0$ be a given, fixed total budget, and let $m > n + 1$ be the number of exploration actions. (We will specify how m is chosen later.) The budget B_{epr} spent on exploration under a uniform exploration policy π and the budget B_{ept} remaining for exploitation are given by

$$B_{\text{epr}} = c_{\text{epr}} m, \qquad\qquad B_{\text{ept}} = B - B_{\text{epr}}.$$

During exploration, each action yields a sample of the form

$$X_{\text{epr},\ell} = \left(1, X_\ell^{(1)}, \cdots, X_\ell^{(n)}, f(Y_\ell)\right)^T \qquad\qquad \ell \in [m],$$

where the subscript ℓ is a sampling index. For $S \subseteq [n]$, let $X_{\text{epr},\ell}|_S \in \mathbb{R}^{s+1}$ (including the intercept term) and $X_{\text{epr},\ell}|_Y \in \mathbb{R}$ be the restriction of $X_{\text{epr},\ell}$ to model S and $f(Y)$, respectively. The coefficient vector β_S in (2.2) can be estimated by a standard least squares procedure:

$$\widehat{\beta}_S = Z_S^\dagger X_{\text{epr}}|_Y \tag{3.3}$$

where

$$Z_S = \left(X_{\text{epr},1}|_S, \cdots, X_{\text{epr},m}|_S\right)^T, \qquad X_{\text{epr}}|_Y = \left(X_{\text{epr},1}|_Y, \ldots X_{\text{epr},m}|_Y\right)^T$$

is the design matrix and data vector, respectively, and $Z_S^\dagger$ is the Moore-Penrose pseudoinverse of Z_S. For simplicity, we will assume that the design matrix has full rank in the following discussion, so that $Z_S^\dagger := \left(Z_S^T Z_S\right)^{-1} Z_S^T$. In exploitation, one selects an action $a_{\text{ept}}(S)$ for some $S \subseteq [n]$ and uses (2.3) to build an MC estimator for $\mathbb{E}[f(Y)]$. The true coefficient vector β_S is unknown but can be replaced by the estimate $\widehat{\beta}_S$, yielding the Linear Regression Monte-Carlo (LRMC) estimator associated with model S:

$$\mathsf{LRMC}_S = \frac{1}{N_S} \sum_{\ell \in [N_S]} X_{S,\ell}^T \widehat{\beta}_S, \tag{3.4}$$

where N_S is the number of affordable samples to exploit model S:

$$N_S = \left\lfloor \frac{B_{\text{ept}}}{c_{\text{ept}}(S)} \right\rfloor = \left\lfloor \frac{B - c_{\text{exp}} m}{c_{\text{ept}}(S)} \right\rfloor \qquad\qquad c_{\text{ept}}(S) := \sum_{i \in S} c_i,$$

and $X_{S,\ell}$ are i.i.d. samples of X_S which are independent of the samples in the exploration stage. Here we choose not to reuse samples from the exploration phase during the exploitation process for convenience of analysis.

REMARK 3.1. *A similar analysis of the LRMC estimator that reuses or recycles the exploration data can be carried out but with more complicated notation. For most applications that we consider, where $m/N_S \ll 1$ (since the high-fidelity model is often substantially more expensive than the low-fidelity emulators), reuse has little impact on the estimator in general. Regardless of analysis, one can always recycle exploration samples in a practical setting; our numerical results show that recycling exploration samples has negligible impact on the performance of our procedure for the examples we have tested. See Appendix A for a theoretical justification, and Figure 2 in Section 7 for numerical evidence.*

We will show in Section 4 that for fixed $S \subseteq [n]$, LRMC_S defined in (3.4) is, almost surely, a consistent estimator for $\mathbb{E}[f(Y)]$, and we will provide a convergence rate.

3.2. Loss function. The loss function used in bandit learning is called regret, which is often defined by the reward difference between a policy and an oracle. In our case, it is more convenient to define loss as a quantity that we wish to minimize. Note that the output of a uniform exploration policy π is an LRMC estimator (3.4), where the selected model S satisfies $\pi_{m+1} = a_{\text{ept}}(S)$. One way to measure the immediate quality of (3.4) is through the following conditional mean-squared error (MSE) on $\widehat{\beta}_S$:

$$\mathsf{MSE}_S|_{\widehat{\beta}_S} = \mathbb{E}\left[\left(\mathsf{LRMC}_S - \mathbb{E}[f(Y)]\right)^2 \big| \widehat{\beta}_S \right]. \tag{3.5}$$

The LRMC_S alone, despite being an unbiased estimator (which is easily verified using independence), is not a sum of i.i.d. random variables unless conditioned on $\widehat{\beta}_S$. Once conditioned, LRMC_S becomes a sum of i.i.d. random variables that is a biased estimator for $\mathbb{E}[f(Y)]$. Define the following statistics of X_S,

$$x_S = \mathbb{E}[X_S], \qquad\qquad \Sigma_S = \mathrm{Cov}[X_S].$$

By writing $\mathsf{MSE}_S|_{\widehat{\beta}_S}$ using the bias-variance decomposition, we obtain

$$\begin{aligned} \mathsf{MSE}_S|_{\widehat{\beta}_S} &= (x_S^T(\widehat{\beta}_S - \beta_S))^2 + \mathbb{V}\left[\mathsf{LRMC}_S \big| \widehat{\beta}_S\right] \\ &= (\widehat{\beta}_S - \beta_S)^T x_S x_S^T (\widehat{\beta}_S - \beta_S) + \frac{1}{N_S} \widehat{\beta}_S^T \Sigma_S \widehat{\beta}_S, \end{aligned} \tag{3.6}$$

where $\mathbb{V}[\cdot|\widehat{\beta}_S]$ is the $\widehat{\beta}_S$-conditional variance operator. Note that (3.6) decomposes the loss incurred in the exploration and exploitation phases by the bias term and variance term, respectively. The *average conditional MSE* of the LRMC is defined as (3.6) averaged over the randomness of the model noise ε_S in exploration:

$$\begin{aligned} \overline{\mathsf{MSE}}_S|_{\widehat{\beta}_S} &:= \mathbb{E}_{\varepsilon_S}\left[\mathsf{MSE}_S|_{\widehat{\beta}_S}\right] \\ &= \frac{1}{N_S}\left[\beta_S^T \Sigma_S \beta_S + \sigma_S^2 \,\mathrm{tr}\big(\Sigma_S (Z_S^T Z_S)^{-1}\big)\right] + \sigma_S^2 \,\mathrm{tr}\big(x_S x_S^T (Z_S^T Z_S)^{-1}\big). \end{aligned} \tag{3.7}$$

Note that $\overline{\mathsf{MSE}}_S|_{\widehat{\beta}_S}$ defined in (3.7) is random, and one could alternatively define it by averaging all the randomness (both ε_S and Z_S). However, this approach is more difficult to analyze due to the term $\mathbb{E}[(Z_S^T Z_S)^{-1}]$. In the following discussion, we will use (3.7) as the *loss function* associated with the uniform exploration policy π to examine its average quality in terms of estimating $\mathbb{E}[f(Y)]$.

4. Consistency of the LRMC estimator. In this section, we will show that the LRMC estimators constructed in the previous section are consistent estimators for $\mathbb{E}[f(Y)]$. The result can be derived as a corollary of a nonasymptotic estimate of the convergence rate of the $\widehat{\beta}_S$-conditional MSE. The following definition will be used in the subsequent analysis:

DEFINITION 4.1 (α-Orlicz norm). *The α-Orlicz ($\alpha \geq 1$) norm of a random variable W is defined as*

$$\|W\|_{\psi_\alpha} := \inf\{C > 0 : \mathbb{E}[\exp(|W|^\alpha / C^\alpha)] \leq 2\}.$$

THEOREM 4.2. *Fix $S \subseteq [n]$. Suppose X_S satisfies*

$$\max\left\{ \sup_{\|\theta\|_2=1} \mathbb{E}[\langle X_S, \theta\rangle^4]^{1/4}, \|\|X_S\|_2\|_{\psi_1} \right\} \leq K < \infty \qquad K \geq 1, \tag{4.1}$$

where $\|\cdot\|_{\psi_1}$ is the 1–Orlicz norm, and $\Lambda_S := \mathbb{E}[X_S X_S^T] = x_S x_S^T + \Sigma_S$ is invertible. Then, for large m, with probability at least $1 - m^{-2}$,

$$\mathsf{MSE}_S|_{\widehat{\beta}_S} \lesssim \frac{\beta_S^T \Sigma_S \beta_S}{N_S} + (s+1)\sigma_S^2 \frac{\log m}{m}, \tag{4.2}$$

where the implicit constant in $\lesssim$ is independent of m.

Proof. See Appendix B. ∎

Theorem 4.2 implies the following consistency result of the LRMC estimator, which follows immediately from (4.2) and an application of the Borel-Cantelli lemma:

COROLLARY 4.3 (consistency of the LRMC). *Let $\{B_k\}_{k=1}^\infty$ satisfy $\lim_{k\uparrow\infty} B_k = \infty$, and let m_k, N_k be the respective exploration round and exploitation samples satisfying $\min(m_k, N_k) \to \infty$. For fixed $S \subseteq [n]$, denote $\mathsf{LRMC}_{S,k}$ the LRMC estimator under budget B_k, and $\widehat{\beta}_{S,k}$ the corresponding estimator for β_S. Then, for almost every sequence $\{\widehat{\beta}_{S,k}\}_{k=1}^\infty$, $\mathsf{LRMC}_{S,k} \xrightarrow{\mathbb{P}} \mathbb{E}[f(Y)]$ as $k \to \infty$.*

Without a nonasymptotic convergence rate, a strong consistency result for the LRMC estimator can be established using existing results in [38, Theorem 1]. In fact, a sufficient condition for the estimator $\widehat{\beta}_S$ (hence $\widehat{\sigma}_S^2$) to be strongly consistent is that the largest and smallest eigenvalues of $(Z_S^T Z_S)^{-1}$, $\lambda_{\max}$ and $\lambda_{\min}$, satisfy $\lambda_{\min} \to \infty$ and $\log \lambda_{\max}/\lambda_{\min} \to 0$ as $m \to \infty$ a.s., which is verifiable under the condition (4.1).

5. Algorithms. In this section, we first analyze the asymptotic loss associated with uniform exploration policies. Then, we propose an adaptive ETC (AETC) algorithm based on the estimated average MSEs, which selects the optimal exploration round and the best model to commit for a large budget almost surely. For convenience, we will ignore all integer rounding effects in the rest of the discussion.

5.1. Oracle loss of uniform exploration policies. A uniform exploration policy spends the first m rounds on exploration and then chooses a fixed model S among the subsets of $[n]$ for exploitation. Fixing B and for each S, the optimal exploration $m_S(B)$ balances the terms in expression (3.7). The best uniform exploration policy is the one that stops exploring at time $m_S(B)$ and then picks model S for exploitation, where model S has the smallest average conditional MSE computed at the optimal exploration round $m_S(B)$. To develop computable expressions, we will be mainly concerned with the case when B tends to infinity.

To find the optimal m for fixed S, note that for sufficiently large m, the law of large numbers tells us that almost surely from (3.7),

$$\begin{aligned}\overline{\mathsf{MSE}}_S|_{\widehat{\beta}_S} &= \frac{1}{N_S}\left(\beta_S^T\Sigma_S\beta_S + \sigma_S^2\operatorname{tr}\big(\Sigma_S(Z_S^TZ_S)^{-1}\big)\right) + \sigma_S^2\operatorname{tr}\big(x_Sx_S^T(Z_S^TZ_S)^{-1}\big)\\ &\simeq \frac{c_{\mathrm{ept}}(S)}{B-c_{\mathrm{epr}}m}\left(\beta_S^T\Sigma_S\beta_S + \frac{1}{m}\sigma_S^2\operatorname{tr}\big(\Sigma_S\Lambda_S^{-1}\big)\right) + \frac{1}{m}\sigma_S^2\operatorname{tr}\big(x_Sx_S^T\Lambda_S^{-1}\big)\\ &\simeq \frac{\beta_S^T\Sigma_S\beta_S}{B-c_{\mathrm{epr}}m}c_{\mathrm{ept}}(S) + \frac{\sigma_S^2\operatorname{tr}\big(x_Sx_S^T\Lambda_S^{-1}\big)}{m}, \qquad (5.1)\end{aligned}$$

where $A_1(m)\simeq A_2(m)$ means that $A_1(m)/A_2(m)\to 1$ as $m\to\infty$ with probability 1. Denoting

$$k_1(S) = c_{\mathrm{ept}}(S)\beta_S^T\Sigma_S\beta_S \qquad k_2(S) = \sigma_S^2\operatorname{tr}\big(x_Sx_S^T\Lambda_S^{-1}\big) \le (s+1)\sigma_S^2, \qquad (5.2)$$

the asymptotically best m for exploiting model S can be found by minimizing (5.1) :

$$m_S = \operatorname*{arg\,min}_{1<m<B/c_{\mathrm{epr}}} \frac{k_1(S)}{B-c_{\mathrm{epr}}m} + \frac{k_2(S)}{m} = \frac{B}{c_{\mathrm{epr}} + \sqrt{\frac{c_{\mathrm{epr}}k_1(S)}{k_2(S)}}}, \qquad (5.3)$$

where the latter equality can be computed explicitly since (5.1) is a strictly convex function of m in its domain. The $\overline{\mathsf{MSE}}_S|_{\widehat{\beta}_S}$ corresponding to $m=m_S$ is

$$\overline{\mathsf{MSE}}^*_S|_{\widehat{\beta}_S} = \frac{(\sqrt{k_1(S)}+\sqrt{c_{\mathrm{epr}}k_2(S)})^2}{B} \propto \left(\sqrt{k_1(S)}+\sqrt{c_{\mathrm{epr}}k_2(S)}\right)^2. \qquad (5.4)$$

The best model is the one that has the smallest $\overline{\mathsf{MSE}}^*_S|_{\widehat{\beta}_S}$ under the optimal exploration round:

$$S^* = \operatorname*{arg\,min}_{S\subseteq[n]} \overline{\mathsf{MSE}}^*_S|_{\widehat{\beta}_S} = \operatorname*{arg\,min}_{S\subseteq[n]} \left(\sqrt{k_1(S)}+\sqrt{c_{\mathrm{epr}}k_2(S)}\right)^2, \qquad (5.5)$$

where the right-hand side is assumed to have a unique minimizer. The uniform exploration policy π^* with exploration round count $m=m_{S^*}$ (more precisely $(\lfloor m_{S^*}\rfloor)$) and exploitation action $a_{\mathrm{ept}}(S^*)$ achieves the smallest asymptotic loss among all uniform exploration policies; such a policy is referred to as a *perfect uniform exploration policy*. Note that determination of this optimal policy requires oracle information, thus cannot be directly applied in practice. We will provide a solution for this in the next section.

REMARK 5.1. *For every $S\subseteq[n]$, one can verify that the numerator in (5.4) is bounded by $2(k_1(S)+c_{epr}k_2(S))$. On the other hand, the MC estimator using the high-fidelity samples alone has MSE of a similar form with numerator $c_0\mathbb{V}[f(Y)]$ (Section*

2.2). Taking the quotient of the two quantities yields

$$\frac{2(k_1(S)+c_{epr}k_2(S))}{c_0\mathbb{V}[f(Y)]} \overset{(5.2)}{\leq} \frac{2[c_{ept}(S)\beta_S^T\Sigma_S\beta_S+c_{epr}(s+1)\sigma_S^2]}{c_0\mathbb{V}[f(Y)]} \overset{(2.2)}{\leq} 2\left[\frac{c_{ept}(S)}{c_0}+\frac{c_{epr}(s+1)}{c_0}\frac{\sigma_S^2}{\mathbb{V}[f(Y)]}\right].$$

This can be unconditionally bounded by $2(n+1)(n+2)$ *for all* $S\subseteq[n]$*, and can be significantly smaller than* 1 *if both*

$$\frac{c_{ept}(S)}{c_0}\ll 1 \qquad\qquad \frac{\sigma_S^2}{\mathbb{V}[f(Y)]}\ll 1. \tag{5.6}$$

This implies that the best uniform exploration policies are at most a constant (which only depends on n*) worse than the classical MC, and can be significantly better if* (5.6) *is satisfied for at least one* S *(which is often the case in practice).*

5.2. An adaptive ETC algorithm. Finding a uniform exploration policy is impossible without oracle access to $\beta_S,\Sigma_S,x_S,\sigma_S^2$ and Λ_S^{-1}. These quantities, despite being unknown, can be estimated at any particular time t in exploration ($t>s+1$):

(5.7)

$$\widehat{\beta}_S(t)=Z_S^{\dagger}X_{\text{epr}}|_Y \text{ (using the first } t \text{ samples)} \qquad \widehat{x}_S(t)=\frac{1}{t}\sum_{\ell\in[t]}X_{\text{epr},\ell}|_S$$

$$\widehat{\Sigma}_S(t)=\frac{1}{t-1}\sum_{\ell\in[t]}(X_{\text{epr},\ell}|_S-\widehat{x}_S(t))(X_{\text{epr},\ell}|_S-\widehat{x}_S(t))^T$$

$$\widehat{\sigma}_S^2(t)=\frac{1}{t-s-1}\sum_{\ell\in[t]}\left(X_{\text{epr},\ell}|_Y-X_{\text{epr},\ell}|_S\widehat{\beta}_S\right)^2 \qquad \widehat{\Lambda}_S^{-1}(t)=\left(\frac{1}{t}Z_S^TZ_S\right)^{-1}$$

Note that $\widehat{\sigma}_S^2(t)$ is an unbiased estimator for σ_S^2 when the noise is Gaussian. Plugging (5.7) into (3.7) and ignoring the higher order term $\sigma_S^2\operatorname{tr}\big(\Sigma_S(Z_S^TZ_S)^{-1}\big)$ yields the sample estimator for $\overline{\mathsf{MSE}}_S|_{\widehat{\beta}_S}$ with exploration round m at time t:

$$\widehat{\overline{\mathsf{MSE}}}_S|_{\widehat{\beta}_S}(m;t)=\frac{c_{\text{ept}}(S)\widehat{\beta}_S^T(t)\widehat{\Sigma}_S(t)\widehat{\beta}_S(t)}{B-c_{\text{epr}}m}+\frac{\widehat{\sigma}_S^2(t)\operatorname{tr}\Big(\widehat{x}_S(t)\widehat{x}_S^T(t)\widehat{\Lambda}_S^{-1}(t)\Big)}{m}. \tag{5.8}$$

We call (5.8) the *empirical average conditional MSE* of model S, which we subsequently call the empirical MSE. Correspondingly, we define more empirical estimates of quantities that were previously defined in terms of oracle information:

(5.9)

$$k_{1,t}(S)=c_{\text{ept}}(S)\widehat{\beta}_S^T(t)\widehat{\Sigma}_S(t)\widehat{\beta}_S(t), \qquad k_{2,t}(S)=\widehat{\sigma}_S^2(t)\operatorname{tr}\Big(\widehat{x}_S(t)\widehat{x}_S^T(t)\widehat{\Lambda}_S^{-1}(t)\Big),$$

$$m_S(t)=\frac{B}{c_{\text{epr}}+\sqrt{\frac{c_{\text{epr}}k_{1,t}(S)}{k_{2,t}(S)}}}, \qquad \overline{\mathsf{MSE}}^*_S|_{\widehat{\beta}_S}(t)=\frac{(\sqrt{k_{1,t}(S)}+\sqrt{c_{\text{epr}}k_{2,t}(S)})^2}{B}.$$

To mimic the idea of a perfect uniform exploration policy, we start by collecting the smallest necessary number of exploration samples to make the above estimation

possible, which requires $|S|+2$ exploration rounds. For $t \geq |S|+2$, we use the collected exploration data to compute (5.7), (5.8), (5.9), from which we can estimate the best empirical MSE for each model S . At this point, if $m_S(t) > t$, then more exploration is needed for model S, and the optimal empirical MSE is approximately $\widehat{\mathsf{MSE}}^*_S|_{\widehat{\beta}_S}(t) = \widehat{\mathsf{MSE}}_S|_{\widehat{\beta}_S}(m_S(t);t)$. Otherwise, the best trade-off point has passed and stopping immediately yields the minimal empirical MSE, which equals $\widehat{\mathsf{MSE}}_S|_{\widehat{\beta}_S}(t;t)$. An algorithm could use the estimated optimal expected MSEs to determine which model is currently most favorable for exploitation. But another exploration round should transpire if $m_S(t)$ for this favorable model is larger than the current step t.

During implementation, we use the above procedure once t exceeds a relatively small number, i.e., the maximum number of regressors plus one. The estimation error of the second term in (5.8) is of constant order and cannot be ignored at the beginning. If it becomes pathologically close to zero during the initial stages of the algorithm, exploration may terminate after only a few rounds. To combat this deficiency, we employ a common regularization in bandit learning: For every S, we additively augment $k_{2,t}(S)$ with a small regularization parameter α_t, which decays to 0 as $t \to \infty$. For large B, adding α_t encourages the algorithm to explore at the beginning, but the encouragement becomes negligible as $\alpha_t \to 0$. Putting the ideas together yields the following adaptive ETC (AETC) algorithm for multifidelity approximation:

Algorithm 5.1 AETC algorithm for multifidelity approximation (single-valued case)

Input: B: total budget, c_i: cost parameters, $\alpha_t \downarrow 0$: regularization parameters
Output: $(\pi_t)_t$

1: compute the maximum exploration round $M = \lfloor B/c_{\text{epr}} \rfloor$
2: **for** $t \in [n+2]$ **do**
3: $\quad \pi_t = a_{\text{epr}}([n])$
4: **end for**
5: **while** $n+2 \leq t \leq M$ **do**
6: $\quad$ **for** $S \subseteq [n]$ **do**
7: $\quad\quad$ compute $k_{1,t}(S), k_{2,t}(S)$ using (5.7), (5.9), and set $k_{2,t}(S) \leftarrow k_{2,t}(S) + \alpha_t$
8: $\quad\quad$ compute $m_S(t)$ using (5.9)
9: $\quad\quad$ compute the optimal expected MSE using (5.8): $h_S(t) = \widehat{\mathsf{MSE}}_S|_{\widehat{\beta}_S}(m_S(t) \vee t; t)$
10: $\quad$ **end for**
11: $\quad$ find the optimal model $S^*(t) = \arg\min_{S \subseteq [n]} h_S(t)$
12: $\quad$ **if** $m_{S^*(t)}(t) > t$ **then**
13: $\quad\quad \pi_{t+1} = a_{\text{epr}}([n])$ and $t \leftarrow t+1$
14: $\quad$ **else**
15: $\quad\quad \pi_t = a_{\text{ept}}(S^*(t))$ and $t \leftarrow M+1$
16: $\quad$ **end if**
17: **end while**

REMARK 5.2. *Algorithm 5.1 is an adaptive version of the ETC algorithm (Algorithm 2.1). To better comprehend Algorithm 5.1, and in particular the policy actions π that it implicitly defines, note that the first 'for' loop (step 2-4) involves collecting a minimum number of samples to make the parameters that will be used estimable. The 'while' loop (step 5-17) determines the exploration rate adaptively based on the*

estimated optimal mean-squared errors. (Exploration/exploitation rate refers to the number of samples used for exploration/exploitation in AETC.) In particular, the inner 'for' loop (step 6-10) computes for each $S \subseteq [n]$ the current expected optimal mean-squared error, and step 11 selects the best model based on the estimated values. The 'if – else' procedure (step 12-16) then decides if more exploration is needed based on the information of the selected best model.

5.3. AETC consistency and optimality. Algorithm 5.1 ensures that both exploration and exploitation go to infinity as $B \to \infty$. (See Appendix C.) As a consequence, almost surely, the model chosen by Algorithm 5.1 for exploitation converges to S^* and the corresponding exploration is asymptotically optimal.

THEOREM 5.3. *Let $m(B)$ be the exploration round chosen by Algorithm 5.1 under budget B, and $S(B)$ be the model for exploitation, i.e., $S(B) = S^*(m(B))$. Suppose (4.1) holds uniformly for all $S \subseteq [n]$, and $\lim_{t\to\infty} \alpha_t = 0$. Then, with probability 1,*

$$\lim_{B\to\infty} \frac{m(B)}{m_{S^*}} = 1, \tag{5.10a}$$

$$\lim_{B\to\infty} S(B) = S^*, \tag{5.10b}$$

where S^* and m_{S^*} are the model choice and exploration round count given oracle information defined in (5.5) and (5.3), respectively.

Proof. See Appendix C. ∎

Theorem 5.3 concludes that as the budget goes to infinity, both the exploration round count and the chosen exploitation model will converge to the optimal ones given by the perfect uniform exploration policies for almost every realization. However, this does not imply that the loss of the corresponding policy π asymptotically matches the loss of the perfect uniform exploration policy. Indeed, the formula defining the loss in (3.7) assumes m is deterministic and cannot be applied when $m = m(B)$ is chosen in an adaptive (and random) manner.

Selecting α_t is often considered an art in bandit learning, and for us is mostly used for theoretical analysis. In practice, we observe that it is sufficient to choose α_t as a sequence with exponential decay in t.

REMARK 5.4. *Let $N(B)$ denote the number of affordable samples for exploiting under budget B. (5.10a) implies that both $m(B)$ and $N(B)$ diverge as B goes to infinity. In addition, one can see from the proof that $S^*(t) = S^*$ for sufficiently large t almost surely. These combined with the strong law of large numbers imply that the estimators produced by Algorithm 5.1 are almost surely consistent.*

6. Vector-valued responses and efficient estimation.

6.1. Vector-valued high-fidelity models. We now generalize the results in the previous section to the case of vector-valued responses, i.e., $k_0 > 1$. Let

$$f(Y) = (f^{(1)}(Y), \cdots, f^{(k_0)}(Y))^T \in \mathbb{R}^{k_0}$$

denote the response vector. The linear model assumption in (2.2) is modified as follows:

$$f(Y) = \beta_S X_S + \varepsilon_S, \tag{6.1}$$

where $\varepsilon_S = (\varepsilon_S^{(1)}, \cdots, \varepsilon_S^{(k_0)})^T \sim (0, \Gamma_S)$ is a centered multivariate sub-Gaussian distribution with covariance matrix Γ_S, and $\beta_S = (\beta_S^{(1)}, \cdots, \beta_S^{(k_0)})^T \in \mathbb{R}^{k_0 \times (s+1)}$ is the

coefficient matrix for model S. For fixed S, one may use (3.3) to estimate the coefficients $\beta_S^{(j)}$ by $\widehat{\beta}_S^{(j)}$ for $j \in [k_0]$, and the corresponding LRMC estimators can be built similarly as in (3.4). To generalize the MSE, we consider a common class of quadratic risk functionals to measure the quality of the LRMC estimators. Let $Q \in \mathbb{R}^{q\times l}$, and let $\widehat{\theta}$ be an estimator for $\theta \in \mathbb{R}^l$ where q is arbitrary. The *Q-risk* of $\widehat{\theta}$ is defined as

$$\mathsf{Risk}(\widehat{\theta}) = \mathbb{E}\left[\left\|Q(\widehat{\theta}-\theta)\right\|_2^2\right] = \mathbb{E}\left[(\widehat{\theta}-\theta)^T Q^T Q(\widehat{\theta}-\theta)\right], \tag{6.2}$$

where we suppress the notational dependence of Risk on Q. We use the following conditional Q-risk on the estimated coefficient matrix $\widehat{\beta}_S$ in place of (3.5):

$$\mathsf{Risk}_S|_{\widehat{\beta}_S} := \mathsf{Risk}(\mathsf{LRMC}_S)|_{\widehat{\beta}_S} = \mathbb{E}\left[\|Q(\mathsf{LRMC}_S - \mathbb{E}[f(Y)])\|_2^2 \,|\widehat{\beta}_S\right] \tag{6.3}$$

where above $Q \in \mathbb{R}^{q\times k_0}$ for arbitrary q. A similar result as Theorem 4.2 can be obtained for the LRMC estimator under (6.3):

THEOREM 6.1. *Under the same conditions as in Theorem 4.2 and assumption (6.1), the following result holds: For large m, with probability at least $1-m^{-2}$,*

$$\mathsf{Risk}_S|_{\widehat{\beta}_S} \lesssim \frac{1}{N_S}\operatorname{tr}\left(\Sigma_S \beta_S^T Q^T Q \beta_S\right) + (s+1)\operatorname{tr}\left(Q\Gamma_S Q^T\right)\frac{\log m}{m}, \tag{6.4}$$

where the implicit constant in $\lesssim$ is independent of m.

Other results in the previous section (including Algorithm 5.1 and Theorem 5.3) can also be generalized to the vector-valued case; since the ideas are similar, we do not restate them here. An extended discussion of the results and the proof of Theorem 6.1 can be found in Appendix D. The corresponding algorithm is Algorithm D.1.

6.2. Efficient estimation. Estimation for vector-valued responses is more challenging due to the presence of high-dimensional parameters. Nevertheless, we are only interested in some functionals of the high-dimensional parameters, i.e., $\operatorname{tr}\left(Q\Gamma_S Q^T\right)$ for instance instead of Γ_S itself. For these functionals, the plug-in estimators are relatively accurate for $t \gtrsim s+1$. See Appendix D for a discussion that theoretically establishes this point.

7. Numerical simulations. In this section we demonstrate performance of the AETC algorithms (Algorithms 5.1 and D.1). The regularization parameters α_t are set as $\alpha_t = 4^{-t}$ in all simulations, except in one case (first plot in Figure 5) where we investigate how the accuracy of AETC depends on α_t. We will utilize four methods to solve multifidelity problems:

- (MC) Classical Monte Carlo, where the entire budget is expended over the high-fidelity model.
- (MFMC) The multifidelity Monte Carlo procedure from [46], which is provided with oracle correlation information in Table 2 and Table 3.
- (AETC) Algorithm 5.1 (for scalar responses) or D.1 (for vector responses).
- (AETC-re) Algorithm 5.1 or D.1, but we recycle samples from exploration when computing exploitation estimates. See Remark 3.1.

To evaluate results, we compute and report an empirical mean-squared error over 500 samples. Since both the AETC and AETC-re produce random estimators due to the exploration step, the experiment (including both exploration and exploitation) is repeated 200 times with the 0.05-0.50-0.95-quantiles recorded.

7.1. Multifidelity finite element approximation for parametric PDEs. In the first example, we investigate the performance of our approach on multifidelity parametric solutions of linear elastic structures. We consider two scenarios associated with different geometries of the domain, namely *square* and *L-shape* structures. The geometry, boundary conditions, and loading in these structures are shown in Figure 1.

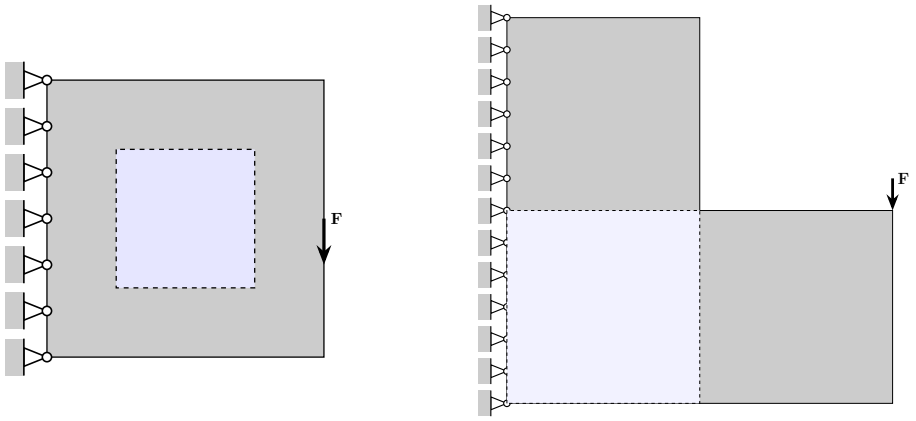

FIG. 1. *Geometry, boundary conditions, and loading for two linear elastic structures, the square and the L-shape. The regions in which the vector-valued quantities are investigated are highlighted with the light blue color.*

The model is an elliptic PDE that governs motion in linear elasticity. In a parametric setting, we introduce the parametric/stochastic version of this governing equation. In a bounded spatial domain D with boundary ∂D, the parametric elliptic PDE is

$$\begin{cases} \nabla \cdot (E(\boldsymbol{p}, \boldsymbol{x}) \nabla u(\boldsymbol{p}, \boldsymbol{x})) = f(\boldsymbol{x}) & \forall (\boldsymbol{p}, \boldsymbol{x}) \in \mathcal{P} \times D \\ u(\boldsymbol{p}, \boldsymbol{x}) = 0 & \forall (\boldsymbol{p}, \boldsymbol{x}) \in \mathcal{P} \times \partial D \end{cases} \tag{7.1}$$

where E, the elasticity matrix dependent on state variables $\boldsymbol{x}$, is parameterized with the random variables $\boldsymbol{p}$, and f is the forcing function. We assume that $\boldsymbol{p} \in \mathcal{P}$ is a d-dimensional random variable with independent components $\{p_i\}_{i=1}^d$. The solution of this parametric PDE is the displacement $u \equiv u(\boldsymbol{p}, \boldsymbol{x}) : \mathcal{P} \times D \to \mathbb{R}$. The displacement is used to compute the scalar quantity of interest, the *compliance*, which is the measure of elastic energy absorbed in the structure as a result of loading,

$$\mathsf{cpl} \coloneqq \int_D \nabla u(\boldsymbol{p}, \boldsymbol{x})^T E(\boldsymbol{p}, \boldsymbol{x}) \nabla u(\boldsymbol{p}, \boldsymbol{x}) d\boldsymbol{x} \tag{7.2}$$

The linear elastic structure is subjected to plane stress conditions and the uncertainty is considered in the material properties, namely the elastic modulus, which manifests in the model through the random parameters $\boldsymbol{p}$. To model the uncertainty, we consider a random field for the elastic modulus via a Karhunen-Loéve (KL) expansion:

$$E(\boldsymbol{p}, \boldsymbol{x}) = E_0(\boldsymbol{x}) + \delta \left[\sum_{i=1}^{d} \sqrt{\lambda_i} E_i(\boldsymbol{x}) p_i \right] \tag{7.3}$$

where $\delta = 0.5, E_0(\boldsymbol{x}) = 1$ are constants, and the random variables p_i are uniformly distributed on $[-1, 1]$ and the eigenvalues λ_i and basis functions E_i are taken from the analytical expressions for the eigenpairs of an exponential kernel on $D = [0, 1]^2$. The Poisson's ratio is $\nu = 0.3$, and we initially use $d = 4$ parameters. We solve the partial differential equation (7.1) for each fixed $\boldsymbol{p}$ via the finite element method with standard bilinear square isotropic finite elements on a rectangular mesh.

In this example, we form a multifidelity hierarchy through mesh coarsening: Consider $n = 7$ mesh resolutions with mesh sizes $h = \{1/(2^{8-L})\}_{L=1}^{7}$ where L denotes the level. The mesh associated with $L = 1$ yields the most accurate model (highest fidelity), which is taken as the high-fidelity model in our experiments. To utilize the notation presented earlier in this article, our potential low-fidelity regressors is formed from the compliance computed form various discretizations,

$$X^{(i)} = \mathsf{cpl}^{(i)} \in \mathbb{R}, \qquad i \in [6],$$

where $\mathsf{cpl}^{(i)}$ is the compliance of the solution computed using a solver with mesh size $h = 2^{i-7}$. We provide more details about the discretization and the uncertainty model of this section in Appendix E.

The cost for each model is the computation time, which we take to be inversely proportional to the mesh size squared, i.e., h^2. (This corresponds to employing a linear solver of optimal linear complexity.) We normalize cost so that the model with the lowest fidelity has unit cost. The total budget B ranges from 10^5 to 4×10^5, and our simulations increment the budget over this range by 0.5×10^5 units. The budget range considered here is sufficient to communicate relevant results. We refer to [60] for a further study of the square domain over a larger budget range of 10^5 to 10^7.

7.1.1. Scalar high-fidelity output. In the first experiment, the response variable is chosen as the compliance of the high-fidelity model, i.e.,

$$f(Y) = \mathsf{cpl}^{(0)} \in \mathbb{R}.$$

Therefore, for AETC we use Algorithm 5.1. The oracle statistics of $f(Y)$ and $X^{(i)}$ are computed over 50000 independent samples, and shown in illustrated in Table 2.

Models	$f(Y)$	$X^{(1)}$	$X^{(2)}$	$X^{(3)}$	$X^{(4)}$	$X^{(5)}$	$X^{(6)}$
$\mathrm{Corr}(\cdot, f(Y))$	1	0.998	0.992	0.976	0.940	0.841	-0.146
Mean	9.641	9.197	8.749	8.287	7.782	7.141	6.160
Standard deviation	0.127	0.113	0.099	0.086	0.072	0.052	0.027
Cost	4096	1024	256	64	16	4	1

Models	$f(Y)$	$X^{(1)}$	$X^{(2)}$	$X^{(3)}$	$X^{(4)}$	$X^{(5)}$	$X^{(6)}$
$\mathrm{Corr}(\cdot, f(Y))$	1	0.999	0.995	0.980	0.932	0.733	-0.344
Mean	25.940	24.256	22.902	21.180	19.054	16.072	12.275
Standard deviation	0.290	0.242	0.195	0.149	0.104	0.061	0.070
Cost	4096	1024	256	64	16	4	1

TABLE 2

*Oracle information of $f(Y)$ and $X^{(i)}$ (approximated with 3 digits) in the case of the square domain (**top**) and the L-shape domain (**bottom**).*

In both cases, the expensive models are more correlated with the high-fidelity model than the cheaper ones. Such hierarchical structure is often a necessary assumption for the implementation of multifidelity methods, the MFMC, for instance.

Accuracy results for various multifidelity procedures are shown in the first two plots in Figure 2. As the budget goes to infinity, the model selected by AETC converges to the minimizer given by (5.5), which can be explicitly computed using oracle information. Particularly, solutions to (5.5) in the case of the square domain and the L-shape domain are respectively as follows:

- Square domain: $f(Y) \sim X^{(4)} + X^{(5)} + X^{(6)} + \text{intercept}$
- L-shape domain: $f(Y) \sim X^{(3)} + X^{(4)} + X^{(5)} + X^{(6)} + \text{intercept}$.

The regression coefficients for the square domain and the L-shape domain are 6.151 ($X^{(4)}$), -6.509 ($X^{(5)}$), 1.444 ($X^{(6)}$), -0.640 (intercept) and 5.537 ($X^{(3)}$), -6.606 ($X^{(4)}$), 2.559 ($X^{(5)}$), -0.437 ($X^{(6)}$), -1.210 (intercept), respectively. Note that there is a canceling effect (change in sign of coefficients) for the more expensive regressors due to high correlations. Neither limiting model above will be used by the MFMC estimator as the relationship between cost and correlation does not satisfy the assumption [46, condition (20)]. For every budget level B under our test, we compute the empirical probability (as a percentage) that the limiting model is selected by AETC and plot this in the rightmost panel in Figure 2.

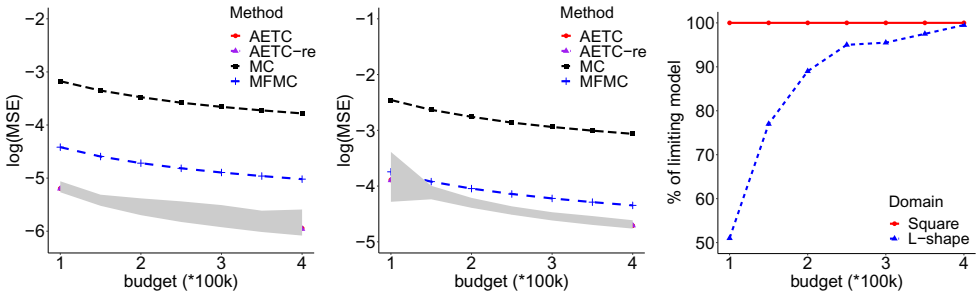

FIG. 2. *Comparison of the* ($\log_{10}$) *mean-squared error of the LRMC estimator given by the AETC algorithm, the AETC algorithm reusing the exploration samples (AETC-re), the MC estimator, and the MFMC estimator as the total budget increases from* 10^5 *to* 4×10^5 *in the case of the square domain* (***left***) *and the L-shape domain* (***middle***). *The* 0.05-0.50-0.95-*quantiles are plotted for the LRMC estimator to measure its uncertainty. We also compute the probability (plotted as a percentage) that the AETC algorithm selects the limiting model given by* (5.5) (***right***).

Figure 2 shows that for both domain geometries, the AETC algorithm outperforms both the MC and the MFMC by a notable margin, even though the latter has access to oracle correlation statistics that are not provided to AETC. Little difference between AETC and AETC-re is visible, which is not surprising since the number of the exploitation samples is much larger than the number of exploration rounds. The mean-squared error of AETC is smaller in the case of the square domain than in the L-shape domain under the same budget, for which a possible explanation is that the variance of the surrogate models in the former is smaller than in the latter (Table 2), making the exploration procedure more efficient in the square domain case. Finally, note that as the budget goes to infinity, the mean-squared error of the AETC algorithm decays to zero, and the frequency of the AETC exploiting the limiting model given by (5.5) converges to 1, verifying the asymptotic results in Theorem 5.3. The statistics of the learned regression coefficients of the limiting model in both cases when $B = 4 \times 10^5$ are provided in Figure 3.

In this example, despite a deterministic nonlinear relationship between $f(Y)$ and $X^{(i)}$, the conditional expectation $\mathbb{E}[f(Y)|X_S]$ does seem to admit a reasonable ap-

proximation using an appropriately constructed linear combinations of the low-fidelity model outputs in the selected subsets by AETC. In particular, it can be seen from Figure 3 that the regression coefficients learned by AETC on the limiting model, which are estimated based on less than 25 random exploration samples in both cases, are concentrated around their true values computed using the oracle dataset (50000 independent samples). The concentration effect is more prominent in the square domain case due to a larger exploration rate.

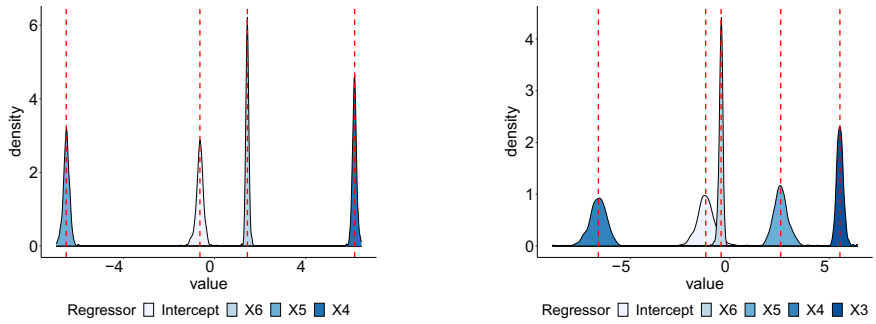

FIG. 3. *Densities of the estimated coefficients of the limiting model learned by AETC under budget $B = 4 \times 10^5$ in 200 experiments in the case of the square domain (**left**) and the L-shape domain (**right**). Red dashed lines correspond to the true values of the coefficients estimated using the oracle dataset.*

In the rest of this section, we focus for simplicity on the square domain case to further investigate the performance of AETC. We will take a much larger KL truncation dimension $d = 20$ to inject more uncertainty into the models. To better demonstrate the behavior of the algorithms in the asymptotic regime, we use an increased budget range compared to the previous simulation, i.e., the total budget B ranges from 2×10^5 to 12×10^5, with budget increment 2×10^5 units. The oracle statistics of $f(Y)$ and $X^{(i)}$ are computed over 50000 independent samples, and shown in Table 3:

Models	$f(Y)$	$X^{(1)}$	$X^{(2)}$	$X^{(3)}$	$X^{(4)}$	$X^{(5)}$	$X^{(6)}$
Corr$(\cdot, f(Y))$	1	0.998	0.989	0.966	0.914	0.789	-0.135
Mean	9.645	9.200	8.752	8.290	7.784	7.142	6.160
Standard deviation	0.137	0.120	0.104	0.089	0.073	0.052	0.027
Cost	4096	1024	256	64	16	4	1

TABLE 3

Oracle information of $f(Y)$ and $X^{(i)}$ (approximated with 3 digits) in the case of the square domain with randomness dimension $d = 20$.

We first conduct a standard analysis for AETC, including an average mean-squared error comparison with the other two methods and the convergence to the limiting model. As before, we use oracle statistics to compute the limiting model via (5.5), which in this case is given by $f(Y) \sim X^{(3)} + X^{(4)} + X^{(5)} + X^{(6)} + \text{intercept}$. Note that this model is different from the one in the previous example where $d = 4$. A plausible explanation, as can be inspected from Table 2 and 3, is that the low-fidelity models are less correlated with the high-fidelity model when d is large, thus AETC leans toward utilizing more low-fidelity models to achieve equally accurate predictions. Moreover, we include two additional plots to summarize the statistics (median) of the

exploration/exploitation rate in AETC as opposed to the optimal rates (5.4) computed using the oracle statistics, as well as a plot of the estimation given by AETC along 100 trajectories at different budget levels. The results are reported in Figure 4.

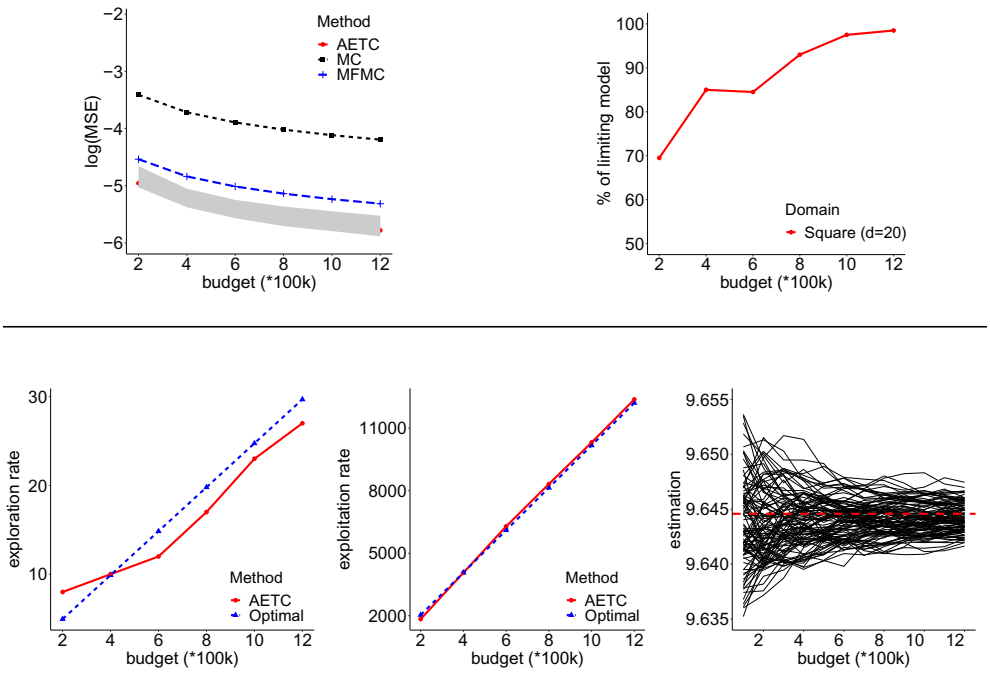

FIG. 4. *Comparison of the* ($\log_{10}$) *mean-squared error of the LRMC estimator given by the AETC algorithm, the MC estimator, and the MFMC estimator as the total budget increases from* 2×10^5 *to* 12×10^5 *in the case of the square domain with KL truncation parameter* $d = 20$ ***(top left)***. *The* 0.05-0.50-0.95*-quantiles are plotted for the LRMC estimator to measure its uncertainty. We also compute the probability (plotted as a percentage) that the AETC algorithm selects the limiting model given by* (5.5) ***(top right)***, *and the median of the exploration* ***(bottom left)*** *and exploitation* ***(bottom middle)*** *rate in AETC at different budget. Finally, we provide the estimated values for* $\mathbb{E}[f(Y)]$ *of the LRMC estimation in AETC at different budget along 100 random trajectories, with the red dashed line referring to the true value of* $\mathbb{E}[f(Y)]$ ***(bottom right)***.

Figure 4 shows that AETC consistently outperforms the other two methods as in Figure 2 and the expected model convergence. Both the exploration and exploitation rates in AETC asymptotically match the optimal rates computed using the oracle statistics, verifying the statement (5.10a) in Theorem 5.3. Moreover, by plotting the AETC estimation along 100 random trajectories (with multiple evaluations at different budgets), we notice that the LRMC estimator in the AETC algorithm converges to $\mathbb{E}[f(Y)]$ as the budget goes to infinity. This observation is consistent with Remark 5.4.

We next investigate how the performance of AETC depends on the regularization parameters α_t. In addition to the value $\alpha_t = 4^{-t}$ previously used, we consider two alternative choices $\alpha_t = 2^{-t}$ and $\alpha_t = 8^{-t}$ corresponding to a more and less active exploration strategy, respectively, in the early stage of learning. We apply AETC with different regularization parameters to the same training dataset under different total budgets and record the corresponding average mean squared errors. The simulation results are reported in Figure 5.

It can be seen from Figure 5 that for a large budget, the three exponential-decaying choices for α_t yield similar accuracy results. For a small budget, however, the larger α_t is slightly more accurate, likely due to a more aggressive early-stage exploration enforced by the regularization.

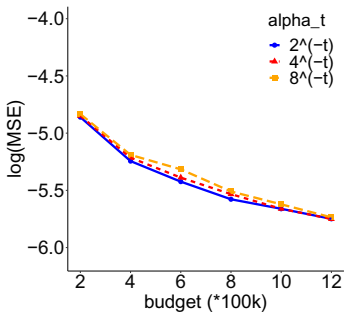

FIG. 5. *Comparison of the average ($\log_{10}$) mean-squared error of the AETC algorithm at three different regularization parameters ($\alpha_t = 2^{-t}, 4^{-t}, 8^{-t}$) as the total budget increases from 2×10^5 to 12×10^5 in the case of the square domain with KL truncation $d = 20$.*

7.1.2. Vector-valued high-fidelity output. The same experiment is repeated when $f(Y)$ is taken as a vector-valued response, and hence we utilize Algorithm D.1. We take $f(Y)$ to be defined by 9 randomly selected components of the magnitude of the discrete solution (displacement) given by the high-fidelity model within a region shown with the highlighted (blue) color in Figure 1, i.e.,

$$f(Y) := \bar{\boldsymbol{u}}^{(0)}|_T \qquad \bar{\boldsymbol{u}}^{(0)} = \sqrt{\left(\boldsymbol{u}_x^{(0)}\right)^2 + \left(\boldsymbol{u}_y^{(0)}\right)^2} \in \mathbb{R}_+^{2601}$$

where T is a subset of coordinates in the highlighted region in Figure 1 with $|T| = 9$, and the arithmetic operations above are taken componentwise. The full spatial domain of the structures depicted in Figure 1 is in $[0,1]^2$, and the highlighted regions are squares with the size 0.5×0.5 in the center and lower left corner of the square and L-shape structures, respectively. In order to compute the vector-valued quantity at the same points across all resolutions, we evaluate $\bar{\boldsymbol{u}}$ on a fixed 51×51 mesh across different resolutions, where the evaluations are accomplished by using the continuous solution from the finite element approximation.

For the square domain, T is taken as a randomly selected 3×3 block of pixels from the 51×51 solution vector corresponding to the highlighted region in Figure 1. For the L-shape domain, T is taken as 9 randomly sampled components from the 51×51-dimensional solution vector corresponding to the highlighted region in Figure 1. For both cases, Q is the identity matrix I_9, so the Q-risk defined in (6.3) is simply the sum of the mean-squared error of each response. We show oracle information of the correlations between $X^{(i)}$ and $f^{(j)}(Y)$, $i \in [6], j \in [9]$ in Figure 6.

For both domain structures, cheaper models have some strong correlations with the high-fidelity model, implying that the cost-correlation hierarchical structure (an assumption for the MFMC) is violated. We thus compare only the MC estimator with the AETC (for vector-valued high-fidelity output). The oracle limiting models to which AETC will converge are:

- Square domain: $f(Y) \sim X^{(4)} + X^{(5)} + X^{(6)} + \text{intercept}$
- L-shape domain: $f(Y) \sim X^{(3)} + X^{(4)} + X^{(5)} + X^{(6)} + \text{intercept}$.

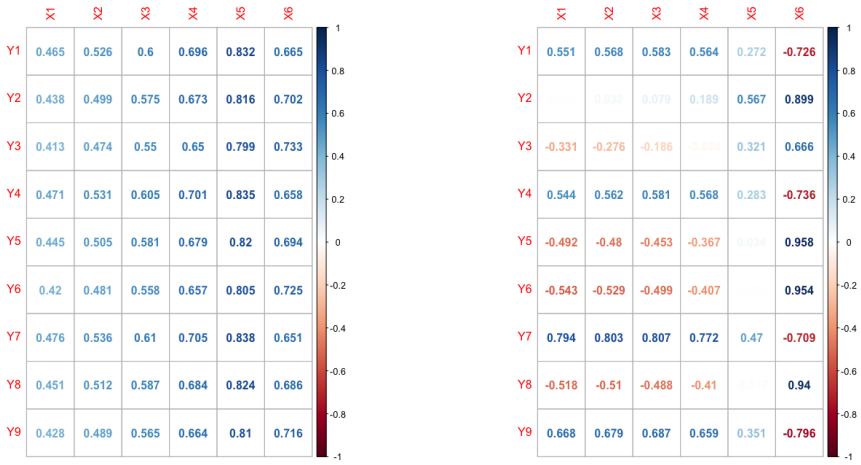

FIG. 6. *Correlations between* $X^{(i)}$ *and* $f^{(j)}(Y)$, $i \in [6], j \in [9]$ *in the case of the square domain* (***left***) *and the L-shape domain* (***right***). *The entry associated to indices (Xi, Yj) is the value of* $Corr(X^{(i)}, f^{(j)}(Y))$.

The details of the results are given in Figure 7, which are consistent with the conclusion drawn from Figure 2.

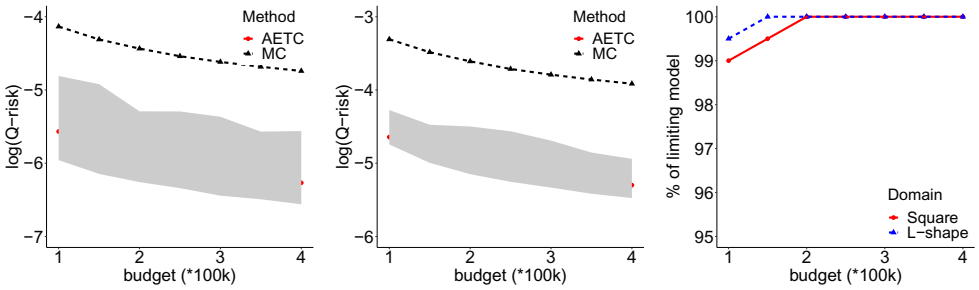

FIG. 7. *Comparison of the* ($\log_{10}$) *Q-risk of the LRMC estimator given by the AETC algorithm and the MC estimator as the total budget increases from* 10^5 *to* 4×10^5 *in the case of the square domain* (***left***) *and the L-shape domain* (***middle***). *The* 0.05-0.50-0.95-*quantiles are plotted for the LRMC estimator to measure its uncertainty. We also compute the probability (plotted as a percentage) that the AETC algorithm selects the limiting model given by* (D.11) (***right***).

7.2. Mixture of reduced-order models. In this example, we consider a multi-fidelity problem using the same model as in the previous section on the square domain, but we generate different types of reduced-order models (or surrogate models) which approximate the compliance of the solution to (7.1). We will use the compliance computed via the Finite Element Method on a $2^6 \times 2^6$ mesh as the ground truth, i.e., the surrogate model $X^{(1)}$ in the first table in Table 2 is treated as the high-fidelity model Y. We explore two classes of choices for reduced-order models:

- Gaussian process (GP) emulators. Low-fidelity regressors $X^{(i)}$ for $i \in [6]$ are generated as the mean of GP emulators built on compliance data from Y. We use an exponential covariance kernel and optimize hyperparameters

by maximizing the log-likelihood. For $n_T = 10, 100, 1000$ training points in parameter space, this defines models $X^{(1)}$, $X^{(3)}$, $X^{(5)}$, respectively. We then select non-optimal hyperparameters with the same training data $n_T = 10, 100, 1000$, which defines models $X^{(2)}$, $X^{(4)}$, $X^{(6)}$, respectively. The cost of these models is given by the cost of training, optimization, and evaluation the GP averaged over 2×10^4 different values of $\boldsymbol{p}$. We perform each experiment five times and report the average time (cost) for each model in Table 4.

- Projection-based model reduction with proper orthogonal decomposition (POD). We generate k POD basis functions from high-fidelity displacement data, use this to form a rank-k Galerkin projection of the finite element formulation, and models $X^{(i)}$ for $i = (7, \ldots, 12)$ are the compliances computed from the projected systems of rank $k = (1, 2, 3, 4, 5, 10)$, respectively. The cost of this procedure is taken as *only* the cost of solving the rank-k projected system and does *not* include the time required to collect POD training data or the time required to compute the POD modes.

More details about the experimental setup above are given in Appendix F. The oracle statistics and costs for the GP and POD low-fidelity models are shown in Table 4. Note that here we have not normalized the cost relative to the low-fidelity model.

Models	$f(Y)$	$X^{(1)}$	$X^{(2)}$	$X^{(3)}$	$X^{(4)}$	$X^{(5)}$
$\mathrm{Corr}(\cdot, f(Y))$	1	0.993	0.414	1-2e-05	0.401	1-1e-06
Mean	9.197	9.195	9.147	9.197	9.045	9.197
Standard deviation	0.113	0.122	0.556	0.113	0.335	0.113
Cost	9233.69	0.31	0.31	2.31	2.33	30.79
(5.4) with $S = \{i\}$	–	4.202	203.025	0.358	214.676	3.574

Models	$X^{(6)}$	$X^{(7)}$	$X^{(8)}$	$X^{(9)}$	$X^{(10)}$	$X^{(11)}$	$X^{(12)}$
$\mathrm{Corr}(\cdot, f(Y))$	1-2e-04	0.999	1-2e-04	1-5e-05	1-8e-06	≈ 1	≈ 1
Mean	9.196	9.189	9.194	9.195	9.197	9.197	9.197
Standard deviation	0.114	0.113	0.113	0.113	0.113	0.113	0.113
Cost	31.29	0.18	0.45	0.68	0.85	1.03	2.02
(5.4) with $S = \{i\}$	6.226	0.732	0.243	0.180	0.137	**0.119**	0.230

TABLE 4

Oracle information of $f(Y)$ and $X^{(i)}$. The numerator of the asymptotic average conditional MSE estimate (5.4) is also shown.

We now apply the (scalar response) AETC algorithm to this multifidelity setup, with the total budget ranging from 10^5 to 2×10^5, incremented by 0.2×10^5 in our experiments. We have 12 low-fidelity models in total, and exhausting all of them for selection would require complexity on the order of $2^{12} - 1$. Since the AETC algorithm generally produces an efficient combination of relatively cheap regressors, we set the maximal number of regressors in the model to be 5 to accelerate computation, i.e., we explore only for $s = |S| \leq 5$. (We could have used the full model, which would not incur too much an increase in computation time but would require taking more exploration samples to start with, which is a waste of resources.)

The performance of AETC is compared to the direct MC estimator applied to $f(Y)$, and the results are reported in the first plot in Figure 8. The figure shows that AETC is more efficient than the MC estimator by a substantial margin in terms of the mean-squared error. To better see how this is reflected in practice, we fix the budget to be $B = 10^5$, run 200 experiments of both AETC and the MC, and compute

the difference between the estimated values of $\mathbb{E}[f(Y)]$ and the ground truth. This is illustrated in the second plot in Figure 8. Since this multifidelity setup contains surrogates obtained from different types of methods, we investigate which are chosen by AETC for exploitation. This information is given in the last plot in Figure 8. The most frequent models chosen by the AETC are between model $X^{(10)}$ and $X^{(11)}$, both of which have near-perfect correlation with $f(Y)$ with only a moderate cost. In fact, $X^{(11)}$ is also the oracle limiting model given by (5.5), although the convergence is rather slow due to the competitor $X^{(10)}$, which has an extremely high correlation but a slightly cheaper cost. The other models which have been selected by the AETC are the other POD models as well as their combinations with models $X^{(1)}$ and $X^{(2)}$. These models correspond to the lowest fidelity models in each method of approximation and are often cheap to sample from with reasonably high correlation.

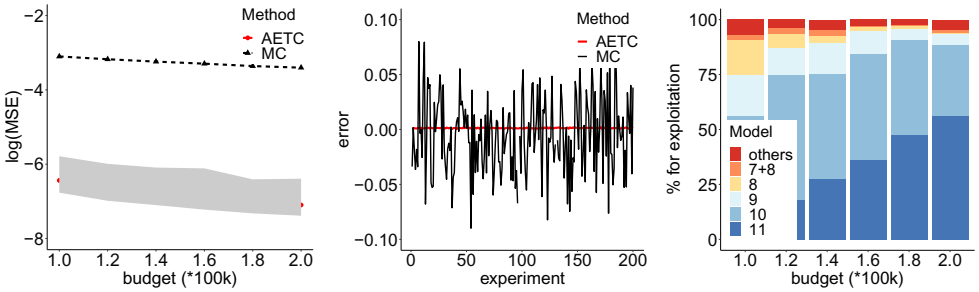

FIG. 8. *The ($\log_{10}$) mean-squared error of the estimators as the budget increases, with the* 0.05-0.5-0.95 *quantiles plotted for the AETC algorithm to measure the uncertainty in the exploration (**left**). Comparison of the estimation errors of MC and the AETC algorithm by plotting instances of the error over each of* 200 *experiments for a fixed total budget $B = 10^5$ (**middle**). Empirical probability (plotted as a percentage) of certain models being selected by AETC for exploitation (**right**).*

8. Conclusions. We have proposed a novel algorithm for the multifidelity problem based on concepts from bandit learning. Our proposed AETC procedures, Algorithms 5.1 and D.1, operate under the assumption of a linear relationship between low-fidelity features and a high-fidelity output. Their exploration phase expends resources to learn about model relationships to discover an effective linear relationship. The exploitation phase leverages the cost savings by regressing on low-fidelity features to produce an LRMC estimate of the high-fidelity output. Under the linear model assumption, we show the consistency of the LRMC, whose MSE can be partitioned into a component stemming from exploration, and another from exploitation. This partition allows us to construct the AETC algorithm, which is an adaptive procedure operating under a specified budget that decides how much effort to expend in exploration versus exploitation, and also identifies an effective linear regression low-fidelity model. We show that, for a large budget, our AETC algorithms explore and exploit optimally.

The main advantage of our approach is that no *a priori* statistical information or hierarchical structure about models is required, and very little knowledge about model relationships is needed: AETC needs only identification of which model is the trusted high-fidelity one, along with a specification of the cost of sampling each model. The algorithm proceeds from this information alone, which is a distinguishing feature of our approach compared to alternative multifidelity and multilevel methods, and can

be used to tackle situations when a natural cost versus accuracy hierarchy is difficult to identify, and when relationships between models are not known.

Acknowledgement. We would like to thank the anonymous referees for their very helpful comments which significantly improved the presentation of the paper.

Appendices.

A. The LRMC reusing the exploration data. The goal of this section is to provide a quantitative discussion that accompanies Remark 3.1. Our analysis in this paper computes coefficients $\widehat{\beta}_S$ in exploitation using only the N_S samples that are taken during exploitation. This analysis neglects the possibility of reusing the m samples from exploration in the estimation of $\widehat{\beta}_S$. All our numerical results fall into a regime where $m \ll N_S$, corresponding to the case when there are many more exploitation samples (of low-fidelity models) than exploration rounds (which require querying the high-fidelity model). This section demonstrates that in this multifidelity regime of interest, recycling of exploration samples during exploitation has negligible impact.

Fix a model $S \subseteq [n]$. The LRMC estimator with the exploration samples reused is defined as

$$\mathsf{LRMC}_{S,\mathrm{re}} = \frac{1}{N_S+m}\left(\underbrace{\sum_{\ell\in[N_S]} X_{S,\ell}^T}_{\text{exploitation samples}} + \underbrace{\sum_{j\in[m]} X_{\mathrm{epr},j}^T|_S}_{\text{exploration samples reused}}\right)\widehat{\beta}_S. \tag{A.1}$$

It follows from a similar computation as in (3.6) that the conditional mean-squared error of $\mathsf{LRMC}_{S,\mathrm{re}}$ on the exploration data (including both the exploration samples and the model noise) is

$$\frac{N_S}{(N_S+m)^2}\widehat{\beta}_S\Sigma_S\widehat{\beta}_S + \left(\widetilde{X}_m^T\widehat{\beta}_S - x_S^T\beta_S\right)^2, \tag{A.2}$$

where

$$\begin{aligned}\widetilde{X}_m &= \frac{N_S}{N_S+m}x_S + \frac{1}{N_S+m}\sum_{j\in[m]} X_{\mathrm{epr},j}|_S \\ &\overset{(5.7)}{=} x_S + \underbrace{\frac{m}{N_S+m}(\widehat{x}_S(m) - x_S)}_{:=\Delta}.\end{aligned}$$

Note $\Delta = o(x_S)$ as $m \to \infty$ due to the law of large numbers. Averaging the model

noise in (A.2) yields the average conditional MSE of $\mathsf{LRMC}_{S,\mathrm{re}}$:

$$
\begin{aligned}
&\mathbb{E}_{\varepsilon_S}[(\text{A.2})] \\
&\overset{(3.7),(B.1)}{=} \frac{N_S}{(N_S+m)^2}\left[\beta_S^T\Sigma_S\beta_S + \sigma_S^2\,\mathrm{tr}\big(\Sigma_S(Z_S^TZ_S)^{-1}\big)\right] + \mathbb{E}_{\varepsilon_S}[(x_S^T(\widehat{\beta}_S-\beta_S))^2] \\
&\quad + 2\mathbb{E}_{\varepsilon_S}[x_S^T(\widehat{\beta}_S-\beta_S)\Delta^T\widehat{\beta}_S] + \mathbb{E}_{\varepsilon_S}[(\Delta^T\widehat{\beta}_S)^2] \\
&\simeq \frac{N_S}{(N_S+m)^2}\beta_S^T\Sigma_S\beta_S + \sigma_S^2x_S^T(Z_S^TZ_S)^{-1}x_S + 2\sigma_S^2x_S^T(Z_S^TZ_S)^{-1}\Delta \\
&\quad + \left(\frac{m}{N_S+m}\right)^2\left[(\widehat{x}_S-x_S)^T\beta_S\right]^2 \\
&\simeq \frac{N_S}{(N_S+m)^2}\beta_S^T\Sigma_S\beta_S + \sigma_S^2x_S^T(Z_S^TZ_S)^{-1}x_S + \left(\frac{m}{N_S+m}\right)^2\left[(\widehat{x}_S-x_S)^T\beta_S\right]^2 \\
&\simeq \frac{N_S}{(N_S+m)^2}\beta_S^T\Sigma_S\beta_S + \frac{1}{m}\sigma_S^2x_S^T\Lambda_S^{-1}x_S + \left(\frac{m}{N_S+m}\right)^2\left[(\widehat{x}_S-x_S)^T\beta_S\right]^2. \qquad \text{(A.3)}
\end{aligned}
$$

Thus, one can verify that

$$
\begin{aligned}
&\frac{1}{N_S}\beta_S^T\Sigma_S\beta_S + \frac{1}{m}\sigma_S^2\,\mathrm{tr}\big(x_Sx_S^T\Lambda_S^{-1}\big) \\
&\le \left(1+\frac{m}{N_S}\right)^2\left[\frac{N_S}{(N_S+m)^2}\beta_S^T\Sigma_S\beta_S + \frac{1}{m}\sigma_S^2x_S^T\Lambda_S^{-1}x_S\right] \\
&\le \left(1+\frac{m}{N_S}\right)^2\cdot(\text{A.3}). \qquad \text{(A.4)}
\end{aligned}
$$

While according to (5.1), the leftmost term in (A.4) converges to $\overline{\mathsf{MSE}}_S|_{\widehat{\beta}_S}$ almost surely. Thus, providing $m/N_S \ll 1$, reusing exploration samples results in the estimate (A.3), which asymptotically has a negligible advantage over the estimate $\overline{\mathsf{MSE}}_S|_{\widehat{\beta}_S}$ that does not reuse samples.

B. Proof of Theorem 4.2. We start by conditioning on Z_S. Note that

$$
\widehat{\beta}_S - \beta_S = Z_S^\dagger\eta_S, \qquad \text{(B.1)}
$$

where η_S is the model noise vector with each component being the model noise ε_S in the corresponding exploration sample, i.e., $\eta_S/\sigma_S \in \mathbb{R}^m$ is an isotropic sub-Gaussian random vector with $\mathbb{E}[\eta_S\eta_S^T] = \sigma_S^2I_m$. Substituting (B.1) into (3.6) and applying the Cauchy-Schwarz inequality ($|2\langle x,y\rangle| \le 2\|x\|_2\|y\|_2 \le \|x\|_2^2 + \|y\|_2^2$) yields

$$
\begin{aligned}
\mathsf{MSE}_S|_{\widehat{\beta}_S} &= \frac{1}{N_S}\left(\beta_S^T\Sigma_S\beta_S + \eta_S^T(Z_S^\dagger)^T\Sigma_SZ_S^\dagger\eta_S + 2\beta_S^T\Sigma_SZ_S^\dagger\eta_S\right) + \eta_S^T(Z_S^\dagger)^Tx_Sx_S^TZ_S^\dagger\eta_S \\
&\le \frac{2}{N_S}\left(\beta_S^T\Sigma_S\beta_S + \eta_S^T(Z_S^\dagger)^T\Sigma_SZ_S^\dagger\eta_S\right) + \eta_S^T(Z_S^\dagger)^Tx_Sx_S^TZ_S^\dagger\eta_S \\
&\le \frac{2}{N_S}\left(\beta_S^T\Sigma_S\beta_S + \|\Sigma_S\|_2\|Z_S^\dagger\eta_S\|_2^2\right) + \|B_S\eta_S\|_2^2 \qquad B_S = \sqrt{x_Sx_S^T}Z_S^\dagger. \qquad \text{(B.2)}
\end{aligned}
$$

Both $\|Z_S^\dagger\eta_S\|_2^2$ and $\|B_S\eta_S\|_2^2$ are the quadratic forms of sub-Gaussian random vectors and can be bounded with high probability using the Hanson-Wright inequality [59,

Theorem 6.3.2]:

$$\mathbb{P}\left(\|Z_S^\dagger \eta_S\|_2^2 > C_1\sigma_S^2\|Z_S^\dagger\|_F^2 \log m\right) \le \frac{1}{3m^2}$$
$$\mathbb{P}\left(\|B_S\eta_S\|_2^2 > C_1\sigma_S^2\|B_S\|_F^2 \log m\right) \le \frac{1}{3m^2}, \tag{B.3}$$

where C_1 is an absolute constant depending only on the sub-Gaussian norm of ε_S/σ_S, and

$$\|Z_S^\dagger\|_F^2 = \operatorname{tr}\Big(Z_S^\dagger (Z_S^\dagger)^T\Big) = \frac{1}{m}\operatorname{tr}\big((m^{-1}Z_S^T Z_S)^{-1}\big)$$
$$\|B_S\|_F^2 = \operatorname{tr}\big(B_S^T B_S\big) = \frac{1}{m}\operatorname{tr}\Big(\big(m^{-1}Z_S^T Z_S\big)^{-1} x_S x_S^T\Big). \tag{B.4}$$

Since $Z_S^T Z_S$ is a sum of i.i.d. outer products, appealing to a deviation result in [41, Theorem 2.1], we obtain that under assumption (4.1) and for $m > s+1$,

$$\mathbb{P}\left(\left\|\frac{1}{m}Z_S^T Z_S - \Lambda_S\right\|_2 > C_2\frac{K^2\log^5 m}{\sqrt{m}}\right)$$
$$= \mathbb{P}\left(\left\|\frac{1}{m}\sum_{\ell\in[m]} X_{\mathrm{epr},\ell}|_S X_{\mathrm{epr},\ell}^T|_S - \mathbb{E}[X_S X_S^T]\right\|_2 > C_2\frac{K^2\log^5 m}{\sqrt{m}}\right) < \frac{1}{3m^2}, \tag{B.5}$$

where C_2 is an absolute constant. Combining (B.5) with the matrix perturbation equality for an invertible matrix A

$$(A+\Delta A)^{-1} = A^{-1} - A^{-1}\Delta A A^{-1} + o(\|\Delta A\|_2)$$

yields with high probability,

$$\big(m^{-1}Z_S^T Z_S\big)^{-1} = \Lambda_S^{-1} + E_S \qquad \|E_S\|_2 \lesssim \frac{K^2\log^5 m}{\sigma_{\min}^2(\Lambda_S)\sqrt{m}}, \tag{B.6}$$

where $\sigma_{\min}(\Lambda_S)$ is the smallest singular value of Λ_S. Putting (B.6), (B.5), (B.4), (B.3) and (B.2) together, we have with probability at least $1-m^{-2}$,

$$\begin{aligned}
\mathsf{MSE}_S|_{\widehat{\beta}_S} &\lesssim \frac{2}{N_S}\left[\beta_S^T\Sigma_S\beta_S + C_1\sigma_S^2\|\Sigma_S\|_2\operatorname{tr}\left(\Lambda_S^{-1} + \frac{K^2\log^5 m}{\sigma_{\min}^2(\Lambda_S)\sqrt{m}}I_s\right)\frac{\log m}{m}\right] \\
&\quad + C_1\sigma_S^2\operatorname{tr}\left(\Lambda_S^{-1}x_S x_S^T + \frac{K^2\log^5 m}{\sigma_{\min}^2(\Lambda_S)\sqrt{m}}x_S x_S^T\right)\frac{\log m}{m} \\
&\lesssim \frac{1}{N_S}\beta_S^T\Sigma_S\beta_S + \sigma_S^2\operatorname{tr}\big(\Lambda_S^{-1}x_S x_S^T\big)\frac{\log m}{m} \\
&\le \frac{1}{N_S}\beta_S^T\Sigma_S\beta_S + \sigma_S^2\operatorname{tr}\big(\Lambda_S^{-1}(x_S x_S^T + \Sigma_S)\big)\frac{\log m}{m} \\
&= \frac{1}{N_S}\beta_S^T\Sigma_S\beta_S + (s+1)\sigma_S^2\frac{\log m}{m}.
\end{aligned} \tag{B.7}$$

This finishes the proof.

C. Proof of Theorem 5.3. We first show that $m(B)$ diverges as $B \to \infty$ almost surely. To this end, it suffices to show that with probability 1,

$$\sup_{t>n+1} \max_{S\subseteq[n]} \widehat{\beta}_S^T(t)\widehat{\Sigma}_S(t)\widehat{\beta}_S(t) < \infty. \tag{C.1}$$

Indeed, if (C.1) is true, by definition (5.9), for almost every realization ω, there exists an $L(\omega) < \infty$ such that

$$\sup_{t>n+1} \max_{S\subseteq[n]} k_{1,t}(S;\omega) < L(\omega), \tag{C.2}$$

where ω is included to stress the quantity's dependence on realization. The exploration stopping criterion of Algorithm 5.1 requires that

$$m(B;\omega) \geq m_{S^*(t;\omega)}(t;\omega) = \frac{B}{c_{\text{epr}} + \sqrt{\frac{c_{\text{epr}} k_{1,t}(S^*(t;\omega);\omega)}{k_{2,t}(S^*(t;\omega);\omega)}}} \overset{\text{(C.2)}}{\geq} \frac{B}{c_{\text{epr}} + \sqrt{\frac{c_{\text{epr}} L(\omega)}{\alpha_{m(B;\omega)}}}}$$

Note that $m(B;\omega)$ is nondecreasing in B. If $m(B;\omega)$ did not diverge, then there would exist an integer $C > 0$ such that $\sup_B m(B;\omega) < C$. Meanwhile, this also suggests that $\inf_B \alpha_{m(B;\omega)} > \alpha_C > 0$, forcing the right-hand side of the above inequality to diverge; hence $m(B;\omega)$. A contradiction. Thus,

$$\lim_{B\to\infty} m(B;\omega) = \infty. \tag{C.3}$$

To show (C.1), note that $\widehat{\Sigma}_S(t)$ converges to Σ_S almost surely due to the strong law of large numbers. For $\widehat{\beta}_S$, the sub-exponential assumption on the distribution of X_S for all $S \subseteq [n]$ (condition (4.1)) ensures (B.5) (with m replaced by t) for all $S \subseteq [n]$, which combined with the Borel-Cantelli lemma implies that with probability 1, $\lambda_{\min,S}(t) \to \infty$ and

$$\log \lambda_{\max,S}(t) = o(\lambda_{\min,S}(t)) \qquad\qquad \forall S \subseteq [n],$$

where $\lambda_{\max,S}(t)$ and $\lambda_{\min,S}(t)$ are respectively the largest and smallest eigenvalue of $Z_S^T Z_S$. Appealing to [38, Theorem 1], with probability 1, $\widehat{\beta}_S(t) \to \beta_S$ as $t \to \infty$ for all $S \subseteq [n]$. Thus we have proved (C.1).

We now work with a fixed realization ω along which $m(B;\omega) \to \infty$ as $B \to \infty$, and all estimators in (5.7) converge to the true parameters as $t \to \infty$. Recall the empirical MSE estimated for model S in round t, as a function of the exploration round m:

$$\widehat{\overline{\mathsf{MSE}}}_S|_{\widehat{\beta}_S}(m;t) = \frac{k_{1,t}(S)}{B - c_{\text{epr}} m} + \frac{k_{2,t}(S)}{m}.$$

Take $\delta < 1/2$ sufficiently small. Since (5.5) is assumed to have a unique minimizer, by consistency and a continuity argument, there exists a sufficiently large $T(\delta;\omega)$ so that, for all $t \geq T(\delta,\omega)$,

$$\max_{(1-\delta)m_{S^*} \leq m \leq (1+\delta)m_{S^*}} \widehat{\overline{\mathsf{MSE}}}_{S^*}|_{\widehat{\beta}_{S^*}}(m;t) < \min_{S\subseteq[n], S\neq S^*} \min_{0<m<B/c_{\text{epr}}} \widehat{\overline{\mathsf{MSE}}}_S|_{\widehat{\beta}_S}(m;t) \tag{C.4}$$

$$1 - \delta \leq \frac{m_S(t;\omega)}{m_S} \leq 1 + \delta \qquad\qquad \forall S \subseteq [n]. \tag{C.5}$$

Since m_S scales linearly in B and $m(B;\omega)$ diverges as $B \uparrow \infty$, there exists a sufficiently large $B(\delta;\omega)$ such that for $B > B(\delta;\omega)$,

$$\min\left\{\min_{S\subseteq[n]} m_S, m(B;\omega)\right\} > 2T(\delta;\omega). \tag{C.6}$$

(C.5), (C.6) and $\delta < 1/2$ together imply that, at $T(\omega)$, for all $S \subseteq [n]$,

$$m_S(T(\delta;\omega);\omega) \overset{\text{(C.5)},\delta<1/2}{\geq} \frac{m_S}{2} \overset{\text{(C.6)}}{>} T(\delta;\omega). \tag{C.7}$$

Combining (C.7) with (C.4) yields that, at $T(\omega)$, the algorithm chooses S^* as the optimal model, and the corresponding estimated optimal exploration rate is larger than $T(\delta;\omega)$. By the design of our algorithm, more exploration is needed, and (C.4) and (C.5) further ensure that S^* will be chosen in the subsequent exploration until the stopping criterion is met, proving (5.10a). Moreover, when the algorithm stops,

$$1 - \delta \leq \frac{m(B;\omega)}{m_{S^*}} \leq 1 + \delta + \frac{1}{m_{S^*}}.$$

Taking $B \to \infty$ followed by $\delta \to 0$ yields (5.10b), finishing the proof.

D. Proof of Theorem 6.1. In this section, we provide some omitted details of the analysis in the case of vector-valued high-fidelity models. We first complete the proof of Theorem 6.1, which establishes a similar nonasymptotic convergence rate of the LRMC estimator under the Q-risk. Then, we give a detailed description of AETC algorithm for vector-valued high-fidelity models, and examine the estimation efficiency in the exploration phase. We end this section by providing a numerical example where very high-dimensional high-fidelity output is considered.

For simplicity, we assume that noises $\varepsilon_S^{(1)}, \cdots, \varepsilon_S^{(k_0)}$ are jointly Gaussian; the sub-Gaussian case can be discussed similarly using analogous concentration inequalities.

D.1. Proof of Theorem 6.1. Let $r_S = \operatorname{rank}(\Gamma_S)$ and $Q \in \mathbb{R}^{q\times k_0}$. Similar to the calculation in (B.2), we first rewrite $\mathsf{Risk}_S|_{\widehat{\beta}_S}$ as

$$\begin{aligned}\mathsf{Risk}_S|_{\widehat{\beta}_S} &= \frac{1}{N_S}\operatorname{tr}\Big(\Sigma_S\widehat{\beta}_S^T Q^T Q\widehat{\beta}_S\Big) + x_S^T(\widehat{\beta}_S-\beta_S)^T Q^T Q(\widehat{\beta}_S-\beta_S)x_S \\ &\leq \frac{2}{N_S}\Big[\operatorname{tr}\big(\Sigma_S\beta_S^T Q^T Q\beta_S\big) + \operatorname{tr}\Big(\Sigma_S(\widehat{\beta}_S-\beta_S)^T Q^T Q(\widehat{\beta}_S-\beta_S)\Big)\Big] \\ &\quad + x_S^T(\widehat{\beta}_S-\beta_S)^T Q^T Q(\widehat{\beta}_S-\beta_S)x_S. \end{aligned} \tag{D.1}$$

Conditional on Z_S,

$$Q(\widehat{\beta}_S-\beta_S) = Q(\varepsilon_{S,1},\cdots,\varepsilon_{S,m})(Z_S^\dagger)^T, \tag{D.2}$$

where $\varepsilon_{S,\ell} \sim \mathcal{N}(0,\Gamma_S)$ are i.i.d. random vectors representing the noise vector in the ℓ-th exploration sample. Since $r_S = \operatorname{rank}(\Gamma_S)$, by the properties of multivariate normal distributions, there exists a $P_S \in \mathbb{R}^{k_0\times r_S}$ such that

$$\varepsilon_{S,\ell} = P_S\xi_{S,\ell} \qquad\qquad P_S P_S^T = \Gamma_S \tag{D.3}$$

where $\xi_{S,\ell}$ are i.i.d. normal distributions $\mathcal{N}(0, I_{r_S})$. Thus,

$$(\widehat{\beta}_S-\beta_S)^T Q^T Q(\widehat{\beta}_S-\beta_S) = Z_S^\dagger(\xi_{S,1},\cdots,\xi_{S,m})^T P_S^T Q^T Q P_S(\xi_{S,1},\cdots,\xi_{S,m})(Z_S^\dagger)^T. \tag{D.4}$$

Note that $(\xi_{S,1}, \cdots, \xi_{S,m})$ is a Gaussian random matrix, and multiplying it on the left by a unitary matrix yields a matrix whose rows are i.i.d. $\mathcal{N}(0, I_m)$. Thus, taking the eigendecomposition $P_S^T Q^T Q P_S = V \mathrm{diag}(\lambda_1, \cdots, \lambda_{r_S}) V^*$ and plugging it into (D.4) yields

$$(\widehat{\beta}_S - \beta_S)^T Q^T Q (\widehat{\beta}_S - \beta_S)^T = \sum_{i=1}^{r_S} \lambda_i Z_S^\dagger g_i g_i^T (Z_S^\dagger)^T \tag{D.5}$$

where g_i are i.i.d. $\mathcal{N}(0, I_m)$. Substituting (D.5) into (D.1) yields

$$\mathsf{Risk}_S|_{\widehat{\beta}_S} \leq \frac{2}{N_S} \left(\mathrm{tr}\big(\Sigma_S \beta_S^T Q^T Q \beta_S\big) + \|\Sigma_S\|_2 \sum_{i=1}^{r_S} \lambda_i \|Z_S^\dagger g_i\|_2^2 \right) + \sum_{i=1}^{r_S} \lambda_i \|B_S g_i\|_2^2,$$

where B_S is as defined in (B.2). For $\|Z_S^\dagger g_i\|_2^2$ and $\|B_S g_i\|_2^2$, the proof of Theorem 4.2 tells us that for large m, with probability at least $1 - m^{-2}$,

$$\|Z_S^\dagger g_i\|_2^2 \lesssim \frac{\log m}{m} = o(1) \qquad \|B_S g_i\|_2^2 \lesssim (s+1)\frac{\log m}{m} \qquad \forall i \in [r_S]. \tag{D.6}$$

The proof is complete by observing $\sum_{i=1}^{r_S} \lambda_i = \mathrm{tr}\big(P_S^T Q^T Q P_S\big) = \mathrm{tr}\big(Q \Gamma_S Q^T\big)$.

D.2. AETC for vector-valued high-fidelity models. We now describe the analog of Algorithm 5.1. The only modification occurs when we estimate the model variance structure. Instead of estimating the model variance σ_S^2 in (5.7), we need to estimate the model covariance matrix Γ_S, which can be computed as the sample covariance of the residuals:

$$\widehat{\Gamma}_S(t) = \frac{1}{t-s-1} \sum_{\ell \in [t]} \Big(f(Y_\ell) - \widehat{\beta}_S X_{\mathrm{epr},\ell}|_S \Big) \Big(f(Y_\ell) - \widehat{\beta}_S X_{\mathrm{epr},\ell}|_S \Big)^T. \tag{D.7}$$

Similar empirical estimates in (5.9) can be defined as

$$\begin{aligned} \tilde{k}_{1,t}(S) &= c_{\mathrm{ept}}(S)\, \mathrm{tr}\Big(\widehat{\Sigma}_S(t) \widehat{\beta}_S^T(t) Q^T Q \widehat{\beta}_S(t) \Big) \\ \tilde{k}_{2,t}(S) &= \mathrm{tr}\Big(Q \widehat{\Gamma}_S(t) Q^T \Big)\, \mathrm{tr}\Big(\widehat{x}_S(t) \widehat{x}_S^T(t) \widehat{\Lambda}_S^{-1}(t) \Big) \end{aligned} \tag{D.8}$$

and

$$\tilde{m}_S(t) = \frac{B}{c_{\mathrm{epr}} + \sqrt{\frac{c_{\mathrm{epr}} \tilde{k}_{1,t}(S)}{\tilde{k}_{2,t}(S)}}}, \quad \widehat{\overline{\mathsf{Risk}}}_S|_{\widehat{\beta}_S}(m;t) = \frac{\tilde{k}_{1,t}(S)}{B - c_{\mathrm{epr}} m} + \frac{\tilde{k}_{2,t}(S)}{m}, \tag{D.9}$$

$$\overline{\mathsf{Risk}}_S^*|_{\widehat{\beta}_S}(t) = \frac{\left(\sqrt{\tilde{k}_{1,t}(S)} + \sqrt{c_{\mathrm{epr}} \tilde{k}_{2,t}(S)} \right)^2}{B}. \tag{D.10}$$

The analog of Algorithm 5.1 for vector-valued high-fidelity models can be summarized as follows:

Algorithm D.1 AETC algorithm for multifidelity approximation (vector-valued case)

Input: B: total budget, c_i: cost parameters, $\alpha_t \downarrow 0$: regularization parameters
Output: $(\pi_t)_t$

compute the maximum exploration round $M = \lfloor B/c_{\mathrm{epr}} \rfloor$
for $t \in [n+2]$ **do**
  $\pi_t = a_{\mathrm{epr}}([n])$
end for
while $n+2 \le t \le M$ **do**
  for $S \subseteq [n]$ **do**
    compute $\tilde{k}_{1,t}(S), \tilde{k}_{2,t}(S)$ using (5.7), (D.7) and (D.8), and set $\tilde{k}_{2,t}(S) \leftarrow \tilde{k}_{2,t}(S) + \alpha_t$
    compute $\tilde{m}_S(t)$ using (D.9)
    compute the optimal Q-risk using (D.9): $h_S(t) = \widehat{\overline{\mathsf{Risk}}}_S|_{\widehat{\beta}_S}(\tilde{m}_S(t) \vee t; t)$
  end for
  find the optimal model $S^*(t) = \arg\min_{S \subseteq [n]} h_S(t)$
  if $\tilde{m}_{S^*(t)}(t) > t$ **then**
    $\pi_{t+1} = a_{\mathrm{epr}}([n])$ and $t \leftarrow t+1$
  else
    $\pi_{t+1} = a_{\mathrm{ept}}(S^*(t))$ and $t \leftarrow M+1$
  end if
end while

Applying a similar argument as the proof of Theorem 5.3, one can show that the exploitation model produced by Algorithm D.1 converges almost surely to

$$\tilde{S}^* = \arg\min_{S \subseteq [n]} \left(\sqrt{\tilde{k}_{1,t}(S)} + \sqrt{c_{\mathrm{epr}} \tilde{k}_{2,t}(S)} \right)^2 \tag{D.11}$$

with optimal exploration as $B \to \infty$, where

$$\tilde{k}_1(S) = c_{\mathrm{ept}}(S) \operatorname{tr}\big(\Sigma_S(t) \beta_S^T(t) Q^T Q \beta_S(t)\big)$$
$$\tilde{k}_2(S) = \operatorname{tr}\big(Q \Gamma_S(t) Q^T\big) \operatorname{tr}\big(x_S(t) x_S^T(t) \Lambda_S^{-1}(t)\big).$$

D.3. Efficient estimation. For large k_0, both β_S and Γ_S are high-dimensional, which may not admit a good global estimation for small exploration rate m. However, the parameters $\tilde{k}_1(S)$ and $\tilde{k}_2(S)$ used in decision-making are scalar-valued and only involve marginals of the high-dimensional parameters. In the following theorems, we will justify that relative accuracy of the plug-in estimators for both quantities defined in (D.8) is independent of k_0 under suitable assumptions.

THEOREM D.1. *Under the same condition as Theorem 6.1 and for large* $t > \max\{s+1, 5\}$, *it holds with probability at least* $1 - t^{-2}$ *that*

$$\frac{\left| \operatorname{tr}\left(Q \widehat{\Gamma}_S(t) Q^T\right) - \operatorname{tr}\big(Q \Gamma_S Q^T\big) \right|}{\operatorname{tr}(Q \Gamma_S Q^T)} \lesssim \frac{\log t}{\sqrt{t-s-1}}, \tag{D.12}$$

where the implicit constant in $\lesssim$ *is universal.*

For the estimation of $\operatorname{tr}\big(Q \beta_S \beta_S^T Q\big)$, for convenience we consider a model S without intercept:

THEOREM D.2. *Assume that (4.1) holds for the 2-Orlicz norm, and*

$$\frac{\operatorname{tr}\big(Q\Gamma_S Q^T\big)}{\operatorname{tr}\big(Q\beta_S\beta_S^T Q\big)} \lesssim \mathcal{O}(1). \tag{D.13}$$

Suppose that S does not include the intercept term and the corresponding covariance matrix Σ_S is nonsingular. Under the same condition as Theorem 6.1 and for large t, it holds with probability at least $1-t^{-2}$ that

$$\frac{\Big|\operatorname{tr}\Big(\widehat{\Sigma}_S(t)\widehat{\beta}_S^T(t)Q^TQ\widehat{\beta}_S(t)\Big) - \operatorname{tr}\big(\Sigma_S\beta_S^TQ^TQ\beta_S\big)\Big|}{\operatorname{tr}\big(\Sigma_S\beta_S^TQ^TQ\beta_S\big)} \lesssim \kappa(\Sigma_S)\sqrt{\frac{\log t}{t}}, \tag{D.14}$$

where $\kappa(\Sigma_S)$ is the condition number of Σ_S, and the implicit constant is independent of t and k_0.

REMARK D.3. *Under additional assumption on the Frobenius norm of $Q\beta_S$ and its submatrices, similar results can be obtained for S containing the intercept, with $\kappa(\Sigma_S)$ replaced by $\kappa(M_{1,1}(\Sigma_S))$, where $M_{i,j}(\cdot)$ denotes the (i,j) minor of a matrix.*

REMARK D.4. *The constant $\kappa(\Sigma_S)$ in (D.14), despite depending on the singular values of Σ_S, is independent of k_0. This combined with (D.12) implies that the estimation in Algorithm D.1 is relatively efficient regardless of the response dimension k_0.*

Proof of Theorem D.1. Rewriting (D.7) using (D.2) and (6.1), we have

$$\begin{aligned} Q\widehat{\Gamma}_S(t)Q^T &= \frac{1}{t-s-1}Q(\varepsilon_{S,1},\cdots,\varepsilon_{S,t})(I_t - Z_SZ_S^\dagger)(\varepsilon_{S,1},\cdots,\varepsilon_{S,t})^TQ^T \\ &\overset{(D.3)}{=} \frac{1}{t-s-1}QP_S(\xi_{S,1},\cdots,\xi_{S,t})(I_t - Z_SZ_S^\dagger)(\xi_{S,1},\cdots,\xi_{S,t})^TP_S^TQ^T. \end{aligned} \tag{D.15}$$

Note that $I_t - Z_SZ_S^\dagger$ is a $(t-s-1)$-dimensional (random) orthogonal projection matrix, i.e., $I_t - Z_SZ_S^\dagger = UU^T$ for some $U \in \mathbb{R}^{t\times(t-s-1)}$ with orthonormal columns. Thus, by the rotational invariance of multivariate normal distributions,

$$\begin{aligned} &(\xi_{S,1},\cdots,\xi_{S,t})(I_t - Z_SZ_S^\dagger)(\xi_{S,1},\cdots,\xi_{S,t})^T \\ &\overset{\mathcal{D}}{=} (\zeta_{S,1},\cdots,\zeta_{S,t-s-1})(\zeta_{S,1},\cdots,\zeta_{S,t-s-1})^T, \end{aligned} \tag{D.16}$$

where $\zeta_{S,\ell}, \ell \in [t-s-1]$ are (fixed) i.i.d. normal vectors $\mathcal{N}(0, I_{r_S})$, and $\overset{\mathcal{D}}{=}$ denotes equality in distribution. Substituting (D.16) into (D.15) and taking the trace yields

$$\begin{aligned} \operatorname{tr}\Big(Q\widehat{\Gamma}_S(t)Q^T\Big) &\overset{\mathcal{D}}{=} \frac{1}{t-s-1}\operatorname{tr}\big(QP_S(\zeta_{S,1},\cdots,\zeta_{S,t-s-1})(\zeta_{S,1},\cdots,\zeta_{S,t-s-1})^TP_S^TQ^T\big) \\ &= \frac{1}{t-s-1}\operatorname{tr}\big((\zeta_{S,1},\cdots,\zeta_{S,t-s-1})^TP_S^TQ^TQP_S(\zeta_{S,1},\cdots,\zeta_{S,t-s-1})\big) \\ &= \frac{1}{t-s-1}\sum_{i=1}^{t-s-1}\zeta_{S,i}^TP_S^TQ^TQP_S\zeta_{S,i} \\ &= \frac{1}{t-s-1}\boldsymbol{\zeta}_{S,t-s-1}^T\boldsymbol{\Omega}_{t-s-1}\boldsymbol{\zeta}_{S,t-s-1} \end{aligned}$$

where $\boldsymbol{\zeta}_{S,t-s-1} = (\zeta_{S,1}^T, \cdots, \zeta_{S,t-s-1}^T)^T$ is a standard multivariate normal vector in $\mathbb{R}^{r_S(t-s-1)}$ and

$$\boldsymbol{\Omega}_{t-s-1} = \underbrace{\begin{pmatrix} P_S^T Q^T Q P_S & & \\ & \ddots & \\ & & P_S^T Q^T Q P_S \end{pmatrix}}_{t-s-1 \text{ diagonal blocks}} \tag{D.17}$$

Apply the Hanson-Wright inequality [59, Theorem 6.2.1] to $\boldsymbol{\zeta}_{S,t-s-1}^T \boldsymbol{\Omega}_{t-s-1} \boldsymbol{\zeta}_{S,t-s-1}$ and we yield that for every $t > s+1$ and $\delta \geq 4\|\boldsymbol{\Omega}_{t-s-1}\|_F$,

$$\begin{aligned} &\mathbb{P}\left(|\boldsymbol{\zeta}_{S,t-s-1}^T \boldsymbol{\Omega}_{t-s-1} \boldsymbol{\zeta}_{S,t-s-1} - (t-s-1)\operatorname{tr}\big(P_S^T Q^T Q P_S\big)| > \delta\right) \\ &\quad \leq 2\exp\left(-\frac{C\delta}{4\|\boldsymbol{\Omega}_{t-s-1}\|_F}\right), \end{aligned}$$

where $C \leq 1$ is an absolute constant, and 4 comes from an upper bound for the square of the sub-Gaussian norm of the standard normal distribution. For $t > \max\{s+1, 5\}$, taking

$$\delta = C^{-1}\|\boldsymbol{\Omega}_{t-s-1}\|_F \log\big(2t^2\big)$$

yields that with probability at least $1 - t^{-2}$,

$$\left|\boldsymbol{\zeta}_{S,t-s-1}^T \boldsymbol{\Omega}_{t-s-1} \boldsymbol{\zeta}_{S,t-s-1} - (t-s-1)\operatorname{tr}\big(P_S^T Q^T Q P_S\big)\right| \leq C^{-1}\log\big(2t^2\big)\|\boldsymbol{\Omega}_{t-s-1}\|_F$$

Dividing both sides by $(t-s-1)\operatorname{tr}\big(Q\Gamma_S Q^T\big)$ and using $\operatorname{tr}\big(P_S^T Q^T Q P_S\big) = \operatorname{tr}\big(Q\Gamma_S Q^T\big)$ finishes the proof of (D.12):

$$\begin{aligned} \frac{\left|\operatorname{tr}\Big(Q\widehat{\Gamma}_S(t)Q^T\Big) - \operatorname{tr}\big(Q\Gamma_S Q^T\big)\right|}{\operatorname{tr}(Q\Gamma_S Q^T)} &\leq \frac{\|\boldsymbol{\Omega}_{t-s-1}\|_F}{\operatorname{tr}(Q\Gamma_S Q^T)\sqrt{t-s-1}} \frac{C^{-1}\log\big(2t^2\big)}{\sqrt{t-s-1}} \\ &= \frac{\|Q\Gamma_S Q^T\|_F}{\operatorname{tr}(Q\Gamma_S Q^T)} \frac{C^{-1}\log\big(2t^2\big)}{\sqrt{t-s-1}} \\ &\leq \frac{3C^{-1}\log t}{\sqrt{t-s-1}}. \end{aligned} \tag{D.18}$$

The proof is complete. ∎

Proof of Theorem D.2. By the triangle inequality and Cauchy-Schwarz inequality,

$$\begin{aligned} &\left|\operatorname{tr}\Big(\widehat{\Sigma}_S(t)\widehat{\beta}_S^T(t)Q^T Q\widehat{\beta}_S(t)\Big) - \operatorname{tr}\big(\Sigma_S\beta_S^T Q^T Q\beta_S\big)\right| \\ \leq\ &\left|\operatorname{tr}\Big(\widehat{\Sigma}_S(t)\widehat{\beta}_S^T(t)Q^T Q\widehat{\beta}_S(t)\Big) - \operatorname{tr}\Big(\widehat{\Sigma}_S(t)\beta_S^T Q^T Q\beta_S\Big)\right| \\ &+ \left|\operatorname{tr}\Big(\widehat{\Sigma}_S(t)\beta_S^T Q^T Q\beta_S\Big) - \operatorname{tr}\big(\Sigma_S\beta_S^T Q^T Q\beta_S\big)\right| \\ \leq\ &\|\widehat{\Sigma}_S(t)\|_F\|\widehat{\beta}_S^T(t)Q^T Q\widehat{\beta}_S(t) - \beta_S^T Q^T Q\beta_S\|_F + \|\widehat{\Sigma}_S(t) - \Sigma_S\|_F\|\beta_S^T Q^T Q\beta_S\|_F \\ \leq\ &\|\widehat{\Sigma}_S(t)\|_F\|\widehat{\beta}_S^T(t)Q^T Q\widehat{\beta}_S(t) - \beta_S^T Q^T Q\beta_S\|_F + \|\widehat{\Sigma}_S(t) - \Sigma_S\|_F \operatorname{tr}\big(\beta_S^T Q^T Q\beta_S\big) \\ \leq\ &\sqrt{s}\left(\|\widehat{\Sigma}_S(t)\|_2\|\widehat{\beta}_S^T(t)Q^T Q\widehat{\beta}_S(t) - \beta_S^T Q^T Q\beta_S\|_F + \|\widehat{\Sigma}_S(t) - \Sigma_S\|_2 \operatorname{tr}\big(\beta_S^T Q^T Q\beta_S\big)\right). \end{aligned} \tag{D.19}$$

Let a_i and b_i denote the i-th column of $Q\beta_S$ and $Q\widehat{\beta}_S(t)$, respectively, i.e., $Q\beta_S = (a_1, \cdots, a_s)$ and $Q\widehat{\beta}_S(t) = (b_1, \cdots, b_s)$. Then,

$$\begin{aligned}
&\|\widehat{\beta}_S^T(t)Q^TQ\widehat{\beta}_S(t) - \beta_S^TQ^TQ\beta_S\|_F^2 \\
&= \sum_{i,j\in[s]} \left(\langle a_i, a_j\rangle - \langle b_i, b_j\rangle\right)^2 \\
&= \sum_{i,j\in[s]} \left(\langle a_i - b_i, a_j\rangle + \langle b_i, a_j - b_j\rangle\right)^2 \\
&\le 2\sum_{i,j\in[s]} \left(\langle a_i - b_i, a_j\rangle^2 + \langle b_i, a_j - b_j\rangle^2\right) \\
&\le 2\sum_{i,j\in[s]} \left(\|a_i - b_i\|_2^2\|a_j\|_2^2 + \|b_i\|_2^2\|a_j - b_j\|_2^2\right) \\
&\le 2\sum_{i,j\in[s]} \left(\|a_i - b_i\|_2^2\|Q\beta_S\|_2^2 + \|Q\widehat{\beta}_S(t)\|_2^2\|a_j - b_j\|_2^2\right) \\
&\le 4s\left(\|Q\beta_S\|_2 + \|Q\widehat{\beta}_S(t)\|_2\right)^2\|Q\beta_S - Q\widehat{\beta}_S(t)\|_F^2.
\end{aligned}$$

Substituting this into (D.19) yields

$$\begin{aligned}
&\left|\operatorname{tr}\left(\widehat{\Sigma}_S(t)\widehat{\beta}_S^T(t)Q^TQ\widehat{\beta}_S(t)\right) - \operatorname{tr}\left(\Sigma_S\beta_S^TQ^TQ\beta_S\right)\right| \\
&\le 2s\|\widehat{\Sigma}_S(t)\|_2\left(\|Q\beta_S\|_2 + \|Q\widehat{\beta}_S(t)\|_2\right)\|Q\beta_S - Q\widehat{\beta}_S(t)\|_F \\
&\quad + \sqrt{s}\|\widehat{\Sigma}_S(t) - \Sigma_S\|_2\operatorname{tr}\left(\beta_S^TQ^TQ\beta_S\right). \qquad \text{(D.20)}
\end{aligned}$$

Under the strengthened assumption of (4.1), we can obtain a better bound for the relative error of the sample covariance estimator $\widehat{\Sigma}_S$ by appealing to [59, Exercise 9.2.5]: With probability at least $1 - 1/2t^2$, $\|\widehat{\Sigma}_S(t) - \Sigma_S\|_2/\|\Sigma_S\|_2 \lesssim \sqrt{\log t/t}$, i.e., $\|\Sigma_S\|_2 + \|\widehat{\Sigma}_S(t)\|_2 \lesssim \|\Sigma_S\|_2$. On the other hand,

$$\begin{aligned}
\|Q\beta_S - Q\widehat{\beta}_S(t)\|_F &= \sqrt{\operatorname{tr}\left((\widehat{\beta}_S(t) - \beta_S)^TQ^TQ(\widehat{\beta}_S(t) - \beta_S)\right)} \\
&\overset{\text{(D.5),(D.6)}}{\lesssim} \sqrt{\operatorname{tr}(Q\Gamma_SQ^T)}\sqrt{\frac{\log t}{t}}, \qquad \text{(D.21)}
\end{aligned}$$

where (D.21) holds with probability at least $1 - 1/2t^2$ for large t. In this case we can verify that,

$$\begin{aligned}
\|Q\beta_S\|_2 + \|Q\widehat{\beta}_S(t)\|_2 &\le 2\|Q\beta_S\|_F + \|Q\widehat{\beta}_S(t) - Q\beta_S\|_F \qquad \text{(D.22)} \\
&\lesssim \|Q\beta_S\|_F\left(2 + \sqrt{\frac{\operatorname{tr}(Q\Gamma_SQ^T)}{\operatorname{tr}\left(Q\beta_S\beta_S^TQ\right)}\frac{\log t}{t}}\right) \\
&\overset{\text{(D.13)}}{\lesssim} \|Q\beta_S\|_F = \sqrt{\operatorname{tr}(\beta_SQ^TQ\beta_S)}.
\end{aligned}$$

Thus, with probability at least $1 - t^{-2}$,

$$\begin{aligned}
&\left|\operatorname{tr}\left(\widehat{\Sigma}_S(t)\widehat{\beta}_S^T(t)Q^TQ\widehat{\beta}_S(t)\right) - \operatorname{tr}\left(\Sigma_S\beta_S^TQ^TQ\beta_S\right)\right| \\
&\lesssim \|\Sigma_S\|_2\sqrt{\operatorname{tr}(\beta_SQ^TQ\beta_S)}\sqrt{\operatorname{tr}(Q\Gamma_SQ^T)}\sqrt{\frac{\log t}{t}} + \operatorname{tr}\left(\beta_S^TQ^TQ\beta_S\right)\|\Sigma_S\|_2\sqrt{\frac{\log t}{t}} \\
&\overset{(D.13)}{\lesssim} \|\Sigma_S\|_2\operatorname{tr}\left(\beta_S^TQ^TQ\beta_S\right)\sqrt{\frac{\log t}{t}}. \qquad \text{(D.23)}
\end{aligned}$$

Dividing (D.23) by $\operatorname{tr}\left(\Sigma_S\beta_S^TQ^TQ\beta_S\right)$ finishes the proof:

$$\begin{aligned}
\frac{\left|\operatorname{tr}\left(\widehat{\Sigma}_S(t)\widehat{\beta}_S^T(t)Q^TQ\widehat{\beta}_S(t)\right) - \operatorname{tr}\left(\Sigma_S\beta_S^TQ^TQ\beta_S\right)\right|}{\operatorname{tr}\left(\Sigma_S\beta_S^TQ^TQ\beta_S\right)} &\lesssim \frac{\|\Sigma_S\|_2\operatorname{tr}\left(\beta_S^TQ^TQ\beta_S\right)}{\operatorname{tr}\left(\Sigma_S\beta_S^TQ^TQ\beta_S\right)}\sqrt{\frac{\log t}{t}} \\
&\le \kappa(\Sigma_S)\sqrt{\frac{\log t}{t}}. \qquad \square
\end{aligned}$$

E. Numerical simulation of Section 7.1. In Section 7.1, the partial differential equation (7.1) for a fixed $\boldsymbol{p}$ is frequently solved with the Finite Element Method. We use standard bilinear square isotropic finite elements on a rectangular mesh. After spatial discretization, this results in a linear system in the form of

$$\boldsymbol{K}\boldsymbol{u} = \boldsymbol{f} \qquad \text{(E.1)}$$

where $\boldsymbol{K}$ and $\boldsymbol{f}$ are stiffness matrix and force vector, respectively, and $\boldsymbol{u}$ is the vector of nodal displacements. Our output quantity of interest is the scalar compliance, defined as

$$\mathsf{cpl} = \boldsymbol{u}^T\boldsymbol{K}\boldsymbol{u}. \qquad \text{(E.2)}$$

In this example, we form a multifidelity hierarchy through coarsening of the discretization: Consider $n = 7$ mesh resolutions with mesh sizes $h = \{1/(2^{8-L})\}_{L=1}^7$ where L denotes the level. According to this hierarchy of meshes, the mesh associated with $L = 1$ yields the most accurate model (highest fidelity), which is taken as the high-fidelity model in our experiments. The Poisson's ratio is $\nu = 0.3$. The finite element computations in this paper are performed in MATLAB using part of the publicly available code for topology optimization [4].

To model the uncertainty, we consider a random field for the elastic modulus via the Karhunen-Loéve (KL) expansion (7.3), where $\delta = 0.5, E_0 = 1$ are constants. The random variables p_i are uniformly distributed on $[-1, 1]$ and the eigenvalues λ_i and basis functions E_i are taken from the analytical expressions for the eigenpairs of an exponential kernel on $D = [0, 1]^2$. In one dimension, i.e., $D_1 = [0, 1]$, the eigenpairs are

$$\lambda_i^{1D} = \frac{2}{w_i^2 + 1} \quad b_i^{1D} = A_i(\sin(w_i\boldsymbol{x}) + w_i\cos(w_i\boldsymbol{x})) \qquad i \in \mathbb{N} \qquad \text{(E.3)}$$

where w_i are the positive ordered solutions to

$$\tan(w) = \frac{2w}{w^2 - 1}, \qquad \text{(E.4)}$$

See, e.g., [54]. In our simulations, we use the approximation $w_i \approx i\pi$, which is the asymptotic behavior of these solutions. We generate two-dimensional $D = [0,1]^2$ eigenpairs via tensorization of the one-dimensional pairs,

$$\lambda_{ij} = \lambda_i^{1D}\lambda_j^{1D} \quad E_{ij} = b_i^{1D} \otimes b_j^{1D} \qquad i,j \in \mathbb{N} \tag{E.5}$$

where $\otimes$ denotes the tensor product. In this example, $d = 4$ corresponds to the tensor-product indices $(i,j) \in [2] \times [2]$. When $d = 20$, we consider the first 20 terms of the total-order polynomial indices $\{\{0,0\},\{10,01\},\{20,11,02\},\}$ To compute the elastic modulus for different resolutions, we evaluate the analytical basis functions in (E.3) at different resolutions.

F. Low-fidelity models of Section 7.2. Models $X^{(i)}$ for $i \in [6]$ in Section 7.2 are formed as GP predictors. We construct a GP emulator for C as a function of the parameters $\boldsymbol{p}$, and use the GP mean as a low-fidelity emulator. Given $\mathcal{D}^\circ = \{(\boldsymbol{p}_i^\circ, \boldsymbol{Y}(p_i^\circ))\}_{i=1}^{n_T}$ training samples with $\boldsymbol{P}^\circ = \{\boldsymbol{p}_i^\circ\}_{i=1}^{n_T}$ the sampling nodes and $\boldsymbol{Y}^\circ = \{\boldsymbol{Y}(p_i^\circ)\}_{i=1}^{n_T}$ the observational data, let $\boldsymbol{K}_{GP}$ denote a kernel matrix with entries $[\boldsymbol{K}_{GP}]_{ij} = k(\boldsymbol{p}_i, \boldsymbol{p}_j, \boldsymbol{\theta})$ for a given kernel function $k(\boldsymbol{p}, \boldsymbol{p}') : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is the kernel function and $\boldsymbol{\theta}$ is a hyperparameter that tunes the kernel. We adopt a standard GP training procedure, which determines $\boldsymbol{\theta}$ by maximizing a log-likelihood objective function,

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}) &= \log p(\boldsymbol{Y}|\boldsymbol{P}) = \log \mathcal{N}(\boldsymbol{Y}|\boldsymbol{0}, \boldsymbol{K}_{GP}(\boldsymbol{P},\boldsymbol{\theta}) + \lambda \boldsymbol{I}) \\ &= \frac{1}{2}\boldsymbol{Y}^T(\boldsymbol{K}_{GP} + \lambda \boldsymbol{I})^{-1}\boldsymbol{Y} + \frac{1}{2}\log|\boldsymbol{K}_{GP} + \lambda \boldsymbol{I}| + \frac{n}{2}\log(2\pi) \end{aligned} \tag{F.1}$$

where $\boldsymbol{I}$ is the identity matrix and $\lambda \geq 0$ is a constant. The parameter λ is introduced to model the effect of noise in the data, but in this example we do not consider any noise and therefore we set $\lambda = 0$. We choose an exponential kernel function,

$$k(\boldsymbol{p}, \boldsymbol{p}') = \exp(-\|\boldsymbol{p} - \boldsymbol{p}'\|_2/\theta) \tag{F.2}$$

with a scalar hyperparameter θ. Once the optimal hyperparameter θ^* is obtained from optimization of the log likelihood (F.1), we compute the approximate GP sample on the test node $\boldsymbol{p}^*$ via

$$X(p^*) = \bar{\boldsymbol{Y}}^\circ + \boldsymbol{K}_{i\circ}^T \boldsymbol{K}_{GP}^{-1}(\boldsymbol{Y}^\circ - \bar{\boldsymbol{Y}}^\circ), \tag{F.3}$$

where $\boldsymbol{K}_{i\circ} = k(\boldsymbol{p}^*, \boldsymbol{P}^\circ)$ is a vector obtained by evaluating the kernel function with the training samples $\boldsymbol{P}^\circ$ and the particular test sample $\boldsymbol{p}^*$, and $\bar{\boldsymbol{Y}}^\circ$ is the sample mean of the training data.

Our training data are generated via compliance samples from the high-fidelity model. The fidelity of each GP emulator in this case are determined by the number of training samples n_T as well as the optimal/nonoptimal choice of hyperparameters. We generate three fidelities by choosing $n_T = 10,\ 100,\ 1000$ training samples and compute the optimal hyperparameter θ via the minimization of (F.1). The approximated models from the optimal hyperparameter calculation are indexed by $X^{(1)}$, $X^{(3)}$, $X^{(5)}$ for $n_T = 10, 100, 1000$, respectively. We then generate three more models with the same number of training samples as before, but with a nonoptimal hyperparameter as $\theta^* \leftarrow 0.1\theta^*$. We index these nonoptimal models as $X^{(2)}$, $X^{(4)}$, $X^{(6)}$. The cost of all 6 of these models is given by the cost of training, optimization, and evaluation of (F.3) averaged over 2×10^4 different values of $\boldsymbol{p}^*$.

Models $X^{(i)}$ for $i = 7, \ldots, 12$ are defined via projection-based model reduction using proper orthogonal decomposition (POD). Let the matrix

$$\boldsymbol{U}^H = [\boldsymbol{u}^H(\boldsymbol{p}^{(1)})|\boldsymbol{u}^H(\boldsymbol{p}^{(2)})|\ldots|\boldsymbol{u}^H(\boldsymbol{p}^{(n)})] \in \mathbb{R}^{N_h \times N_s}$$

be comprised of high-fidelity nodal displacement vectors on parametric samples $\boldsymbol{P} = \{\boldsymbol{p}^{(i)}\}_{i=1}^{N_s}$. In this notation, and the current example, N_h is the number of high-fidelity finite element degrees of freedom and N_s is the number of parametric samples. We form emulators by projecting $\boldsymbol{u}^H$ onto k basis vectors collected as columns vectors into a matrix $\boldsymbol{V}_k$,

$$\boldsymbol{u}^H(\boldsymbol{p}^{(i)}) \approx \boldsymbol{V}_k \boldsymbol{w}(\boldsymbol{p}^{(i)}), \qquad \boldsymbol{V}_k \in \mathbb{R}^{N_h \times k}. \tag{F.4}$$

As in typical POD approaches, we choose $\boldsymbol{V}_k$ as the dominant k left-singular vector of $\boldsymbol{U}^H$. The POD coefficients $\boldsymbol{w}$ are computed as a Galerkin projection of the original high-fidelity finite element system,

$$(\boldsymbol{V}_k^T \boldsymbol{K} \boldsymbol{V}_k)\boldsymbol{w} \quad = \quad \boldsymbol{V}_k^T \boldsymbol{f}. \tag{F.5}$$

Using the notation $\boldsymbol{K}_k \equiv \boldsymbol{V}_k^T \boldsymbol{K} \boldsymbol{V}_k$, we compute an approximate compliance via $\mathsf{cpl} = \boldsymbol{U}^T \boldsymbol{K} \boldsymbol{U} \simeq \boldsymbol{w}^T \boldsymbol{K}_k \boldsymbol{w}$.

We generate six more models, $X^{(7)}, \ldots, X^{(12)}$, using the procedure above by making five choices for the reduced dimension k: $k = \{1, 2, 3, 4, 5, 10\}$. We ignore the cost of generating the matrix $\boldsymbol{U}^H$ and take as cost only the computation time for solving the linear system $\boldsymbol{K}_k \boldsymbol{w} = \boldsymbol{F}_k$ as well as the approximate compliance calculation $\mathsf{cpl} \simeq \boldsymbol{w}^T \boldsymbol{K}_k \boldsymbol{w}$ averaged 2×10^4 realizations of $\boldsymbol{p}$ as in the GP case.

G. A larger scale problem. Under the same setup as the numerical experiments in Section 7.1, we now apply Algorithm D.1 to a larger scale problem where the high-fidelity output is high-dimensional. Fix the budget as $B = 2 \times 10^5$, and let

$$f(Y) = \bar{\boldsymbol{u}}^{(0)} \in \mathbb{R}_+^{2601} \qquad Q = I_{2601}.$$

In this case, the number of affordable samples by the (high-fidelity) MC estimator is $\lfloor B/4096 \rfloor = 48$. We compare AETC and MC algorithms for approximating $\mathbb{E}[f(Y)]$. In our experiment, the number of exploration rounds m chosen by the AETC algorithm in the case of the square domain and the L-shape domain is 26 and 12, respectively. The respective exploitation models are $f(Y) \sim X^{(4)} + X^{(5)} + X^{(6)} + \text{intercept}$ and $f(Y) \sim X^{(3)} + X^{(4)} + X^{(5)} + X^{(6)} + \text{intercept}$, and the affordable exploitation samples for the chosen model are 2762 and 1581, respectively. $\frac{\mathrm{tr}\left(Q\Gamma_S Q^T\right)}{\mathrm{tr}\left(Q\beta_S\beta_S^T Q\right)}$ in both cases are of order 10^{-6}, which satisfies the assumption (D.13). The condition number of the $(1,1)$ minor of Σ_S are 712.03 and 6486.57 for the square and the L-shape domain, respectively. Although these relatively large quantities appear in the worst-case estimate (D.14), we observe that the performance of the algorithm in practice is more accurate than these estimates suggest.

We investigate exploitation error conditioned on the exploration above: We apply the trained model (the model selected during exploration) for exploitation 500 times, and for each estimate we compute the total error (computed in relation to a ground truth MC run over 50000 independent high-fidelity samples, which is visualized in Figure 9), i.e., the squared ℓ_2 norm of the difference between $\mathbb{E}[f(Y)]$ and its estimate. We then find the 0.05-0.5-0.95 quantiles of this scalar total error, and plot the spatial

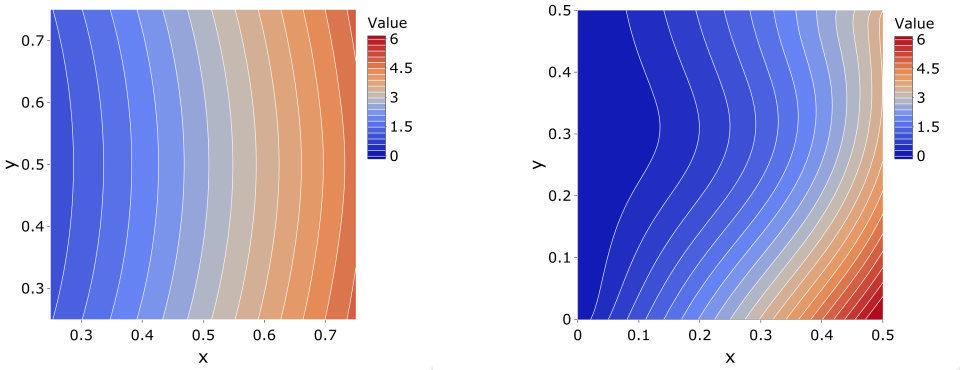

FIG. 9. *Visualization of the ground truth* $\mathbb{E}[f(Y)]$ *computed by the MC over* 50000 *independent samples in the case of the square domain (**Left**) and the L-shape domain (**Right**).*

pointwise error at the realization identified by the total error quantiles. The results are compared to the budget-B MC estimator in Figure 10. Within the uncertainty, the AETC algorithm enjoys a superior performance overall, on both geometries.

REFERENCES

[1] N. M. ALEXANDROV, J. J. E. DENNIS, R. M. LEWIS, AND V. TORCZON, *A trust-region framework for managing the use of approximation models in optimization*, Structural Optimization, 15 (1998), pp. 16–23.

[2] N. M. ALEXANDROV, R. M. LEWIS, C. R. GUMBERT, L. L. GREEN, AND P. A. NEWMAN, *Approximation and model management in aerodynamic optimization with variable-fidelity models*, Journal of Aircraft, 38 (2001), pp. 1093–1101.

[3] V. ANANTHARAM, P. VARAIYA, AND J. WALRAND, *Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays-part ii: Markovian rewards*, IEEE Transactions on Automatic Control, 32 (1987), pp. 977–982.

[4] E. ANDREASSEN, A. CLAUSEN, M. SCHEVENELS, B. S. LAZAROV, AND O. SIGMUND, *Efficient topology optimization in matlab using 88 lines of code*, Structural and Multidisciplinary Optimization, 43 (2011), pp. 1–16.

[5] J.-Y. AUDIBERT AND S. BUBECK, *Best arm identification in multi-armed bandits*, in COLT-23th Conference on Learning Theory-2010, 2010, pp. 13–p.

[6] P. AUER, N. CESA-BIANCHI, AND P. FISCHER, *Finite-time analysis of the multiarmed bandit problem*, Machine learning, 47 (2002), pp. 235–256.

[7] P. AUER, N. CESA-BIANCHI, Y. FREUND, AND R. E. SCHAPIRE, *Gambling in a rigged casino: The adversarial multi-armed bandit problem*, in Proceedings of IEEE 36th Annual Foundations of Computer Science, IEEE, 1995, pp. 322–331.

[8] P. AUER, N. CESA-BIANCHI, Y. FREUND, AND R. E. SCHAPIRE, *The nonstochastic multiarmed bandit problem*, SIAM Journal on Computing, 32 (2002), pp. 48–77, https://doi.org/10.1137/s0097539701398375, https://doi.org/10.1137%2Fs0097539701398375.

[9] P. AUER AND R. ORTNER, *UCB revisited: Improved regret bounds for the stochastic multi-armed bandit problem*, Periodica Mathematica Hungarica, 61 (2010), pp. 55–65, https://doi.org/10.1007/s10998-010-3055-6, https://doi.org/10.1007%2Fs10998-010-3055-6.

[10] P. BENNER, S. GUGERCIN, AND K. WILLCOX, *A survey of projection-based model reduction methods for parametric dynamical systems*, SIAM Review, 57 (2015), pp. 483–531.

[11] D. BOUNEFFOUF AND I. RISH, *A survey on practical applications of multi-armed and contextual bandits*, arXiv preprint arXiv:1904.10040, (2019).

[12] P. BRATLEY, B. L. FOX, AND L. E. SCHRAGE, *A Guide to Simulation*, Springer, 1987.

[13] S. BUBECK, N. CESA-BIANCHI, ET AL., *Regret analysis of stochastic and nonstochastic multi-armed bandit problems*, Foundations and Trends® in Machine Learning, 5 (2012), pp. 1–122.

[14] S. BUBECK, N. CESA-BIANCHI, AND S. M. KAKADE, *Towards minimax policies for online linear*

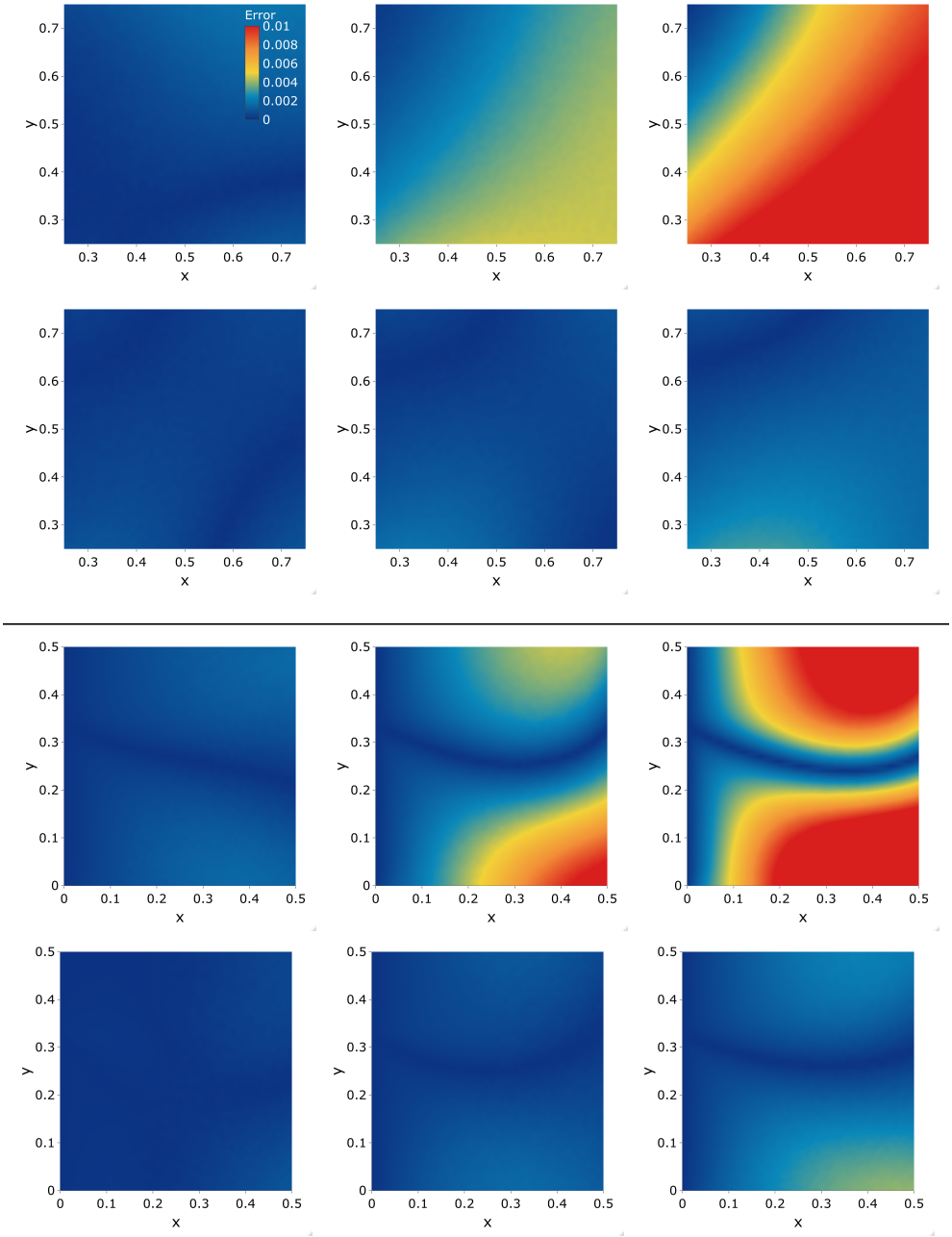

FIG. 10. *Error comparison of the estimated* $\mathbb{E}[f(Y)]$ *given by the MC and the AETC algorithms. Top two rows: square domain. Bottom two rows: L-shape domain. Within each two-row block, the top row corresponds to budget-B MC, and the bottom row to budget-B AETC. (Left, middle, right) columns: pointwise spatial errors corresponding to* $(0.05, 0.5, 0.95)$ *quantiles of the total scalar* ℓ_2 *error. The color limits are uniform for every plot, and are quantified in the top-left plot.*

optimization with bandit feedback, in Conference on Learning Theory, JMLR Workshop and Conference Proceedings, 2012, pp. 41–1.

[15] S. Bubeck, R. Munos, and G. Stoltz, *Pure exploration in multi-armed bandits problems*, in International conference on Algorithmic learning theory, Springer, 2009, pp. 23–37.

[16] C.-C. Chang and C.-J. Lin, *Libsvm: A library for support vector machines*, ACM Trans. Intell. Syst. Technol., 2 (2011), pp. 1–27.

[17] J. A. Christen and C. Fox, *Markov chain monte carlo using an approximation*, J. Comput. Graph. Statist., 14 (2005), pp. 795–810.

[18] C. Cortes and V. Vapnik, *Support-vector networks*, Machine Learning, 20 (1995), pp. 273–297.

[19] V. Dani, T. P. Hayes, and S. M. Kakade, *Stochastic linear optimization under bandit feedback*, (2008).

[20] M. Eldred, A. Giunta, and S. Collis, *Second-order corrections for surrogate-based optimization with model hierarchies*, 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Multidisciplinary Analysis Optimization Conferences, (2004), pp. A550–A591.

[21] A. I. J. Forrester and A. J. Keane, *Recent advances in surrogate-based optimization*, Progr. Aerospace Sci., 36 (2009), pp. 50–79.

[22] A. I. J. Forrester, A. Sobester, and A. J. Keane, *Engineering Design via Surrogate Modelling: A Practical Guide*, Wiley, 2008.

[23] C. Fox and G. Nicholls, *Sampling conductivity images via mcmc*, The Art and Science of Bayesian Image Analysis, University of Leeds, 14 (1997), pp. 91–100.

[24] W. A. Fuller, *Sampling statistics*, vol. 560, John Wiley & Sons, 2011.

[25] A. Garivier and E. Kaufmann, *Optimal best arm identification with fixed confidence*, in Conference on Learning Theory, PMLR, 2016, pp. 998–1027.

[26] M. B. Giles, *Multilevel monte carlo path simulation*, Operations research, 56 (2008), pp. 607–617.

[27] M. B. Giles, *Multilevel monte carlo methods.*, Acta Numer., 24 (2015), pp. 259–328.

[28] J. C. Gittins, *Bandit processes and dynamic allocation indices*, Journal of the Royal Statistical Society: Series B (Methodological), 41 (1979), pp. 148–164.

[29] A. A. Gorodetsky, G. Geraci, M. S. Eldred, and J. D. Jakeman, *A generalized approximate control variate framework for multifidelity uncertainty quantification*, Journal of Computational Physics, 408 (2020), p. 109257, https://doi.org/10.1016/j.jcp.2020.109257, https://doi.org/10.1016%2Fj.jcp.2020.109257.

[30] A. A. Gorodetsky, J. D. Jakeman, G. Geraci, and M. S. Eldred, *MFNets: MULTI-FIDELITY DATA-DRIVEN NETWORKS FOR BAYESIAN LEARNING AND PREDICTION*, International Journal for Uncertainty Quantification, 10 (2020), https://doi.org/10.1615/Int.J.UncertaintyQuantification.2020032978.

[31] S. Gugercin, A. C. Antoulas, and C. Beattie, *H-2 model reduction for large-scale linear dynamical systems*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 609–638.

[32] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*, Methuen, London, 1964.

[33] J. S. Hesthaven, G. Rozza, and B. Stamm, *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*, Springer, 2016.

[34] T. Jaksch, R. Ortner, and P. Auer, *Near-optimal regret bounds for reinforcement learning.*, Journal of Machine Learning Research, 11 (2010).

[35] K. Kandasamy, G. Dasarathy, J. Schneider, and B. Poczos, *The multi-fidelity multi-armed bandit*, arXiv preprint arXiv:1610.09726, (2016).

[36] P.-S. Koutsourelakis, *Accurate uncertainty quantification using inaccurate computational models*, SIAM Journal on Scientific Computing, 31 (2009), pp. 3274–3300.

[37] T. Lai and H. Robbins, *Asymptotically efficient adaptive allocation rules*, Advances in Applied Mathematics, 6 (1985), pp. 4–22, https://doi.org/10.1016/0196-8858(85)90002-8, https://doi.org/10.1016%2F0196-8858%2885%2990002-8.

[38] T. L. Lai and C. Z. Wei, *Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems*, The Annals of Statistics, 10 (1982), pp. 154–166, https://doi.org/10.1214/aos/1176345697, https://doi.org/10.1214%2Faos%2F1176345697.

[39] T. Lattimore and C. Szepesvári, *Bandit algorithms*, Cambridge University Press, 2020.

[40] A. J. Majda and B. Gershgorin, *Quantifying uncertainty in climate change science through empirical information theory*, Proc. Natl. Acad. Sci. USA, 107 (2010), pp. 14958–14963.

[41] S. Mendelson and A. Pajor, *On singular values of matrices with independent rows*, Bernoulli, 12 (2006), pp. 761–773, https://doi.org/10.3150/bj/1161614945, https://doi.org/10.3150%2Fbj%2F1161614945.

[42] A. Narayan, C. Gittelson, and D. Xiu, *A stochastic collocation algorithm with multifidelity models*, SIAM Journal on Scientific Computing, 36 (2014), pp. A495–A521.
[43] B. L. Nelson, *On control variate estimators*, Computers & Operations Research, 14 (1987), pp. 219–225.
[44] L. W. Ng and K. Willcox, *Monte-carlo information-reuse approach to aircraft conceptual design optimization under uncertainty*, J. Aircraft, 53 (2016), pp. 427–438.
[45] B. Peherstorfer, D. Pflüger, and H.-J. Bungartz, *Density estimation with adaptive sparse grids for large data sets*, Proceedings of the 2014 SIAM International Conference on Data Mining, 36 (2014), pp. 443–451.
[46] B. Peherstorfer, K. Willcox, and M. Gunzburger, *Optimal model management for multifidelity monte carlo estimation*, SIAM Journal on Scientific Computing, 38 (2016), pp. A3163–A3194, https://doi.org/10.1137/15m1046472, https://doi.org/10.1137%2F15m1046472.
[47] B. Peherstorfer, K. Willcox, and M. Gunzburger, *Survey of multifidelity methods in uncertainty propagation, inference, and optimization*, SIAM Review, 60 (2018), pp. A550–A591.
[48] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, Massachusetts Institute of Technology, 2006.
[49] G. Rozza, D. B. P. Huynh, and A. T. Patera, *Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations*, Arch. Comput. Methods Engrg., 15 (2008), pp. 229–275.
[50] P. Rusmevichientong and J. N. Tsitsiklis, *Linearly parameterized bandits*, Mathematics of Operations Research, 35 (2010), pp. 395–411, https://doi.org/10.1287/moor.1100.0446, https://doi.org/10.1287%2Fmoor.1100.0446.
[51] D. Schaden and E. Ullmann, *Asymptotic analysis of multilevel best linear unbiased estimators*, arXiv preprint arXiv:2012.03658, (2020).
[52] D. Schaden and E. Ullmann, *On multilevel best linear unbiased estimators*, SIAM/ASA Journal on Uncertainty Quantification, 8 (2020), pp. 601–635, https://doi.org/10.1137/19m1263534, https://doi.org/10.1137%2F19m1263534.
[53] L. Sirovich, *Turbulence and the dynamics of coherent structures*, Quart. Appl. Math., 45 (1987), pp. 561–571.
[54] A. L. Teckentrup, P. Jantsch, C. G. Webster, and M. Gunzburger, *A multilevel stochastic collocation method for partial differential equations with random input data*, SIAM/ASA Journal on Uncertainty Quantification, 3 (2015), pp. 1046–1074.
[55] W. R. Thompson, *On the likelihood that one unknown probability exceeds another in view of the evidence of two samples*, Biometrika, 25 (1933), p. 285, https://doi.org/10.2307/2332286, https://doi.org/10.2307%2F2332286.
[56] L. Tran-Thanh, A. Chapman, J. E. Munoz De Cote Flores Luna, A. Rogers, and N. R. Jennings, *Epsilon–first policies for budget–limited multi-armed bandits*, (2010).
[57] L. Tran-Thanh, A. Chapman, A. Rogers, and N. Jennings, *Knapsack based optimal policies for budget–limited multi–armed bandits*, in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 26, 2012.
[58] V. Vapnik, *Statistical Learning Theory*, Wiley, 1998.
[59] R. Vershynin, *High-dimensional probability: An introduction with applications in data science*, vol. 47, Cambridge university press, 2018.
[60] Y. Xu and A. Narayan, *Budget-limited distribution learning in multifidelity problems*, arXiv preprint arXiv:2105.04599, (2021).
[61] Y. Zhang, N. H. Kim, C. Park, and R. T. Haftka, *Multifidelity surrogate based on single linear regression*, AIAA Journal, 56 (2018), pp. 4944–4952, https://doi.org/10.2514/1.j057299, https://doi.org/10.2514%2F1.j057299.
[62] Q. Zhao, *Multi-armed bandits: Theory and applications to online learning in networks*, Synthesis Lectures on Communication Networks, 12 (2019), pp. 1–165.