Rare Event Detection by Acquisition-Guided Sampling

Huiling Liao, Xiaoning Qian, *Senior Member, IEEE*, Jianhua Z. Huang, and Peng Li, *Fellow, IEEE*

***Abstract*— Motivated by the challenges in detecting extremely rare failures for sophisticated specifications in circuit design, we consider the problem of detecting regions of interest (ROIs) that consist of specifications with the value of a complex target function for the system performance being below or above a certain pre-specified threshold. Though Bayesian optimization (BO) has been applied to this problem, it is not effective in identifying multiple ROIs as it was originally designed for global optimization and tends to focus on searching the area where the global optimum is most likely to be. In this work, we propose a sampling strategy for fast ROI detection within a limited number of target function evaluations. The sampling distribution is designed so that the probability of a specification being sampled is proportional to the corresponding value of the acquisition function. Such an acquisition-guided sampling algorithm promotes a wider search of the sample space and a simpler incorporation of different criteria to determine the specifications to be evaluated next. To further improve the performance, we propose a new design of the acquisition function and two modifications of existing acquisition functions. Numerical studies on synthetic functions and a real-world circuit design application demonstrate that the proposed method can enjoy a stronger exploration ability provided by sampling and achieve faster ROI detection with higher coverage.**

***Note to Practitioners*—This study considers the extremely rare failure detection in automated circuit design, fabrication, packaging, and verification. Obtaining enough observations of interest within a given budget of evaluations is challenging due to the scarcity of extremely rare failures. Bayesian optimization (BO) has been adopted to tackle this problem, but it may not achieve satisfactory coverage of multiple failure regions, since its goal is to find the global optimum of the target function. In this paper, we propose a sampling-based rare event detection strategy tailored to efficiently detect regions of interest (ROIs) with high coverage, along with newly designed acquisition functions incorporating the pre-specified threshold and ideas of experimental design. This sampling strategy has greater robustness to the choice of acquisition function. Also, since multiple queries can be easily obtained through sampling, various criteria can be easily incorporated for determining the next batch of evaluation specifications without much increase in computational complexity.**

***Index Terms*— Sequential design of experiment, Bayesian optimization, sampling, region detection.**

Received 5 May 2024; revised 16 August 2024; accepted 27 September 2024. Date of publication 16 October 2024; date of current version 21 March 2025. This article was recommended for publication by Associate Editor D. Wang and Editor T. Nishi upon evaluation of the reviewers' comments. This work was supported in part by the National Science Foundation (NSF) under Award 1553281, Award 1956313, Award 2119103, and Award 2215573. *(Corresponding author: Huiling Liao.)*

Huiling Liao is with the Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616 USA (e-mail: hliao13@iit.edu).

Xiaoning Qian is with the Department of Electrical and Computer Engineering and the Department of Computer Science and Engineering, Texas A&M University, College Station, TX 77840 USA, and also with the Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11973 USA (e-mail: xqian@ece.tamu.edu).

Jianhua Z. Huang is with the School of Data Science, The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China (e-mail: jhuang@cuhk.edu.cn).

Peng Li is with the Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106 USA (e-mail: lip@ece.ucsb.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/TASE.2024.3475951, provided by the authors.

Digital Object Identifier 10.1109/TASE.2024.3475951

I. Introduction

RARE event detection in complex systems, natural or human-engineered, is challenging due to the immense search space and complex behavior. For example, automated fabrication and packaging of electronics require delicate specifications and protocol designs with high reliability. Another example is circuit verification, an essential part of the process to ensure the quality of circuit design, where the techniques of rare event detection are needed to make sure that the design meets all the requirements of interest while having stable high performance and extremely low failure rate.

One research area that is related to the rare event detection problem is failure rate estimation, for which several statistical methods have been proposed (e.g. MixIS [1], SSS [2], Statistical Blockade [3]). Although they are useful in certain industry practice, assumptions imposed in these methods for an accurate estimation of the failure rate can be unrealistic, especially when the failure rate is extremely low, since a large number of evaluations are required while the evaluation is usually expensive.

Another related area is structural reliability analysis of complex engineering systems, where the failure probability of a system is estimated to assess the effects of various uncertainties which may arise from natural variability, operating conditions, or simply because of an incomplete or lack of knowledge [4]. To improve the performance of the analytical approach (e.g. FORM and SORM [5], [6]), active learning has been adopted and active learning reliability (ALR) has surged to solve complicated structure reliability problems at affordable cost with the help of a surrogate model for a limit state function (e.g. [7], [8], [9], [10]).

The other related research area is boundary detection given a threshold or accurate set estimation. One of such approaches, known as *excursion set*, considers the exceedance probability defined as $P(\sup_{\mathbf{x}\in\Omega} f(\mathbf{x}) \geq T)$ for a given threshold T of the target function $f(\mathbf{x})$ with the corresponding excursion sets defined as $\Gamma_T(f, \Omega) = \{\mathbf{x} \in \Omega : f(x) \geq T\}$. For a more detailed introduction to the excursion set, readers can refer to [11]. However, most existing studies aim for the accurate estimation of excursion sets (e.g. [12], [13], [14]) and uncertainty quantification (e.g. [15]), both necessarily requiring a large sample size and with goals diverge from ours.

Different from these works with goals of accurate failure probability estimation, prediction model estimation, or the most probable point of failure localization, the current study deals with the situation of extremely rare events when an accurate estimation of failure rate or set recovery is impossible given a realistic number of observations. We focus on methods that are capable of achieving fast detection and high coverage of ROIs within a limited budget, in response to the urgent demand from circuit verification.

Bayesian optimization (BO), a sequential design strategy for global optimization of black-box functions that are expensive to evaluate, has been adopted for the failure detection problem in the context of analog and mixed-signal (AMS) design verification in [16] and [17]. The authors tackled the failure detection problem by comparing the global optimum found by BO with a pre-specified threshold value for inference. They also developed a new acquisition function named pBOHC designed for better exploration and exploitation trade-offs to leverage distance between specifications to encourage the widespread of locations to be evaluated.

It is reasonable to expect that, after detecting a specification of interest (SOI), the algorithm should shift to exploring alternative potential regions to detect as many ROIs as possible rather than persistently evaluating its neighbor to find the optimum. As BO tends to search the area where the global optimum is most likely to be, it usually has a lower coverage rate and slower detection of multiple ROIs for multi-modal scenarios. Consequently, direct application of BO is not the most suitable approach for our formulated ROI detection problem.

In this work, we develop a sampling strategy tailored for fast detection and high coverage of ROIs within a given budget of target function evaluations, with the help of customized acquisition functions. The main contributions of this work can be summarized as follows:

1) We propose the acquisition-guided sampling (AcqS) algorithm to naturally achieve the exploration and exploitation trade-off for the ROI detection problem, where the sampling distribution is suitably designed based on customized acquisition functions.
2) We develop and evaluate new acquisition functions tailored for fast detection and maximal coverage of ROIs.
3) Hamiltonian Monte Carlo (HMC) sampling strategy is adopted for efficient acquisition-guided sampling.
4) Add-on modules including pre-screening and subset selection are designed to encourage the diversity of selected suggestions over the whole feasible space while avoiding redundant exploration.

This article is organized as follows. In Section II, we provide the formulation of the problem and a brief overview of Bayesian optimization. The methodology development is discussed in Section IV, followed by numerical studies of synthetic test functions and a real-world circuit design application in Section V and VI, respectively. Finally, a discussion of the proposed method and future research is given in Section VII.

II. Problem Formulation

Let $\mathbf{x}$ denote a D-dimensional vector representing the specification in the input space $\Omega \subset \mathbb{R}^D$, and $y(\mathbf{x})$ be the corresponding (possibly noisy) response. Denote the set of initial observations of size n_0 as $\mathcal{D}_0 = \{(\mathbf{x}_i, y_i); i = 1, \ldots, n_0\}$ and $T \in \mathbb{R}$ is the pre-specified threshold. We further define the specification of interest (SOI) $\mathbf{x}^*$ as a specification satisfying:

$$y(\mathbf{x}^*) < T, \text{ with } \mathbf{x}^* \in \Omega, \tag{1}$$

if the lower range of the performance value is of interest (a straightforward modification to $y(\mathbf{x}^*) > T$ is needed if the higher range is of interest, and a corresponding modification of our methodology should follow). We can then define the complete set of SOIs of a response $y(\cdot)$ as

$$\Gamma_{y,T} = \{\mathbf{x} \in \Omega : y(\mathbf{x}) < T\}. \tag{2}$$

With the continuity assumption on the black-box function, this is an open set as a union of disjointed non-empty open subsets, namely the regions of interest (ROIs) denoted as $\Gamma_T^{(k)}, k = 1, \ldots, n_r$ with n_r being the total number of ROIs, which can be written as follows:

$$\Gamma_{y,T} = \cup_{k=1}^{n_r} \Gamma_T^{(k)}, \quad \text{where } \Gamma_T^{(i)} \cap \Gamma_T^{(j)} = \emptyset, \ \forall i \neq j \in \{1, \ldots, n_r\}. \tag{3}$$

However, due to the complex structure of the objective function, the explicit form for $\Gamma_T^{(k)}$ might not be available and it will get harder with the growth of dimension. As our goal is to detect as many rare event regions as possible, we propose to detect representative SOIs in disjoint regions over the input space, which can provide more information for analysis of the design or relationship between different attributes and further modifications. In other words, in this study, rather than recovering the open set $\Gamma_{y,T}$ or the corresponding boundary detection, we focus on ROI identification via detecting representative SOIs in disconnected subsets $\Gamma_T^{(k)}, k = 1, \ldots, n_r$.

III. Bayesian Optimization for ROI Detection

Bayesian optimization (BO) was introduced to analog and mixed-signal (AMS) design verification and failure detection for the first time in [16] and they further developed an acquisition function for rare event detection in [17] that includes the distance to existing points to encourage search in the unexplored area. Consider that when the objective function is too complex to analyze or develop a comprehensive understanding with limited observations. If we are only interested in verification, we can easily determine the existence of the rare failure event by comparing the global optimum of

Algorithm 1 BO for ROI Detection

Input : Initial data $\mathcal{D}_0$, budget n_b, number of queries n_s, surrogate model $f(\cdot)$, acquisition $\alpha(\cdot)$.
Build the initial surrogate model: $f(\mathbf{x}|\mathcal{D}_0)$;
Initialize the counter: $t = 1$;
while $t < n_b$ **do**
  Determine queries via optimization:
$$\mathbf{x}_t = \text{argmax}_{\mathbf{x}}\alpha(\mathbf{x}|\mathcal{D}_{t-1}, f);$$
  Add $\mathcal{D}_t$ with the newly evaluated data:
$$y_t = y(\mathbf{x}_t),\ \mathcal{D}_t = \mathcal{D}_{t-1} \cup \{\mathbf{x}_t, y_t\};$$
  Update model and counter: $f(\mathbf{x}|\mathcal{D}_t)$, $t = t + 1$;
end
Compare $y_{\min} = \min_{t\in\{1,\ldots,n_b\}} y_t$ with T to conclude.

the black-box function with the pre-specified threshold, thus solving the verification problem. In other words, if

$$\inf_{\mathbf{x}\in\Omega} y(\mathbf{x}) < T, \tag{4}$$

it is straightforward to conclude that the failure event exists if we can pinpoint the most extreme case with a performance value smaller than the threshold. This observation suggests a solution for this problem by formulating it as a global optimization problem given a limited evaluation budget. BO has been carefully designed for sample-efficient global optimization for objective functions that are either expensive or difficult to evaluate, which gains popularity after its application in hyperparameter tuning in machine learning [18]. Its properties meet some requirements of our ROI detection problem, and the workflow of BO for ROI detection is provided in Algorithm 1. There are two essential elements in classical BO: a statistical surrogate model $f(\cdot)$ and an acquisition function $\alpha(\cdot)$.

First of all, since the original objective function is either too complex to directly work with or unavailable, a surrogate model $f(\cdot)$ is proposed to approximate the behavior of the objective function. The typical assumptions on smoothness and continuity allow the utilization of more information. To take into account the model uncertainty, we often resort to a probabilistic model, among which the Gaussian process (GP) stands out as the most common choice. Without the loss of generality, we illustrate the idea by considering a single objective Bayesian Optimization. Given a set of observations after the t-th round of evaluations denoted by $\mathcal{D}_t$, GP is a generative model characterized by a mean function μ_0 and a covariance function Σ_0 as follows,

$$\begin{aligned} f|\mathbf{x} &\sim \mathcal{GP}(\mu_0, \Sigma_0), \\ y|\mathbf{x}, f, \sigma_0^2 &\sim \mathcal{N}(f, \sigma_0^2), \end{aligned} \tag{5}$$

with σ_0^2 being the intrinsic noise variance. The collection of training and testing evaluations are assumed to be jointly Gaussian distributed. We can derive their joint distribution and further compute the predictive posterior distribution through Bayes' rule. In the above equations, the mean function $\mu_0(\cdot)$ is usually chosen to be a constant function for simplicity; and the covariance function is generally determined by a kernel function $\mathcal{K}(\mathbf{x}, \mathbf{x}')$, the common choices of which include the squared exponential and the Matérn kernel function utilized in this study. There are two hyperparameters involved, namely one for smoothness ν and one for length scale ℓ. The first one is usually chosen to be $\nu = 2.5$ for a twice differential kernel and we decide ℓ through maximal likelihood estimation after adopting a gamma prior to the length scale parameter. Also, the noise term for the Gaussian process regression is inferred based on observations.

Given a GP surrogate model, various acquisition functions $\alpha(\cdot)$ have been developed to address the exploration and exploitation trade-off. By utilizing the posterior predictive distribution under the surrogate model, acquisition functions are usually designed to balance the two objectives for selecting the next point to evaluate, namely either considering the points with the best performance given the current knowledge (exploitation) or trying new options that can give better performance in the future at the cost of an exploitation opportunity (exploration). Commonly used acquisition functions include Expected Improvement (EI, [19]), Probability of Improvement (PoI, [20], [21]), Upper Confidence Bound (GP-UCB or GP-LCB, [22]), Entropy Search (ES, [23]) and their variants (e.g. [24], [25], [26], [27], [28]). Tree-Structured Parzen Estimator (TPE, [29], [30]) is a variant of the BO methods, modeling $p(y|\mathbf{x})$ through $p(y)$ and $p(\mathbf{x}|y)$ and maximizing the acquisition function via optimization over the ratio of the conditional distributions of x given $y \geq \gamma$ and $y < \gamma$ for some γ as a threshold. However, depending on the design of acquisition functions and the shape of input space, the optimization over acquisition can be tough and get stuck in the local regions when it is non-convex or multi-modal. To enhance the algorithm efficiency, there is often a preference for batch suggestion, which requires multiple optimizations with different starting points. Although parallelization can be implemented, it still requires substantial computational resources. In addition, in our ROI detection problem, efficient BO only focuses on detecting one ROI where the global optimum falls and does not guarantee the effective identification of all the desired ROIs due to the inherent focus of BO on global optimization. As it has been studied in [31], [32], sampling can be more efficient under certain conditions, and it can naturally provide a batch of suggesting points as well as encouraging exploration beyond certain local regions. Thus, in this study, we propose an acquisition-guided sampling algorithm for our ROI detection.

IV. Proposed: Acquisition-Guided Sampling

A. Overview

The complete workflow of our proposed AcqS method is given in the pseudo-code Algorithm 2, which can be compared with Algorithm 1 that uses BO directly for ROI detection. After selecting certain specifications for evaluation, both algorithms fit a surrogate model. The two algorithms diverge in subsequent steps. While Algorithm 1 determines the next specification for evaluation by maximizing an acquisition function, our new Algorithm 2 draws a sample of specifications from a suitably designed sampling distribution.

Algorithm 2 AcqS Algorithm for ROI Detection

Input : Initial data $\mathcal{D}_0$, budget n_b, surrogate model $f(\cdot)$, acquisition $\alpha(\cdot)$, number of candidates J, number of queries n_s at each iteration

Build the initial surrogate model: $f(\mathbf{x}|\mathcal{D}_0)$;

Initialize the counters: $t = 0, n_0 = 0$;

while $n_t < n_b$ **do**

Determine queries via acquisition-guided sampling:

$$\mathbf{X}_t \in \mathbb{R}^{J\times D} \sim p(\mathbf{x}) \propto g(\alpha(\mathbf{x}, \boldsymbol{\theta}_t)).$$

Pre-Screening / Combination with other criterion:

$$n_s^* = \min(n_s, n_b - n_t),$$
$$\mathbf{X}_t \in \mathbb{R}^{n_s^*\times D} \leftarrow \text{Pre-Screening}(\mathbf{X}_t).$$

Update $\mathcal{D}_t$ and $\mathcal{M}_t$ with new observations:

$$\mathbf{y}_t \leftarrow y(\mathbf{X}_t),$$
$$\mathcal{D}_t \leftarrow \mathcal{D}_{t-1} \cup \{\mathbf{X}_t, \mathbf{y}_t\},\ \ \mathcal{M}_t \leftarrow \text{SubsetSelect}(\mathcal{D}_t).$$

Update model and counters:

$f|\mathcal{M}_t,\ \ n_t = n_{t-1} + n_s^*,\ \ t = t + 1;$

end

Return the set of SOIs: $\mathcal{F} = \{\mathbf{x} \in \mathcal{D}_t : y(\mathbf{x}) < T\}$.

Our sampling distribution is specified using an acquisition function so that specifications with high acquisition values have higher probabilities of being sampled (Section IV-B). Tailor-made acquisition functions are designed to better serve our goal of high ROI coverage (Section IV-C). After a set of specifications are sampled, we further introduce some optional add-on modules that can be seamlessly integrated for performance improvement in Section IV-D. A pre-screening step is introduced to select the most promising point to evaluate (Section IV-D1). This gives the additional flexibility of using different criteria for selecting next point to evaluate. Moreover, reweighting or subset selection can be adopted to fit the surrogate model and adjust the attention assigned to the potential ROIs (Section IV-D2).

B. Design of the Sampling Distribution

A critical element of our AcqS algorithm is the specification of the sampling distribution $p(\mathbf{x})$. Without loss of generality, we focus on the rare failure detection scenario where the regions of $\mathbf{x}$'s with objective function values smaller than a predetermined threshold are of interest.

For simplicity, we take the expected improvement (EI) as an example, which is defined as the expectation of the performance improvement described by an increment function

$$I(\mathbf{x}; u) = \max\{0, u - f(\mathbf{x})\} \tag{6}$$

with u to be specified. Let $\mathbf{x}^*$ denote the observed specification with the lowest performance value achieved given n_t observations and denote $f(\mathbf{x}^*) = f_t^*$. A common choice for u is f_t^* or $f_t^* - \eta$, where $\eta > 0$ is introduced as a small positive value to enlarge the difference for performance improvement. With $u = f(\mathbf{x}^*)$ and a GP surrogate model for f, we have the classic EI defined as

$$\begin{aligned}\alpha_{ei}(\mathbf{x}) &= \mathbb{E}(I(\mathbf{x}); f(\mathbf{x}^*)) = \mathbb{E}[\max\{0, f_t^* - f(\mathbf{x})\}] \\ &= (f_t^* - \mu(\mathbf{x}))\Phi\Big(\frac{f_t^* - \mu(\mathbf{x})}{\sigma(\mathbf{x})}\Big) + \sigma(\mathbf{x})\phi\Big(\frac{f_t^* - \mu(\mathbf{x})}{\sigma(\mathbf{x})}\Big).\end{aligned} \tag{7}$$

There are hyperparameters for the surrogate model $f(\cdot)$ such as ν and ℓ mentioned earlier for the kernel function, which can be estimated through the maximum likelihood estimation, and we do not include them in $f(\cdot)$ for notational simplicity and to distinguish them from other hyperparameters required in the specification of the acquisition function.

The rationale of using the EI as an acquisition function in BO is that one prefers a specification with a higher EI value when selecting the next point for evaluation as it has higher improvement in expectation of the customized increment function. To extend this idea in our sampling-based approach, we define the probability density of the sampling distribution to be proportional to the EI value. More precisely, our acquisition-guided sampling scheme is defined as

$$\mathbf{x} \sim p(\mathbf{x}) = \alpha_{ei}(\mathbf{x})/Z_t \propto \alpha_{ei}(\mathbf{x}), \tag{8}$$

where Z_t is a normalization constant such that $p(\mathbf{x})$ integrates to one for ensuring a valid probability density.

For other acquisition functions that can take negative values such as GP-LCB defined as $\alpha_{lcb}(\mathbf{x}, \theta_t) = \mu(\mathbf{x}) - \theta_t\sigma(\mathbf{x})$ with $\theta_t \in \mathbb{R}$, a monotonic function $g(\cdot)$ will be introduced to ensure a non-negative density value over the sampling space. Thus, for any given acquisition $\alpha(\cdot)$, our acquisition-guided sampling scheme can be expressed in the general form as follows,

$$\mathbf{x} \sim p(\mathbf{x}) = g(\alpha(\mathbf{x}, \boldsymbol{\theta}_t))/Z_t \propto g(\alpha(\mathbf{x}, \boldsymbol{\theta}_t)). \tag{9}$$

Here $\boldsymbol{\theta}_t$ represents a vector of hyperparameters for the acquisition function, and it is allowed to be adaptively chosen and dependent on the observations, allowing incorporation of desired information and adjustable properties such as annealing and tempering. Although there is flexibility in choosing an appropriate monotonic function for $g(\cdot)$, we use the exponential function in our study.

Furthermore, we can introduce additional pre-screening criteria to restrict the sampling space (see Section IV-D1 below). To this end, the sampling density can be specified as

$$p(\mathbf{x}) \propto \underbrace{g(\alpha(\mathbf{x}, \boldsymbol{\theta}_t))}_{\text{Acquisition-guided}} \times \underbrace{1(\mathbf{x}\text{ satisfies other criteria})}_{\text{Pre-screening}}. \tag{10}$$

Various techniques are available for sampling from a distribution with an unknown normalizing constant, such as slice sampling [33], Metropolis-Hasting Monte Carlo (MC) sampling [34], [35], and Hamiltonian MC (HMC) sampling [36]. We opt for HMC, because preserving the value of Hamiltonian is one of its well-known features, and its intrinsic searching pattern enables the sampling distribution to be fixed at a certain energy level specified by the maximal acquisition value, aligning with our objective of identifying ROIs with the maximal coverage.

C. Acquisition Functions for ROI Detection

To better serve our purpose of achieving the maximal coverage of ROIs, we propose two modifications of the existing acquisition functions in BO and a new acquisition function. They are specifically designed to use the performance target to guide the searching process, encouraging the wider spread of candidate specifications, and eventually improving the performance of ROI detection.

a) Threshold-Guided Search: In the definition of ROI (see (1) – (3)), there is a predetermined threshold T. We would like to make use of this knowledge to help achieve the desired maximal coverage of ROIs. To this end, we take $u = T$ in the increment function (6) so that it becomes

$$I(\mathbf{x}; T) = \max\{0, T - f(\mathbf{x})\}, \tag{11}$$

and use it to define the EI. While the classic EI given in (7) is designed for solving a minimization problem, this modification of the EI is more suitable to the detection of ROIs.

b) Distance-Adaptive Acquisition: To encourage maximal coverage of ROIs in our sampling process, we would like to give more weights to points that are away from the points that we have already evaluated. We propose the following adjustment to the sampling density

$$p(\mathbf{x}) \propto \exp\{\alpha(\mathbf{x}, \boldsymbol{\theta}_t) \times T_t(\mathbf{x})\}. \tag{12}$$

Motivated by the minimax and maximin design in sequential design for experiments and output-weighted acquisition [37], we propose the following distance adjustment term,

$$T_t(\mathbf{x}) = \left(\prod_{i=1}^{n_t} d(\mathbf{x} - \mathbf{x}_i)\right)^{1/p}, \tag{13}$$

where $d(\mathbf{x} - \mathbf{x}_i)$ is the (non-negative) distance between a candidate specification $\mathbf{x}$ and the observed specification $\mathbf{x}_i$, $i \in \mathcal{M}_t = \{1, \ldots, n_t\}$, and $d(\cdot)$ is a metric that can be chosen according to different problem settings. We use the Euclidean distance and $p = 2$ in this study. This idea is similar to the additional high-coverage term in the pBOHC acquisition function proposed in [16] where $p = n_t$ and a specifically designed distance metric is used. The distance adjustment term prefers specifications in unexplored areas by them given higher weights, therefore helping to reach high coverage of ROIs.

c) Proposed Acquisition Design - Probability Ratio (PR): We observed that, for commonly used acquisition functions such as EI, the function values at different locations can be very close, and as a consequence, our acquisition function motivated sampling distribution (i.e., (8)) may get close to a uniform distribution when the number of observations used for training the surrogate model is small, making its use less effective. Here we design a new acquisition function to deal with this issue. Consider the failure detection problem, and suppose that a failure occurs if the objective function is below certain threshold. Let ζ be a value chosen for probability improvement, which can be but not limited to f_t^* the the smallest objective function value given n_t evaluations as in Algorithm 2 and the threshold T. We define a new acquisition function named the probability ratio (PR),

$$\alpha_{PR}(\mathbf{x}, \zeta) = \frac{\text{Prob}(f(\mathbf{x}) < \zeta)}{\text{Prob}(f(\mathbf{x}) \geq \zeta)}. \tag{14}$$

It is an application of the transformation $g(x) = x/(1-x)$ to the probability of improvement, $\text{Prob}(f(\mathbf{x}) < f_t^*)$. We found empirically that this acquisition function can usually lead to a sampling distribution that is less uniformly distributed.

D. Add-On Modules

In our proposed sampling workflow for rare event detection, multiple add-on modules can be integrated to further improve sampling efficiency and algorithm performance.

1) Pre-Screening: At each application of the acquisition-guided sampling in Algorithm 2, multiple candidates can be obtained. A pre-screening step is introduced to select the most promising candidates for performing the actual, expensive evaluation. Various selection criteria that can be considered are listed below.

- Distance-related criteria: The distances between suggestions and their distances to existing points along with the principle of minimax, maximin, can be considered to encourage the coverage of the variatidon space for diversity (e.g. [38], [39], [40], [41]). Moreover, various divergence or distance measures between distributions can also be used as a criterion when the information gain or divergence is of greater interest (e.g. [42], [43], [44]).
- Acquisition functions: Acquisition functions different from the one used to define the sampling distribution can provide new measurements for the performance of points from various perspectives.
- Objectives for multi-modal optimization: As the multi-modal optimization problem targets finding as many local optima as possible and the corresponding derivatives can be adopted as additional criteria to guide the local search (e.g. [45], [46]), the target function combining multiple objectives can be utilized to encourage exploration.

2) Subset Selection: The quality of observations can dramatically affect the performance of the model, as noise and redundant information can lead to biased estimation or computational difficulty. Thus, we resort to subset selection, a common technique in machine learning which usually contains the process of learning the necessary patterns or target-related information from a dataset. The main purpose of subset selection in our acquisition-guided sampling strategy is to help remove redundant information and reweight SOIs for surrogate model fitting. Additionally, it can reduce computational complexity due to the growth of sample size in GP modeling. Various strategies can be used to address this issue, including both supervised and unsupervised subset selection. Motivated by [31] and [47], we adopt a response-based sampling scheme to generate a representative subset as follows:

(1) Divide the range of responses $\{y_i\}_{i=1}^{n_t}$ to K disjoint intervals $S_1, \ldots, S_K$ (referred to as the response intervals). Let $|S_k|, k = 1, \ldots, K$ denote the number of observations in each interval.

TABLE I
EVALUATION METRICS OF ROI DETECTION PERFORMANCE

Notation	Definition (in one replicate)	Description
t_i	$\sum_{k=1}^{t_i} 1\{f(\mathbf{x}_k) < T, \mathbf{x}_k \in \Gamma_T^{\leq i}\} = i$ or $\sum_{k=1}^{t_i} 1\{f(\mathbf{x}_k) < T\} = i$	Average 1-st time hitting the i-th ROI or SOI
r_i	$1\{\text{finding } \Gamma_T^{(i)}\}$	Average rate of detecting the i-th ROI
$N_n^r,\ N_n$	$\sum_{k=1}^{n} 1\{f(\mathbf{x}_k) < T,\ \mathbf{x}_k \in \Gamma_T^{(i)}\}$ or $\sum_{k=1}^{n} 1\{f(\mathbf{x}_k) < T\}$	Average number of ROIs/SOIs until the n-th evaluation
A_n	$\sum_{i=1}^{N_n^r} Conv_i$	Average length $(D = 1)$/area $(D > 1)$ of the convex hull generated by the SOIs in disjointed ROIs in the n-th iteration
$y_{min},\ y_{max}$	$\min_{i\in\{1,\ldots,n_b\}} y_i,\ \ \max_{i\in\{1,\ldots,n_b\}} y_i$	The optimal performance value achieved in n_b evaluations

(2) Sample an equal number of samples n^* from each S_k uniformly to obtain the new representative set $\mathbf{X}^* = (\mathbf{X}_1^*, \ldots, \mathbf{X}_K^*)$ with $\mathbf{X}_k^* \in \mathbb{R}^{n^* \times D}$ for model fitting.

In this study, this subsampling method is used to keep the implementation relatively simple and ensure the diversity of points in the subset. Moreover, by sampling an equal number of points from each response interval, we achieve a reweighing of available points and a reassignment of attention.

V. SIMULATION RESULTS

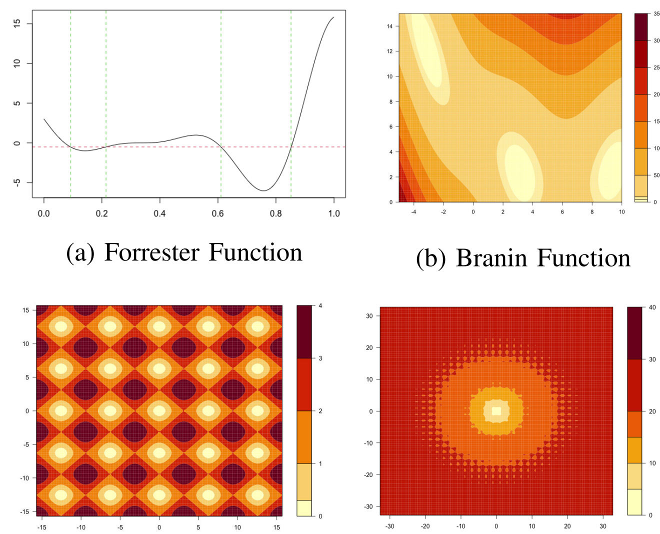
We evaluate our AcqS approach for ROI detection and compare it with existing methods, including uniform random sampling, the classical GP-based BO, and TPE. For BO, different acquisition functions including EI, PoI, UCB, and Thompson sampling (TS) are considered. The initial dataset is composed of specifications with responses far away from threshold-defining ROIs. We report the performance statistics based on $M = 30$ simulation runs for each setting.

A. Evaluation Metrics

To illustrate the performance of ROI detection of all the methods for comparison, we use the following four metrics for evaluation in terms of detection time for efficiency, percentage of detecting the i-th ROI, and the number of ROIs found for coverage and the area of the convex hull generated by the identified SOIs within each ROI for illustration of the exploration ability. First, t_i is the 1-st time hitting the i-th ROI when ROIs are known in advance; when no available information about ROIs is available, t_i represents the i-th time locating the i-th SOI, with a smaller t_i indicating faster detection. Another metric, r_i, evaluates the percentage of finding the i-th ROI. N_n^r represents the number of detected ROIs until the n-th iteration of evaluations in either sampling or BO procedures. A_n measures the area of the convex hulls that constructed the points within the ROIs until the n-th iteration, with a higher value indicating better coverage and exploration with the ability to get out of certain small regions. When boundaries for ROIs are not explicitly available, we check the optimal value achieved y_{min} instead of A_n. The descriptions of these metrics are summarized in Table I.

B. Benchmark Optimization Functions

We first benchmark the performance on standard black-box optimization test functions, including the 1-D Forrester, 2-D Branin, 2-D Cos25, and Ackley functions with dimensions of 6 and 10. Figure 1 includes the contour figures for these functions with a 2-D Ackley function given as an illustration for its high-dimensional behavior. The empty entry in the tables displayed in the following section is for the case where no detection is available.



(a) Forrester Function (b) Branin Function (c) Cos25 Function (d) 2-D Ackley Function

Fig. 1. Synthetic Functions: (a) Forrester, (b) Branin, (c) Cos25, and (d) 2-D Ackley and their ROIs with threshold $T = -0.5, 5, 0.3,$ and 5, respectively. (a) Two ROIs are the intervals bounded by the green dashed lines for the Forrester function; in (b)(c)(d) ROIs are colored in light yellow.

a) Forrester Function [48]: It is a 1-D benchmark synthesis function defined on $[0, 1]$ as $f(x) = (6x - 2)^2 \sin(12x - 4),\ x \in [0, 1]$. It is multi-modal with a global optimum and a local optimum. We set the threshold to be $T = -0.5$, which gives two ROIs to be detected. This function is used to illustrate the difference between the ROI detection problem and global optimization. The behavior of the test function and the two ROIs is shown in Figure 1(a).

The left panel of Table II summarizes the performance comparison results for random sampling, TPE, BO and AcqS with EI, PoI and UCB. When random sampling is adopted, the relatively large number of ROIs and the wide detected intervals indicate the exploration ability of sampling without additional information. TPE provides better performance than random sampling regarding the first detection time for both

ROIs as more information is utilized. With EI and PoI as the acquisition functions, it is clear that our proposed AcqS method outperforms classical BO as it can find the ROIs faster with higher coverage. Furthermore, the larger interval detected for ROIs by our sampling illustrates better exploration ability.

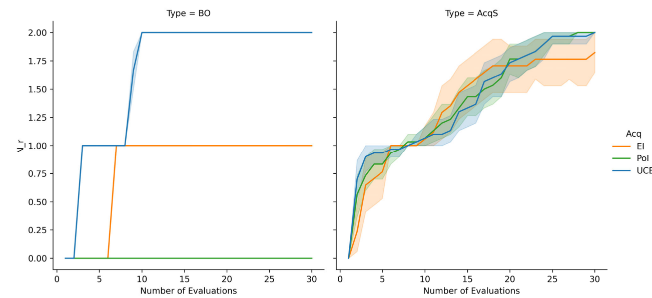
Compared to all baselines, our proposed approach is able to identify more specifications more quickly in the target ROIs. Both BO and our sampling strategy with UCB perform well in this case with full coverage of ROIs where the sampling method provides the fastest first hitting time of the 1-st ROI while BO has the smaller 2-nd ROI detection time. In addition, it is shown in Figure 2(a) with the number of evaluations on the x-axis and the average number of detected ROIs on the y-axis that the proposed acquisition-guided sampling achieves a similar performance with different choices of acquisition functions. As is displayed in the right panel for our AcqS approach, the curves generally climb faster and tend to achieve higher coverage than those in the left panel using BO. In other words, it is more robust with respect to the choice of acquisition functions with EI, PoI and UCB represented by orange, green, and blue, respectively. More importantly, it converges faster to ROIs at the beginning and is able to detect all target ROIs within the limited budget.

b) Branin Function: It is a 2-D benchmark function defined on $x_1 \in [-5, 10]$, $x_2 \in [0, 15]$ with three global minima with form $f(\mathbf{x}) = a(x_2 - bx_1^2 + cx_1 - r)^2 + s(1-t)\cos(x_1) + s$, with $(a, b, c, r, s, t) = (1, 5.1/(4\pi^2), 5/\pi, 6, 10, 1/(8\pi))$ We set the threshold $T = 5$, resulting in three ROIs around $(-\pi, 12.275)$, $(\pi, 2.275)$, and $(9.42478, 2.475)$, illustrated as yellow regions in Figure 1(b).

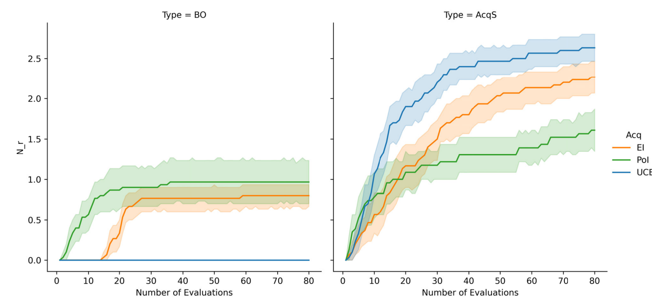
The numerical results are provided in Table II. With the same acquisition function, AcqS can generally achieve a higher coverage of ROIs, according to $r_{1:3}$, N_{80}^r, and A_{80}, as well as faster detection based on $t_{1:3}$. We further investigate the growth trend of the average number of detected ROIs using BO on the left panel and AcqS on the right in Figure 2(b). It is obvious that BO gets stuck at certain regions and encounters difficulty in detecting SOIs in disjointed ROIs as the curves barely go up after finding one ROI while AcqS with various acquisition functions have a higher average number of ROIs detected. Our sampling strategy with different acquisition functions provides more stable performance and rapid detection as trends by AcqS with different acquisitions in the right panel show consistently faster and better detection performances than those in the left panel by BO.

c) Cos25 Function: It is a 2-D benchmark function with 25 global minima denoted as the light yellow regions in Figure 1(c). BO-based methods can easily get stuck in one of these ROIs, which makes it challenging to achieve the higher rare event detection coverage. We set $T = 0.3$. As shown in Table III, the number of ROIs detected by our sampling methods is generally higher than that obtained by classical BO with the same acquisition function. Figure 3 shows the growth of the number of ROIs detected and the first hitting time for the i-th ROI. Though BO with EI can achieve a pretty fast detection for the first ROI, it is challenging for it to explore more over the space. The performance of our sampling methods with different acquisition functions is relatively consistent, showing that acquisition-guided sampling is more robust to the choice of acquisition functions and can contribute to faster detection and more stable growth of the average number of detected ROIs.

d) Ackley Function: It is another widely studied benchmark function for optimization. Cases with two dimensions

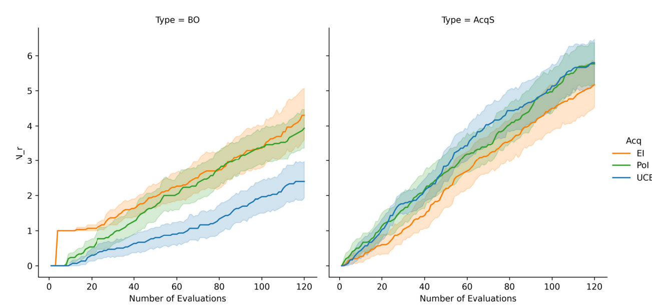


(a) Forrester Test Function (T = -0.5)

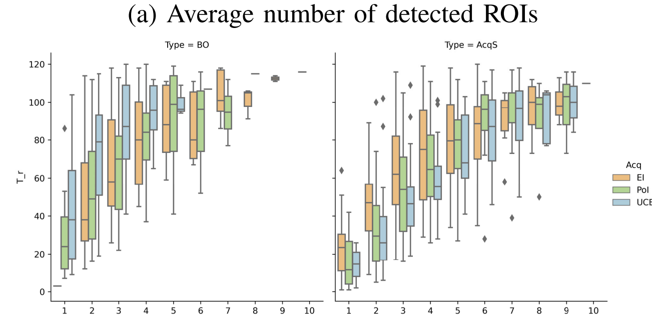


(b) Branin Test Function (T = 5)

Fig. 2. Average number of detected ROIs for (a) Forrester and (b) Branin function using BO and AcqS.



(a) Average number of detected ROIs



(b) Average 1-st hitting time for the ith ROI

Fig. 3. The average number of detected ROIs and 1-st time hitting the i-th ROI for Cos25 function using BO and AcqS.

TABLE II
RESULTS FOR RANDOM SAMPLING, TPE, BO AND ACQS WITH EI, POI, AND UCB FOR THE FORRESTER AND BRANIN FUNCTION

Forrester ($T = -0.5$)						Branin ($T = 5$)						
Type	Acq	$t_1(r_1)$	$t_2(r_2)$	N^r_{30}	A_{30}	Type	Acq	$t_1(r_1)$	$t_2(r_2)$	$t_3(r_3)$	N^r_{80}	A_{80}
Random		4.10(1.0)	9.33(.90)	1.90	0.20	Random		9.93(1.0)	33.43(1.0)	41.50(.67)	2.67	1.28
TPE		2.83(1.0)	7.89(.90)	1.90	0.23	TPE		11.63(1.0)	32.36(.83)	33.44(.30)	2.13	2.63
BO	TS	3.43(1.0)	9.33(.20)	1.20	0.08	BO	TS	9.67(1.0)	24.53(.57)	45.12(.27)	1.83	2.30
BO	EI	6.00(1.0)		1.00	0.01	BO	EI	20.96(.80)			0.80	0.71
AcqS	EI	2.65(1.0)	14.50(.82)	1.82	0.22	AcqS	EI	12.27(1.0)	28.79(.93)	47.90(.33)	2.27	1.64
BO	PoI					BO	PoI	6.41(.73)	19.00(.20)	15.00(.03)	0.97	0.54
AcqS	PoI	2.13(1.0)	16.50(1.0)	2.00	0.25	AcqS	PoI	10.00(1.0)	35.92(.57)	74.00(.04)	1.61	4.26
BO	UCB	2.00(1.0)	8.33(1.0)	2.00	0.13	BO	UCB					
AcqS	UCB	1.60(1.0)	16.90(1.0)	2.00	0.26	AcqS	UCB	7.17(1.0)	15.87(1.0)	37.20(.67)	2.67	3.49

TABLE III
RESULTS FOR RANDOM SAMPLING, TPE, BO AND ACQS WITH EI, POI AND UCB FOR COS25 FUNCTION

Function		Cos25 ($T = 0.3$)			
Type	Acq	$t_1(r_1)$	$t_9(r_9)$	N^r_{120}	y_{min}
Random		25.00(1.0)	85.5(.07)	5.17	0.05
TPE		17.25(1.0)	103.50(.10)	5.90	0.03
BO	TS	17.77(1.0)	94.40(.17)	5.97	0.02
BO	EI	3.00(1.0)	112.50(.07)	4.30	0.03
AcqS	EI	22.77(1.0)	99.67(.10)	5.20	0.04
BO	PoI	27.20(1.0)		3.97	0.02
AcqS	PoI	15.97(1.0)	97.33(.10)	5.77	0.03
BO	UCB	44.44(.90)		2.43	0.12
AcqS	UCB	14.23(1.0)	100.00(.13)	5.80	0.04

TABLE IV
RESULTS FOR RANDOM SAMPLING, TPE, BO AND ACQS WITH EI, POI AND UCB FOR 6-D AND 10-D ACKLEY FUNCTION

Function		Ackley-6 ($T = 10$)			Ackley-10 ($T = 15$)		
Type	Acq	$t_1(r_1)$	N^r_{120}	y_{min}	$t_1(r_1)$	N^r_{120}	y_{min}
Random				17.70			19.50
TPE		85.00(.03)	0.03	14.45			17.52
BO	TS	90.69(.43)	0.43	9.79	74.80(.33)	0.50	14.95
BO	EI			19.33			21.57
AcqS	EI	35.86(.25)	0.32	11.22	52.33(.70)	1.87	13.89
BO	PoI			20.18			18.45
AcqS	PoI	70.64(.37)	0.37	10.57	54.68(.73)	1.77	14.16
BO	UCB			15.33			21.57
AcqS	UCB	65.25(.27)	0.27	11.30	53.33(.70)	0.87	14.18

$D = 6, 10$ are considered in our study for comparison in higher dimensional settings. Similar trends as discussed for the previous benchmark functions can be easily observed in these two settings. The performance of our proposed sampling approach has more potential for successful detection of the ROIs and achieving better sample efficiency for ROI detection.

C. Proposed Modifications for Acquisition Functions

The numerical results for the 2-D synthetic Branin function with the number of ROIs being 3 are displayed below for illustration. Additional results for other benchmark optimization functions are given in the supplementary file. We denote the acquisition with threshold-guided modification using the subscript α_T and adaptive temperature modification with adp_α. The first 8 rows of Table V present the results for BO with EI, PoI, UCB, and their variants with the proposed acquisition modifications as discussed in Section IV-C. By comparing the performance of methods that use the original EI and PoI with those that utilize the threshold-guided modification $\text{EI}_T, \text{PoI}_T$, the detection rates of the three ROIs with such a modification are generally higher, indicating a higher coverage of the detected ROIs, regardless of using BO or our AcqS. The first time hitting the 1-st ROI given by EI with the threshold-guided modification is obviously smaller, indicating faster detection via the introduction of threshold T. EI_T and PoI_T also outperform EI and PoI in BO, respectively, in terms of both the number of detected ROIs and the area of the convex hull given by the points in the detected ROIs, which indicates a better exploration of the space. EI_T and PoI_T provide a higher N^r_{80} for our sampling strategy. Comparing the sampling methods with their BO counterparts, the number of ROIs found by the AcqS is generally higher, and the results given by different acquisition functions show more robust performance.

Similar performance improvement can be observed via the performance comparison as well with the introduction of adaptive temperature based on the distances. BO with adp_EI achieves an obvious improvement for the 1-st hitting time for the 1-st ROI and its coverage, as well as detection of the remaining two ROIs. The sampling algorithm with adp_EI also improves from that with EI as the acquisition function in terms of $t_i, r_i, i = 1, 2, 3$ and N^r_{80}. BO and the sampling method with adp_PoI both obtain significant performance enhancement in all five metrics as well, indicating a wider spread of the evaluated specifications, higher coverage, and faster detection.

Regardless of the choice of acquisition function, the performances provided by AcqS approach are close to each other, meaning it is more robust and flexible to the selection of acquisition functions in ROI detection compared to their BO counterparts. The combination of two modifications can also show their strength as adp_PoI_T had the best 1-st hitting

TABLE V

RESULTS FOR BO AND ACQS WITH EI, POI, UCB AND ITS VARIANTS FOR BRANIN FUNCTION WITH $T = 5$

Acq	Type	$t_1(r_1)$	$t_2(r_2)$	$t_3(r_3)$	N_{80}^r	A_{80}
EI	BO	20.96(.80)			0.80	0.71
	AcqS	12.27(1.0)	28.79(.93)	47.90(.33)	2.27	1.64
PoI	BO	6.41(.73)	19.00(.20)	15.00(.03)	0.97	0.54
	AcqS	10.00(1.0)	35.92(.57)	74.00(.04)	1.61	4.26
EI_T	BO	8.03(1.0)	11.71(.23)		1.23	2.12
	AcqS	8.48(1.0)	23.32(.93)	50.44(.59)	2.52	1.95
PoI_T	BO	9.47(1.0)	29.08(.83)	44.33(.40)	2.23	1.70
	AcqS	9.56(1.0)	21.78(1.0)	51.86(.58)	2.58	3.09
adp_EI	BO	6.23(1.0)	48.17(1.0)	72.67(.30)	2.30	4.68
	AcqS	10.93(1.0)	28.17(.97)	46.47(.50)	2.47	1.56
adp_PoI	BO	8.50(1.0)	49.43(.93)	70.22(.30)	2.23	4.29
	AcqS	7.20(1.0)	17.93(1.0)	46.31(.53)	2.53	4.70
adp_EI_T	BO	6.30(1.0)	30.20(1.0)	55.40(.50)	2.50	6.09
	AcqS	10.44(1.0)	23.28(1.0)	43.18(.44)	2.44	2.01
adp_PoI_T	BO	12.63(1.0)	35.79(.97)	44.38(.43)	2.40	2.13
	AcqS	6.00(1.0)	10.00(1.0)	27.00(.50)	2.50	5.19
UCB	BO					
	AcqS	7.17(1.0)	15.87(1.0)	37.20(.67)	2.67	3.49
adp_UCB	BO					
	AcqS	9.00(1.0)	13.00(1.0)	30.00(1.0)	3.00	6.23

time for the first two ROIs. For UCB and its corresponding variants for the Branin test function, our sampling method generally provides better results for the coverage and area of convex hulls generated by points within ROIs, indicating better exploration over the space and demonstrating the strength of our proposed modifications and acquisition functions in ROI detection. In this setting, adp_UCB achieves the highest coverage to identify all 3 ROIs and the largest area generated by the detect specifications. Overall, with the introduction of the adjusted distance term and threshold information, modified acquisitions can be used to achieve better ROI detection.

D. Proposed Acquisition Function: Probability Ratio

Simulation results for Branin function are provided for comparison of PR and its variants used in BO and sampling approaches for illustration. Table VI shows that the incorporation of PR with other modifications has achieved faster detection and higher coverage. The adjusted distance term can encourage better exploration over the space and jump out of certain local regions. There is a higher success rate for finding all ROIs with modified PR acquisition functions. AcqS generally has better performance in ROI detection. PR can converge to the ROIs faster with the risk of getting stuck in the small local region around the first ROI. The combination of proposed modification designs with PR can integrate their strengths together, leading to better ROI detection performance.

E. ROI Detection Under Various Rare Degree

The rare degree can be controlled by the choice of threshold T. When the threshold is set as a lower bound for the desired design performance, the smaller it gets, the harder the problem becomes, as the ROIs are smaller, and the problem becomes more similar to the global optimization. The following additional experiments for Branin synthetic functions with different T's illustrate the performance difference for BO and the acquisition-guided sampling strategy. We can observe similar numerical patterns with different thresholds T's, our proposed sampling strategy is consistently better while the problem becomes more challenging as the rare degree increases (lower T). For example, for Branin at $T = 1$ and 3, the evaluation metrics obtained by our sampling strategy are better than those obtained by BO-based methods in general, achieving $N_{80}^r = 1.87$ and 2.50 values based on UCB. According to the pairwise comparison, our AcqS approach usually detected a higher number of ROIs and a larger volume of the convex hulls generated by the SOIs detected, indicating a stronger exploration ability and less likely to be stuck in local regions.

TABLE VI

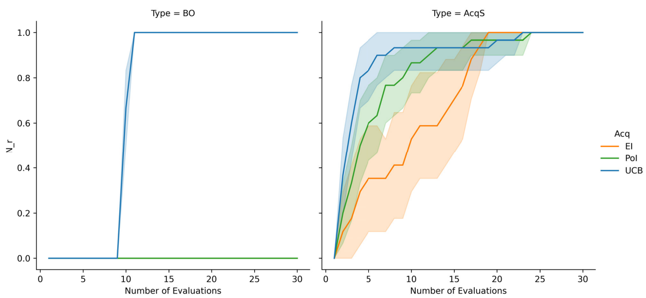
COMPARISON OF PR AND ITS VARIANTS IN BO AND ACQS WITH EI, POI, AND UCB FOR THE BRANIN WITH $T = 5$

Acq	BO				
	$t_1(r_1)$	$t_2(r_2)$	$t_3(r_3)$	N_{80}^r	A_{80}
EI	20.96(.80)			0.80	0.71
PoI	6.41(.73)	19.00(.20)	15.00(.03)	0.97	0.54
UCB					
PR	7.22(.60)	21.86(.23)		0.83	0.28
PR_T	15.27(1.0)	32.33(.90)	50.50(.40)	2.30	1.74
adp_PR	6.83(1.0)	44.78(.90)	63.62(.27)	2.17	4.24
adp_PR_T	12.43(1.0)	36.48(.90)	50.27(.37)	2.27	2.54
Acq	**AcqS**				
	$t_1(r_1)$	$t_2(r_2)$	$t_3(r_3)$	N_{80}^r	A_{80}
EI	12.27(1.0)	28.79(.93)	47.90(.33)	2.27	1.64
PoI	10.00(1.0)	35.92(.57)	74.00(.04)	1.61	4.26
UCB	7.17(1.0)	15.87(1.0)	37.20(.67)	2.67	3.49
PR	9.83 (1.0)	25.17(.40)	32.00(.07)	1.47	4.31
PR_T	10.53(1.0)	25.89(.93)	45.28(.60)	2.53	1.66
adp_PR	12.80(1.0)	32.61(.95)	57.12(.43)	2.38	1.36
adp_PR_T	9.25(1.0)	25.00(1.0)	41.11(.75)	2.75	3.92

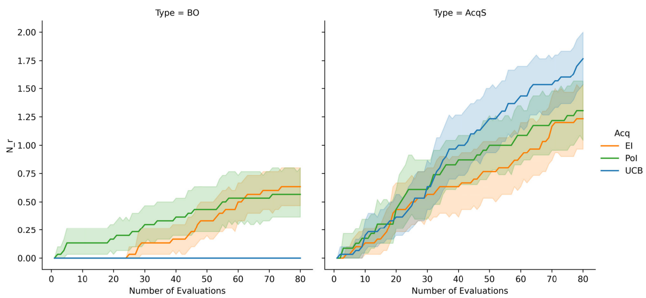
TABLE VII

RESULTS FOR BRANIN FUNCTION WITH DIFFERENT T'S

Type	Acq	T	$t_1(r_1)$	$t_2(r_2)$	$t_3(r_3)$	N_{80}^r	A_{80}
Random		3	25.67(1.0)	37.38(.80)	50.73(.50)	2.30	0.12
TPE			14.40(1.0)	34.72(.60)	24.33(.10)	1.70	1.03
BO	TS		12.47(1.0)	29.41(.57)	47.75(.27)	1.83	1.33
BO	EI		33.81(.70)			0.70	0.40
AcqS	EI		17.07(1.0)	34.85(.87)	49.57(.23)	2.10	0.57
BO	PoI		7.50(.67)	36.00(.13)		0.80	0.27
AcqS	PoI		12.30(1.0)	40.83(.52)	75.00(.04)	1.57	2.06
BO	UCB						
AcqS	UCB		10.63(1.0)	26.90(.97)	46.50(.53)	2.50	1.43
Type	**Acq**	**T**	$t_1(r_1)$	$t_2(r_2)$	$t_3(r_3)$	N_{80}^r	A_{80}
Random		1	40.43(.77)	51.33(.20)	56.00(.03)	1.00	0.00
TPE			35.26(.90)	47.80(.17)		1.07	0.02
BO	TS		17.43(1.0)	36.00(.53)	55.50(.20)	1.73	0.30
BO	EI		47.95(.63)			0.63	0.03
AcqS	EI		31.32(.83)	54.82(.37)	32.00(.03)	1.23	0.02
BO	PoI		28.60(.50)	48.00(.07)		0.57	0.02
AcqS	PoI		25.71(.91)	51.00(.39)		1.30	0.11
BO	UCB						
AcqS	UCB		28.52(.97)	53.91(.77)	58.75(.13)	1.87	0.03



(a) Forrester Test Function (T = -5)



(b) Branin Test Function (T = 1)

Fig. 4. Average number of detected ROIs for Forrester and Branin using BO and AcqS with proposed modifications.

VI. Circuit Design Verification

Data-driven approaches by integrating machine learning in traditional circuit design and verification provide new solutions to sample-efficient automated analog and mixed-signal (AMS) design. We implement our AcqS approach to AMS design for a two-stage differential amplifier with 14 distinct design parameters, the schematic illustration of which is in Figure 5. The circuit simulator used is Synopsys HSpice, and the circuit is designed under a commercial 90nm CMOS technology. The experiments are executed over an Intel(R) Core(TM) i5-6500 CPU with a clock speed of 3.2GHz. A figure of merit (FOM) is used to obtain the circuit performance measure from the simulator, which is defined as a combination of four responses to quantify the performance of design, including unity gain frequency (UGF), gain, common-mode rejection ratio (CMRR) and power consumption represented by the quiescent current (Id): FOM $=$ UGF $+$ gain $+$ CMRR $-$ Id. A higher FOM is desired and according to [49], FOM $\geq$ 5.20 can be rare to find. We, thus, specify $T = 5.0$. EI, PoI, and UCB are chosen for comparison. In addition to the previous evaluation metrics, we also report $\bar{y}_{max}$ and $\bar{y}_r$, representing the average of maximal FOM of the value detected and that of the desired detected specifications, respectively.

As for EI, PoI and their variants, Table VIII shows that including the threshold information and adjusted distance term generally can help improve the detection time for both BO and sampling methods. AcqS method can detect more SOIs, while the modified PoI-type acquisition for BO can achieve faster detection. The performance values reached are comparable for BO and AcqS. The sampling method guided by EI with the threshold and adjusted distance term has the highest $\bar{y}_r = 5.50$. EI_T-, PoI- and adp_PoI_T-guided AcqS can also achieve rather high FOMs being 5.24, 5.29 and 5.20, which are over

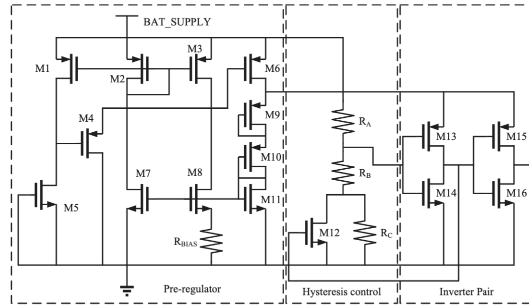


Fig. 5. A two-stage differential amplifier circuit diagram.

TABLE VIII

Results for BO and AcqS With EI, PoI, UCB, PR and Its Variants in Circuit Design Application With $T = 5.0$

Acq	Modification	Type	$t_1(r_1)$	N_{30}	$\bar{y}_{max}$	$\bar{y}_r$
EI		BO	16.83(.20)	0.23	4.87	5.17
		AcqS	20.00(.27)	0.27	4.92	5.18
	T	BO	11.00(.03)	0.03	4.76	5.17
		AcqS	17.50(.47)	0.57	5.03	5.24
	adp	BO	20.00(.10)	0.10	4.68	5.12
		AcqS	13.50(.33)	0.43	4.93	5.12
	adp + T	BO	10.33(.10)	0.10	4.68	5.17
		AcqS	18.67(.10)	0.10	4.89	5.50
PoI		BO	14.11(.30)	0.67	4.96	5.13
		AcqS	14.11(.30)	0.67	4.97	5.29
	T	BO	15.50(.20)	0.90	4.87	5.07
		AcqS	12.75(.13)	0.13	4.88	5.10
	adp	BO	9.00(.10)	0.13	4.72	5.17
		AcqS	19.40(.17)	0.17	4.86	5.05
	adp + T	BO	8.00(.03)	0.03	4.62	5.04
		AcqS	11.00(.13)	0.13	4.90	5.20
UCB		BO	8.67(.10)	0.13	4.59	5.16
		AcqS	18.00(.27)	0.30	4.92	5.21
	adp	BO	13.00(.07)	0.07	4.29	5.02
		AcqS	18.67(.20)	0.20	4.92	5.18
PR		BO	15.22(.30)	0.77	4.96	5.21
		AcqS	16.67(.30)	0.30	5.04	5.42
	T	BO	17.33(.30)	1.53	4.90	5.10
		AcqS	19.73(.37)	0.40	4.98	5.17
	adp	BO	24.00(.07)	0.07	4.71	5.07
		AcqS	15.80(.33)	0.37	4.95	5.23
	adp + T	BO	24.00(.07)	0.07	4.68	5.18
		AcqS	19.00(.27)	0.27	4.92	5.19

or at our specified threshold $T = 5.20$. As for UCB and its variant adp_UCB, the corresponding sampling method usually has a higher probability of SOI detection while the BO-based counterpart can probably do better in hitting time due to the feature of optimization. AcqS with UCB did better in $\bar{y}$ as $\bar{y}_r = 5.21$, indicating the searching ability of sampling.

The results by new PR-based acquisition functions are shown in the last 8 rows in Table VIII. PR-guided AcqS has the highest average maximal FOM at 5.04. The mean FOM of those successful specifications $\bar{y}_r = 5.42$, ranking the second best. As for detection time, PR and its variants have rather similar behavior with sampling methods. This type of acquisition function has relatively good performance in terms of the average of best FOM values compared to EI, PoI, and UCB. Overall, we can see that the number of detected SOIs by AcqS is consistently better than that given by BO with EI, PoI, and UCB. The newly proposed PR can greatly improve BO and has the potential to achieve better performance.

VII. Conclusion and Future Work

In this paper, we study the rare event detection problem. Rather than global optimization of a complex, possibly black-box target function as the original goal of BO, we focus on fast detection and high coverage of ROIs given a limited evaluation budget. Thus, we reformulate the problem as ROI detection via detecting representative SOIs in disjoint ROIs and propose the acquisition-guided sampling (AcqS) algorithm.

Despite using different types of acquisition functions, acquisition-guided sampling can provide more stable ROI detection results than the original BO methods. Numerical results for synthetic test functions and a real-world application of circuit design show a significant improvement of our AcqS approach over existing methods in terms of the ROI coverage and interval/area of the convex hull generated by detected SOIs among disconnected ROIs. This study also provides a flexible framework for ROI detection which can easily incorporate various add-on modules for better performance, such as screening with various criteria for multiple sampled candidates, and subset selection for fitting the surrogate model. This study also shows that our AcqS approach is more robust to the choice of acquisition functions when the number of observations is small.

Because of the nature of the ROI detection problem, our AcqS approach deals with the exploration-exploitation tradeoff differently from existing BO methods. Since the sampling density of AcqS is defined using a given acquisition function, it inherits its exploration-exploitation trade-off property. However, sampling naturally enhances the exploration ability of our AcqS, which is necessary to discover as many as possible ROIs. Moreover, since the ROI detection problem treats all specifications with values below the pre-determined threshold equally, exploitation has a different meaning from what it is for BO. Several of our proposed acquisition functions are specially designed to satisfy the needs of ROI detection.

For future research, better acquisition function design and balanced strategies considering local and global search at different stages can be investigated to improve the sampling efficiency in ROI detection. Relaxing the Gaussian assumption is also under consideration, as a broader set of models can be included with reasonable computational complexity.

References

[1] R. Kanj, R. Joshi, and S. Nassif, "Mixture importance sampling and its application to the analysis of SRAM designs in the presence of rare failure events," in *Proc. 43rd Annu. Conf. Design Autom.*, 2006, pp. 69–72.

[2] S. Sun and X. Li, "Fast statistical analysis of rare circuit failure events via Bayesian scaled-sigma sampling for high-dimensional variation space," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Sep. 2015, pp. 1–4.

[3] A. Singhee and R. A. Rutenbar, "Statistical blockade: Very fast statistical simulation and modeling of rare circuit events and its application to memory design," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 28, no. 8, pp. 1176–1189, Aug. 2009.

[4] M. Moustapha, S. Marelli, and B. Sudret, "Active learning for structural reliability: Survey, general framework and benchmark," *Struct. Saf.*, vol. 96, May 2022, Art. no. 102174.

[5] K. Breitung, "Asymptotic approximations for probability integrals," *Probabilistic Eng. Mech.*, vol. 4, no. 4, pp. 187–190, Dec. 1989.

[6] B. Fiessler, H.-J. Neumann, and R. Rackwitz, "Quadratic limit states in structural reliability," *J. Eng. Mech. Division*, vol. 105, no. 4, pp. 661–676, Aug. 1979.

[7] B. Echard, N. Gayton, and M. Lemaire, "AK-MCS: An active learning reliability method combining Kriging and Monte Carlo simulation," *Struct. Saf.*, vol. 33, no. 2, pp. 145–154, Mar. 2011.

[8] A.-K. E. Haj and A.-H. Soubra, "Improved active learning probabilistic approach for the computation of failure probability," *Struct. Saf.*, vol. 88, Jan. 2021, Art. no. 102011.

[9] X. Yue, Y. Wen, J. H. Hunt, and J. Shi, "Active learning for Gaussian process considering uncertainties with application to shape control of composite fuselage," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 1, pp. 36–46, May 2020.

[10] C. Lee, X. Wang, J. Wu, and X. Yue, "Failure-averse active learning for physics-constrained systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 20, no. 4, pp. 2215–2226, May 2004.

[11] R. J. Adler and J. E. Taylor, *Random Fields and Geometry*, vol. 80. Cham, Switzerland: Springer, 2007.

[12] C. Chevalier, "Fast uncertainty reduction strategies relying on Gaussian process models," Ph.D. thesis, Inst. Math. Statist. Actuarial Sci., Universität Bern, Bern, Switzerland, 2013.

[13] C. Chevalier, J. Bect, D. Ginsbourger, E. Vazquez, V. Picheny, and Y. Richet, "Fast parallel kriging-based stepwise uncertainty reduction with application to the identification of an excursion set," *Technometrics*, vol. 56, no. 4, pp. 455–465, Oct. 2014.

[14] D. Azzimonti, D. Ginsbourger, C. Chevalier, J. Bect, and Y. Richet, "Adaptive design of experiments for conservative estimation of excursion sets," *Technometrics*, vol. 63, no. 1, pp. 13–26, Jan. 2021.

[15] D. Azzimonti, J. Bect, C. Chevalier, and D. Ginsbourger, "Quantifying uncertainties on excursion sets under a Gaussian random field prior," *SIAM/ASA J. Uncertainty Quantification*, vol. 4, no. 1, pp. 850–874, Jan. 2016.

[16] H. Hu, P. Li, and J. Z. Huang, "Parallelizable Bayesian optimization for analog and mixed-signal rare failure detection with high coverage," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2018, pp. 1–8.

[17] H. Hu, P. Li, and J. Z. Huang, "Enabling high-dimensional Bayesian optimization for efficient failure detection of analog and mixed-signal circuits," in *Proc. 56th ACM/IEEE Design Autom. Conf. (DAC)*, Jun. 2019, pp. 1–6.

[18] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Proc. 25th Int. Conf. Neural Inf. Process. Syst.*, vol. 2, 2012, pp. 2951–2959.

[19] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *J. Global Optim.*, vol. 13, no. 4, pp. 455–492, Dec. 1998.

[20] H. J. Kushner, "A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise," *J. Basic Eng.*, vol. 86, no. 1, pp. 97–106, Mar. 1964.

[21] J. Mockus, "On Bayesian methods for seeking the extremum," in *Proc. IFIP Tech. Conf.*, 1974, pp. 400–404.

[22] N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger, "Gaussian process optimization in the bandit setting: No regret and experimental design," 2009, *arXiv:0912.3995*.

[23] P. Hennig and C. J. Schuler, "Entropy search for information-efficient global optimization," *J. Mach. Learn. Res.*, vol. 13, no. 6, pp. 1809–1837, 2012.

[24] F. Berkenkamp, P. Angela Schoellig, and A. Krause, "No-regret Bayesian optimization with unknown hyperparameters," *J. Mach. Learn. Res.*, vol. 20, no. 50, pp. 1–24, 2019.

[25] R. B. Christianson and R. B. Gramacy, "Robust expected improvement for Bayesian optimization," *IISE Trans.*, vol. 56, no. 12, pp. 1294–1306, Dec. 2024.

[26] Z. Xu, Y. Guo, and J. H. Saleh, "Efficient hybrid Bayesian optimization algorithm with adaptive expected improvement acquisition function," *Eng. Optim.*, vol. 53, no. 10, pp. 1786–1804, Oct. 2021.

[27] Z. Wang and S. Jegelka, "Max-value entropy search for efficient Bayesian optimization," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 3627–3635.

[28] D. Hernández-Lobato, J. Hernandez-Lobato, A. Shah, and R. Adams, "Predictive entropy search for multi-objective Bayesian optimization," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1492–1501.

[29] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Proc. 24th Int. Conf. Neural Inf. Process. Syst.*, 2011, pp. 2546–2554.

[30] J. Bergstra, D. Yamins, and D. Cox, "Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures," in *Proc. Int. Conf. Mach. Learn.*, 2013, pp. 115–123.

[31] Y.-A. Ma, Y. Chen, C. Jin, N. Flammarion, and M. I. Jordan, "Sampling can be faster than optimization," *Proc. Nat. Acad. Sci. USA*, vol. 116, no. 42, pp. 20881–20885, Oct. 2019.

[32] X. Gao, M. Gürbüzbalaban, and L. Zhu, "Global convergence of stochastic gradient Hamiltonian Monte Carlo for nonconvex stochastic optimization: Nonasymptotic performance bounds and momentum-based acceleration," *Oper. Res.*, vol. 70, no. 5, pp. 2931–2947, Sep. 2022.

[33] R. M. Neal, "Slice sampling," *Ann. Statist.*, vol. 31, no. 3, pp. 705–767, Jun. 2003.

[34] G. O. Roberts and R. L. Tweedie, "Exponential convergence of Langevin distributions and their discrete approximations," *Bernoulli*, vol. 2, no. 4, p. 341, Dec. 1996.

[35] O. Stramer and R. Tweedie, "Langevin-type models I: Diffusions with given stationary distributions and their discretizations," *Methodol. Comput. Appl. Probab.*, vol. 1, pp. 283–306, Jan. 1999.

[36] R. M. Neal, "MCMC using Hamiltonian dynamics," 2012, *arXiv:1206.1901*.

[37] A. Blanchard and T. Sapsis, "Bayesian optimization with output-weighted optimal sampling," *J. Comput. Phys.*, vol. 425, Jan. 2021, Art. no. 109901.

[38] M. E. Johnson, L. M. Moore, and D. Ylvisaker, "Minimax and maximin distance designs," *J. Stat. Planning Inference*, vol. 26, no. 2, pp. 131–148, Oct. 1990.

[39] M. D. Morris and T. J. Mitchell, "Exploratory designs for computational experiments," *J. Stat. Planning Inference*, vol. 43, no. 3, pp. 381–402, Feb. 1995.

[40] V. R. Joseph, E. Gul, and S. Ba, "Maximum projection designs for computer experiments," *Biometrika*, vol. 102, no. 2, pp. 371–380, Jun. 2015.

[41] T. J. Santner, B. J. Williams, W. I. Notz, T. J. Santner, B. J. Williams, and W. I. Notz, *Space-Filling Designs for Computer Experiments*. Cham, Switzerland: Springer, 2018.

[42] J. Burbea and C. R. Rao, "Entropy differential metric, distance and divergence measures in probability spaces: A unified approach," *J. Multivariate Anal.*, vol. 12, no. 4, pp. 575–596, Dec. 1982.

[43] A. Ullah, "Entropy, divergence and distance measures with econometric applications," *J. Stat. Planning Inference*, vol. 49, no. 1, pp. 137–162, Jan. 1996.

[44] K. T. Abou-Moustafa and F. P. Ferrie, "A note on metric properties for some divergence measures: The Gaussian case," in *Proc. Asian Conf. Mach. Learn.*, 2012, pp. 1–15.

[45] K. Deb and A. Saha, "Multimodal optimization using a bi-objective evolutionary algorithm," *Evol. Comput.*, vol. 20, no. 1, pp. 27–62, 2012.

[46] J. Knowles, D. Corne, and K. Deb, *Multiobjective Problem Solving From Nature: From Concepts to Applications*. Cham, Switzerland: Springer, 2007.

[47] Y. Chen and N. Zhang, "Optimal subsampling for large sample ridge regression," 2022, *arXiv:2204.04776*.

[48] A. Sobester, A. Forrester, and A. Keane, *Engineering Design via Surrogate Modelling: A Practical Guide*. Hoboken, NJ, USA: Wiley, 2008.

[49] N. S. K. Somayaji, H. Hu, and P. Li, "Prioritized reinforcement learning for analog circuit optimization with design knowledge," in *Proc. 58th ACM/IEEE Design Autom. Conf. (DAC)*, Dec. 2021, pp. 1231–1236.

[50] E. Marinari and G. Parisi, "Simulated tempering: A new Monte Carlo scheme," *Europhys. Lett.*, vol. 19, no. 6, pp. 451–458, Jul. 1992.

[51] H. Chernoff, "Sequential design of experiments," *Ann. Math. Statist.*, vol. 30, no. 3, pp. 755–770, Sep. 1959.

[52] J. Kim and S. Choi, "On local optimizers of acquisition functions in Bayesian optimization," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Cham, Switzerland: Springer, 2020, pp. 675–690.

[53] J. M. Hernández-Lobato, M. W. Hoffman, and Z. Ghahramani, "Predictive entropy search for efficient global optimization of black-box functions," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, vol. 1, 2014, pp. 918–926.

[54] M. Betancourt, "A conceptual introduction to Hamiltonian Monte Carlo," 2017, *arXiv:1701.02434*.

[55] M. Jamil and X.-S. Yang, "A literature survey of benchmark functions for global optimization problems," 2013, *arXiv:1308.4008*.

[56] R. M. Neal, "Sampling from multimodal distributions using tempered transitions," *Statist. Comput.*, vol. 6, no. 4, pp. 353–366, Dec. 1996.

[57] S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, May 1983.

[58] V. Nguyen and M. A. Osborne, "Knowing the what but not the where in Bayesian optimization," in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 7317–7326.

[59] J. Liu, J. Paisley, M.-A. Kioumourtzoglou, and B. Coull, "Accurate uncertainty estimation and decomposition in ensemble learning," in *Proc. 33rd Int. Conf. Neural Inf. Process. Syst.*, 2019, pp. 8952–8963.

[60] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas, "Taking the human out of the loop: A review of Bayesian optimization," *Proc. IEEE*, vol. 104, no. 1, pp. 148–175, Aug. 2015.

[61] R. Jordan, D. Kinderlehrer, and F. Otto, "The variational formulation of the Fokker–Planck equation," *SIAM J. Math. Anal.*, vol. 29, no. 1, pp. 1–17, Jan. 1998.

[62] M. D. Hoffman and A. Gelman, "The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1593–1623, 2014.

[63] A. Durmus, E. Moulines, and E. Saksman, "On the convergence of Hamiltonian Monte Carlo," 2017, *arXiv:1705.00166*.

[64] M. M. Graham and A. J. Storkey, "Continuously tempered Hamiltonian Monte Carlo," 2017, *arXiv:1704.03338*.

[65] S. Livingstone, M. Betancourt, S. Byrne, and M. Girolami, "On the geometric ergodicity of Hamiltonian Monte Carlo," *Bernoulli*, vol. 25, no. 4A, pp. 3109–3138, Nov. 2019.

[66] A. Wibisono, "Sampling as optimization in the space of measures: The Langevin dynamics as a composite optimization problem," in *Proc. Conf. Learn. Theory*, 2018, pp. 2093–3027.

[67] Z. Gao, J. Tao, Y. Su, D. Zhou, X. Zeng, and X. Li, "Fast statistical analysis of rare failure events with truncated normal distribution in high-dimensional variation space," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 41, no. 3, pp. 789–793, Mar. 2022.

[68] A. Singhee, J. Wang, B. H. Calhoun, and R. A. Rutenbar, "Recursive statistical blockade: An enhanced technique for rare event simulation with application to SRAM circuit design," in *Proc. 21st Int. Conf. VLSI Design (VLSID)*, 2008, pp. 131–136.

[69] W. Wu, W. Xu, R. Krishnan, Y.-L. Chen, and L. He, "REscope: High-dimensional statistical circuit simulation towards full failure region coverage," in *Proc. 51st ACM/EDAC/IEEE Design Autom. Conf. (DAC)*, Jun. 2014, pp. 1–6.

[70] H. Shin, S. X.-D. Tan, G. Shi, and E. Tlelo-Cuautle, "Rare event diagnosis by iterative failure region locating and elite learning sample selection," in *Proc. 16th Latin-Amer. Test Symp. (LATS)*, Mar. 2015, pp. 1–5.

[71] M. Seeger, "Gaussian processes for machine learning," *Int. J. Neural Syst.*, vol. 14, no. 2, pp. 69–106, 2004.

[72] V. R. Joseph, "Space-filling designs for computer experiments: A review," *Qual. Eng.*, vol. 28, no. 1, pp. 28–35, Jan. 2016.

[73] E. T. Jaynes, "Information theory and statistical mechanics," *Phys. Rev.*, vol. 106, no. 4, pp. 620–630, May 1957.

[74] N. Bou-Rabee, A. Eberle, and R. Zimmer, "Coupling and convergence for Hamiltonian Monte Carlo," *Ann. Appl. Probab.*, vol. 30, no. 3, pp. 1209–1250, Jun. 2020.

[75] R. J. Adler, "On excursion sets, tube formulas and maxima of random fields," *Ann. Appl. Probab.*, vol. 10, no. 1, pp. 1–74, Feb. 2000.

[76] L. Yin, Y. Deng, and P. Li, "Simulation-assisted formal verification of nonlinear mixed-signal circuits with Bayesian inference guidance," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 32, no. 7, pp. 977–990, Jul. 2013.

[77] W. Lyu et al., "An efficient Bayesian optimization approach for automated optimization of analog circuits," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 6, pp. 1954–1967, Jun. 2018.

[78] S.-K.-S. Fan, C.-Y. Hsu, D.-M. Tsai, F. He, and C.-C. Cheng, "Data-driven approach for fault detection and diagnostic in semiconductor manufacturing," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 4, pp. 1925–1936, Oct. 2020.

[79] P. I. Frazier, "A tutorial on Bayesian optimization," 2018, *arXiv:1807.02811*.

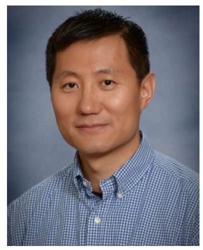
[80] A. G. Zhilinskas, "Single-step Bayesian search method for an extremum of functions of a single variable," *Cybernetics*, vol. 11, no. 1, pp. 160–166, 1976.

[81] M. Moustapha, S. Marelli, and B. Sudret, "A generalized framework for active learning reliability: Survey and benchmark," 2021, *arXiv:2106.01713*.

Huiling Liao received the B.Econ. degree in statistics from the Renmin University of China, Beijing, China, in 2015, and the Ph.D. degree in statistics from Texas A&M University, College Station, TX, USA, in 2022. She was a Post-Doctoral Associate with the Division of Biostatistics and Health Data Science, University of Minnesota, Minneapolis, MN, USA. She is currently an Applied Mathematics Department Assistant Professor with Illinois Institute of Technology, Chicago, IL, USA. Her research interests include machine learning, causal inference, and graph and network data.

Xiaoning Qian (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Yale University, New Haven, CT, USA. He is currently a Professor with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA. He is also affiliated with the Department of Computer Science and Engineering. He holds a joint appointment with the Brookhaven National Laboratory, Upton, NY, USA. His recent honors include the National Science Foundation CAREER Award, Texas A&M Engineering Experiment Station (TEES) Senior Faculty Fellow, and the Montague-Center for Teaching Excellence Scholar at Texas A&M University. His research interests include machine learning and Bayesian computation and their applications in materials and life sciences, computational network biology, and biomedical signal and image analysis.

Jianhua Z. Huang received the B.S. and M.S. degrees in probability and statistics from Beijing University, Beijing, China, in 1989 and 1992, respectively, and the Ph.D. degree in statistics from the University of California at Berkeley, Berkeley, CA, USA, in 1997. He worked with the Wharton School, University of Pennsylvania, and Texas A&M University, before his current position as a Presidential Chair Professor with The Chinese University of Hong Kong, Shenzhen. His research interests include statistics and interdisciplinary research. He is currently a fellow of American Statistical Association and the Institute of Mathematical Statistics and an Elected Member of the International Statistical Institute. He was a recipient of the 2022 INFORMS Impact Prize.



Peng Li (Fellow, IEEE) received the Ph.D. degree from Carnegie Mellon University in 2003. He is currently a Professor of Electrical and Computer Engineering with the University of California at Santa Barbara. His research interests include integrated circuits and systems, electronic design automation, brain-inspired computing, and applied machine learning. His work has been recognized by the ICCAD Ten Year Retrospective Most Influential Paper Award, four IEEE/ACM Design Automation Conference (DAC) Best Paper Awards, and the Best Paper Awards from ICCAD, ICCD, and ASAP. He served as the Vice President for Technical Activities of the IEEE Council on Electronic Design Automation (CEDA).