Efficient Region of Interest Detection under Guidance

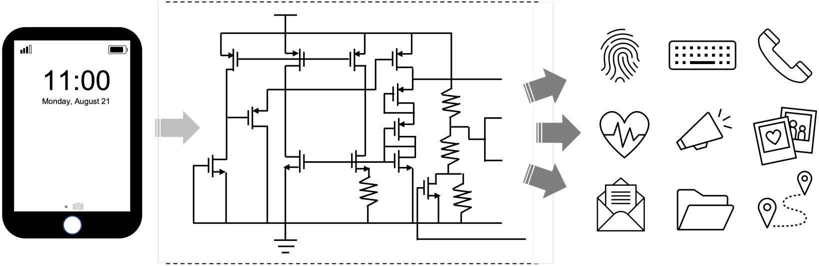
Huiling LIAO
hliao13@iit.edu || RE 234C

Department of Applied Mathematics,
Illinois Institute of Technology

April 9, 2025

Motivation

"Does this proposed design do what is intended? "



$\Rightarrow$ Design v.s. **Verification**

Circuit design

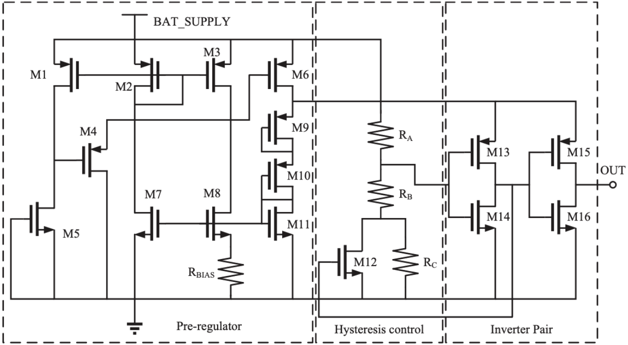


Figure: AMS design for a two-stage differential amplifier circuit diagram.

A figure of merit (FOM) is defined as a combination of four responses to quantify the performance of design.

Motivation

Verification:

- Configurations with extremely low failure rate is preferred
  - e.g. light intensity, the minimum intensity > some threshold?
- Difficulties
  - Large number of variables included
  - Unknown relationships between parameters and effects on the performance
  - Computationally expensive
  - Required extremely low failure rate

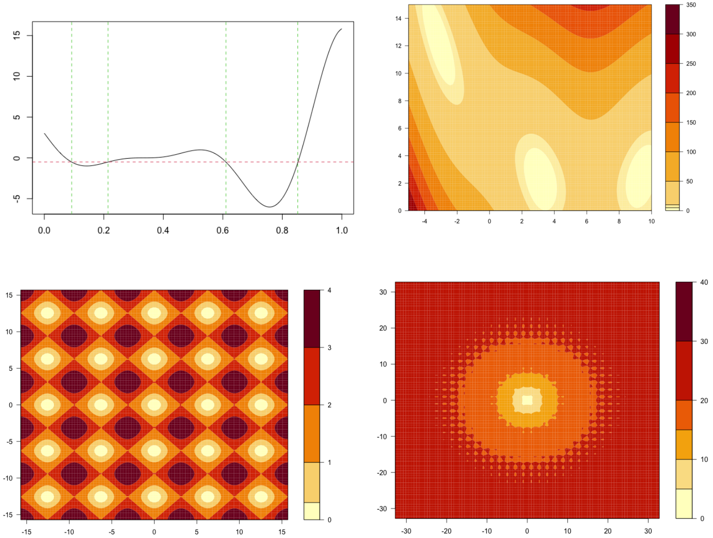
Background

Verification:

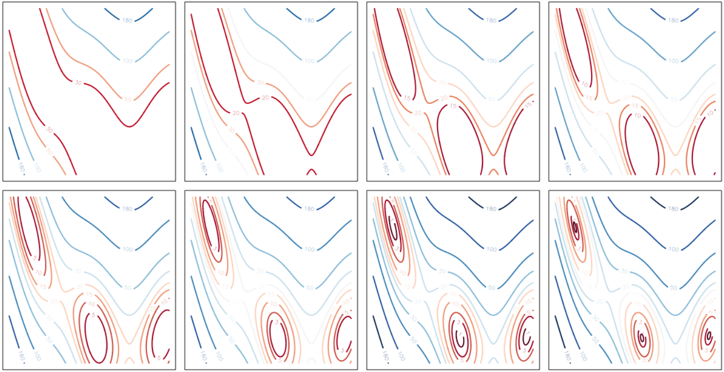
- Methods for failure rate estimation usually sampling-based and assume
  - The region of failure is somehow known
  - Both success and failure cases are observed

When failure is extremely rare, **failure detection** itself becomes a much more fundamental problem than failure rate estimation.

Synthetic Functions



Intuitive Idea



Problem Setting: Start with 1-dim response

Given a $D-$dimension parameter space $\Omega \subset \mathbb{R}^D$, the performance value $y(\mathbf{x})$ at $\mathbf{x} \in \Omega \subset \mathbb{R}^D$ can be determined by some hard-to-evaluate simulator/measurement.

A point can be regarded as a failure/success if

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

where T is the targeted specification (the smaller the worse).

Properties of $y(\cdot)$:

- Highly nonlinear in high dimensional space
- Expensive in terms of simulation or measurement

Problem Setting

Goal: To detect the failure within a certain number of evaluations

- The smaller number of evaluations, the better (Better budget)
- The more, the better (Coverage of ROIs)
- The faster, the better (first hitting time for each region)

△ If in each iteration: we can evaluate the location which is *of the highest potential* to have the “best” performance, we might be able to achieve this goal.

Problem Formulation

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. We can then define the complete set of SOIs of a response $y(\cdot)$ as

$$\begin{aligned} \Gamma_{y,T} &= \{\mathbf{x} \in \Omega : y(\mathbf{x}) < T\}. \\ &= \cup_{k=1}^{n_r} \Gamma_T^{(k)}, \\ \text{where } \Gamma_T^{(i)} \cap \Gamma_T^{(j)} &= \emptyset, \ \forall i \neq j \in \{1, \dots, n_r\}. \end{aligned} \tag{2}$$

$\Rightarrow$ Detect representative SOIs in disjoint regions as fast as more as possible under some constraints

Potential Evaluation Metrics

Notation	Description
t_i	Average 1-st time hitting the i-th ROI or SOI
r_i	Average rate of detecting the i-th ROI
$N_n^r,\ N_n$	Average number of ROIs/SOIs until the n-th evaluation
A_n	Average length $(D = 1)$/area $(D > 1)$ of the convex hull generated by the SOIs in disjointed ROIs in the n-th iteration
$y_{min},\ y_{max}$	The optimal performance value achieved in n_b evaluations

Relationship with existing methods

- **Failure rate estimation** (e.g. MixIS [1], SSS [2], Statistical Blockade [3])

Relationship with existing methods

- **Failure rate estimation** (e.g. MixIS [1], SSS [2], Statistical Blockade [3])

- **Structural reliability analysis** of complex engineering systems (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], FORM and SORM [5, 6]) + active learning (e.g. [7–10]).

Relationship with existing methods

- **Failure rate estimation** (e.g. MixIS [1], SSS [2], Statistical Blockade [3])

- **Structural reliability analysis** of complex engineering systems (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], FORM and SORM [5, 6]) + active learning (e.g. [7–10]).

- **Excursion set**: considers the exceedance probability for a given threshold T of the target function $f(\mathbf{x})$ with the corresponding excursion sets defined as

$$\Gamma_T(f, \Omega) = \{\mathbf{x} \in \Omega : f(x) \geq T\}$$

Relationship with existing methods

- **Failure rate estimation** (e.g. MixIS [1], SSS [2], Statistical Blockade [3])
- **Structural reliability analysis** of complex engineering systems (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], FORM and SORM [5, 6]) + active learning (e.g. [7–10]).
- **Excursion set**: considers the exceedance probability for a given threshold T of the target function $f(\mathbf{x})$ with the corresponding excursion sets defined as

$$\Gamma_T(f, \Omega) = \{\mathbf{x} \in \Omega : f(x) \geq T\}$$

- **Bayesian optimization**: [11, 12] reformulate as global optimization and compare

Relationship with existing methods

- **Failure rate estimation** (e.g. MixIS [1], SSS [2], Statistical Blockade [3])
- **Structural reliability analysis** of complex engineering systems (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], FORM and SORM [5, 6]) + active learning (e.g. [7–10]).
- **Excursion set**: considers the exceedance probability for a given threshold T of the target function $f(\mathbf{x})$ with the corresponding excursion sets defined as

$$\Gamma_T(f, \Omega) = \{\mathbf{x} \in \Omega : f(x) \geq T\}$$

- **Bayesian optimization**: [11, 12] reformulate as global optimization and compare

However, most existing studies aim for the accurate estimation of excursion sets (e.g. [13–15]) and uncertainty quantification (e.g. [16]), both necessarily requiring a large sample size and with goals diverge from ours.

Bayesian Decision Making

1. Specify the probability space of all the possible outcomes and decisions
2. Determine the probability distribution of outcomes given each decision option
3. Define a utility function mapping outcome onto real numbers
4. Compute the expected utility as a function of a given decision and choose the decision with the best expected utility.

Bayesian Optimization

A sequential learning strategy to globally optimize function

$$\min_{x \in \Omega} f(\mathbf{x}), \text{ or } \max_{x \in \Omega} f(\mathbf{x})$$

f typically have the following properties:

- Continuous
- Expensive to evaluate (time, computational resource, cost, etc)
- "Black box", lack known structure, hard to optimize
- Only observe (noisy) $f(\mathbf{x})$, no first- or second-order derivatives

Bayesian Optimization: Algorithm

To solve this problem, the BO algorithm is as following:

Algorithm Bayesian Optimization for Rare Event Detection

Choose a **surrogate model** for modeling f;

while *stopping criterion is not met* **do**

  Determine the next point to sample by **acquisition function** ; `/* inner optim. */`

  Add the newly sampled data to the set of observations;

  Update model;

Compare $y_{\min} = \min_{t \in \{1,\ldots,n_b\}} y_t$ with T to conclude.

- Two essential components $\begin{cases} \textbf{Surrogate model} \\ \textbf{Acquisition function} \end{cases}$

Bayesian optimization: Surrogate model

Gaussian process regression is a typical choice of the surrogate statistical model. Assume $\mathbf{X}_{1:n} = [\mathbf{x}_1, \dots, \mathbf{x}_n], \mathbf{x}_i \in \mathbb{R}^D, i = 1, \dots, n$, and their corresponding function values are denoted as $\mathbf{f} = [f(\mathbf{x}_1), \dots, f(\mathbf{x}_n)]^T$.

$$\mathbf{f}(\mathbf{X}) \sim \mathcal{GP}(\mu_0(\mathbf{X}), \Sigma_0(\mathbf{X}, \mathbf{X}))$$

Bayesian optimization: Surrogate model

Gaussian process regression is a typical choice of the surrogate statistical model. Assume $\mathbf{X}_{1:n} = [\mathbf{x}_1, \ldots, \mathbf{x}_n], \mathbf{x}_i \in \mathbb{R}^D, i = 1, \ldots, n$, and their corresponding function values are denoted as $\mathbf{f} = [f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)]^T$.

$$\mathbf{f}(\mathbf{X}) \sim \mathcal{GP}(\mu_0(\mathbf{X}), \Sigma_0(\mathbf{X}, \mathbf{X}))$$

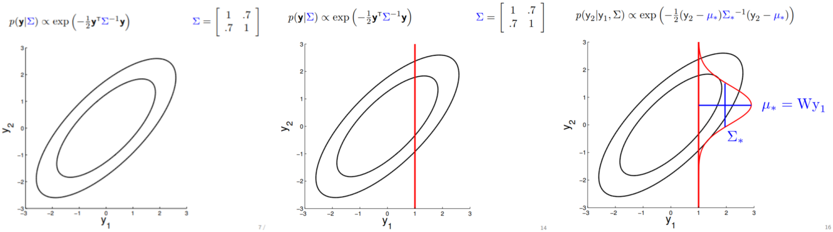
"A Gaussian process is a collection of random variables, any finite number of which have consistent Gaussian distributions."

Bayesian optimization: Surrogate model

Gaussian process regression is a typical choice of the surrogate statistical model. Assume $\mathbf{X}_{1:n} = [\mathbf{x}_1, \ldots, \mathbf{x}_n], \mathbf{x}_i \in \mathbb{R}^D, i = 1, \ldots, n$, and their corresponding function values are denoted as $\mathbf{f} = [f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)]^T$.

$$\mathbf{f}(\mathbf{X}) \sim \mathcal{GP}(\mu_0(\mathbf{X}), \Sigma_0(\mathbf{X}, \mathbf{X}))$$

"A Gaussian process is a collection of random variables, any finite number of which have consistent Gaussian distributions."



Bayesian optimization: Acquisition function

Exploration v.s. Exploitation

(point with high uncertainty) (point with best expected value)

- LCB:

$$\text{LCB}(\mathbf{x}|\mathcal{D}_n) = \mu(\mathbf{x}|\mathcal{D}_n) - \kappa\sigma(\mathbf{x}|\mathcal{D}_n)$$

- Expected Improvement (EI):

$$\text{EI}(\mathbf{x}|\mathcal{D}_n) = E_n[(f^* - f(\mathbf{x}))_+|\mathcal{D}_n]$$

- Probability of Improvement (PI):

$$\text{PoI}(\mathbf{x}|\mathcal{D}_n) = P(f(\mathbf{x}) < f^*|\mathcal{D}_n)$$

- (Maximal/Predicted) Entropy Search

Bayesian optimization: Acquisition function

- **Expected Improvement (EI)**:
  Suppose we have one additional evaluation to perform, then our choice after this evaluation will be either the newly evaluated point $\mathbf{x}$ or the previously best point $\mathbf{x}_n^*$ with function value f_n^*. The improvement of this point is

$$(f_n^* - f(\mathbf{x}))_+ = \begin{cases} f_n^* - f(\mathbf{x}), & \text{if } f(\mathbf{x}) < f_n^* \\ 0 & \text{if } f(\mathbf{x}) >= f_n^* \end{cases}$$

$$\begin{aligned} \implies \mathrm{EI}(\mathbf{x}|\mathcal{D}_n) &= E_n[(f^* - f(\mathbf{x}))_+|\mathcal{D}_n] \\ &= \int (f^* - f(\mathbf{x}))_+ p(f(\mathbf{x})|\mathcal{D}_n) df(\mathbf{x}) \\ &= \sigma(\mathbf{x}) \left(z(\mathbf{x})\Phi(z) + \phi(z(\mathbf{x})) \right), \text{ where } z(\mathbf{x}) = \frac{f_n^* - \mu(\mathbf{x})}{\sigma(\mathbf{x})} \end{aligned}$$

Research problems of interest

Overall goal: with a budget constraint, design an algorithm to detect as many ROIs as possible.

Research problems of interest

Overall goal: with a budget constraint, design an algorithm to detect as many ROIs as possible.

- Without the constraint of sample size, can we propose an algorithm to achieve a high coverage of ROIs through

Research problems of interest

Overall goal: with a budget constraint, design an algorithm to detect as many ROIs as possible.

- Without the constraint of sample size, can we propose an algorithm to achieve a high coverage of ROIs through
  - (possibly experimental design without an accurate recovery of excursion set, such as qMC with different discrepancy, shrinking level set, partition/reweighting different parts of space)
  - In the framework of Bayesian Optimization, design the proper acquisition function
  - Sampling or Optimization (design distribution)

Research problems of interest

Overall goal: with a budget constraint, design an algorithm to detect as many ROIs as possible.

- Without the constraint of sample size, can we propose an algorithm to achieve a high coverage of ROIs through
  - (possibly experimental design without an accurate recovery of excursion set, such as qMC with different discrepancy, shrinking level set, partition/reweighting different parts of space)
  - In the framework of Bayesian Optimization, design the proper acquisition function
  - Sampling or Optimization (design distribution)
- Develop the corresponding theoretical guarantee
  - (through reaching different components of excursion set for Gaussian random field, first hitting time, ...

Research problems of interest

Overall goal: with a budget constraint, design an algorithm to detect as many ROIs as possible.

- Without the constraint of sample size, can we propose an algorithm to achieve a high coverage of ROIs through
  - (possibly experimental design without an accurate recovery of excursion set, such as qMC with different discrepancy, shrinking level set, partition/reweighting different parts of space)
  - In the framework of Bayesian Optimization, design the proper acquisition function
  - Sampling or Optimization (design distribution)
- Develop the corresponding theoretical guarantee
  - (through reaching different components of excursion set for Gaussian random field, first hitting time, . . .
- With limited evaluation time, how can we extend the design?

Research problems of interest

Overall goal: with a budget constraint, design an algorithm to detect as many ROIs as possible.

- Without the constraint of sample size, can we propose an algorithm to achieve a high coverage of ROIs through
  - (possibly experimental design without an accurate recovery of excursion set, such as qMC with different discrepancy, shrinking level set, partition/reweighting different parts of space)
  - In the framework of Bayesian Optimization, design the proper acquisition function
  - Sampling or Optimization (design distribution)
- Develop the corresponding theoretical guarantee
  - (through reaching different components of excursion set for Gaussian random field, first hitting time, . . .
- With limited evaluation time, how can we extend the design?
- Multiple objectives: y being multivariate