Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

MULTI-FIDELITY SIMULATION OPTIMIZATION WITH LEVEL SET APPROXIMATION USING PROBABILISTIC BRANCH AND BOUND

David D. Linz

University of Washington
Department of Industrial
and Systems Engineering
Seattle, WA 98195-2650, USA

Hao Huang

Yuan Zu University
Department of Industrial Engineering
and Management
Taoyuan City 320, TAIWAN

Zelda B. Zabinsky

University of Washington
Department of Industrial and Systems Engineering
Seattle, WA 98195-2650, USA

ABSTRACT

Simulated systems are often described with a variety of models of different complexity. Making use of these models, algorithms can use low complexity, "low-fidelity" models or meta-models to guide sampling for purposes of optimization, improving the probability of generating good solutions with a small number of observations. We propose an optimization algorithm that guides the search for solutions on a high-fidelity model through the approximation of a level set from a low-fidelity model. Using the Probabilistic Branch and Bound algorithm to approximate a level set for the low-fidelity model, we are able to efficiently locate solutions inside of a target quantile and therefore reduce the number of high-fidelity evaluations needed in searches. The paper provides an algorithm and analysis showing the increased probability of sampling high-quality solutions within a low-fidelity level set. We include numerical examples that demonstrate the effectiveness of the multi-fidelity level set approximation method to locate solutions.

1 INTRODUCTION

Simulation models have been shown to be widely applicable in a number of fields that involve complex and random systems; subsequently optimizing simulation models can be a computationally expensive problem. However, in many contexts simulated systems have a range of different models with varying complexities, or have simpler meta-models that track the behavior of the system (Barton and Meckesheimer 2006). As such, experts looking to optimize a system can make use of lower-complexity models to guide their analysis of high-complexity systems.

Under a multi-fidelity approach to simulation optimization, a high-fidelity model provides an accurate response at a high computational cost, whereas a low-fidelity model, or a meta-model, has a marginally less reliable response at a lower computational cost. Since both models represent the same system, high and low-fidelity outputs are presumed to have a strong similarity. Using this relationship, multi-fidelity optimization employs the low-fidelity simulation (or a meta-model) to guide the optimization of a high-fidelity model, reducing the computational time required to generate good or optimal solutions to the problem (Xu et al. 2014).

The multi-fidelity approach has become popular in a number of applications (Molina-Cristóbal et al. 2010, Huang et al. 2015, Chiu et al. 2016, Qiu et al. 2016). Similarly, multi-fidelity approaches have taken a variety of forms. One approach emphasizes efficient sampling of the high-fidelity models using an ordinal transformation based on a low-fidelity model (Xu et al. 2014, Li et al. 2015, Li et al. 2016, Xu et al. 2016). Alternative methods have used multi-fidelity models as a basis for constructing meta-models for optimizing expensive high-fidelity functions (Forrester et al. 2007). However, to date, simulation optimization methods have yet to explore the use of multi-fidelity models for partitioning algorithms to focus searching a high-fidelity function's domain for optimization.

Our paper extends the multi-fidelity approach using the Probabilistic Branch and Bound (*PBnB*) (Huang and Zabinsky 2013) algorithm to create a level set approximation of a low-fidelity function to focus the sampling of a high-fidelity function. Specifically, the algorithm creates a level set approximation for some best $\delta \in (0,1)$ percent of solutions on the low-fidelity model in a common domain. Given a large amount of similarity between the behavior of low-fidelity and high-fidelity models, the developed level set for the low-fidelity model will likely contain many good points when evaluated with the high-fidelity model. Therefore, the low-fidelity model can be used to significantly improve high-fidelity searches for optimal or near-optimal solutions.

Results are outlined as follows. Section 2 describes the problem statement along with the basic assumptions concerning the relationship between low and high-fidelity models. We subsequently outline the algorithm in Section 3 with theorems describing the increased probability of sampling points in the target set presented in Section 4. We present numerical results on selected test functions in Section 5. A general discussion of results including possible next steps is included in Section 6.

2 OVERLAPPING LEVEL SETS OF MULTI-FIDELITY OPTIMIZATION PROBLEMS

We consider a high-fidelity black box function f_H with an accompanying minimization problem

$$min_x \ f_H(x) \tag{1}$$

$$\text{s.t. } x \in S$$

where $S \subset \mathbb{R}^n$ and S is defined by box-constraints. We are interested in sampling solutions in some "target" set that is comprised of points that fall into some low, δ_{target} quantile of f_H ($0 \leq \delta_{target} \leq 1$), namely $x \in L_H(\delta_{target}, S)$ where $L_H(\delta_{target}, S) = \{x : f_H(x) \leq y_H(\delta_{target}, S)\}$ and $y_H(\delta_{target}, S)$ is the δ_{target}-quantile of the domain S, or explicitly:

$$y_H(\delta_{target}, S) = argmin_{y \in \{f_H(x): x \in S\}} \{P(f_H(X) \leq y) \geq \delta_{target}\}$$

where X is uniformly distributed on S. In addition to the high-fidelity function f_H, we also have a low-fidelity function such that f_L, defined on the same domain. We note that for quantile level, δ, $\delta = \frac{\nu(L_H(\delta,S))}{\nu(S)}$.

In order for the low-fidelity model to give some indication of where the target set resides in S, we assume that there is some δ_{shared} level set, such that $\delta_{shared} \geq \delta_{target}$, where both functions overlap. Specifically, let $L_L(\delta_{shared}, S)$ be the level set for the low-fidelity model and define the overlap as:

$$\gamma = \frac{\nu\left(L_H(\delta_{shared}, S) \bigcap L_L(\delta_{shared}, S)\right)}{\nu\left(L_H(\delta_{shared}, S)\right)} \tag{2}$$

where ν is a volume measure. We assume that the overlap ratio γ is of some sufficient size $0 < \gamma \leq 1$.

Assumption 1. *We assume a minimum value of the overlap ratio:*

$$\gamma \geq 1 + \delta_{target} - \frac{\delta_{target}}{\delta_{shared}}. \tag{3}$$

Alternatively, the assumption could be restated in terms of our desired δ_{target}. For a given overlap γ at some δ_{shared}, the assumption will indicate a minimum value for the δ_{target} quantile:

$$\delta_{target} \geq \frac{(1-\gamma)\cdot\delta_{shared}}{(1-\delta_{shared})}.$$

Assumption 1 indicates that the high-fidelity function behavior has a significant amount of overlap at a given δ_{shared} quantile. We define concentration on some given set σ as being the ratio of volume from the set $L_H(\delta_{target}, S)$ in the set σ, or $\frac{\nu(L_H(\delta_{target},S)\bigcap\sigma)}{\nu(\sigma)}$. Domains with high concentration of points will be more efficiently searched by random search algorithms. Based on the assumption, we show that the level set $L_L(\delta_{shared}, S)$ has a greater "concentration" of points from $L_H(\delta_{target}, S)$ than S in Section 4.

We can visualize improved concentration in the example shown in Figure 1. In Figure 1(a) we represent high and low-fidelity level sets on a two-dimensional domain. The high-fidelity model shares a significant amount of its δ_{shared} level set with the low-fidelity model. Due to the significant amount of overlap, there is a large concentration of points from δ_{target} level set of the high-fidelity model within the δ_{shared} level set of the low-fidelity model. Figure 1(b) illustrates high and low-fidelity function responses with corresponding level sets.

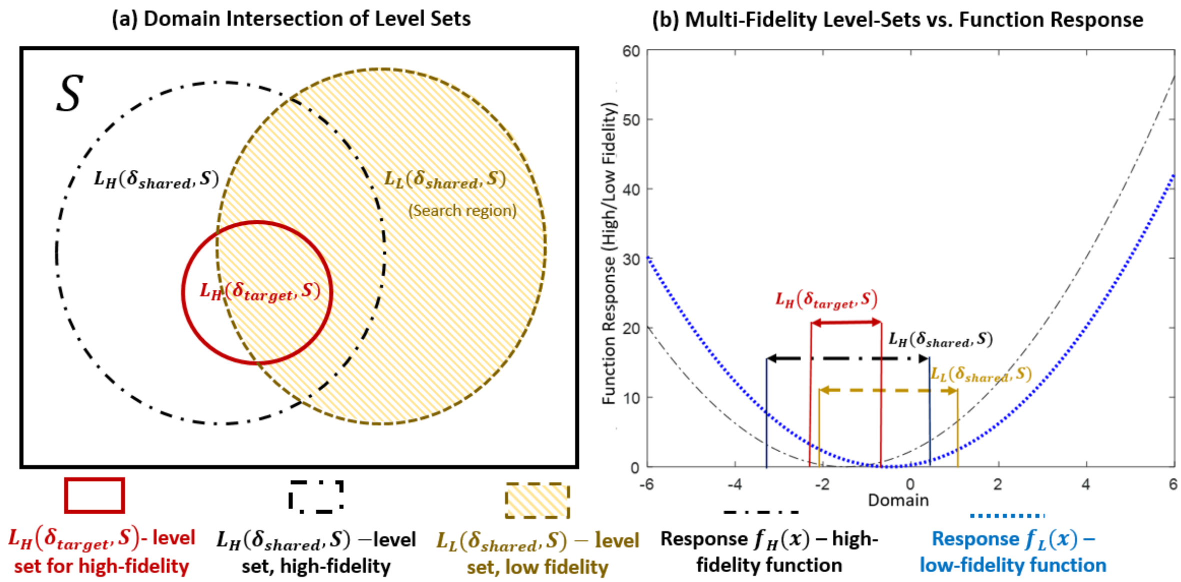

Figure 1: A graph of intersecting level sets between a high-fidelity function and a low-fidelity function on a common domain. In (a), $L_L(\delta_{shared}, S)$ is outlined with a dashed line, $L_H(\delta_{shared}, S)$ is outlined with a dash-dot line, and $L_H(\delta_{target}, S)$ is outlined with a solid line. In (b), the level sets are shown relative to their function responses.

Since we can evaluate f_L much faster than f_H, a fast approximation of $L_L(\delta_{shared}, S)$ level set will enable an algorithm to be more likely to sample points in $L_H(\delta_{target}, S)$ than sampling from the entire domain.

3 OPTIMIZATION ALGORITHM USING LOW-FIDELITY MODELS TO GUIDE HIGH-FIDELITY SAMPLING

The general multi-fidelity optimization approach exploits the computationally inexpensive approximation of a low-fidelity level set to subsequently guide sampling of the high-fidelity function. To accomplish this, we make use of the Probabilistic Branch and Bound (*PBnB*) algorithm for level set approximation outlined by (Huang and Zabinsky 2013). Using the low fidelity function, another search algorithm can optimize the high-fidelity function on the resulting level-set approximation which will likely contain a high concentration of good solutions. We outline the method for searching a domain for high-fidelity solutions within a target level set in algorithm 1:

Algorithm 1. ***High-Fidelity Random Search Optimization through Low-Fidelity Level Set Generation***

For a given low-fidelity budget N_{low} function evaluations and a high-fidelity budget level N_{high} function evaluations

1. ***Initialize Parameters:*** *Select B, δ_{shared}, δ_{target}, ε, and α parameters.*
2. ***Run PBnB****: Implement a level set approximation using the PBnB algorithm on the low-fidelity function, f_L, on the domain S for k iterations until the algorithm reaches N_{low} function evaluations. The PBnB algorithm outputs three disjoint subregions, Σ_k^p, Σ_k^m, and Σ_k^c (corresponding respectively to the pruned, maintained, and undecided sub-regions such that after k iterations*
$$S = \Sigma_k^p \cup \Sigma_k^m \cup \Sigma_k^c.$$
Let $\tilde{\Sigma}_k = S \backslash \Sigma_k^p = \Sigma_k^m \cup \Sigma_k^c$ be the "level set approximation" at iteration k such that $\tilde{\Sigma}_k \approx L_L(\delta_{shared}, S)$.
3. ***Search High Fidelity Model:*** *Run a search algorithm for N_i function-evaluations of the high fidelity model f_H on each sub-region σ_i in Σ_k, where $N_i = \frac{\nu(\sigma_i)}{\nu(\tilde{\Sigma}_k)} \cdot N_{high}$.*
4. ***Output:*** *Rank the N_{high} candidate points by their high-fidelity function values and return the best δ_{target} percent.*

The approximation of the $L_L(\delta_{shared}, S)$, in Step 2, is accomplished by the PBnB algorithm which repeatedly estimates the quantile and then determines whether to maintain or prune regions based on whether they fall within the estimated level set. The *PBnB* algorithm requires parameters: B, the number of divisions across a given dimension to create a new set of sub-regions when branching; α, a performance confidence level; ε, an upper bound of the volume incorrectly estimated; and, δ_{shared}, the quantile of the level set that needs to be approximated. The fully specified algorithm can be found in Huang and Zabinsky (2013).

4 ANALYSIS SHOWING INCREASED CONCENTRATION OF TARGET SOLUTIONS ON LOW-FIDELITY LEVEL SETS

We use Assumption 1 to develop theorems concerning the concentration of points from $L_H(\delta_{target}, S)$ that can be found in $L_L(\delta_{shared}, S)$ (or an approximation of $L_L(\delta_{shared}, S)$). We expect the efficiency of the random search in Step 4 of Algorithm 1 to be generally improved by a higher concentration of points in $L_L(\delta_{shared}, S)$ since the probability of uniformly sampling a point from the target level set will be equal to the concentration.

Theorem 1. *The volume intersection of the $L_H(\delta_{target}, S)$ and the $L_L(\delta_{shared}, S)$ level sets has a lower bound, such that:*

$$\nu(L_H(\delta_{target}, S) \bigcap L_L(\delta_{shared}, S)) \geq \nu(S) \cdot (\delta_{target} - \delta_{shared} \cdot (1-\gamma)). \tag{4}$$

Proof. We can express the target level set for the high fidelity model as a union of two disjoint subsets $L_H(\delta_{target}, S) = (L_H(\delta_{target}, S) \bigcap L_L(\delta_{shared}, S)) \bigcup (L_H(\delta_{target}, S) \backslash L_L(\delta_{shared}, S))$. Therefore we can write:

$$\nu\left(L_H(\delta_{target}, S) \bigcap L_L(\delta_{shared}, S)\right) + \nu\left(L_H(\delta_{target}, S) \backslash L_L(\delta_{shared}, S)\right) = \nu\left(L_H(\delta_{target}, S)\right)$$

$$\nu\left(L_H(\delta_{target}, S) \bigcap L_L(\delta_{shared}, S)\right) = \nu\left(L_H(\delta_{target}, S)\right) - \nu\left(\left(L_H(\delta_{target}, S) \backslash L_L(\delta_{shared}, S)\right)\right)$$

Since $\nu\left((L_H(\delta_{shared},S)\backslash L_L(\delta_{shared},S))\right) \geq \nu\left((L_H(\delta_{target},S)\backslash L_L(\delta_{shared},S))\right)$, we can lower bound the expression with:

$$\nu\left(L_H(\delta_{target},S)\bigcap L_L(\delta_{shared},S)\right) \geq \nu\left(L_H(\delta_{target},S)\right) - \nu\left((L_H(\delta_{shared},S)\backslash L_L(\delta_{shared},S))\right).$$

We can then write $(L_H(\delta_{shared},S)\backslash L_L(\delta_{shared},S))$ in terms of the intersection $L_H(\delta_{shared},S)\bigcap L_L(\delta_{shared},S)$

$$\begin{aligned}
&\nu\left(L_H(\delta_{target},S)\right) - \nu\left((L_H(\delta_{shared},S)\backslash L_L(\delta_{shared},S))\right)\\
&= \nu\left(L_H(\delta_{target},S)\right) - \nu\left(L_H(\delta_{shared},S)\backslash\left(L_H(\delta_{shared},S)\bigcap L_L(\delta_{shared},S)\right)\right)\\
&= \nu\left(L_H(\delta_{target},S)\right) - \left(\nu\left(L_H(\delta_{shared},S)\right) - \nu\left(L_H(\delta_{shared},S)\bigcap L_L(\delta_{shared},S)\right)\right)\\
&= \nu\left(L_H(\delta_{target},S)\right) - \nu\left(L_H(\delta_{shared},S)\right) + \nu((L_H(\delta_{shared},S)\bigcap L_L(\delta_{shared},S))
\end{aligned}$$

and replacing from the definition in (2) we get

$$= \nu\left(L_H(\delta_{target},S)\right) - \nu\left(L_H(\delta_{shared},S) + \gamma\cdot\nu(L_H(\delta_{shared},S)\right)$$

Using the relationship between quantile and domain volume, we substitute to get:

$$\begin{aligned}
&= \delta_{target}\cdot\nu\left(S\right) - \delta_{shared}\cdot\nu\left(S\right) + \gamma\cdot\delta_{shared}\cdot\nu(S)\\
&= \nu\left(S\right)\cdot\left(\delta_{target} - \delta_{shared}\cdot(1-\gamma)\right)
\end{aligned}$$

■

Corollary 1. *Given Assumption 1, the ratio of the concentration of target level set of the high-fidelity function inside $L_L(\delta_{shared},S)$ is greater than δ_{target} that is,*

$$\frac{\nu(L_H(\delta_{target},S)\bigcap L_L(\delta_{shared},S))}{\nu(L_L(\delta_{shared},S))} \geq \delta_{target}.$$

Proof. Starting with Assumption 1,

$$\gamma \geq 1 + \delta_{target} - \frac{\delta_{target}}{\delta_{shared}}$$

and rearranging the terms, we get

$$\frac{\delta_{target} - \delta_{shared}\cdot(1-\gamma)}{\delta_{shared}} \geq \delta_{target}.$$

Noting that $\frac{1}{\delta_{shared}} = \frac{\nu(S)}{\nu(L_L(\delta_{shared},S))}$, we get:

$$\delta_{target} \leq \frac{\delta_{target} - \delta_{shared}\cdot(1-\gamma)}{\delta_{shared}} = \frac{\nu(S)}{\nu(L_L(\delta_{shared},S))}\left(\delta_{target} - \delta_{shared}\cdot(1-\gamma)\right).$$

From the relationship in Theorem 1 we get

$$\delta_{target} \leq \frac{\nu(S)}{\nu(L_L(\delta_{shared},S))}\left(\delta_{target} - \delta_{shared}\cdot(1-\gamma)\right) \leq \frac{\nu(L_H(\delta_{target},S)\bigcap L_L(\delta_{shared},S))}{\nu(L_L(\delta_{shared},S))}$$

which demonstrates the proposition.

■

Corollary 1 provides a theoretical lower bound on the concentration of points from the level set $L_H(\delta_{target}, S)$ in the low-fidelity level set $L_L(\delta_{shared}, S)$ which under conditions of Assumption 1 will be greater than the concentration of points from $L_H(\delta_{target}, S)$ in the domain S (δ_{target} by definition). The last theorem extends this result to characterize a lower bound on the concentration of points from $L_H(\delta_{target}, S)$ inside the level set approximation $\tilde{\Sigma}_k$ generated by the PBnB algorithm.

Theorem 2. *After the application of the PBnB algorithm for k iterations, the concentration of points from the δ_{target} level set in the level set approximation $\tilde{\Sigma}_k$ has a lower bound,*

$$\frac{\nu\left(\tilde{\Sigma}_k \bigcap L_H(\delta_{target}, S)\right)}{\nu(\tilde{\Sigma}_k)} \geq \frac{(\delta_{target} - \delta_{shared} \cdot (1-\gamma)) - \varepsilon}{1 - \frac{\nu(\Sigma_k^P)}{\nu(S)}}$$

with probability $(1-\alpha)^4$, where Σ_k^P is the set of pruned sub-regions after k iterations of the PBnB algorithm.

Proof. By definition of the level set approximation $\tilde{\Sigma}_k$, we note that

$$\nu\left(\tilde{\Sigma}_k\right) = \nu(S) - \nu\left(\Sigma_k^P\right). \tag{5}$$

$$\nu\left(\tilde{\Sigma}_k \bigcap L_H(\delta_{target}, S)\right) = \nu\left(L_H(\delta_{target}, S) \backslash \Sigma_k^P\right) = \nu\left(\left(L_H(\delta_{target}, S) \bigcap L_L(\delta_{shared}, S)\right) \backslash \Sigma_k^P\right)$$

This is strictly

$$\begin{aligned} &= \nu\left(L_H(\delta_{target}, S) \bigcap L_L(\delta_{shared}, S) \backslash \left(\Sigma_k^P \bigcap L_L(\delta_{shared}, S)\right)\right) \\ &= \nu\left(L_H(\delta_{target}, S) \bigcap L_L(\delta_{shared}, S)\right) - \nu\left(\Sigma_k^P \bigcap L_L(\delta_{shared}, S)\right) \end{aligned}$$

and applying Theorem 1 we get

$$\geq \nu(S) \cdot (\delta_{target} - \delta_{shared} \cdot (1-\gamma)) - \nu\left(\Sigma_k^P \bigcap L_L(\delta_{shared}, S)\right).$$

Theorem 6 from (Huang and Zabinsky 2013) which states that at iteration k of the PBnB algorithm, the volume of the incorrectly pruned regions does not exceed ε with probability $(1-\alpha)^4$, that is,

$$P\left(\nu\left(L_L(\delta_{shared}, S) \bigcap \Sigma_k^P\right) \leq \varepsilon\right) \geq (1-\alpha)^4,$$

and combined with above, results in

$$\nu(S) \cdot (\delta_{target} - \delta_{shared} \cdot (1-\gamma)) - \nu\left(\Sigma_k^P \bigcap L_L(\delta_{shared}, S)\right) \geq \nu(S) \cdot (\delta_{target} - \delta_{shared} \cdot (1-\gamma)) - \varepsilon \tag{6}$$

with probability $(1-\alpha)^4$. Combining (4) with (6) we obtain

$$\nu\left(\tilde{\Sigma}_k \bigcap L_H(\delta_{target}, S)\right) \geq \nu(S) \cdot (\delta_{target} - \delta_{shared} \cdot (1-\gamma)) - \varepsilon$$

and dividing by $\nu\left(\tilde{\Sigma}_k\right)$ and using (5), we get

$$\frac{\nu\left(\tilde{\Sigma}_k \bigcap L_H(\delta_{target}, S)\right)}{\nu(\tilde{\Sigma}_k)} \geq \frac{(\delta_{target} - \delta_{shared} \cdot (1-\gamma)) - \varepsilon}{1 - \frac{\nu(\Sigma_k^P)}{\nu(S)}}$$

with probability $(1-\alpha)^4$. ■

Theorems 1 and 2 imply that given values of α and ε, we can get a level set approximation $\tilde{\Sigma}_k$ with a significantly higher concentration of solutions in the target quantile, $L_H(\delta_{target}, S)$, than in S.

5 NUMERICAL TESTS

We explore the utility of level set approximation for multi-fidelity optimization with numerical tests on common test functions. Similar to previous multi-fidelity models, the use of numerical test functions allows an examination of optimization performance for closely related "high" fidelity and "low" fidelity functions. While the proposed algorithm could be used for mixed integer and continuous variables, the initial tests focus on problems within a continuous domain without noise. First, we examine the one dimensional test example used in Xu et al. (2016), with the next section looking at "high" fidelity and "low" fidelity functions based on a set of common multi-dimensional non-convex test functions from Ali et al. (2005).

5.1 Test Example in One Dimension

The algorithm is initially tested on the function used in Xu et al. (2016), where the high-fidelity function can be expressed as a one-dimensional analytical function:

$$f_H(x) = \frac{sin^6(0.09\pi x)}{2^2\left(2^{2((x-10)/80)^2}\right)} + 0.1 \cdot cos(0.5\pi) + 0.5 \cdot \left(\frac{x-40}{60}\right) + 0.4 \cdot sin\left(\frac{x+10}{100}\pi\right). \tag{7}$$

For purposes of this paper's focus on continuous domains, we use the domain $S = [0, 100]$. We subtract the original function from the maximum 1.4277 to create a positive value minimization problem with a global minimum of 0. The corresponding low-fidelity function to the high-fidelity function in (7) is developed by dropping all but the first terms, so that $f_L(x) = \frac{sin^6(0.09\pi x)}{2^2\left(2^{2((x-10)/80)^2}\right)}$.

For analysis, we approximate the overlap between $L_H(\delta_{shared}, S)$ and $L_L(\delta_{shared}, S)$ for $\delta_{shared} = 0.3$ as $\hat{\gamma}$, such that

$$\hat{\gamma} \approx \frac{\nu(L_H(\delta_{shared}, S) \bigcap L_L(\delta_{shared}, S))}{\nu(L_H(\delta_{shared}, S)}. \tag{8}$$

We determine $\hat{\gamma} = 0.90$, for this example, by dividing the domain into 100 regions and approximating whether each falls into $L_H(\delta_{shared}, S)$ or $L_L(\delta_{shared}, S)$ based on a uniform sample of 8000 points. Therefore after applying Algorithm 1 with adequate level set approximation, we are guaranteed to improve the concentration of points within target-quantiles greater than 0.04 based on Corollary 1.

Setting $\delta_{shared} = 0.3$, $B = 2, \alpha = 0.1$, $N_{low} = 50000$, we run the PBnB algorithm to generate an level set approximation $\tilde{\Sigma}_k$ of $L_L(\delta_{shared}, S)$. Subsequently, 10000 high-fidelity evaluations are taken both from $\tilde{\Sigma}_k$ (as specified in Step 4. of Algorithm 1) and from S, uniformly.

Conducting 30 experiments, we can obtain the number of points that are sampled inside a range of target quantiles $(10\%, 5\%, 1\%)$ relative to the number of high-fidelity function evaluations. Figure 2 plots the number of points sampled within three different quantiles on the "high-fidelity" function based on sampling from the entire domain, S, (solid lines) and sampling from $\tilde{\Sigma}_k$ (broken lines). Figure 2 demonstrates that, for an equivalent number of sample points, sampling on $\tilde{\Sigma}_k$ results in significantly more points being found inside the 10%, 5% 1% target quantiles.

We observe generally that sampling on the level set approximation $\tilde{\Sigma}_k$ will improve the probability of sampling points inside the $L_H(\delta_{target}, S)$ for $\delta_{target} \geq 0.04$ by Corollary 1. However, from Figure 2, we notice that sampling points inside $L_H(\delta_{target}, S)$ with $\delta_{target} = 0.01$ is improved even without the theoretical guarantee.

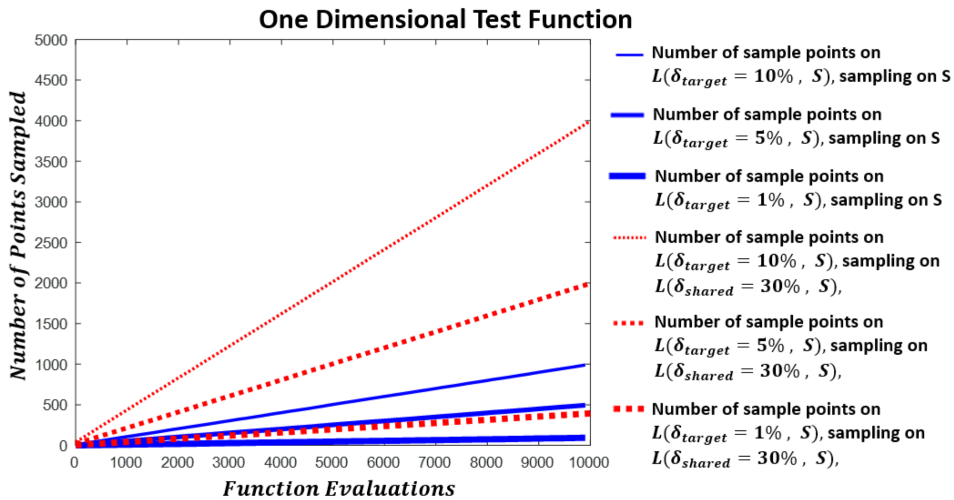

Figure 2: Number of sample points within target quantiles ($\delta_{target} = 10\%, 5\%, 1\%$) from either the level set approximation $\tilde{\Sigma}_k$ with $\delta_{shared} = 0.3$ (broken line), or random sampling on the entire domain S (solid line) for the one-dimensional test function in (7). The vertical axis shows the number of points sampled, averaged across 30 experiments. The horizontal axis tracks the number of high-fidelity function evaluations.

5.2 Test Examples in Multiple Dimensions

We further examine a selection of common non-convex test functions. To reliably control the volume of overlap between different level sets, we construct a high and low fidelity model based on an offset value in the low-fidelity test function.

$$f_L(x + \chi) = f_H(x) \tag{9}$$

where χ is a constant offset. For our tests, we examine three common non-convex test functions used in Ali et al. (2005), which can be summarized with their selected offsets.

1. **Rosenbrock Function:** $f_L(x) = \sum_{i=1}^{n-1} [100 \cdot (x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$ in two dimensions on the domain $S = \{[-2,2],[-2,2]\}$ with a true minimum at $(1,1)$ with value 0. We use an offset $\chi = (-0.2, -0.2)$ which results in an estimated overlap in two dimensions as $\hat{\gamma} = 0.80$ at $\delta_{shared} = 0.3$, $\hat{\gamma} = 0.69$ at $\delta_{shared} = 0.20$, and $\hat{\gamma} = 0.44$ at $\delta_{shared} = 0.10$.
2. **Sinusoidal Function:** $f_L(x) = -[A \cdot \prod_{i=1}^{n} sin(x_i - z) + \prod_{i=1}^{n} sin(B \cdot (x_i - z))]$ with $A = 2.5$, $B = 5$, $z = 30$ in two dimensions on the domain $S = \{[0,180],[0,180]\}$ with a true minimum at $(90,90)$ with value -3.5. We use an offset $\chi = (-5, -5)$ which results in an estimated overlap in two dimensions as $\hat{\gamma} = 0.81$ at $\delta_{shared} = 0.20$, $\hat{\gamma} = 0.78$ at $\delta_{shared} = 0.3$, and $\hat{\gamma} = 0.65$ at $\delta_{shared} = 0.10$.
3. **Griewank Function (Problem):** $f_L(x) = 1 + \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} cos\left(\frac{x_i}{\sqrt{x_i}}\right)$ with the constraints $\{[-5,5],[-5,5]\}$ With true optimum at $f(x^*) = 0$ with $x^* = (0, \ldots, 0)$. There are an unknown number of local minimum. We use an offset $\chi = (-0.2, -0.2)$ which results in an estimated overlap in two dimensions as $\hat{\gamma} = 0.88$ at $\delta_{shared} = 0.3$, $\hat{\gamma} = 0.85$ at $\delta_{shared} = 0.20$, and $\hat{\gamma} = 0.76$ at $\delta_{shared} = 0.10$.

The $\hat{\gamma}$ values for each of the test functions and offsets are computed using a mesh division method (50 divisions for each dimension with 2000 sample points from each sub-region). Based on Corollary 1, we can guarantee sampling within $L_L(\delta_{shared} = 0.3, S)$ will increase the chance of candidate points in $L_H(\delta_{target}, S)$

for $\delta_{target} \geq 0.08$ for the Rosenbrock, for $\delta_{target} \geq .08$ for the Sinusoidal function, and $\delta_{target} \geq 0.05$ for the Griewank function.

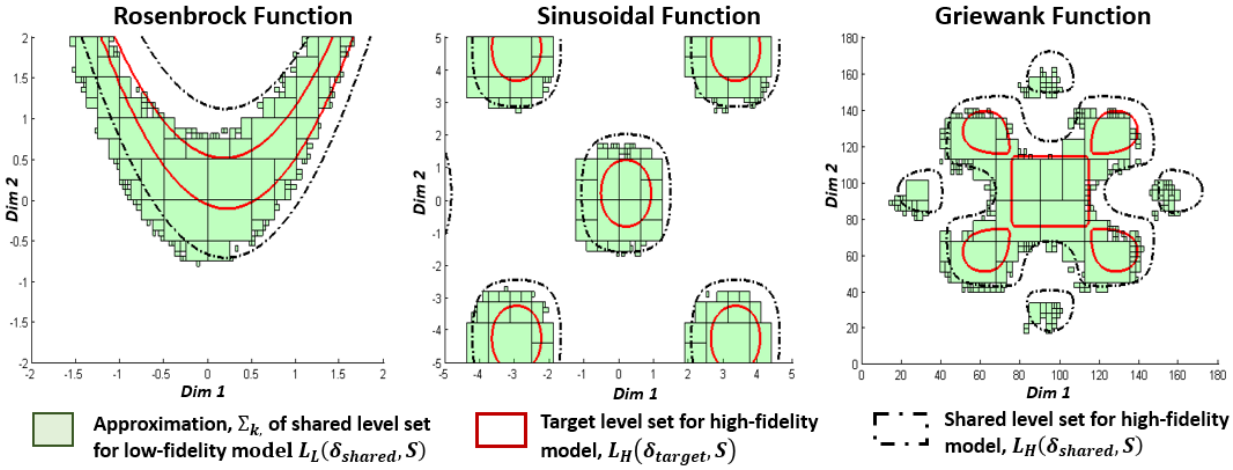

Figure 3: A plot of the intersecting level sets between the high-fidelity and low-fidelity models for the Rosenbrock, Sinusoidal, and Griewank functions. The $\delta_{shared} = 0.3$ level set approximation for the low-fidelity model is inside the light shaded area, the $\delta_{shared} = 0.3$ level set for the high-fidelity model is outlined by the broken line, and the $\delta_{target} = 0.1$ level set for the high-fidelity model is shown inside the solid line.

Using Algorithm 1 with $\delta_{shared} = 0.3$ we can develop an approximate level set using the same parameters from section 5.1. This results in a level set approximation of hyper-rectangles $\tilde{\Sigma}_k$ for each of the described test functions that can be illustrated in Figure 3. Figure 3 plots the $\delta_{shared} = 0.3$ level sets for the high-fidelity function (inside-the dash-dot line), the approximated level set for the low-fidelity function (inside the shaded rectangles), and the high-fidelity target level set (for $\delta_{target} = 0.1$). We can see that, in all three functions, there is significant overlap between high-fidelity target level sets and the approximated level sets.

As in the one-dimensional test, we compare the number of generated candidate solutions in $L_H(\delta_{target}, S)$ when sampling uniformly from the approximated level set, $\tilde{\Sigma}_k$, versus uniform sampling on the entire domain S. In Figure 4 we plot the number of points found in the $10\%, 5\%, 1\%$ target quantiles. This comparison demonstrates that sampling on the approximated level set improves the number of generated candidate solutions for the selected test functions, practically, even for δ_{target} levels that are lower than specified in Assumption 1.

We also examine the effect of using Algorithm 1 to find minimum function values using a variety of different δ_{shared}-levels. In Figure 5, we plot the minimum values sampled from a variety of approximated level sets with *different* δ_{shared} quantile levels. The average log minimum values (averaged over 30 experiments) are plotted versus the number of function evaluations and compared with random sampling on the entire domain. The plots in Figure 5 demonstrate that the sampling from *any* of the approximated level sets improves the minimum value found over uniform sampling on the domain. While the Rosenbrock function shows that sampling on each approximated level set performs similarly for different values of δ_{shared}, the approximated level sets using lower δ_{shared} values improve the sampling for the Sinusoidal and Griewank functions.

6 DISCUSSION

The paper outlines an approach to multi-fidelity optimization by approximating a low-fidelity level set in order to improve the sampling of a high-fidelity function. Using the PBnB algorithm to approximate a level set, the algorithm conducts a high-fidelity random search in a generated set of hyper-rectangles

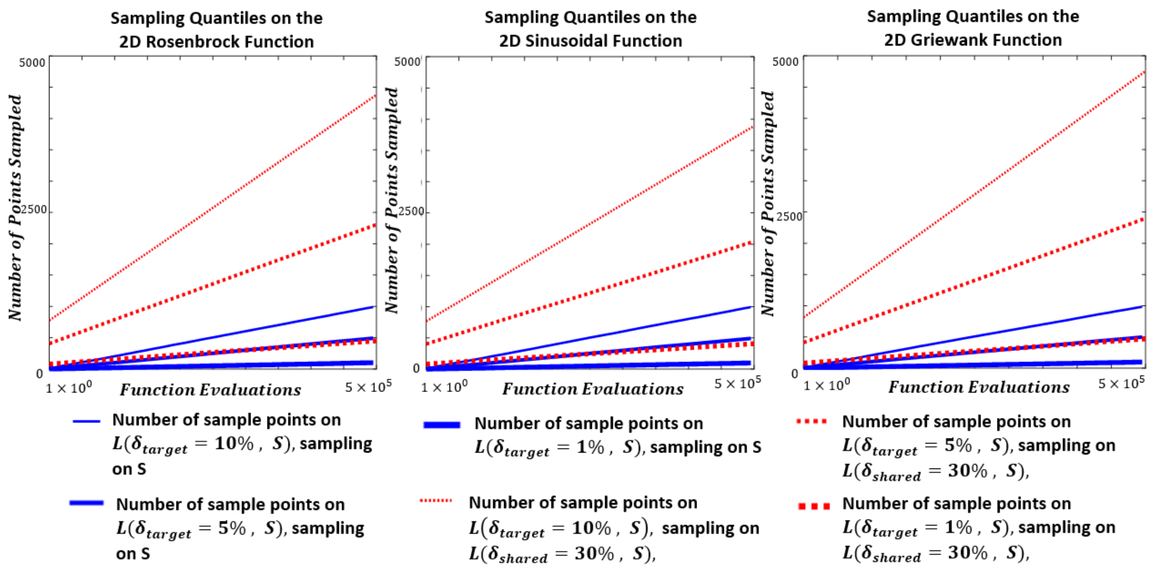

Figure 4: Number of sampled points within target quantiles ($\delta_{target} = 10\%, 5\%, 1\%$) from either the level set approximation $\tilde{\Sigma}_k$ generated by the PBnB algorithm with $\delta_{shared} = 0.3$ (broken line), or random sampling on the entire domain S (solid line) for the Rosenbrock, Sinusoidal, and Griewank Functions. The vertical axis shows the number of points sampled, averaged across 30 experiments. The horizontal axis tracks the number of high-fidelity function evaluations.

that approximate the level set of given low-fidelity model. Given sufficiently overlapping level sets, the concentration of points inside the approximated level set (and therefore the efficiency of a random search performed there) will be improved.

The usefulness of this algorithm is demonstrated by its application to pairs of low-fidelity and high-fidelity test functions. In all cases (with a selection of the δ_{shared} that creates a large amount of overlap between the level sets of high-fidelity and low-fidelity models), the efficiency of the algorithm in sampling candidate points from a target level set is improved.

The largest challenge with the implementation of the algorithm is selecting a δ_{shared} where both the high and low-fidelity level sets overlap. However, given that the a low-fidelity model will likely have broadly similar behavior to the high-fidelity model with a sufficiently high δ_{shared}, there is likely to be significant amount of overlap when δ_{shared} is not small.

One potential extension is to use alternate searches on both the high and low-fidelity models. First using the low-fidelity model to prune away areas of the domain that are unfit (that do not fall within a very high δ_{shared} level set) and subsequently using the high-fidelity model to determine whether it is likely that a smaller δ_{shared} level set can be used to further focus sampling. This also might provide an opportunity to use a shifting value of δ_{shared} based on observed properties of the high-fidelity function. Another extension could explore the possibility of modifying the the low-fidelity model based on statistical tests for correlation between low-fidelity and high-fidelity models within various sub-regions which could account for regions of the domain where the two models have poor or negative correlation.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation (NSF) under Grant CMMI-1632793.

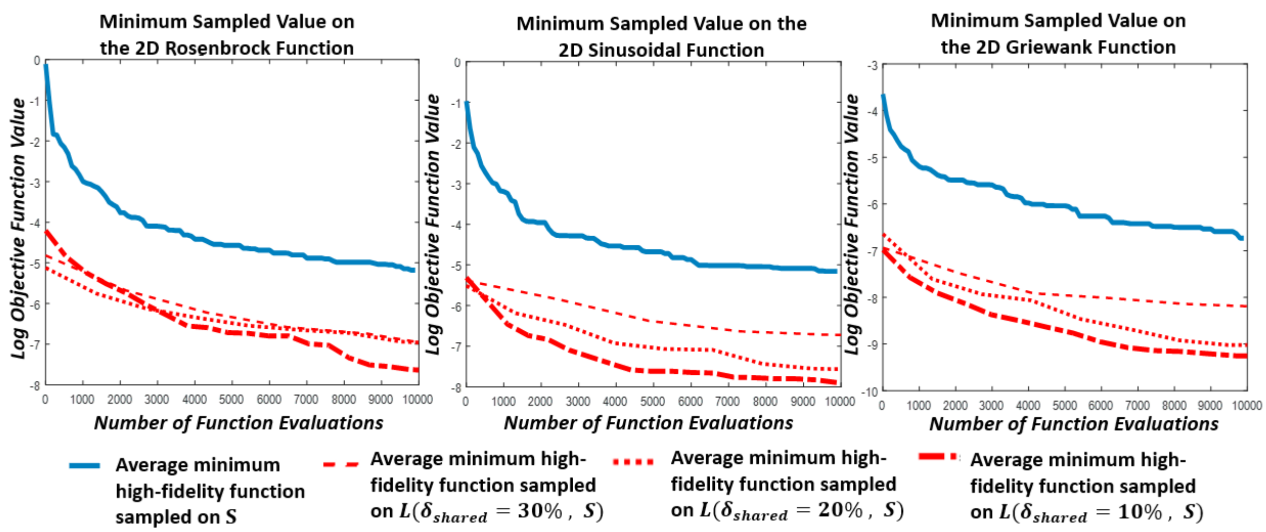

Figure 5: Minimum objective function value (log) averaged across 30 experiment replications for uniform sampling on the entire domain (solid lines) and uniform sampling from the level set approximation $\tilde{\Sigma}_k$ for δ_{shared} = 0.3, 0.2, and 0.1 respectively. Three test functions are examined: Rosenbrock, Sinusoidal, and Griewank.

REFERENCES

Ali, M. M., C. Khompatraporn, and Z. B. Zabinsky. 2005. "A Numerical Evaluation of Several Stochastic Algorithms on Selected Continuous Global Optimization Test Problems". *Journal of Global Optimization* 31 (4): 635–672.

Barton, R. R., and M. Meckesheimer. 2006. *Chapter 18 Metamodel-Based Simulation Optimization*, Volume 13. Elsevier.

Chiu, C.-c., J. T. Lin, and S. Zhang. 2016. "Improving the Efficiency of Evolutionary Algorithms for Large-Scale Optimization with Multi-Fidelity Models". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. Frazier, R. Szechtman, E. Zhour, T. Huschka, and S. Chick, 815–826. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Forrester, A. I. J., A. Sóbester, and A. J. Keane. 2007. "Multi-Fidelity Optimization via Surrogate Modelling". *Proceedings: Mathematical, Physical and Engineering Sciences* 463 (2088): pp. 3251–3269.

Huang, E., J. Xu, S. Zhang, and C. H. Chen. 2015. "Multi-fidelity Model Integration for Engineering Design". *Procedia Computer Science* 44 (C): 336–344.

Huang, H., and Z. B. Zabinsky. 2013. "Adaptive Probabilistic Branch and Bound with Confidence Intervals for Level Set Approximation". In *Proceedings of the 2013 Winter Simulation Conference*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, 4146–4157. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Li, H., Y. Li, L. H. Chew, and E. Peng. 2015. "Multi-Objective Multi-Fidelity Optimization with Ordinal Transformation and Optimal Sampling". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. Chan, I. Moon, T. Roeder, C. Macal, and M. Rossetti, Volume 1, 1689–1699. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Li, H., G. Pedrielli, C.-h. Chen, M. Chen, L. H. Lee, and E. P. Chew. 2016. "V-Shaped Sampling Based on Kendall-Distance to Enhance Optimization with Ranks". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. Frazier, R. Szechtman, E. Zhour, T. Huschka, and S. Chick, 671–681. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Molina-Cristóbal, A., P. R. Palmer, B. A. Skinner, G. T. Parks, A. Molina-Cristobal, P. R. Palmer, B. A. Skinner, and G. T. Parks. 2010. "Multi-fidelity Simulation Modelling in Optimization of a Submarine Propulsion System". *2010 IEEE Vehicle Power and Propulsion Conference, VPPC 2010*:1–6.

Qiu, Y., J. Song, and Z. Liu. 2016. "A Simulation Optimization on the Hierarchical Healthcare Delivery System Patient Flow Based on Multi-Fidelity Models". *International Journal of Production Research* 54 (21): 6478–6493.

Xu, J., S. Zhang, and C.-h. Chen. 2014. "Efficient Multi-Fidelity Simulation Optimization". In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk, S. Diallo, I. Ryzhov, L. Yilmaz, S. Buckley, and J. Miller, 3940–3951. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Xu, J., S. Zhang, E. Huang, and C.-h. Chen. 2016. "MO 2 TOS: Multi-Fidelity Optimization with Ordinal Transformation and Optimal Sampling". *Asia-Pacific Journal of Operational Research* 33 (3): 1650017:1–26.

AUTHOR BIOGRAPHIES

DAVID D. LINZ is a PhD candidate in the Department of Industrial and Systems Engineering at the University of Washington. His research interests include stochastic optimization, and simulation optimization with applications to healthcare strategy. His e-mail address is ddlinz@uw.edu.

HAO HUANG is an Assistant Professor in the Department of Industrial Engineering and Management at the Yuan Ze University at Taoyuan, Taiwan. His research interests include simulation optimization and healthcare applications. His e-mail address is haohuang@saturn.yzu.edu.tw.

ZELDA B. ZABINSKY is a Professor in the Department of Industrial and Systems Engineering at the University of Washington, with adjunct appointments in the departments of Electrical Engineering, Mechanical Engineering, and Civil and Environmental Engineering. She is an IIE Fellow. She has published numerous papers in the areas of global optimization, algorithm complexity, and optimal design of composite structures. Professor Zabinsky's research interests are in global optimization under uncertainty for complex systems. Her email address is zelda@u.washington.edu.