Beyond the Pareto Efficient Frontier: Constraint Active Search for Multiobjective Experimental Design

Gustavo Malkomes 1 Bolong Cheng 1 Eric Hans Lee 1 Michael McCourt 1

Abstract
Many problems in engineering design and simulation require balancing competing objectives under the presence of uncertainty. Sample-efficient multiobjective optimization methods focus on the objective function values in metric space and ignore the sampling behavior of the design configurations in parameter space. Consequently, they may provide little actionable insight on how to choose designs in the presence of metric uncertainty or limited precision when implementing a chosen design. We propose a new formulation that accounts for the importance of the parameter space and is thus more suitable for multiobjective design problems; instead of searching for the Pareto-efficient frontier, we solicit the desired minimum performance thresholds on all objectives to define regions of satisfaction. We introduce an active search algorithm called Expected Coverage Improvement (ECI) to efficiently discover the region of satisfaction and simultaneously sample diverse acceptable configurations. We demonstrate our algorithm on several design and simulation domains: mechanical design, additive manufacturing, medical monitoring, and plasma physics.

1. Introduction
Accelerated design and optimization via machine learning is receiving increasing interest in multiple science and engineering disciplines, such as materials design, drug discovery, and chemical engineering (Forrester et al., 2008; Negoescu et al., 2011; Molesky et al., 2018). Using numerical simulation to study the impact of design decisions prior to manufacturing has become a common strategy to reduce the number of fabrications, thus reducing the substantial cost of development cycles.

In recent years, active learning has gained popularity in scientific communities. These adaptive methods more efficiently identify promising candidates compared to brute-force or evolutionary algorithms (Song et al., 2018; Attia et al., 2020; Haghaniifar et al., 2020; Duris et al., 2020). Bayesian optimization (BO) is arguably the most prominent example of active learning, where the goal is to seek the optimizer \( \text{arg max}_x f(x) \) of an expensive-to-evaluate black-box objective \( f \).

However, in most real world applications, multiple competing objectives need to be investigated (Koziel et al., 2014; Singh et al., 2016). Optimality of multiobjective problems is usually defined by the concept of Pareto efficiency, where no individual objective can be improved without loss in at least one other objective. Most work on optimizing multiple objectives focuses on searching for the Pareto efficient frontier (and the corresponding parameters). In general, for an \( m \)-objective problem, the (nearly) Pareto efficient solutions are assumed to be a dense set of points on an \((m - 1)\)-dimensional manifold (Konakovic-Lukovic et al., 2020). Unfortunately, using the Pareto frontier as the measurement of success is limiting in engineering and design applications; candidates near the Pareto frontier, unaccounted for in most optimization literature, are still of scientific interest in the real world (del Rosario et al., 2020).

In practice, design objectives can also be formulated as inequality constraints, where the decision makers seek to find design choices that satisfy known performance thresholds on each objective. This formulation is related to the topic of level set estimation (LSE) in the single objective setting. Its goal is to determine the decision boundary of parameter space that separates it above and below the threshold on the objective (Gotovos et al., 2013). However, the notion of level set does not naturally extend to the multiobjective settings; a naive approach of identifying the level set for each objective and taking the intersection is not computationally efficient. In addition, most LSE algorithms focus on searching for the decision boundary; in the multiobjective setting, this behavior cannot reveal the tradeoff among objectives, since there are not enough samples within the region of interest.

Figure 1 illustrates the limitations of multiobjective BO and
2. Related Work

2.1. Multiobjective Bayesian Optimization

Bayesian optimization (BO) is a popular method for optimizing expensive, black-box objective functions. BO consists of two components: a probabilistic model, to model the objective function $f$, and an acquisition function, to determine which parameters $x$ to sample next. In this section, we focus on reviewing the literature of BO in the multiobjective setting. We refer the readers to Frazier (2018) for a comprehensive review of BO in the single objective setting.

There are several strategies at adapting BO to multiobjective optimization. The simplest approach is linear scalarization, where single objective BO strategies optimize a linear combination of the objectives. Knowles (2006) first proposes using a random scalarization of the objective functions to search for the entire frontier. This method is further improved in Paria et al. (2020), which models each individual objective with a separate Gaussian process (GP).

A related approach involves reformulating the problem as a constrained optimization problem, commonly known as the $\varepsilon$-constraint method (Haghaniifar et al., 2020). In particular, this formulation leverages prior works from constrained Bayesian optimization literature (Gelbart, 2015; Gardner et al., 2014). In constrained BO, a probabilistic model is formed for each constraint and used as a multiplicative penalty term in the acquisition function (Letham et al., 2019). Recent works on constrained BO borrow ideas...
from the mathematical optimization community such as the 
trust region (Eriksson & Poloczek, 2021) and Alternating 
Directions Methods of Multiplier (Ariafar et al., 2019).

A separate line of research in multiobjective BO improves 
the hypervolume of the Pareto efficient frontier directly 
(Emmerich et al., 2011; Konakovic-Lukovic et al., 2020; 
Daulton et al., 2020). Belakaria et al. (2019) adapts the 
entropy search acquisition function for the multiobjective 
setting, by maximizing the entropy of the Pareto frontier. 
Belakaria et al. (2020) and Fernández-Sánchez et al. (2021) 
build upon this work to incorporate additional black-box 
inequality constraints. Almost all multiobjective BO liter-
ature uses the hypervolume of the Pareto frontier as their 
performance criterion. While hypervolume provides a 
single numerical value for the ease of comparison, using it 
as the sole criterion of success misses the goal of multi-
objective optimization — understanding tradeoffs between 
parameters and objectives.

2.2. Level Set Estimation

The goal of level set estimation is to determine the regions 
where the objective function lies above or below a known 
threshold level. Bryan et al. (2005) first uses a GP-based 
active learning approach to tackle this problem through the 
straddle heuristic, defined as

\[ \text{STRADDLE}(x) = 1.96 \sqrt{\hat{v}(x)} - |f(x) - \tau|, \]

where \( \tau \) is the threshold of the level set and \( v \) is the model 
uncertainty. Gotovos et al. (2013) generalizes the straddle 
heuristic and provides theoretical guarantees on the classifi-
cation quality. Zanette et al. (2019) maximizes the expected 
increase in the volume of the predicted level set. Iwazaki 
et al. (2020) considers the LSE problem with input uncer-
tainty, where the input parameter values are subject to per-
turbation from a known Normal distribution. This work 
is related to research in safe BO (Sui et al., 2015; 2018), 
where only configurations exceeding the threshold with high 
probability are sampled.

Almost all LSE methods focus on identifying the decision 
boundary. To our best knowledge, Wang et al. (2018) is the 
closest example of learning and diversely sampling from a 
region of interest (in the single objective setting). It divides 
the sampling procedure into two disjoint algorithms: one 
determines the level set and the other samples diversely 
from the interior of the learned level set.

The notion of a level set does not naturally extend beyond 
more than one objective. Bryan & Schneider (2008) only 
considers the multiobjective problem in a composite setting, 
where a single threshold is used for a linear combination of 
the objective functions. One naive and inefficient approach 
to extend existing LSE algorithms to multiobjective prob-
lems is to estimate the level set of each objective separately 
and then intersect each region of interest.

2.3. Active Search

Our proposed work is also related to the topic of active 
search (Garnett et al., 2012). Active search can be seen as 
a special case of Bayesian optimization, where one has binary 
observations and cumulative reward. The goal of active 
search is to sequentially discover members of a rare, desired 
class. Inspecting any element is assumed to be expensive, 
representing, for instance, the cost of performing a real 
world laboratory experiment.

Garnett et al. (2012) explores this problem using Bayesian 
decision theory, introduces a natural utility function for 
active search, and derives the optimal policy for this frame-
work. In general, only myopic (one-step) approximations 
to the optimal policy are computationally tractable. More 
recently, Jiang et al. (2017) proposes a nonmyopic and 
efficient approximation to the optimal policy, which is later 
extended to the batch setting in Jiang et al. (2018). Active 
search, however, is not designed to solve multiobjective 
continuous problems — a problem that we address next.

3. Constraint Active Search

Scientists, engineers, and decision makers seldom rely on 
single-point observations to make final design decisions. 
Instead, they seek a diverse set of design configurations to 
embrace their understanding of each objective function’s 
impact and its respective tradeoffs. As Figure 1 illustrates, 
neither BO nor LSE is well-suited to address these experi-
mental design concerns. The former clusters around optima, 
and the latter clusters around the level set boundary.

We propose an alternative formulation to the multiobjective 
design problem better suited to gather a diverse spectrum of 
design configurations. The goal of this formulation is to 
efficiently cover an unknown, possibly disconnected, set \( S \) 
defined by the objective thresholds \( \tau \) in a sequential fashion. 
We call this problem constraint active search. In Section 5, 
we illustrate how CAS may enable a practitioner to make 
better decisions than those made by exclusively identifying 
the Pareto frontier on an additive manufacturing problem.

This concept is related to del Rosario et al. (2020) which 
advocates that the Pareto frontier itself is insufficient for sci-
entific and engineering settings. That work introduces the 
notion of Pareto shells, the set of near Pareto optimal solu-
tions, as the desired outcome of multiobjective optimization. 
We propose that, when thresholds \( \tau \) are available, studying 
the full region of constraint satisfaction\(^1\) is preferable.

\(^1\)We use the term satisfaction instead of feasibility merely to 
stress that the constraints are over the objectives and not explicitly 
deﬁned over the parameter space.
3.1. Problem Statement

Suppose we want to search for design configurations in a compact search space $\mathcal{X}$. We may judge the quality of a design $\mathbf{x} \in \mathcal{X}$ by evaluating $m$ expensive black-box objective functions $f_1, f_2, \ldots, f_m$, each mapping $\mathcal{X}$ to $\mathbb{R}$. We seek designs $\mathbf{x}$ that yield acceptable performance, defined as $f(\mathbf{x}) \geq \tau$, where $\tau = [\tau_1, \tau_2, \ldots, \tau_m]$. Specifically, we wish to sequentially select configurations from the set:

$$\mathcal{S} = \{ \mathbf{x} \mid f(\mathbf{x}) \geq \tau \},$$

where $f(\mathbf{x}) \geq \tau := f_i(\mathbf{x}) \geq \tau_i$, $i = 1, \ldots, m$. At iteration $t$, we assume we have evaluated $t$ design configurations $\mathbf{X}_t = \{ \mathbf{x}_j \}_{j=1}^t$ and their respective outputs $\mathbf{Y}_t = \{ \mathbf{y}_j \}_{j=1}^t$, where $\mathbf{y}_j = [y_{j1}, y_{j2}, \ldots, y_{jm}]$ is the potentially noisy observation of the objective functions $f_1, f_2, \ldots, f_m$ evaluated at $\mathbf{x}_j$. These observations form the dataset $D_t = (\mathbf{X}_t, \mathbf{Y}_t)$.

We assume we have $m$ independent probabilistic models that capture our prior beliefs about observations $y_{ij} = f_{ij}(\mathbf{x}_i) + \varepsilon_i$ for $i = 1, 2, \ldots, m$, as a probability distribution over $p(y)$, where $\varepsilon_i$ is additive Gaussian noise. We also assume $p(y)$ allows posterior updates when $D$ is observed, i.e., $p(y \mid D)$. Let $Z$ be an indicator variable that tells us if a point $\mathbf{x}$ will satisfy the thresholds, $Z(\mathbf{x}) = 1[\mathbf{y}(\mathbf{x}) \succeq \tau]$. Our model of $\mathbf{y}$ allows us to compute the probability of any point $\mathbf{x}$ belonging to the satisfactory region $\mathcal{S}$ after observing $D_t$, which we denote by $p(Z(\mathbf{x}) = 1 \mid D_t)$. We delay discussing the statistical model’s details until Section 4, as they are immaterial to decision-making.

3.2. Expected Coverage Improvement

Our approach to constraint active search considers an additional resolution parameter $r$; in real-world problems one can often define $r$ in terms of the robustness of the design to perturbations. Typically, most experimental design problems possess a sense of known resolution, such as simulation accuracy or manufacturing precision/tolerance. This presumes that any design configuration within distance $r$ of another does not convey extra information about $\mathcal{S}$.

**Definition 3.1 (Coverage neighborhood).** The coverage neighborhood of any $\mathbf{x}$ is defined as

$$\mathcal{N}_r(\mathbf{x}) = \{ \mathbf{x}' : d(\mathbf{x}, \mathbf{x}') < r \},$$

for an a priori fixed $r \in \mathbb{R}^+$ and an appropriate distance function $d : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}^+$. The coverage neighborhood of a set of points $\mathbf{X}$ is defined as

$$\mathcal{N}_r(\mathbf{X}) = \bigcup_{\mathbf{x} \in \mathbf{X}} \mathcal{N}_r(\mathbf{x}).$$

For Euclidean distance, $\mathcal{N}_r(\mathbf{x})$ is simply a ball of radius $r$.

Constraint active search prioritizes datasets $D$ that cover as much volume of the satisfactory region $\mathcal{S}$ as possible. We express this desire with the following utility function:

$$u(D) = \text{Vol}(\mathcal{N}_r(\mathbf{X}) \cap \mathcal{S}).$$

Intuitively, $u(D)$ measures the total volume of $\mathcal{S}$ that is covered by neighborhood $\mathcal{N}_r(\mathbf{x})$ for all $\mathbf{x} \in \mathbf{X}$. At iteration $t$, the expected improvement of this utility function for sampling a candidate configuration$^2$ $\mathbf{x}$ is:

$$\alpha(\mathbf{x} \mid D_t) = E_{\mathbf{y}} [u(D_{t \cup \{ \mathbf{x}, \mathbf{y} \}}) - u(D_t)]$$

$$= E_{\mathbf{y}} [\text{Vol}(\mathcal{N}_r(\mathbf{x}) \cap \mathcal{S} \cup \mathcal{N}_r(\mathbf{x})) - \mathcal{N}_r(\mathbf{x})] .$$

(1)

We can also refer to (1) as an acquisition function. At iteration $t$, our policy selects the location $\mathbf{x}^*$ that maximizes the acquisition function, i.e., $\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathbf{X}} \alpha(\mathbf{x} \mid D_t)$. We visualize the intuition behind our policy for different

---

$^2$For clarity, we have overloaded the notation for $D_{t \cup \{ \mathbf{x}, \mathbf{y} \}}$ to indicate a pairwise union, which we hope is clearly understood.
realizations of $Z$ in Figure 2. Our acquisition function considers the volume of $\mathcal{S}_Z$ that a new candidate location is expected to cover under our current belief about $Z$.

A common practice in the LSE and active search literature is to work on discrete search spaces: $\mathcal{X}$ could be discretized by sampling a pool of $n$ points. In this discretized setting, (1) simplifies to

$$
\alpha(x \mid D_t) = \mathbb{E}_Z \left[ \sum_{x' \in \mathbb{N}(x) \setminus \mathbb{N}_r(x)} Z(x') \right]
= \sum_{x' \in \mathbb{N}(x) \setminus \mathbb{N}_r(x)} \mathbb{E}_Z [Z(x')]
= \sum_{x' \in \mathbb{N}(x) \setminus \mathbb{N}_r(x)} p(Z(x') = 1 \mid D_t). \tag{2}
$$

In practice, the argmax of $\alpha(x \mid D_t)$ is often nonunique, especially at the start of search. This reflects the intuition that when only a few observations sparsely cover $\mathcal{S}$, multiple—and possibly many—candidates will contribute equally to the expected coverage. Crucially, if this occurs, we break ties among the multiple optimum points by selecting the one furthest from all observations in $X_t$, which encourages selecting a diverse set of points. We call this policy expected coverage improvement (ECI).

### 3.3. Theoretical Analysis of ECI

We present three theoretical results for ECI (proved in the supplement). All assume that $\mathcal{S}$ is fixed a priori — that $\mathcal{S}$ does not change as iteration proceeds.

**Theorem 1.** If we have a budget of one iteration left, then ECI is one-step Bayes-optimal among all feasible policies.

This result is not particularly surprising. The next pair of less trivial theorems concern the convergence of CAS using ECI. Because CAS seeks to sample a diverse set of points, we measure this diversity with fill distance. Fill distance is the standard measure of point diversity in the quasi-Monte Carlo (Joy et al., 1996), experimental design (Pronzato & Müller, 2012), and meshfree approximation (Fasshauer & McCourt, 2015) communities. The convergence of the fill distance of ECI thus describes the limiting behavior of its sample diversity.

Given a set of sample points $X$, the fill distance is formally defined as the following:

$$
\text{FILL}(X, S) = \sup_{x \in S} \min_{x \in X} d(x_j, x). \tag{3}
$$

In Euclidean space, $\text{FILL}(X, S)$ is the radius of the largest empty ball one can fit in $\mathcal{S}$, and measures the spacing of $X$ in $\mathcal{S}$. The smaller a set’s fill, the better distributed it is within $\mathcal{S}$. Minimal fill is defined in the following sense:

$$
\rho = \min_{X \in \mathcal{X}^n} \text{FILL}(X, S).
$$

Low-discrepancy sequences and Latin hypercubes achieve low fill in simple domains. However, computing the minimal fill is generally NP-hard (Pronzato & Müller, 2012).

Having defined fill distance, we can now frame convergence in terms of the limiting behavior of $\text{FILL}(X, S)$ as $n$ approaches infinity. This is formalized as follows.

**Theorem 2.** If $\mathcal{S}$ is dense, ECI produces a dense sequence of evaluation points in $\mathcal{S}$. Equivalently, the limit as $n$ approaches infinity of $\text{FILL}(X, S)$ is zero.

A more interesting remark is that we are able to bound the rate at which ECI fills the set $\mathcal{S}$. We can guarantee that ECI produces a fill bounded above by $4\rho$, and that for a sufficiently large iteration budget, ECI produces a fill that is no more than $2\rho$.

**Theorem 3.** For a fixed and known $\mathcal{S}$, let $\rho = \min_{x \in \mathcal{S}} \text{FILL}(X, S)$. For any $n$, ECI produces a fill bounded above by $4\rho$. Furthermore, there exists a finite $n^*$ for which ECI produces a 2-approximation ratio of $\rho$ for all $n \geq n^*$.

### 4. Experiments

In our experiments, we build an independent Gaussian process (GP) to model each objective (Rasmussen & Williams, 2005). In particular, for a GP model and a single objective,

$$
\rho(Z(x) = 1 \mid D_t) = \int_\tau^\infty \phi(y \mid \mu(x), \nu(x))dy, \tag{4}
$$

where $\phi$ is the Normal PDF, and $\mu(x)$ and $\nu(x)$ are the GP’s posterior mean and variance, respectively, evaluated at $x$. For $m$ objectives modeled by $n$ independent GPs, $\rho(Z(x) = 1 \mid D_t)$ is the product of probabilities associated with each model. Our GPs use a zero mean function and a $C^4$ Matérn covariance kernel with length scales fixed a priori (more information is provided in the supplement).

### 4.1. Baseline Methodologies

In addition to random search (RND), we compare our algorithm to several (modified) baseline algorithms aforementioned in Section 2: $\varepsilon$-constraint BO (Haghighifar et al., 2020), STRADDLE (Bryan et al., 2005), and one-step active search (Garrett et al., 2012). We extend STRADDLE to the multiobjective setting by alternating the objective functions at each iteration. For one-step active search, we maximize the point with highest probability $\rho(Z(x) = 1 \mid D_t)$ at each iteration, which we call ONE-S. In addition, we develop two information theoretic baselines:

**Mutual information (EZ):** We consider the mutual information between $y$ and $Z$. Assuming noiseless observations $f_i(x) = y_i$, this simplifies to the entropy of $Z$:

$$
\text{MI}(y; Z) = H(y) - E_Z[H(y \mid Z)] = H(Z),
$$
Figure 3. Median metric values across 20 repetitions for the additive manufacturing application. We plot the fill distance, number of positive samples, hypervolume, and coverage recall respectively. Shaded region corresponds to values between lower and upper quartiles.

where $H(\cdot)$ is the entropy of a random variable. This would be the one-step optimal strategy that recovers the most information about the boundaries of $S$.

**Entropy inside $S$ (EISR):** We consider the utility:

$$u_{\text{EISR}}(x) = \begin{cases} 0 & Z = 0, \\ H(f \mid x) & Z = 1. \end{cases}$$

$u_{\text{EISR}}$ rewards sampling configurations that are inside $S$ and have high entropy. The corresponding acquisition function can be written as

$$\alpha_{\text{EISR}} = E_Z[u_{\text{EISR}}(x)] = p(Z(x) = 1)H(y).$$

### 4.2. Performance Criteria

We emphasize that no single criterion can sufficiently convey the full strength of any methodology; this is especially true in the multiobjective setting, in which different performance criteria already exist to quantify different algorithmic goals. We hope to impart a nuanced comparison of ECI to existing baselines. To that end, we consider the following four criteria.

**Fill distance:** We want an algorithm that can effectively explore the entire satisfactory region. This is quantified by fill distance, which is defined in Equation 3.

**Positive samples:** The number of satisfactory sampled points is commonly used in active search literature to judge the quality of the algorithm. In our experiments, these are the observed samples inside of $S$.

**Hypervolume:** We measure the hypervolume of region in metric space bounded by the Pareto frontier and the defined thresholds. In particular, we conjecture that algorithms which excel at maximizing the hypervolume may underperform on other criteria, and vice versa.

**Coverage recall:** Finally, we want to combine the idea of maximizing positive samples and minimizing fill distance to measure sample diversity. We propose a recall metric that considers coverage. Given a known $S$ and samples $X$, we define the coverage recall as

$$\text{Coverage recall} = \frac{|N_r(X) \cap S|}{|S|}.$$

Note that the numerator is exactly the utility function that we have proposed in Section 3.1. We will simply refer to this as recall in the following text.

### 4.3. Multiobjective Design Problem Suite

We adopt eight multiobjective engineering design problems presented in the REPROBLEM suite (Tanabe & Ishibuchi, 2020). These synthetic functions are specifically designed to mimic real-world applications and avoid having unrealistic properties of traditional benchmark functions. We follow the naming convention from the original text and detail the problems we selected in the supplement.

### 4.4. Additive Manufacturing

Haghanifar et al. (2020) studies a numerical simulation to help design an additive manufacturing strategy for minimizing the reflection of light at multiple angles of incidence. We conduct the same search for different nanostructures which balance a desire for minimizing normal reflection with reflection at an oblique angle.

The simulation, conducted using the Lumerical FDTD simulator (Lumerical Inc., 2020), has a level of inaccuracy associated with the numerical methods. In addition, there is a level of imprecision associated with the actual manufacturing process which means that the desired design parameters are only realized to limited precision during fabrication. Our goal for this reflection minimization problem is to find nanostructured glass that satisfying $R_{\text{normal}}(x) \leq \tau_{\text{normal}}$ and $R_{\text{oblique}}(x) \leq \tau_{\text{oblique}}$. The design parameters are the dimension of the nanostructures (a cone): maximum cone width, bottom diameter, top diameter, and height of the structure.
Table 1: Selected experimental results, consisting of the mean over 20 independent trials. Hypervolume has been rescaled for ease of viewing: RE33 \( \times 10^{-5} \), Additive \( \times 10^{-2} \), EEG \( \times 10^{2} \), Plasma \( \times 10^{0} \). We denote with bold text the best mean for each problem and each metric. Please refer to the supplementary material for the complete set of experimental results and information regarding the variation around the mean (standard error). For each problem, ECI has the best recall performance.

<table>
<thead>
<tr>
<th>Function</th>
<th>m</th>
<th>d</th>
<th>r</th>
<th>Methods</th>
<th>Fill distance ↓</th>
<th># Positive ↑</th>
<th>Coverage Recall ↑</th>
<th>Hypervolume ↑</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE33</td>
<td>3</td>
<td>4</td>
<td>0.08</td>
<td>RND</td>
<td>0.34</td>
<td>23.80</td>
<td>0.14</td>
<td>5.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ONE-S</td>
<td>0.60</td>
<td>101.55</td>
<td>0.18</td>
<td>7.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EZ</td>
<td>0.26</td>
<td>34.05</td>
<td>0.25</td>
<td>2.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>STRADDLE</td>
<td>0.35</td>
<td>13.70</td>
<td>0.10</td>
<td>2.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BO</td>
<td>0.42</td>
<td>36.55</td>
<td>0.11</td>
<td>8.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EISR</td>
<td>0.24</td>
<td>100.05</td>
<td>0.23</td>
<td>7.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ECI</td>
<td>0.27</td>
<td>100.95</td>
<td>0.73</td>
<td>7.91</td>
</tr>
<tr>
<td>Additive</td>
<td>2</td>
<td>4</td>
<td>35.0</td>
<td>RND</td>
<td>108.22</td>
<td>22.40</td>
<td>0.42</td>
<td>7.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ONE-S</td>
<td>211.86</td>
<td>156.75</td>
<td>0.60</td>
<td>8.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EZ</td>
<td>92.37</td>
<td>106.05</td>
<td>0.89</td>
<td>8.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>STRADDLE</td>
<td>79.30</td>
<td>21.10</td>
<td>0.43</td>
<td>7.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BO</td>
<td>107.84</td>
<td>121.95</td>
<td>0.82</td>
<td>9.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EISR</td>
<td>102.01</td>
<td>121.40</td>
<td>0.98</td>
<td>8.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ECI</td>
<td>34.92</td>
<td>56.00</td>
<td>1.00</td>
<td>8.07</td>
</tr>
<tr>
<td>EEG</td>
<td>2</td>
<td>6</td>
<td>0.25</td>
<td>RND</td>
<td>0.38</td>
<td>4.40</td>
<td>0.58</td>
<td>3.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ONE-S</td>
<td>0.48</td>
<td>31.20</td>
<td>0.50</td>
<td>3.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EZ</td>
<td>0.44</td>
<td>17.95</td>
<td>0.58</td>
<td>3.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>STRADDLE</td>
<td>0.41</td>
<td>2.00</td>
<td>0.38</td>
<td>2.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BO</td>
<td>0.52</td>
<td>4.30</td>
<td>0.35</td>
<td>2.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EISR</td>
<td>0.45</td>
<td>22.85</td>
<td>0.52</td>
<td>3.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ECI</td>
<td>0.38</td>
<td>8.30</td>
<td>0.79</td>
<td>3.65</td>
</tr>
<tr>
<td>Plasma</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>RND</td>
<td>1.02</td>
<td>10.85</td>
<td>0.19</td>
<td>5.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ONE-S</td>
<td>1.01</td>
<td>12.80</td>
<td>0.20</td>
<td>5.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EZ</td>
<td>1.07</td>
<td>10.70</td>
<td>0.17</td>
<td>5.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>STRADDLE</td>
<td>1.07</td>
<td>10.70</td>
<td>0.17</td>
<td>5.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BO</td>
<td>1.04</td>
<td>12.60</td>
<td>0.15</td>
<td>5.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EISR</td>
<td>1.03</td>
<td>10.55</td>
<td>0.14</td>
<td>6.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ECI</td>
<td>0.90</td>
<td>31.25</td>
<td>0.65</td>
<td>6.64</td>
</tr>
</tbody>
</table>

4.5. EEG for Brain Activity Reconstruction

Electroencephalography (EEG) is a non-invasive strategy for reconstructing cerebral electric currents from their resulting electrostatic potential measured on a human’s scalp (Ala et al., 2017). This process consists of solving an inverse problem, requiring localization of the electric dipoles’ locations and moments, subject to the associated model of dipole/potential interaction (Sorrentino & Piana, 2017). The limited spatial resolution of EEG is a well-known problem with its practical viability (Samuelsson et al., 2021), making constraint active search a potentially interesting tool for identifying regions of brain activity.

We use a meshfree solver to generate a scalp potential from a proposed dipole location through a quasi-stationary approximation to Maxwell’s equations (Ala et al., 2015). We then apply our strategy to identify the proposed dipole which most closely matches the true potential created by the true dipole, which is fixed at the start but hidden. We measure scalp potential at 28 locations and try to minimize the 25th and 75th percentiles of absolute deviation between the proposed and true potential; this has the effect of pushing both ends of the interquartile range as low as possible.

4.6. Plasma Physics

A stellarator is device that uses a set of magnetic coils to confine a plasma hot enough to sustain nuclear fusion (Spitzer Jr, 1958). Stellarator coils lack rotational symmetry.

\(^3\)Since the plasma problem is single objective, the hypervolume is computed as the difference between the threshold and the minimum value, \(|\tau - \min y|\).
Beyond the Pareto Efficient Frontier: Constraint Active Search for Multiobjective Experimental Design

Figure 4. Left two panels: difference from the expected design (in CAD) and the fabricated sample of a nanostructure. Right two panels: comparison of acceptable configurations in metric space and in the parameter space. Small variations in the Pareto optimal points can lead to significant discrepancies in the design space. CAS points provide more information about satisfactory design configuration.

and possess notably warped shapes due to the complex, quasi-symmetric magnetic field they must produce.

Recently, (Giuliani et al., 2020) developed a stellarator simulator, PyPlasmaOpt, which significantly streamlines the simulation procedure of previous packages such as StellOpt (Strickler et al., 2003). PyPlasmaOpt parameterizes each coil as a curve in 3D Cartesian coordinates \( \Gamma(\theta) = (x(\theta), y(\theta), z(\theta)) \), where each coordinate admits the following Fourier expansion, e.g., in the case of \( x(\theta) \):

\[
x(\theta) = c_0 + \sum_{k=1}^{\text{order}} s_k \sin(k\theta) + c_k \cos(k\theta).
\]

For each coil, the parameters \( c_k \) and \( s_k \) represent the search space that we must search over. We consider a relatively simple, order one coil which has nine parameters total.

We compute the objective function to be minimized by simulating the stellarator using the given coil shapes (which solves a certain first order, nonlinear ordinary differential equation). The resulting summary information from the simulation is our objective \( f \) and contains three terms:

\[
f(x) = R_{\text{magnetic}}(x) + R_{\text{transform}}(x) + R_{\text{shape}}(x).
\]

The first term quantifies the quasi-symmetry of the magnetic field—the smaller the first term, the more desirable the resulting field. The second term locks the solution into a target rotational transform. The third term penalizes overly complex coils too impractical to manufacture in real life.

4.7. Discussion

We exhaustively tested ECI and other baselines on eleven total problems, but due to limited space, we present a representative subset of our experimental results in this section. Our analysis holds true for the seven additional problems, whose results are in the supplement along with accompanying experimental setups for reproducibility purposes.

Specifically, Table 1 shows the mean value over 20 trials for each performance metric at termination for the RE33 function from the design problem suite and all our real-world applications. These results demonstrate that no algorithm can be the best in all performance metrics. In fact, it highlights each method’s strengths and weaknesses, confirming our thesis that evaluating the performance of multiobjective problems is itself multiobjective.

As expected, \( \text{BO} \) typically outperforms all methods on the hypervolume criterion, which only judges the Pareto frontier. However, it is consistently worse than its peers when judged by diversity (fill distance) and coverage (recall). Our proposed method \( \text{ECI} \), on the other hand, is effective at diverse sampling inside the satisfactory region \( S \); it is almost always the best performing method for coverage recall and fill distance. Unsurprisingly, \( \text{ONE-S} \) excels at accumulating satisfactory samples since it greedily maximizes the number of positive samples by design. These samples are highly clustered, as revealed by the high fill distance and low coverage recall values.

Each policy’s objective is evident when we consider the sample progression. Figure 3 shows the variation on performance as a function of the number of observed samples for the additive manufacturing problem. For instance, \( \text{ECI} \) maximizes the coverage recall and also minimizes the fill distance but stagnates at maximizing hypervolume and number of positives.

Our investigation reveals that different multiobjective algorithms prioritize different performance criteria. Therefore, one can tailor a strategy (perhaps an ensemble of methods) according to each application’s need. If metric optimization is of ultimate relevance, we recommend multiobjective Bayesian optimization. If the decision maker’s goal is to identify rare elements regardless of diversity, then active search is the best choice. However, we argue that selecting a diverse set of satisfactory configurations is the most suited objective for design problems. We provide an example to support this claim in the following section.
5. Constraint Active Search in Practice

This research was motivated, in part, by a collaboration with materials scientists using additive manufacturing to fabricate nanostructured glass with multiple desirable optical and physical characteristics. Multiobjective problems are intrinsically subjective — if we could quantify the tradeoffs between each metric precisely a priori, we would solve a single-objective problem instead. Therefore, multiobjective methodologies provide the decision makers with a set of points at termination. Traditional multiobjective optimization offers a Pareto optimal set of points; our alternative gives a set of diverse points in parameter space that meet the user constraints. The decision maker, a domain expert, must then decide which one or more of these points to promote to physical manufacturing, often considering two key complications:

- Computational and manufacturing imprecision prevent an exact physical realization of the results. See Figure 4 (left two panels) for an example of the CAD model and the associated fabricated component, which suffered from limited manufacturing precision.

- Some metrics are unavailable during optimization — maybe they can only be computed after manufacturing (such as anti-fogging or durability) or require costly expert analysis (such as the supervisor stating expert opinions on a manufacturing scale).

Both of these complications are better addressed through sample diversity in parameter space (not in metric space) — the primary motivation of CAS.

Inaccuracies arising from the transfer of computational results to manufactured outcomes are best dealt with by choosing points most likely to achieve acceptable performance after fabrication. Because CAS explores the parameter space, it has a greater ability to select results far (e.g., in Euclidean sense) from unacceptable outcomes than multiobjective optimization. When only studying the Pareto frontier, no energy is spent exploring the safety of those points from undesirable results. Figure 4 (right) shows the numerical simulation results for the additive manufacturing example; interior points in parameter space are more likely to be selected for physical manufacturing.

For metrics only computable after fabrication, users benefit from a diverse set of designs in parameter space so they may manufacture a select few and still hope to achieve good results on these previously unavailable metrics.

6. Conclusion and Future Work

We introduce constraint active search as a novel methodology to explore multiobjective experimental design options while satisfying minimum performance criteria. One key differentiation between constraint active search and existing multiobjective optimization literature is the emphasis on the design parameters: constraint active search values sample diversity in the satisfactory region.

We propose a new search policy, expected coverage improvement, designed to efficiently improve the coverage of sampled designs over the satisfactory region; theoretical properties of this strategy include a constant approximation ratio to the optimal sample diversity. We define and study several performance criteria when judging the performance of proposed algorithms in multiobjective settings, and we explain the implications of prioritizing each. Using these criteria, we compare our new policy against existing strategies such as Bayesian optimization, level set estimation, and active search on synthetic and real world problems.

There are many open problems in the topic of constraint active search for design and simulation; here we prioritize a few topics for future investigation.

**Distinctness/Precision:** A key element of ECI (as well as the active search baseline) is the choice of the resolution $r$, through which the design precision is defined. This is acceptable, but limited: categorical parameters may exist; different parts of the domain may have different manufacturing precision; some parameters may have different levels of precision than others. Eventually, it would be beneficial to allow for a more abstract definition of design distinctness to support more complicated design circumstances.

**Unknown precision:** In most design/manufacturing settings, the resolution $r$ is known a priori. It is possible, however, that this will need to be discovered over the course of the search process (or, maybe, after the search process). For this situation, our proposed policy would benefit from an adaptive $r$ definition, which might start at a very large value e.g., 10% of the maximum distance between points in the domain, and then scale down as the search progresses.

**Parallelism:** ECI is limited to the sequential setting; extending the search policy to the batch setting and finding a computationally tractable policy is non-trivial.

**Judging performance:** Some of our proposed performance criteria are only computable for benchmark problems where the true satisfactory region can be estimated. In practice, judging the results of a search algorithm is less straightforward. We believe creating more practical performance criteria can accelerate our understanding of this problem.

**Acknowledgements**

We want to thank Sajad Haghanifar and Paul Leu for providing the scanning electron microscope images and motivating the research presented in this paper.
References


Beyond the Pareto Efficient Frontier: Constraint Active Search for Multiobjective Experimental Design


