Supplementary Material: 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]

A. Theoretical Analysis

We first recap relevant important definitions as well as important lemmas. We then prove the three main theoretical results for expected coverage improvement (ECI).

Fill distance is a 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:

Definition 1. *(Fill Distance) Given a set of sample points* $\mathbf{X}$*, the fill distance is formally defined as the following:*

$$\text{FILL}(\mathbf{X}, \mathcal{S}) = \sup_{\mathbf{x}\in\mathcal{S}} \min_{\mathbf{x}_j\in\mathbf{X}} d(\mathbf{x}_j, \mathbf{x}). \tag{1}$$

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

There is a notion of minimal fill, in the following sense:

$$\rho = \min_{\mathbf{X}\in\mathcal{X}^n} \text{FILL}(\mathbf{X}, \mathcal{S}). \tag{2}$$

Low-discrepancy sequences and Latin hypercubes achieve low fill in simple domains. However, computing the minimal fill is generically NP-hard (Pronzato & Müller, 2012). Below, we list two key lemmas regarding fill distance.

Lemma 1. *(Pronzato & Müller, 2012) Given a set* $\mathcal{S}$*, let the greedy solution to the minimal fill (Equation 2) be the algorithm that sequentially adds the point in* $\mathcal{S}$ *furthest away from the closest point in* $\mathbf{X}$*. The greedy solution to the minimal fill achieves an approximation ratio of 2 i.e., the fill of* $\mathbf{X}$ *is bounded above by* 2ρ

Lemma 2. *(Pronzato & Müller, 2012) Let* $\rho(n)$ *be the minimal fill for* n *points in a set* $\mathcal{S}$ *i.e., the solution to Equation 2. Then:*

$$\lim_{n\to\infty} \rho(n) = 0.$$

In other words, as n *grows infinitely large, the minimal fill of* n *points in* $\mathcal{S}$ *decreases to zero.*

Having established these key definitions and lemmas, we now prove the main three theoretical results for ECI. For all three, we assume that $\mathcal{S}$ is a set of m points, known *a priori*. That means we focus only on covering $\mathcal{S}$. We further assume that all points lie in a metric space, and that there is a distance function between u, v, v' such that: i) $d(u, v) = 0$; ii) $d(u, v) = d(v, u)$; iii) $d(u, v) \leq d(u, v') + d(v', v)$.

Theorem 1. *If* $\mathcal{S}$ *is dense, ECI produces a dense sequence of evaluation points in* $\mathcal{S}$*.*

Proof. This is a result of Theorem 3 and Lemma 2. By Theorem 3, ECI produces a set of points bounded above by 4ρ. As n grows to infinity, by Lemma 2, this upper bound therefore goes to zero. This is equivalent to dense sampling. □

Theorem 2. *If we have a budget of one iteration left, then ECI is one-step Bayes-optimal among all feasible policies.*

Proof. By construction, ECI maximizes the improvement in utility. It is therefore one-step Bayes-optimal among all feasible policies. □

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

Proof. This proof is divided into two parts, devoted to the first and second statements, respectively. We first prove the first statement: *For any* n*, ECI produces a fill bounded above by* 4ρ*.*

Let $\mathbf{X}^* = \{\mathbf{x}_i^*\}_{i=1}^n$ be the optimal solution to Equation (2) i.e., the set of points that minimizes the fill distance. Let r be ECI neighborhood radius. Let us assume that $r = 2\rho$.

The optimal solution $\mathbf{X}^*$ covers all points in $\mathcal{S}$ with balls of radius ρ. This creates a partition of $\mathcal{S}$ into n groups, where each $u \in \mathcal{S}$ is mapped to its closest $\mathbf{x}_i^*$. Let $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^n$ be the points selected by ECI. There are two cases:

[1]SigOpt, an Intel company, San Francisco, CA, USA. Correspondence to: Gustavo Malkomes <gustavo.malkomes@intel.com>.

Proceedings of the 38^{th} International Conference on Machine Learning, PMLR 139, 2021.

Case 1: Each $\mathbf{x}_i$ gets mapped to a different optimal group in $\mathbf{X}^*$. In this case, the ECI points cover all points in S with a radius $r = 2\rho$.

Case 2: At least two points $\mathbf{x}_i, \mathbf{x}_j$ are mapped to the same optimal group $\mathbf{x}_k^*$. That means that there is an optimal group $\mathbf{x}_l^*$ that has no ECI point. This contradicts our earlier selection of $\mathbf{X}$ through maximizing ECI; either $\mathbf{x}_i$ or $\mathbf{x}_j$ can cover all points that $\mathbf{x}_k^*$ covers since $r = 2\rho$, and swapping any point in $\mathbf{x}_l^*$ by one of these two points would either increase the coverage or would yield the same coverage but with largest distance from the selected points.

To conclude this first part, we note that we can run the greedy solution to the fill distance to get a good initial value for r: $2\rho \leq r \leq 4\rho$, which gives a final 4-approximation to the optimal fill.

We now prove the second statement: *Furthermore, there exists a finite* $\mathbf{n}^*$ *for which ECI produces a 2-approximation ratio of* ρ *for all* $n \geq \mathbf{n}^*$. ECI is designed to minimize the volume coverage; thus, there will exist some $\mathbf{n}^*$ for which the ECI is zero for all points. When this occurs, we will return to selecting the point furthest away from all observations. Specifically, we can say that ECI's fill distance is bounded above by 2ρ (from Lemma 1) plus a penalty $\varepsilon(n; \mathbf{n}^*)$ associated with the first $\mathbf{n}^*$ samples. Asymptotically this approaches 2ρ as n increases, which gives us an approximation ratio of 2. □

A.1. ECI Time Complexity

Computing a point-wise estimate of ECI is linear with the number of metrics m. Indeed, this is more efficient than hypervolume-based approaches, which scale super-polynomially in m. Assuming we want to use N samples to estimate ECI, the complexity is $\mathcal{O}(mN \max(c, n))$, where c is the cost of making predictions, and n is the number of observations. For GP models, c is quadratic in n.

B. Additional Notes on the Pareto Frontier

To further demonstrate the shortcoming of relying only on the Pareto efficient frontier in multiobjective design problems, we consider the following scenarios. In the first case, consider the objective functions

$$\begin{aligned} f_1(\mathbf{x}) &= \exp(-0.5((\mathbf{x}_1 - 0.2)^2 + (\mathbf{x}_2 - 0.5)^2)), \\ f_2(\mathbf{x}) &= \exp(-0.5((\mathbf{x}_1 - 0.8)^2 + (\mathbf{x}_2 - 0.5)^2)). \end{aligned} \tag{3}$$

In the second case, consider the following objective functions

$$\begin{aligned} f_1(\mathbf{x}) &= \exp(-0.5((\mathbf{x}_1 - 0.2)^2 + 4(\mathbf{x}_2 - 0.5)^2)), \\ f_2(\mathbf{x}) &= \exp(-0.5((\mathbf{x}_1 - 0.8)^2 + 4(\mathbf{x}_2 - 0.5)^2)). \end{aligned} \tag{4}$$

Let us assume the thresholds $\boldsymbol{\tau} = (0.88, 0.88)$. If we try to optimize (3) and (4) over the domain $[0, 1]^2$, we can see that they have the same Pareto efficient frontier. However, the satisfactory regions $\mathcal{S}$ in the parameter space are different. We illustrate this difference in Figure 1. In practice, the different satisfactory regions $\mathcal{S}$ yield very different interpretations in the design parameters for the decision makers.

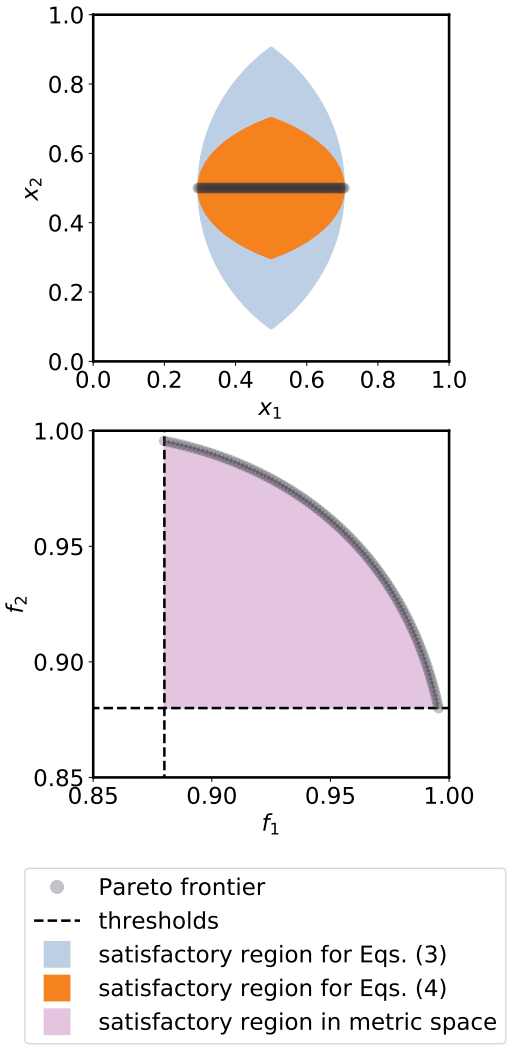

Figure 1. Comparing Pareto efficient frontier and satisfactory regions from equations (3) and (4). The top figure shows the respective $\mathcal{S}$ in the parameter space, the bottom shows the corresponding region in the metric space. Both problems share the same Pareto frontier (labeled in grey circles).

C. Derivation of EZ Policy

Here, we derive the EZ policy. First, we note the symmetric property of mutual information, namely, $\text{MI}(y; z) = \text{MI}(z; y)$. We can express the mutual information quantities

themselves as,

$$\text{MI}(y; z) = H(y) - \mathbb{E}_z\big[H(y \mid z)\big]$$
$$\text{MI}(z; y) = H(z) - \mathbb{E}_y\big[H(z \mid y)\big]$$

Let us look at $\text{MI}(z; y)$ first. Given y, there is no uncertainty in the binary variable z, thus $H(z \mid y) = 0$. This leaves $MI(z; y) = H(z)$. We should get the same conclusion from $MI(y; z)$ as well.

$$\begin{aligned} MI(y; z) &= H(y) - \mathbb{E}_z\big[H(y \mid z)\big] \\ &= H(y) - \sum_{j=\{0,1\}} p(z = j) H(y \mid z = j) \end{aligned} \quad (5)$$

Let $\alpha = \frac{\tau - \mu}{\sigma}$, and let $\phi(\cdot)$ and $\Phi(\cdot)$ be the PDF and CDF of a standard normal distribution respectively. Under the noiseless assumption, we know that $p(z = 0) = p(y \geq \tau) = 1 - \Phi(\alpha)$. We can also derive $H(y|z)$, where

$$H(y|z = 0) = H(y) + \ln(1 - \Phi(\alpha)) + \frac{\alpha\phi(\alpha)}{2(1 - \Phi(\alpha))},$$
$$H(y|z = 1) = H(y) + \ln(\Phi(\alpha)) - \frac{\alpha\phi(\alpha)}{2\Phi(\alpha)}.$$

Substituting these quantities back into Equation (5), we can finally conclude that:

$$\begin{aligned} MI(y; z) &= -(1 - \Phi(\alpha)) \ln(1 - \Phi(\alpha)) - \Phi(\alpha) \ln(\Phi(\alpha)) \\ &= H(z), \end{aligned}$$

as expected. Since $Z(x)$ is a Bernoulli random variable, the maximum of the entropy of a Bernoulli random variable is achieved when $p(Z(x) = 1) = 0.5$. This precisely corresponds to the decision boundary of the $\mathcal{S}$.

D. Experiment Setup

In this section, we detail the setup for all the numerical experiments presented in the main text.

D.1. Modeling decisions and hyperparameters

Budget: For all REPROBLEM experiments, we set the budget b to be $b = \min(100, 40d)$, where d is the number of input dimension. We choose the budgets to be 160, 100, and 130 points for the additive manufacturing, EEG, and plasma physics applications.

Initialization: Before selecting points, each strategy is given the same list of $3d$ randomly selected initial points. We run 20 repetitions for each function, changing the initial list of points on each repetition. The total number of observations is $3d + b$ for each problem.

Rescaling: For the REPROBLEM, we rescale the input domain into $[0, 1]^d$, and the output to be in the range of $[0, 1]$.

Thresholds: For each problem, we choose the thresholds $\boldsymbol{\tau}$ so that the satisfactory region $\mathcal{S}$ is approximate $3 - 5\%$ of the domain $\mathcal{X}$. The exception is the additive manufacturing problem where we use thresholds defined by the original authors. In other words, uniformly sampling from the domain only has $3-5\%$ chance of landing inside of $\mathcal{S}$. We list the $\boldsymbol{\tau}$ *before normalization* for each problem in Section D.2-D.5.

Surrogate models: For all experiments, we use independent GPs to model each objective function. We assume the GPs to have zero mean. Furthermore, we fix the hyperparameters of the GPs *a priori* for all the policies. We randomly select 100 points from the domain and train the GP hyperparameters by minimizing the negative marginal log likelihood across 10 restarts. We use the SciPy SLSQP optimizer to perform the hyperparameter optimization (Virtanen et al., 2020).

Discretization: To facilitate the comparison between policies we discretize the input space with a pool of points. All models are given the same pool of points and none of them can select the same point multiple times. For the REPROBLEM we selected around $5000d$ points. For the additive, EGG, and plasma application problems, we generate around 2500, 15000 and 10000 points, respectively.

Resolution parameter: We choose the r parameters *a priori* for each problem. Our selection is based on a brief analysis of each problem. Specifically, we make sure that each point has on average at least 4 neighbors on the above mentioned pool of points. We list the r values *after normalization* for each problem in subsequent sections D.2-D.5.

D.2. Multiobjective design problem suite

We adopt eight multiobjective engineering design problems presented in the REPROBLEM suite (Tanabe & Ishibuchi, 2020). We select all five two-metric problems; these problems are: four bar truss design, reinforced concrete beam design, pressure vessel design, hatch cover design, and coil compression spring design. We also select three three-metric problems; these are: two bar truss design, welded beam design, and disc brake design. We follow the numbering convention from the source code. The r and $\boldsymbol{\tau}$ parameters for each problem are shown in Table 1.

D.3. Additive manufacturing

We use the uniform nanocone design problems presented in Haghanifar et al. (2020) with the following parameter settings (shown in Table 2). We impose the same linear constraints on the parameters: bottom diameter $\geq$ top diameter, and bottom diameter $\leq$ maximum cone width. We consider the thresholds to be $\tau_{\text{normal}} = 0.3$ and $\tau_{\text{oblique}} = 1.5$.

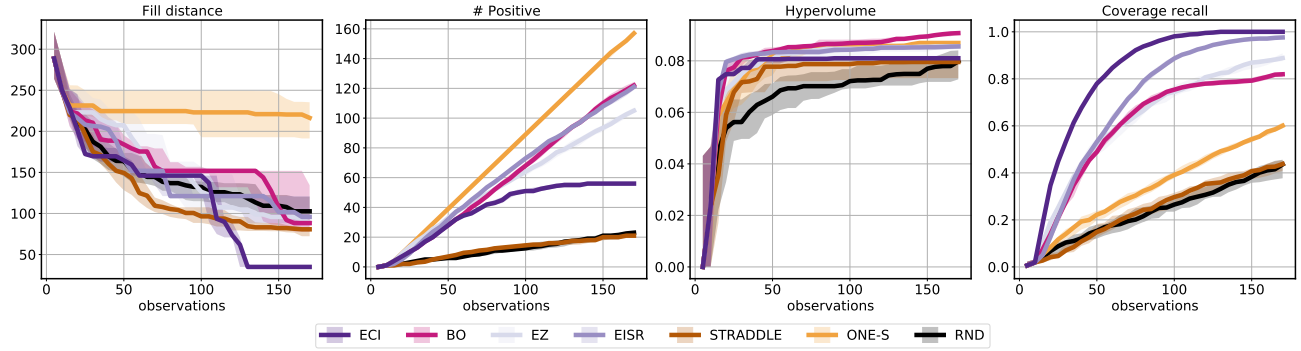

Figure 2. Median metric values across 20 repetitions for the **additive manufacturing** application. We plot the fill distance, number of positive samples, hypervolume, and coverage recall respectively. Shaded region corresponds to values between lower and upper quartiles.

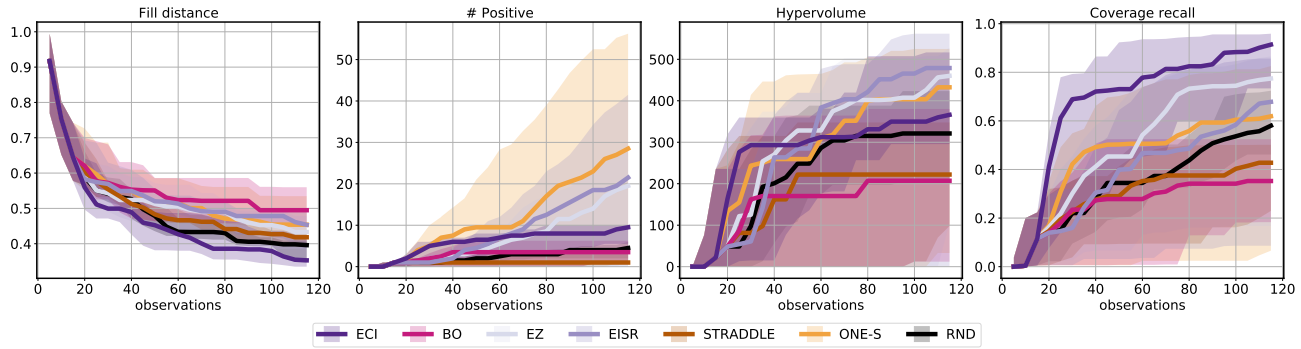

Figure 3. Median metric values across 20 repetitions for the **EEG** application. We plot the fill distance, number of positive samples, hypervolume, and coverage recall respectively. Shaded region corresponds to values between lower and upper quartiles.

Table 1. REPROBLEM description

Name	r	thresholds $\boldsymbol{\tau}$
RE21	0.08	(2000, 0.02)
RE22	0.04	(250, 125)
RE23	0.08	(20000, 50000)
RE24	0.01	(250, 10)
RE25	0.08	(3, 2)
RE31	0.08	(80, 120, 150)
RE32	0.08	(10, 20, 5)
RE33	0.08	(2, 3, 4)

Table 2. Additive manufacturing design parameters

Design Parameters	Bounds (nm)
Max cone width (pitch size)	[1, 400]
Bottom diameter	[1, 400]
Top diameter	[1, 400]
Height	[1, 800]

D.4. EEG for brain activity reconstruction

The domain is simplified from a realistic brain geometry to concentric spheres of size of 0.085 m, 0.092 m and 0.1 m for the outer brain, skull and scalp radii, respectively. The conductivities of the brain, skull and scalp are 0.33 S/m, 0.0125 S/m, 0.33 S/m, respectively. The 28 electrodes on the scalp are placed approximately evenly distributed (selected from a spiral spherical point selection (Saff & Kuijlaars, 1997)) on the top half of the head and above the plane at a 45 degree angle above the nose. The reference point, at which zero potential is defined, was chosen to be $(0, 0, 0.1)$ m, though this choice was arbitrary and should have no impact. To simplify the inverse problem, we consider only a single dipole in the brain.

The method of fundamental solutions (MFS) was used to simulate the scalp potential by coupling three solutions in the brain, skull and scalp generated by a proposed dipole during the iterative optimization process (Ala et al., 2015). The brain, skull and scalp boundaries are populated with 102, 114 and 134 collocations points, respectively, chosen using the spiral strategy. MFS requires fictitious boundaries on which to place source points (using the spiral strategy) outside of the domain where the solution is considered; we created these boundaries for each of the three regions through inflation/deflation. For the coupled solution in the brain, we inflated the brain boundary to a radius of 0.0961 m

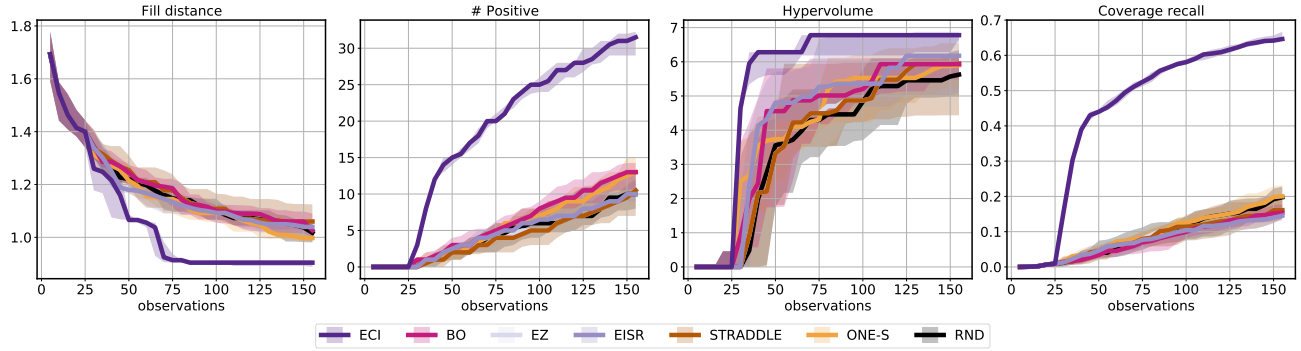

Figure 4. Median metric values across 20 repetitions for the **plasma physics** problem. We plot the fill distance, number of positive samples, hypervolume, and coverage recall respectively. Shaded region corresponds to values between lower and upper quartiles.

and placed 40 source points on it. For the coupled solution in the skull, we deflated the inner boundary to a radius of 0.0791 m and placed 40 source points on it; we also inflated the outer boundary to 0.105 m and placed 45 source points on it. For the coupled solution in the scalp, we deflated the inner boundary to 0.0852 m and placed 45 source points on it; we also inflated the outer boundary to 0.115 m and placed 107 source points on it.

For the empirical study, we placed the dipole at $(0.0, -0.017, 0.051)$ m, with moment $(0.5, 0.4, 0.3)$ A$\cdot$m. We consider the thresholds to be $\boldsymbol{\tau} = (20, 40)$ μV, i.e., dipoles are considered satisfactory if the 25th percentile of the 28 absolute difference values is less than 20 μV and the 75th percentile is less than 100 μV. The optimization domain was defined to be the direct sum of the ball of radius 0.085 m (for the dipole location) and the ball of radius 1 A$\cdot$m(for the dipole moment). The sense of distance (as needed for the fill distance and the resolution parameter) is complicated for this problem, given the complicated domain. Suggested dipoles are considered *too close* for resolution purposes if both the Euclidean distance between dipole locations is less than 0.02 and the Euclidean distance between dipole moments is less than 0.25. The fill distance is the sum of these two Euclidean distances.

D.5. Plasma physics

PyPlasmaOpt paramaterizes 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:

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

For each coil, the parameters c_k and s_k represent the search space that we must search over. In our paper, we assume that $n_{order} = 1$, which provides three parameters for each of three dimensions, resulting in nine parameters total.

The search space of each parameter is given by $[-0.1, 1]$, and so our total search space is simply $[-0.1, 1]^9$

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 general terms:

$$f(\mathbf{x}) = R_{\text{magnetic}}(\mathbf{x}) + R_{\text{transform}}(\mathbf{x}) + R_{\text{shape}}(\mathbf{x}).$$

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.

Our objective function is bounded below by zero and is unbounded from above. We consider the threshold $\tau = 10$, which represents a rather modest deviation from the minimal value.

E. Complete Experiment Results

In this section we present all results for our experiments. We show the mean and standard error for all problems in Tables 3–5. We also plot each metric as we gather more observations for the application problems in Figures 2–4.

E.1. Coverage recall sensitivity

The metric *coverage recall* depends of the choice of parameter resolution r. During evaluation, we compute the coverage of each sampled point considering a particular choice of r. In this section we investigate how the coverage recall metric changes as we vary this parameter. Specifically we will consider the additive manufacturing results.

First, we evaluate different values of r over the range of $[10, 70]$, as a resolution parameter smaller than 10 nm is unrealistic. We display these results on Figure 5. Notice that

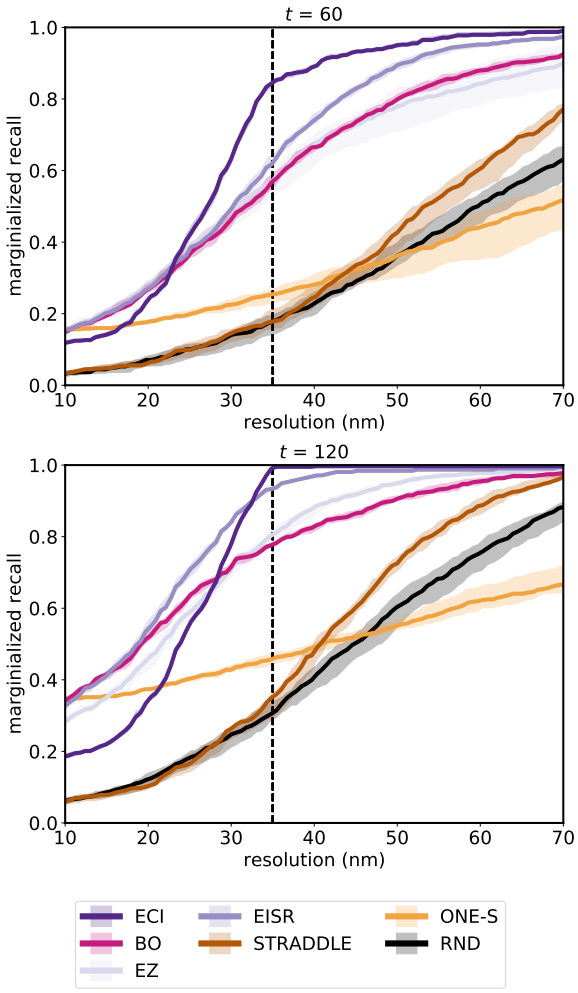

Figure 5. Marginalized recalls for r over $[10, 70]$. The top figure shows the value after 60 samples; the bottom is at the termination. We also label the $r = 35$ with dotted line; this is the r value known a priori for ECI.

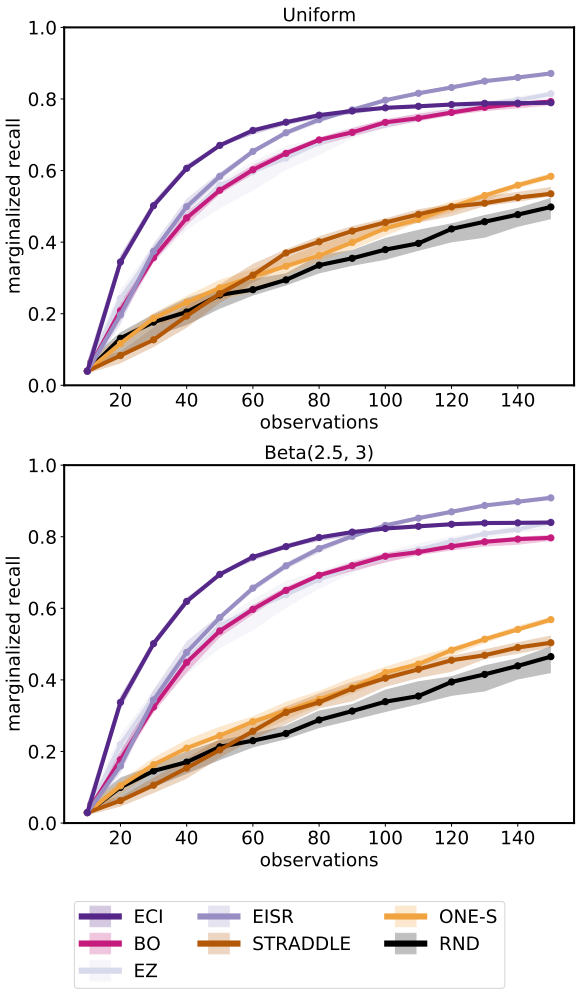

Figure 6. Marginal recall for r from uniform distribution (top) and Beta(2.5, 3) distribution (bottom).

for small r values, greedy methods, e.g., ONE-S, tend to perform well. When r is small, to the point that it only counts the coverage contribution of a single point, this metric converges to the number of positive samples (normalized). For large values of r, all methods will recover perfect recall.

We also consider the marginalized recall for different distribution of r over the range of $[10, 70]$. Figure 6 shows the marginalized recall for r drawn from a uniform distribution and a Beta(2.5, 3) distribution. We observe that ECI has slightly lower marginalized recall than EISR and EZ, since it is more exploratory than these two methods. Therefore, it has much lower recall values when r is low, as we have just shown in Figure 5.

References

Ala, G., Fasshauer, G., Francomano, E., Ganci, S., and McCourt, M. J. The method of fundamental solutions in solving coupled boundary value problems for M/EEG. *SIAM Journal on Scientific Computing*, 37(4): B570–B590, 2015.

Fasshauer, G. E. and McCourt, M. J. *Kernel-based Approximation Methods Using Matlab*. World Scientific, 2015.

Haghanifar, S., McCourt, M., Cheng, B., Wuenschell, J., Ohodnicki, P., and Leu, P. W. Discovering high-performance broadband and broad angle antireflection surfaces by machine learning. *Optica*, 7(7):784–789, Jul 2020. doi: 10.1364/OPTICA.387938.

Joy, C., Boyle, P. P., and Tan, K. S. Quasi-Monte Carlo

Table 3. Application results, consisting of the mean and standard error over 20 independent trials. Best mean value in bold.

Function	Methods	Fill distance ↓	# Positive ↑	Coverage Recall ↑	Hypervolume ↑
glass-4d	RND	108.22 ± 3.83	22.40 ± 0.92	0.42 ± 0.01	7.94E-02 ± 1.32E-03
	ONE-S	211.86 ± 7.77	**156.75 ± 0.83**	0.60 ± 0.00	8.77E-02 ± 3.95E-04
	EZ	92.37 ± 5.00	106.05 ± 0.66	0.89 ± 0.01	8.52E-02 ± 4.35E-04
	STRADDLE	79.30 ± 1.42	21.10 ± 0.54	0.43 ± 0.01	7.78E-02 ± 1.23E-03
	BO	107.84 ± 8.38	121.95 ± 0.73	0.82 ± 0.00	**9.07E-02 ± 1.62E-04**
	EISR	102.01 ± 3.83	121.40 ± 0.76	0.98 ± 0.00	8.56E-02 ± 2.83E-04
	ECI	**34.92 ± 0.00**	56.00 ± 0.47	**1.00 ± 0.00**	8.07E-02 ± 3.65E-04
eeg	RND	**0.38 ± 0.01**	4.40 ± 0.50	0.58 ± 0.03	3.17E+02 ± 3.40E+01
	ONE-S	0.48 ± 0.02	**31.20 ± 6.32**	0.50 ± 0.08	3.45E+02 ± 5.55E+01
	EZ	0.44 ± 0.02	17.95 ± 3.32	0.58 ± 0.08	3.45E+02 ± 5.11E+01
	STRADDLE	0.41 ± 0.01	2.00 ± 0.35	0.38 ± 0.04	2.28E+02 ± 3.78E+01
	BO	0.52 ± 0.02	4.30 ± 0.90	0.35 ± 0.05	2.40E+02 ± 4.03E+01
	EISR	0.45 ± 0.02	22.85 ± 4.72	0.52 ± 0.08	3.55E+02 ± 5.57E+01
	ECI	**0.38 ± 0.02**	8.30 ± 0.87	**0.79 ± 0.06**	**3.65E+02 ± 3.38E+01**
plasma	RND	1.02 ± 0.01	10.85 ± 0.85	0.19 ± 0.01	5.63E+00 ± 2.59E-01
	ONE-S	1.01 ± 0.01	12.80 ± 0.72	0.20 ± 0.01	5.88E+00 ± 1.85E-01
	EZ	1.07 ± 0.02	10.70 ± 1.10	0.17 ± 0.01	5.48E+00 ± 2.76E-01
	STRADDLE	1.07 ± 0.02	10.70 ± 1.10	0.17 ± 0.01	5.48E+00 ± 2.76E-01
	BO	1.04 ± 0.01	12.60 ± 0.77	0.15 ± 0.01	5.74E+00 ± 1.82E-01
	EISR	1.03 ± 0.01	10.55 ± 0.47	0.14 ± 0.01	6.08E+00 ± 2.41E-01
	ECI	**0.90 ± 0.01**	**31.25 ± 0.51**	**0.65 ± 0.00**	**6.64E+00 ± 8.00E-02**

Table 4. REPROBLEM with three metrics, mean and standard error over 20 independent trials. Best mean in bold.

Function	Methods	Fill distance ↓	# Positive ↑	Coverage Recall ↑	Hypervolume ↑
RE31	RND	0.21 ± 0.01	25.50 ± 1.02	0.69 ± 0.02	4.65E-05 ± 1.81E-06
	ONE-S	0.47 ± 0.03	**98.95 ± 0.49**	0.37 ± 0.03	3.78E-05 ± 2.36E-06
	EZ	0.17 ± 0.00	62.45 ± 0.67	0.73 ± 0.01	2.62E-05 ± 2.85E-06
	STRADDLE	0.18 ± 0.00	18.95 ± 0.72	0.51 ± 0.02	3.61E-05 ± 3.46E-06
	BO	0.33 ± 0.01	14.95 ± 0.51	0.24 ± 0.01	5.90E-05 ± 2.21E-07
	EISR	0.17 ± 0.00	88.25 ± 0.41	0.57 ± 0.01	**6.23E-05 ± 8.39E-08**
	ECI	**0.08 ± 0.00**	49.15 ± 0.70	**1.00 ± 0.00**	5.47E-05 ± 7.82E-07
RE32	RND	0.39 ± 0.01	16.75 ± 0.70	0.12 ± 0.00	6.60E-13 ± 1.90E-14
	ONE-S	0.66 ± 0.02	82.20 ± 1.53	0.10 ± 0.01	7.59E-13 ± 7.80E-15
	EZ	0.66 ± 0.02	83.05 ± 1.44	0.11 ± 0.01	**7.74E-13 ± 1.09E-14**
	STRADDLE	**0.36 ± 0.00**	13.10 ± 0.61	0.05 ± 0.00	7.56E-13 ± 1.94E-14
	BO	0.63 ± 0.02	7.40 ± 0.58	0.03 ± 0.00	6.45E-13 ± 4.15E-14
	EISR	0.59 ± 0.02	70.70 ± 2.68	0.09 ± 0.01	7.30E-13 ± 1.40E-14
	ECI	0.57 ± 0.02	**89.00 ± 1.77**	**0.63 ± 0.03**	7.61E-13 ± 8.11E-15
RE33	RND	0.34 ± 0.01	23.80 ± 1.01	0.14 ± 0.00	5.53E-05 ± 1.28E-06
	ONE-S	0.60 ± 0.03	**101.55 ± 0.35**	0.18 ± 0.01	7.03E-05 ± 1.09E-06
	EZ	0.26 ± 0.01	34.05 ± 1.43	0.25 ± 0.00	2.77E-05 ± 2.05E-06
	STRADDLE	0.35 ± 0.01	13.70 ± 0.54	0.10 ± 0.00	2.92E-05 ± 2.05E-06
	BO	0.42 ± 0.01	36.55 ± 0.72	0.11 ± 0.00	**8.65E-05 ± 4.73E-07**
	EISR	**0.24 ± 0.00**	100.05 ± 0.42	0.23 ± 0.00	7.46E-05 ± 5.77E-07
	ECI	0.27 ± 0.00	100.95 ± 0.36	**0.73 ± 0.00**	7.91E-05 ± 1.36E-07

Table 5. REPROBLEM with two metrics, mean and standard error over 20 independent trials. Best mean in bold.

Function	Methods	Fill distance ↓	# Positive ↑	Coverage Recall ↑	Hypervolume ↑
RE21	RND	0.35 ± 0.01	14.05 ± 0.69	0.09 ± 0.00	6.62E-03 ± 3.06E-04
	ONE-S	0.34 ± 0.01	**100.20 ± 0.25**	0.45 ± 0.00	1.04E-02 ± 1.33E-04
	EZ	0.23 ± 0.00	52.20 ± 1.10	0.25 ± 0.00	3.57E-03 ± 3.15E-04
	STRADDLE	0.28 ± 0.00	22.70 ± 0.61	0.09 ± 0.00	3.59E-03 ± 3.04E-04
	BO	0.47 ± 0.01	36.95 ± 0.63	0.09 ± 0.00	**1.30E-02 ± 7.57E-06**
	EISR	**0.20 ± 0.00**	97.55 ± 0.22	0.18 ± 0.00	8.18E-03 ± 1.85E-04
	ECI	0.22 ± 0.00	**100.20 ± 0.25**	**0.69 ± 0.00**	9.05E-03 ± 9.92E-05
RE22	RND	0.17 ± 0.00	26.75 ± 0.91	0.20 ± 0.01	5.80E-07 ± 1.32E-08
	ONE-S	0.36 ± 0.02	32.00 ± 10.18	0.04 ± 0.01	2.61E-07 ± 2.79E-08
	EZ	0.33 ± 0.02	4.85 ± 1.53	0.03 ± 0.00	2.62E-07 ± 2.72E-08
	STRADDLE	0.21 ± 0.00	9.60 ± 0.47	0.06 ± 0.00	2.42E-07 ± 2.60E-08
	BO	0.38 ± 0.01	2.50 ± 0.34	0.02 ± 0.00	2.31E-07 ± 2.73E-08
	EISR	0.20 ± 0.00	9.60 ± 0.43	0.05 ± 0.00	3.49E-07 ± 2.01E-08
	ECI	**0.11 ± 0.00**	**97.10 ± 0.38**	**0.79 ± 0.00**	**7.63E-07 ± 3.82E-09**
RE23	RND	0.40 ± 0.01	21.20 ± 0.85	0.15 ± 0.01	8.28E-04 ± 1.81E-05
	ONE-S	0.67 ± 0.02	**101.35 ± 0.33**	0.09 ± 0.01	9.55E-04 ± 1.54E-05
	EZ	**0.26 ± 0.01**	65.30 ± 1.12	0.31 ± 0.01	8.50E-04 ± 1.44E-05
	STRADDLE	0.27 ± 0.00	25.95 ± 0.79	0.14 ± 0.00	8.14E-04 ± 2.14E-05
	BO	0.47 ± 0.02	53.25 ± 0.97	0.18 ± 0.00	**1.00E-03 ± 9.59E-07**
	EISR	0.30 ± 0.01	96.95 ± 0.43	0.35 ± 0.00	9.75E-04 ± 5.45E-06
	ECI	0.37 ± 0.01	98.80 ± 0.38	**0.82 ± 0.00**	9.00E-04 ± 3.42E-06
RE24	RND	0.10 ± 0.01	12.90 ± 0.96	0.26 ± 0.02	6.09E-03 ± 1.08E-04
	ONE-S	0.07 ± 0.01	**79.60 ± 0.25**	0.77 ± 0.01	7.19E-03 ± 3.25E-05
	EZ	0.04 ± 0.00	39.45 ± 0.43	0.46 ± 0.01	2.81E-03 ± 3.78E-04
	STRADDLE	0.04 ± 0.00	44.45 ± 1.04	0.58 ± 0.01	4.91E-03 ± 2.51E-04
	BO	0.19 ± 0.02	50.95 ± 1.68	0.19 ± 0.01	**7.67E-03 ± 3.65E-06**
	EISR	0.05 ± 0.00	76.95 ± 0.21	0.69 ± 0.01	7.51E-03 ± 4.23E-06
	ECI	**0.02 ± 0.00**	68.15 ± 0.36	**1.00 ± 0.00**	7.05E-03 ± 2.85E-05
RE25	RND	0.24 ± 0.01	11.40 ± 0.71	0.52 ± 0.03	3.14E-09 ± 1.06E-10
	ONE-S	0.42 ± 0.02	**98.20 ± 0.53**	0.42 ± 0.02	3.33E-09 ± 1.62E-10
	EZ	0.39 ± 0.01	98.00 ± 0.47	0.57 ± 0.04	3.64E-09 ± 1.09E-10
	STRADDLE	0.17 ± 0.00	13.65 ± 0.48	0.45 ± 0.01	3.52E-09 ± 4.97E-11
	BO	0.24 ± 0.01	40.80 ± 0.49	0.75 ± 0.01	**4.16E-09 ± 4.57E-12**
	EISR	0.11 ± 0.00	81.40 ± 0.38	0.97 ± 0.00	4.09E-09 ± 1.57E-11
	ECI	**0.08 ± 0.00**	29.70 ± 0.65	**1.00 ± 0.00**	3.59E-09 ± 3.43E-11

methods in numerical finance. *Management Science*, 42 (6):926–938, 1996.

Pronzato, L. and Müller, W. G. Design of computer experiments: space filling and beyond. *Statistics and Computing*, 22(3):681–701, 2012.

Saff, E. B. and Kuijlaars, A. B. Distributing many points on a sphere. *The mathematical intelligencer*, 19(1):5–11, 1997.

Tanabe, R. and Ishibuchi, H. An easy-to-use real-world multi-objective optimization problem suite. *Applied Soft Computing*, 89:106078, 2020. ISSN 1568-4946. doi: https://doi.org/10.1016/j.asoc.2020.106078.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.