



Cite this: DOI: 10.1039/d6dd00123h

# BEACON: a Bayesian optimization inspired strategy for efficient novelty search

Wei-Ting Tang,<sup>ab</sup> Ankush Chakrabarty<sup>c</sup> and Joel A. Paulson<sup>id</sup> \*<sup>ab</sup>

Novelty search (NS) aims to uncover diverse system behaviors through simulation or experiment without requiring a pre-specified scalar objective. This capability is especially relevant to modern discovery problems in chemistry, materials science, and molecular design, where researchers often seek broad coverage of attainable property space rather than a single optimum and where each evaluation may require a costly computation or experiment. For such expensive black-box settings, we propose BEACON, a sample-efficient NS strategy inspired by Bayesian optimization principles. BEACON models the input-to-outcome mapping using multi-output Gaussian processes and selects new inputs by scoring how far plausible posterior outcomes lie from a denoised archive of previously observed outcomes. This gives a distance-based novelty acquisition that accounts for predictive uncertainty and observational noise while operating directly in continuous outcome space, rather than requiring direct optimization over a discretized partition of behaviors. By leveraging efficient posterior sampling together with scalable high-dimensional Gaussian process models, the proposed framework can be extended to settings with large data sets and high-dimensional design variables. We demonstrate BEACON on established benchmark problems together with real-world case studies in materials and molecular discovery. Across these settings, BEACON consistently discovers broader sets of distinct behaviors than several competing baselines under limited evaluation budgets.

Received 15th March 2026

Accepted 22nd June 2026

DOI: 10.1039/d6dd00123h

rsc.li/digitaldiscovery

## 1 Introduction

Across chemistry, materials science, and molecular design, researchers increasingly rely on expensive simulations, data-driven models, and automated experiments to probe complex black-box systems. In many of these settings, the goal is not only to identify a single best candidate, but also to uncover a diverse set of attainable outcomes that can reveal new mechanisms, broaden training data, or expose promising but previously overlooked regions of design space. These needs arise naturally in applications such as molecular discovery, materials screening, and autonomous experimentation, where broad exploration of outcome space can be as valuable as objective-driven optimization.<sup>1–4</sup>

To formalize this setting, consider a black-box map  $f: X \rightarrow O$  that takes user-defined system inputs  $x \in X$  to outcomes (or behaviors)  $f(x) \in O$ . Classical derivative-free search methods, including simplex methods, meta-heuristics, and Bayesian optimization (BO), are designed primarily to optimize some prescribed objective defined over the outcomes; see, for

example.<sup>5–8</sup> This objective-centric viewpoint is highly effective when the user knows what should be maximized or minimized. However, if the real aim is to understand the range of outcomes that the system can realize, then reducing the problem to a single scalar objective can be restrictive and, in some cases, fundamentally misaligned with the scientific goal.

A simple example arises in molecular/materials discovery. Suppose  $f$  maps a candidate design to a vector of experimentally relevant properties, such as solubility, uptake capacity, stability, or selectivity. If the objective is to identify new types of property combinations rather than only the single best design under one fixed score, then standard optimization becomes awkward. Any scalarization will prioritize certain regions of the outcome space and can easily overlook scientifically useful intermediate or rare behaviors. More broadly, the corresponding search landscape is often highly multi-modal and deceptive, making exhaustive exploration prohibitively expensive when each evaluation involves a simulation, synthesis, or physical measurement.

Novelty search (NS) offers a natural alternative by shifting the emphasis from objective maximization to the discovery of previously unseen behaviors.<sup>9,10</sup> Instead of rewarding progress toward a predefined target, NS promotes candidates whose outcomes differ from what has already been observed. In the classical archive-based view, a candidate is novel when its behavior is far from the archive of behaviors discovered so far. This is the sense in which we use the term NS throughout this work. The broader goal

<sup>a</sup>University of Wisconsin–Madison, Department of Chemical and Biological Engineering, Madison, WI, USA. E-mail: joel.paulson@wisc.edu

<sup>b</sup>The Ohio State University, Department of Chemical and Biomolecular Engineering, Columbus, OH, USA

<sup>c</sup>Insulet Corporation, Acton, MA, USA



is to build a discovery campaign that covers a wide range of attainable behaviors under a limited evaluation budget. We later measure this coverage using reachability, but the acquisition function itself operates in the continuous outcome space rather than over a discrete coverage count. This distinction lets BEACON use a continuous, distance-based notion of novelty while still evaluating success in terms of behavior-space coverage.

This perspective has proven useful in automated scientific discovery,<sup>1,2</sup> reinforcement learning,<sup>11</sup> and more recently, autonomous experimentation workflows that explicitly seek previously unobserved phenomena rather than only improved objective values.<sup>12</sup> Most established NS algorithms, however, rely on meta-heuristics such as evolutionary search. Genetic algorithms, for example, have been used to promote behavioral diversity in deceptive landscapes;<sup>13</sup> swarm-based methods have also been adapted for novelty-driven exploration;<sup>14</sup> and MAP-Elites<sup>15</sup> has become a widely used framework for illuminating trade-offs between diversity and performance. While powerful, these methods are typically sample-inefficient and therefore poorly matched to settings in which each function evaluation is costly.

In this work, we draw inspiration from BO to develop a sample-efficient Bayesian novelty search strategy for expensive black-box systems. Rather than placing a surrogate model on a scalar objective, we model the full input-to-outcome relationship using Gaussian processes (GPs).<sup>16</sup> This distinction is important since the surrogate is used to reason directly about the outcomes a user wishes to explore, not to collapse them into a single utility function. As a result, the outcome space can be defined flexibly by the practitioner and may include multiple quantities of interest. In a drug discovery setting, for example, the outcome vector may encode efficacy, solubility, partitioning behavior, synthesizability, or stability, while in a materials setting it may encode multiple uptake or transport properties.

Our main contributions are summarized as follows:

- We introduce a new novelty search algorithm, referred to as BEACON (Bayesian Exploration Algorithm for outCOme Novelty), for noisy, expensive black-box systems. BEACON is designed to improve behavior-space reachability using a limited number of evaluations.
- We propose a Thompson sampling-based acquisition function for NS that scores candidate inputs by the distance between their sampled outcomes and a denoised archive of previously observed outcomes. This gives a continuous NS policy that accounts for predictive uncertainty and observation noise without requiring direct optimization over a fixed behavior partition.
- We extend BEACON to high-dimensional settings through both sparsity-inducing fully Bayesian priors and chemistry-aware surrogate modeling strategies.
- We demonstrate BEACON on a broad set of benchmark and real-world discovery problems, including metal-organic framework discovery for clean energy applications and molecular property discovery with input representations as large as 2133 dimensions.

Taken together, these results show that BEACON can substantially improve the efficiency of novelty-driven exploration in settings that are directly relevant to modern discovery workflows.

The rest of the paper is organized as follows. In Section 2, we define the black-box NS problem considered in this work. In Section 3, we present the required background on BO and multi-output GP regression and introduce key mathematical notation. Section 4 introduces the proposed BEACON algorithm, including the acquisition function, implementation details, and extensions for high-dimensional problems and user-guided behavior constraints. Section 5 presents benchmark comparisons and real-world case studies that illustrate the practical value of BEACON for sample-efficient discovery. Lastly, we provide some concluding remarks in Section 6

## 2 Problem formulation

We consider a vector-valued black-box function

$$f: X \rightarrow O, \quad f = (f^{(1)}, \dots, f^{(n)}),$$

that maps an input space  $X \subset \mathbb{R}^d$  to an  $n$ -dimensional outcome space  $O \subset \mathbb{R}^n$ . In this work, the outcomes are the behaviors we wish to explore. A behavior may be a scalar property, a vector of measured properties, a terminal state, a generated image, or another application-specific representation chosen by the practitioner. The key assumption is that this representation supports a meaningful notion of behavioral similarity, either through a distance metric, a finite partition, or both.

Although the examples in this paper mostly define behaviors using measured or predicted outcomes, the behavior representation can also include known input-derived quantities when they are scientifically relevant. For example, if two chemistries produce similar measured outcomes but differ in cost, synthesis route, reaction conditions, or material availability, these quantities can be included in the behavior representation (or in the distance metric used by the novelty score defined in Section 4). Known quantities do not require GP modeling; they can be appended directly to the sampled outcome representation when evaluating novelty.

We use the term “discovery campaign” to refer to the full sequence of evaluations made under the budget  $T$ . The campaign-level goal is to choose this sequence so that the observed outcomes cover as much of the attainable behavior space as possible, rather than optimizing a single scalar objective. To measure this coverage, we introduce a finite behavior space  $B = \{1, \dots, M\}$  together with behavior regions  $\{C_b\}_{b \in B}$  that partition the relevant outcome space. The behavior assignment map  $\phi: O \rightarrow B$  converts each continuous outcome to its behavior region index, so that  $\phi(y) = b$  if  $y \in C_b$ .

At iteration  $t = 1, \dots, T$ , we query an input  $x_t \in X$  and observe

$$y_t = f(x_t) + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma^2 I_n),$$

where  $\eta_t$  denotes independent Gaussian observation noise. Let

$$I_t = \{\phi(f(x_i)) : i = 1, \dots, t\} \subseteq B$$

be the set of distinct behavior regions reached after  $t$  evaluations. We write  $|I_t|$  for the cardinality of this set, *i.e.*, the number of distinct behavior-region indices reached by iteration  $t$ . We use reachability as one simple measure of campaign progress,



$$\text{Reach}_t = \frac{|I_t|}{|B|}, \quad (1)$$

where  $|B| = M$  is the total number of behavior regions. Reachability is a bin-occupancy measure of coverage. It increases only when the campaign discovers a behavior region that has not been observed before. This makes it useful for novelty-driven discovery because it measures breadth of observed outcomes without requiring a trusted scalar objective or utility function. Other coverage metrics could also be used, depending on the application. For a loss-style convention, one could instead define the “behavior gap” as  $1 - \text{Reach}_t$ , which tracks the fraction of behavior regions that remain unreached and is analogous in spirit to regret-style measures used in bandit and BO settings. Here, we report reachability so that larger values consistently indicate broader coverage.

The corresponding sequential design problem is to choose a sampling policy that discovers as many distinct behavior regions as possible over the evaluation budget. One way to express this goal is through the following problem

$$\max_{\pi} \mathbb{E}_{\pi} \left[ \sum_{t=1}^T \text{Reach}_t \right], \quad (2)$$

where  $\pi$  denotes the sampling policy and the expectation is over any noise and policy randomness. This objective in eqn (2) defines how we evaluate a discovery campaign, but it does not by itself specify the acquisition function used to choose the next input.

This distinction motivates the design of BEACON. A direct one-step reachability acquisition can be derived by maximizing the posterior probability that  $f(x)$  lands in a behavior region not yet reached. Section 1 of the SI derives this acquisition and provides a one-dimensional illustration comparing it with the BEACON novelty acquisition. This direct acquisition is natural when the behavior map  $\phi$  is known in advance and the behavior space is small enough to enumerate. In many early-stage discovery settings, however, the practitioner may know how to represent and compare outcomes before knowing the right partition of outcome space into behaviors. Direct reachability acquisitions also require probability mass calculations over unreached behavior regions, which can become expensive or numerically delicate as the number of regions grows.

BEACON therefore uses NS in the archive-distance sense. Candidate outcomes are preferred when they are far from outcomes already observed. The acquisition operates directly in the continuous outcome space through a user-defined distance metric, while the finite behavior space  $B$  is used to evaluate coverage after samples have been collected. This evaluation space can be constructed using a uniform grid, an  $\varepsilon$ -cover, clustering, or user-defined behavior regions. The choice sets the resolution of the reachability metric. A smaller  $\varepsilon$  (or a finer partition more generally) distinguishes more outcomes and makes reachability stricter, while a coarser partition groups nearby outcomes together. Under the default BEACON acquisition, this choice affects the reported reachability and interpretation of the final discovered set, but not the sampling

decisions themselves. We study the sensitivity of reported reachability to this behavior-space resolution in the SI.

Because  $f$  is expensive to evaluate, exhaustive exploration is generally impractical. We thus use a probabilistic surrogate for the input-to-outcome map. Specifically, BEACON places a multi-output Gaussian process (MOGP) prior on  $f$ . Given observations

$$D_t = \{(x_i, y_i)\}_{i=1}^t,$$

the MOGP induces a posterior distribution  $P(f | D_t)$  over the unknown vector-valued function. This posterior enables BEACON to denoise previously observed outcomes, quantify uncertainty in unobserved regions of input space, and guide sampling toward plausible new regions of outcome space. The framework permits correlated or independent output models; the experiments below use the modeling choices described in Section 3 and the SI.

### 3 Methodological background

This section briefly reviews the two core ingredients underlying BEACON: multi-output Gaussian process (MOGP) surrogate modeling and Bayesian optimization (BO).

#### 3.1 Multi-output Gaussian process surrogate modeling

Gaussian processes (GPs) provide a flexible nonparametric framework for modeling expensive black-box functions.<sup>16</sup> A GP prior is fully specified by a mean function  $\mu$  and a covariance (kernel) function  $\kappa$ . In BEACON, the surrogate is used to model the input-to-outcome map  $f: X \rightarrow O$ , where  $f(x) = (f^{(1)}(x), \dots, f^{(n)}(x))$  may contain one or more outcome coordinates.

The BEACON framework does not require a particular MOGP construction. In most of the experiments reported here, we use independent-output GP models, meaning that each outcome coordinate is modeled with its own scalar GP. This choice keeps the surrogate modeling component simple and makes the empirical results focus on the proposed NS policy rather than on advances in multi-output modeling. Correlated-output models could be substituted when cross-output structure is important, but this is not the focus of the present study.

For conciseness, it is useful to write the vector-valued surrogate using an output-indexed representation. Define

$$h(x, j) = f^{(j)}(x), \quad j \in J := \{1, \dots, n\},$$

and place a scalar GP prior  $h \sim \text{GP}(\mu, \kappa)$  over the extended input space  $X \times J$ ; see, e.g., Kudva *et al.*<sup>17</sup> A general kernel  $\kappa((x, j), (x', j'))$  can encode correlations across outputs. In our experiments, however, we use the block-diagonal kernel

$$\kappa((x, j), (x', j')) = \delta_{jj'} \kappa_j(x, x') \quad (3)$$

where  $\delta_{jj'}$  is the Kronecker delta and  $\kappa_j$  is the scalar kernel for output  $j$ . Thus, eqn (3) is equivalent to fitting one scalar GP per output coordinate.

Suppose we observe the indexed dataset  $A = \{(x_i, j_i, y_i)\}_{i=1}^N$ , where  $y_i$  is a noisy measurement of  $h(x_i, j_i)$ . Under Gaussian



observation noise, the posterior remains a GP,  $h|A \sim \mathcal{G}\mathcal{P}(\mu_A, \kappa_A)$ , with mean and covariance functions<sup>16,18</sup>

$$\mu_A(x, j) = \mu(x, j) + k_A(x, j)^\top (K_A + \sigma^2 I_N)^{-1} (y - \mu_A^{\text{obs}}), \quad (4a)$$

$$\kappa_A((x, j), (x', j')) = \kappa((x, j), (x', j')) - k_A(x, j)^\top (K_A + \sigma^2 I_N)^{-1} k_A(x', j'). \quad (4b)$$

Here,  $y = [y_1, \dots, y_N]^\top$ ,  $\mu_A^{\text{obs}} = [\mu(x_1, j_1), \dots, \mu(x_N, j_N)]^\top$ ,  $k_A(x, j) \in \mathbb{R}^N$  is the vector of covariance values between the test pair  $(x, j)$  and the observed indexed inputs in  $A$ , and  $K_A \in \mathbb{R}^{N \times N}$  is the covariance matrix over the observed indexed inputs.

With the independent-output kernel in eqn (3), these posterior equations decouple across outputs. The posterior mean vector and covariance matrix for  $f(x)$  are then assembled from the scalar GP posteriors for each coordinate. This is the implementation used throughout the numerical studies (except for the MNIST case). In the chemistry and molecular examples, task-specific kernels are used for the relevant scalar GP components; these choices modify  $\kappa_j$  but do not change the BEACON acquisition principle.

It is worth noting that, if output correlations are important, the independent-output GP can be replaced by a correlated multi-output kernel or by a structured high-dimensional output model. For example, in the MNIST case study we use a high-dimensional, multi-task GP model<sup>19</sup> to model the full set of 14  $\times$  14 pixels in the decoded image.

### 3.2 Bayesian optimization

Bayesian optimization (BO) is a sequential strategy for optimizing an expensive black-box function. For exposition, consider the minimization problem

$$x^* = \underset{x \in X}{\operatorname{argmin}} f(x),$$

where  $X$  is the admissible search space and evaluations of  $f$  may be noisy and costly. Classical BO alternates between two steps: fitting a probabilistic surrogate model to the available data and optimizing an acquisition function that quantifies the value of sampling a candidate point next.

Common acquisition functions include upper confidence bound (UCB),<sup>20</sup> expected improvement (EI),<sup>21</sup> and Thompson sampling (TS).<sup>22</sup> TS is a randomized decision rule that draws a sample function  $g$  from the surrogate posterior and then selects the optimizer of that sample. In the minimization setting, this corresponds to choosing  $\arg \min_{x \in X} g(x)$ . TS often performs well in practice and is especially attractive in our setting because it naturally balances exploration and exploitation while remaining easy to adapt beyond classical objective-based optimization.<sup>23</sup>

BEACON borrows this posterior sampling idea from BO but applies it to NS rather than scalar objective optimization. Instead of drawing a posterior sample to identify a likely optimizer of a scalar objective, BEACON draws a posterior sample of the input-to-outcome map and scores candidate inputs by how novel their sampled outcomes are relative to the denoised archive of previously observed outcomes. This adaptation is formally introduced in the next section.

## 4 BEACON method

We now present the proposed BEACON algorithm. We first introduce a Thompson sampling-based acquisition function for novelty search (NS), then summarize the overall algorithm, discuss its efficient optimization, and finally describe extensions for high-dimensional problems and user-guided behavior constraints.

### 4.1 A Thompson sampling acquisition function for novelty search

Our goal is to discover novel system behaviors by exploring unobserved regions of the continuous outcome space  $O$ . A common strategy in the NS literature<sup>9</sup> is to quantify how far a candidate outcome lies from previously observed outcomes. One standard novelty score is the average distance to the  $k$  nearest neighbors in outcome space:

$$\rho(x|D) = \frac{1}{k} \sum_{i=1}^k \operatorname{dist}(f(x), y_i^*), \quad (5)$$

where  $\{y_1^*, \dots, y_k^*\} \subset \{y_1, \dots, y_N\}$  are the  $k$  nearest previously observed outcomes to  $f(x)$  under the user-defined distance metric  $\operatorname{dist}(\cdot)$  over  $O$ .

Although intuitive, directly using eqn (5) is problematic in our setting for two reasons. First,  $f$  is unknown and expensive to query, so  $\rho(x|D)$  cannot be evaluated without performing the experiment or simulation at  $x$ . Second, when observations are noisy, novelty scores based directly on raw outcomes can be distorted by measurement noise. In particular, it may occur that

$$\phi(f(x) + \eta) \neq \phi(f(x)),$$

leading to spurious behavior assignments.

To address these issues, we replace the true black-box map with its MOGP posterior surrogate  $f|D \sim \text{MOGP}(\mu_D, \kappa_D)$ . The surrogate serves two purposes. First, it denoises previously observed outcomes by replacing raw observations with posterior mean predictions. Second, it allows us to account for uncertainty in unobserved regions of the input space through posterior sampling. Using Thompson sampling, we draw a sample function  $g(x) \sim f|D$  and define the BEACON acquisition function as

$$\alpha_{\text{NS}}(x|g, D) = \frac{1}{k} \sum_{i=1}^k \operatorname{dist}(g(x), \mu_D(x_i^*)), \quad (6)$$

where  $\{x_1^*, \dots, x_k^*\} \subset \{x_1, \dots, x_N\}$  are the inputs corresponding to the  $k$  nearest previously sampled outcomes to  $g(x)$ , measured relative to the denoised set  $\{\mu_D(x_1), \dots, \mu_D(x_N)\}$ . This acquisition promotes candidates whose plausible outcomes are far from what has already been observed, while naturally accounting for posterior uncertainty.

The acquisition in eqn (6) does not directly optimize the reachability metric in eqn (1). Instead, it is a continuous archive-distance novelty score used to choose inputs whose plausible outcomes are far from what has already been observed. This design is intentional. A direct reachability



acquisition would require a fixed behavior map and would sum posterior probability mass over behavior regions not yet reached, as derived in the SI (Section 1). Such an acquisition is useful when the behavior partition is known and computationally manageable, but it can be less convenient when the practitioner can define meaningful outcome distances before knowing the right behavior partition. BEACON therefore uses the distance-based novelty score as a general-purpose policy for improving reachability while avoiding direct dependence on the evaluation discretization.

To generate  $g$  efficiently, we use the decoupled posterior sampling approach of Wilson *et al.*,<sup>24</sup> which yields accurate and differentiable approximate GP samples. This is particularly useful because it allows the resulting acquisition function to be optimized with gradient-based methods. In other words, for each BEACON decision step, all random quantities defining the TS are drawn once and then held fixed while the acquisition function is optimized. Thus, the optimizer evaluates a deterministic sampled acquisition surface within that decision step.

We use Thompson sampling because it provides a simple pathwise mechanism for incorporating posterior uncertainty into the novelty score. Once a posterior sample of the input-to-outcome map is drawn, the same archive-distance novelty calculation can be applied regardless of the chosen outcome representation or distance metric. Alternative acquisitions, such as a UCB-style novelty score, are possible but would require additional modeling choices, tuning parameters, or numerical approximations. We do believe, however, that deriving and studying additional acquisition functions for NS is an interesting future work.

## 4.2 Overview of the BEACON algorithm

Algorithm 1 summarizes the proposed BEACON procedure. At each iteration, BEACON fits a MOGP surrogate to the available data, draws a posterior sample, and selects the next query point by maximizing the novelty-based acquisition function in eqn (6). The new observation is then added to the dataset, and the process repeats until the evaluation budget is exhausted. Fig. 1 provides a schematic illustration.

### Algorithm 1 BEACON

- 1: **Input:** Outcome function  $f$ , domain  $X$ , initial dataset  $D_0$ , total budget  $T$ , MOGP prior  $MOGP(\mu, \kappa)$ , neighbor count  $k$ , and outcome distance metric  $\text{dist}(\cdot)$
- 2: **for**  $t = 1$  to  $T$  **do**
- 3:   Fit MOGP posterior:  $MOGP_{t-1} \leftarrow f | D_{t-1}$
- 4:   Sample  $g \sim MOGP_{t-1}$
- 5:    $x_t \leftarrow \arg\max_{x \in X} \alpha_{\text{NS}}(x | g, D_{t-1})$
- 6:   Observe outcome  $y_t = f(x_t) + \eta_t$
- 7:   Update  $D_t \leftarrow D_{t-1} \cup \{(x_t, y_t)\}$
- 8: **end for**
- 9: **Return:** Final dataset  $D_T$

For clarity, Algorithm 1 is written in a standard sequential setting. However, because the method is built around posterior sampling, parallel and asynchronous variants can be developed using ideas from prior work on parallel (distributed) BO.<sup>23</sup>

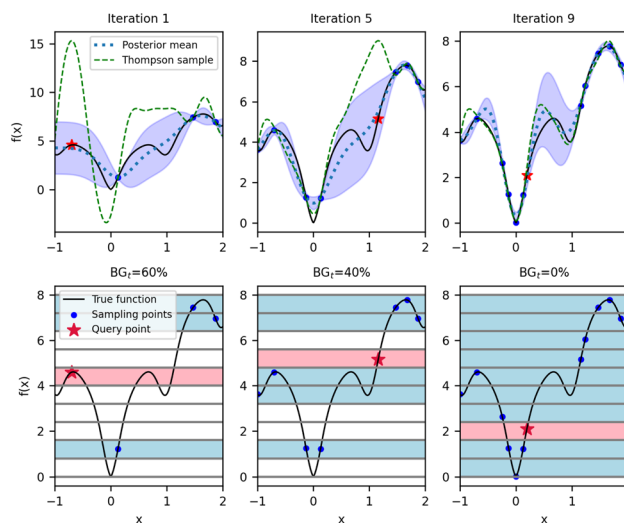


Fig. 1 Visual illustration of BEACON on a one-dimensional test problem. Top: true function (black), GP posterior mean (blue dashed), and a Thompson sample (green dashed) at iterations 1, 5, and 9. Shaded regions indicate posterior uncertainty, and the red star marks the selected query point. Bottom: discretized behavior bins over the outcome space. Blue bands denote previously observed behaviors, while the red band denotes the predicted behavior of the next query. As BEACON progresses, it efficiently increases the reachability metric  $\text{Reach}_t$  by identifying previously unexplored regions. Note that  $BG_t = 1 - \text{Reach}_t$  denotes the behavior gap, which is the complement of reachability.

## 4.3 Gradient-based optimization of acquisition function

Efficiently maximizing  $\alpha_{\text{NS}}$  is essential to the practical performance of BEACON. For continuous input domains, we optimize the acquisition using gradient-based methods such as L-BFGS-B<sup>25</sup> with multiple random restarts.

Although the novelty score depends on a  $k$ -nearest-neighbor operation, it can be written in a form amenable to differentiation:

$$\alpha_{\text{NS}}(x|g, D) = \frac{1}{k} e_k^\top \text{sort}_\uparrow(\{\text{dist}(g(x), \mu_D(x_i))\}_{i=1}^N), \quad (7)$$

where  $e_k$  is a vector whose first  $k$  entries are equal to 1 and whose remaining entries are 0, and  $\text{sort}_\uparrow(\cdot)$  denotes sorting in ascending order. This ascending sort is necessary because novelty is defined using the  $k$  nearest previously observed outcomes. As discussed in Prillo *et al.*,<sup>26</sup> sorting operators admit continuous relaxations that are differentiable almost everywhere, which makes gradient-based optimization practical in this setting.

The cost of evaluating eqn (7) scales linearly with  $N$ , the number of archived outcomes used in the nearest-neighbor calculation at the current decision step. In the experiments considered here,  $N$  remains modest because the evaluation budgets are small and each evaluation is assumed to be expensive. For larger campaigns, the archive can be compressed by retaining representative outcomes rather than every previously sampled outcome, or by using standard nearest-neighbor data structures. For example, if several observations fall in the same behavior region or are very close in outcome space, one



can retain a representative point for that local region of the archive. The effectiveness of this reduction is problem dependent. With fine behavior partitions or higher-dimensional outcome spaces, new evaluations may continue to occupy distinct regions, in which case the representative archive can grow nearly as fast as the total number of samples. Thus, archive management becomes more important as evaluation budgets increase, while remaining a secondary cost in the expensive-evaluation settings emphasized in this work. We do show that BEACON is able to effectively scale to several thousand data points in Section 5 of the SI, especially when taking advantage of efficient large-scale GP implementations.

#### 4.4 Scaling BEACON to high-dimensional problems

A major advantage of BEACON is that it can leverage modern advances in BO surrogate modeling. This is particularly important when the input dimension is large, since GP surrogates generally require additional structure to remain statistically and computationally effective. We consider two complementary strategies.

**4.4.1 SAAS priors.** The sparse axis-aligned subspace (SAAS) prior<sup>27</sup> is a fully Bayesian approach that induces adaptive sparsity in GP models by placing a hierarchical prior over the input lengthscales. This encourages explanations that depend on only a small subset of the input variables, which is beneficial when outcome behavior is driven by a limited number of relevant features. Because of its generality, SAAS provides a useful default option for high-dimensional continuous design spaces. We use this approach in the high-dimensional log  $D$  case study presented later.

**4.4.2 Chemistry-aware GP surrogates.** When the design variables are discrete, structured, or highly correlated (as is common for molecular representations), Euclidean distance-based kernels may be inadequate. In such cases, BEACON can be paired with chemistry-aware GP surrogates built from molecular kernels or structured representations. In our ESOL case study, for example, we use a Tanimoto-fragprint kernel motivated by the GAUCHE framework.<sup>28</sup> More broadly, BEACON is compatible with GP models defined over strings, fingerprints, graphs, and related molecular objects, provided posterior sampling is available; Section 2 of the SI gives the exact surrogate choices used in each case study.

#### 4.5 User-guided BEACON with behavior constraints

In many applications, users have prior knowledge or preferences about which behaviors are worth exploring. For example, an experimentalist may wish to avoid regions that have already been sufficiently characterized or restrict attention to outcome regimes of greatest scientific interest. To incorporate such information, we introduce UG-BEACON, a user-guided variant of BEACON that imposes constraints directly in behavior space.

Let  $\mathcal{G} = \{G_1, \dots, G_M\}$  denote a user-defined partition of the outcome space into non-overlapping regions. This partition can be tailored to the application and need not match the discretization used to compute reachability. Let  $\mathcal{G}_{\text{visited}} \subseteq \mathcal{G}$  denote the subset of regions that are either already explored or

intentionally excluded by the user. UG-BEACON defines the feasible set

$$X_{\text{feas}} = \left\{ x \in X : g(x) \notin \bigcup_{G \in \mathcal{G}_{\text{visited}}} G \right\},$$

where  $g$  is the current Thompson sample (TS) from the posterior model. The next evaluation is then chosen by solving

$$x^* = \underset{x}{\operatorname{argmax}} \alpha_{\text{NS}}(x|g, D) \quad \text{subject to} \quad x \in X_{\text{feas}} \cap X.$$

This modification allows BEACON to avoid redundant sampling and direct exploration toward user-relevant parts of the outcome space, while leaving the underlying surrogate model and posterior-sampling machinery unchanged. UG-BEACON can be viewed as replacing full-range coverage with a targeted notion of coverage over behavior regions that satisfy user-defined preferences. Practical examples and implementation details for constrained behavior discovery are provided in Section 3 of the SI.

## 5 Results and discussion

We compare BEACON against several benchmark methods across synthetic, materials, molecular, reinforcement learning, and user-guided discovery tasks. Across these studies, the goal is behavior-space coverage under a limited evaluation budget. We therefore interpret the results as coverage experiments rather than standard scalar-objective optimization benchmarks.

Unless otherwise stated, all methods begin with 20 points sampled uniformly at random from  $X$  to initialize the models. The baselines include evolutionary novelty search (NS-EA),<sup>10</sup> novelty search based on distance to explored area (NS-DEA),<sup>29</sup> input-space novelty search (NS-FS), maximum posterior variance sampling (MaxVar), Sobol sequence sampling, and random search (RS), depending on which methods are applicable to the problem setting. Full implementation details are provided in Section 2 of the SI. Each method then selects an additional 80–300 evaluations, depending on the problem. All experiments are repeated 20 times unless otherwise noted. Performance is measured using reachability, the fraction of predefined behavior regions discovered after  $t$  evaluations. In all plots, we report the mean reachability together with  $\pm$  one standard deviation across independent replicates. For problems defined over discrete candidate sets, the evolutionary NS baselines are not directly applicable, so comparisons are restricted to methods that can operate in that setting.

Additional problem-specific details and ablation studies on the impact of the behavior space resolution, neighborhood size  $k$ , and observation noise can be found in the SI (Section 4). To demonstrate the flexibility of the user-guided extension, we also consider two constrained discovery tasks using UG-BEACON: an MNIST digit discovery problem and an oil sorbent material design problem.

### 5.1 Synthetic benchmark problems

We first consider scalar-output coverage landscapes adapted from the Ackley, Rosenbrock, and Styblinski–Tang functions.



These functions are commonly used as global optimization benchmarks, but we do not use them here in that context. Instead, the scalar function value is treated as a one-dimensional outcome, and the goal is to cover the range of attainable outcome values. The outcome range is partitioned into 25 equally spaced intervals to define behavior regions. We study each function in three input dimensions,  $d \in \{4, 8, 12\}$ ; complete function definitions/setup details are provided in Section 3 of the SI.

Fig. 2 summarizes the reachability results across these nine problems. BEACON consistently attains the highest reachability and often approaches complete coverage of the discretized behavior space within the allowed budget. The performance gap relative to the baselines generally widens as the input dimension increases, which is consistent with the expected difficulty of exploring high-dimensional spaces using purely input-space sampling or evolutionary heuristics.

To further test the multi-output setting, we also construct a synthetic Multi-Output Plus problem with a long-tailed two-dimensional outcome distribution. This function has six inputs and two outputs (full definition is in Section 3.1 of the

SI). Fig. 3 (top) shows the distribution of outcomes induced by 10 000 uniformly sampled inputs from  $X = [-5, 5]^6$ . Reachability is computed using a  $10 \times 10$  grid over the two-dimensional outcome space. This example provides a controlled vector-valued coverage problem in which the outcome distribution can be visualized directly. It is useful because the attainable outcomes are highly non-uniform, *i.e.*, a large fraction of the probability mass is concentrated near the center, while many behavior regions lie in low-density tails. This structure makes the problem challenging for NS because search procedures can easily oversample dense regions while missing the tails. As shown in Fig. 3 (bottom), BEACON substantially improves coverage of the long-tail outcome regions, reaching nearly 80% reachability, whereas competing methods remain below roughly 40%.

## 5.2 Materials discovery applications

We next consider discovery tasks involving metal–organic frameworks (MOFs), a chemically diverse class of porous crystalline materials relevant to gas storage and separations. Because MOFs span a vast combinatorial design space over

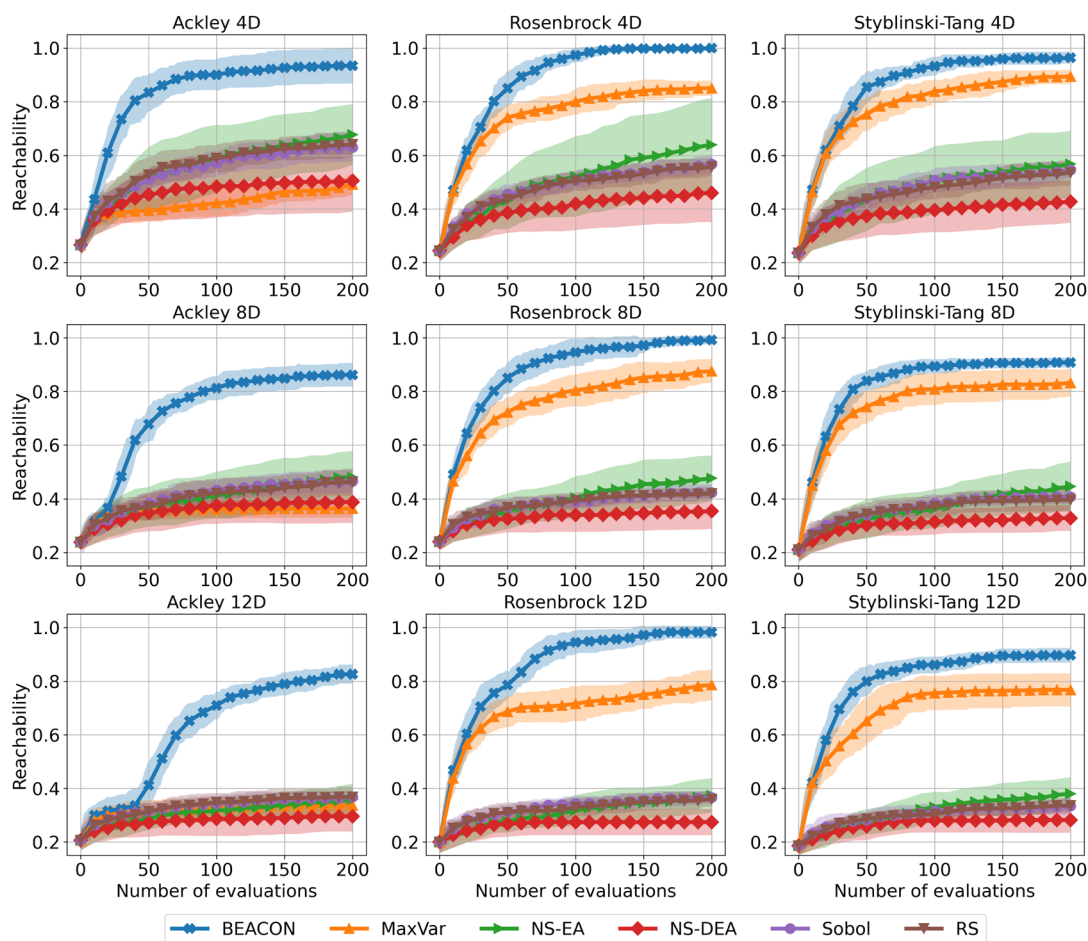


Fig. 2 Reachability results on the synthetic benchmark problems: Ackley (left column), Rosenbrock (middle column), and Styblinski–Tang (right column), with input dimensions 4 (top row), 8 (middle row), and 12 (bottom row). Curves show the mean reachability and shaded bands indicate  $\pm$  one standard deviation across 20 replicates. BEACON consistently achieves the strongest coverage of the outcome space, with particularly large gains in higher dimensions.



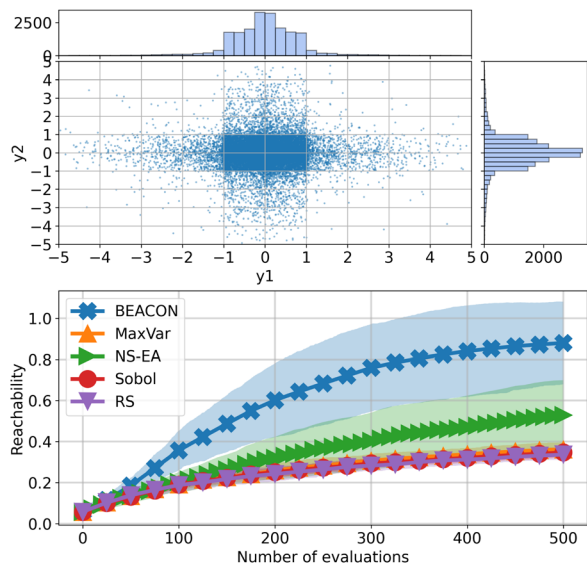


Fig. 3 Results for the multi-output plus test problem. Top: outcome distribution obtained from 10 000 uniformly sampled inputs, illustrating the strongly nonuniform and long-tailed structure of the two-dimensional behavior space. Bottom: reachability *versus* iteration for all methods. BEACON more effectively uncovers low-density outcome regions than the competing baselines.

linkers, metal centers, topologies, and defects, it remains difficult to characterize the full diversity of properties that can arise in practice.<sup>30</sup> Existing MOF databases are also known to exhibit biases that can affect downstream screening and learning studies.<sup>31</sup> These features make MOFs a natural setting for novelty-driven search, especially when evaluations are computationally or experimentally expensive.

In each task, a candidate MOF is represented by a fixed-dimensional descriptor vector  $x \in X \subset \mathbb{R}^d$ , and the property or properties of interest define the outcome  $y = f(x) \in O$ . Descriptor definitions and outcome distributions are provided in Section 3 of the SI.

**5.2.1 Hydrogen uptake capacity.** We first consider a hydrogen storage dataset from ref. 32 containing 98 000 unique MOFs. Each MOF is represented by 7 real-valued descriptors, and the outcome is a scalar hydrogen uptake value.

**5.2.2 Nitrogen uptake capacity.** The second task considers nitrogen uptake in MOFs for gas separation, using the dataset from ref. 33. This dataset contains 5224 MOFs, each represented by 20 descriptors, with scalar nitrogen uptake as the outcome.

**5.2.3 Joint CO<sub>2</sub> and CH<sub>4</sub> uptake.** The third task considers simultaneous exploration of carbon dioxide and methane uptake capacities using the dataset from Moosavi *et al.*<sup>31</sup> with 7000 MOFs. Here, the input consists of 25 descriptors and the outcome is two-dimensional. The corresponding outcome distribution is strongly skewed; see Section 3 of the SI.

Fig. 4 compares BEACON with the applicable baselines on these three tasks. Because the candidate MOFs come from discrete datasets, the evolutionary NS baselines are not included here. Across all three problems, BEACON achieves the

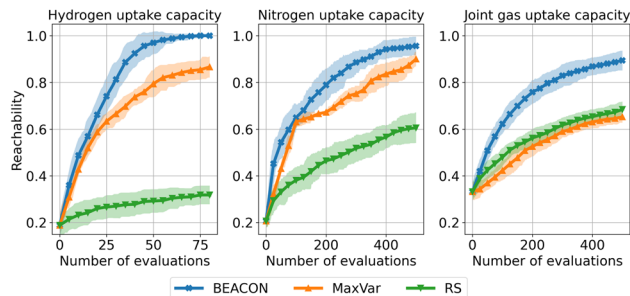


Fig. 4 Reachability results for the MOF discovery tasks: hydrogen uptake (left), nitrogen uptake (middle), and joint CO<sub>2</sub>/CH<sub>4</sub> uptake (right). Because these problems are defined over discrete candidate sets, only methods applicable in that setting are shown. BEACON achieves the highest reachability in all three cases and reaches full coverage within the allowed budget.

strongest reachability and, within the available sampling budget, attains full coverage of the discretized behavior space. These results highlight the value of sample-efficient NS for property-space exploration in chemically large but evaluation-limited settings.

### 5.3 Molecular discovery applications

We next study molecular property discovery problems relevant to pharmaceutical and chemical design. In these settings, the objective is not necessarily to identify a single molecule with the best property value, but rather to discover molecules spanning a broad range of property behaviors. Such diversity can support model development, dataset enrichment, and early-stage screening. Across all tasks, the input space  $X \subset \mathbb{R}^d$  encodes molecular structure using descriptor-based or high-dimensional molecular representations, while the outcome is a scalar molecular property. These examples thus test BEACON in high-dimensional molecular input spaces with scalar behavior-space coverage metrics.

**5.3.1 Water solubility.** We use the dataset from Boobier *et al.*<sup>34</sup> containing 900 small organic molecules, each represented by a 14-dimensional descriptor vector.

**5.3.2 ESOL.** We next consider the ESOL dataset,<sup>35</sup> which contains 1128 molecules with experimentally measured aqueous solubility. Each molecule is represented by a 2133-dimensional binary fragprint vector. For this task, BEACON uses a Tanimoto-fragprint kernel motivated by chemistry-aware GP modeling.<sup>28</sup>

**5.3.3 Log *D*.** Finally, we consider a log *D* dataset containing 2070 molecules from Win *et al.*,<sup>36</sup> each represented by 125 features. To address the higher dimensionality, we pair BEACON with the sparse axis-aligned subspace (SAAS) prior.<sup>27</sup>

In these experiments, reachability is computed over the full observed property range to provide a neutral comparison of broad property-space coverage. For molecular campaigns where the scientific goal is to identify many candidates within a high-performing region, the behavior space can instead be restricted to a preferred property window, weighted toward high-value regions, or discretized more finely in those regions. The user-



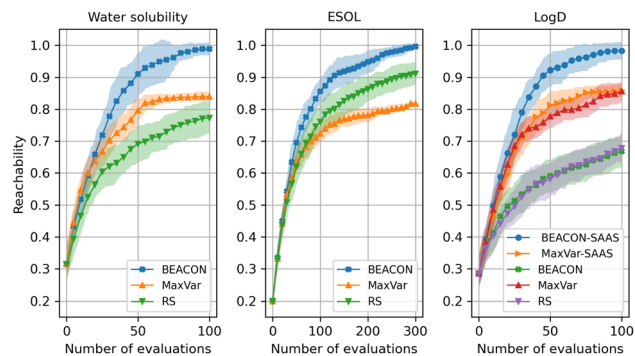


Fig. 5 Reachability results for the molecular discovery tasks: water solubility (left), ESOL (middle), and  $\log D$  (right). BEACON consistently outperforms the benchmark methods, with especially strong gains in the higher-dimensional ESOL and  $\log D$  settings.

guided formulations in Section 5.5 illustrate how BEACON can be adapted to this targeted-coverage setting.

The results in Fig. 5 show that BEACON consistently achieves the highest reachability across all three tasks. The gains are especially pronounced in ESOL and  $\log D$ , where the use of task-appropriate surrogate structure is important for effective search. Overall, these results support the view that BEACON can be naturally combined with chemistry-aware GP models to improve sample-efficient exploration of molecular property space.

#### 5.4 A deceptive reinforcement learning landscape

We next consider a reinforcement learning (RL) maze navigation task with a deliberately deceptive reward landscape. The goal is to control a ball from a fixed start location to a goal location within 300 time steps using a linear control policy with 8 tunable parameters. In this problem, improving reward locally does not necessarily correspond to exploring behaviorally useful parts of the maze; a reward-driven search can therefore become trapped by policies that make partial progress but fail to discover trajectories that reach the goal.

For BEACON and the NS baselines, the behavior outcome is the two-dimensional terminal position of the agent after executing the policy, and novelty is computed in this terminal-state space. Thus, these methods do not directly optimize reward. Instead, they search for policies that lead to new terminal states, which can help overcome the deceptive reward structure by encouraging broader exploration of the maze. In contrast, expected improvement (EI)<sup>37</sup> is included as a reward-driven scalar-objective comparator; its GP surrogate models normalized reward as a function of the controller parameters, and its acquisition seeks policies with high expected reward improvement. We report normalized reward separately as a task-success metric because it measures progress toward the goal and enables comparison with this reward-driven baseline. A schematic of the maze and additional details are given in Section 3.4 of the SI.

As shown in Fig. 6, BEACON substantially outperforms both the novelty-based baselines and the reward-driven methods,

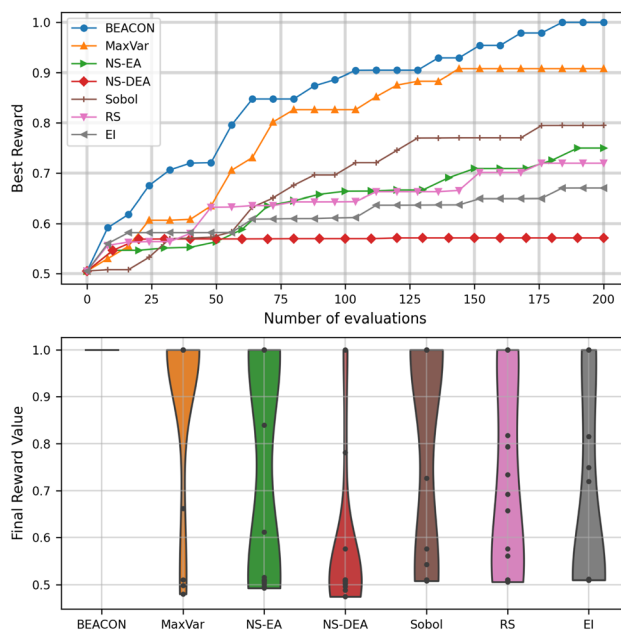


Fig. 6 Results for the RL maze-navigation case study. Top: average best observed reward versus iteration for all methods across 20 replicates. Bottom: violin plot of the best reward obtained at the final iteration. BEACON is the only method that consistently achieves successful maze completion across all replicates.

including EI. The top panel reports the average best observed reward over 20 replicates, while the bottom panel shows the distribution of final best rewards. BEACON is the only method that solves the task across all replicates, demonstrating that sample-efficient NS can be particularly valuable when objective-based search is vulnerable to deception.

#### 5.5 User-guided discovery with behavior constraints

We finally evaluate UG-BEACON on two tasks where user-defined behavior preferences are incorporated directly into the search.

**5.5.1 MNIST digit discovery.** The first example uses the MNIST dataset.<sup>38</sup> The goal is to discover as many distinct handwritten digit classes (0–9) as possible by sampling from the latent space of a trained variational autoencoder (VAE).<sup>39</sup> Each sampled latent point is decoded to a  $14 \times 14$  image, which is treated as the black-box outcome. Thus, BEACON computes novelty in the 196-dimensional pixel-output space (stacked  $14 \times 14$  outputs into a vector) rather than directly in digit-label space. A convolutional neural network (CNN) classifier is then used to assign each generated image to a digit class. UG-BEACON uses these class labels to avoid revisiting already discovered digits during the search, and reachability is computed over the ten digit classes. As shown in Fig. 7, UG-BEACON discovers the full set of digit classes more rapidly than both BEACON and the baseline methods.

**5.5.2 Oil sorbent materials.** As a second example, we consider the discovery of oil sorbent materials under application-driven outcome preferences. Here, the two-dimensional outcome space is partitioned non-uniformly



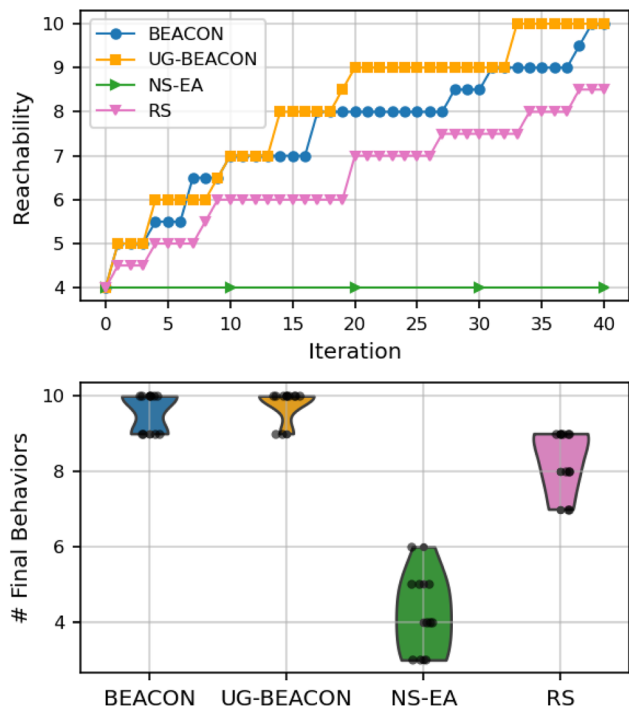


Fig. 7 Results for the MNIST digit discovery case study. Top: average number of unique digit classes discovered *versus* iteration across 10 replicates. Bottom: distribution of the number of discovered classes at the final iteration. UG-BEACON reaches full class coverage more rapidly by incorporating behavior-level constraints based on previously observed digits.

according to utility, *i.e.*, regions corresponding to high adsorption capacity and high strength are discretized more finely, while less relevant regions are coarsened. This allows the user to place greater emphasis on discovery in scientifically/practically valuable regimes. UG-BEACON uses this partition both to focus the search and to avoid redundant sampling of already explored regions. As shown in Fig. 8, UG-BEACON quickly reaches full coverage of the preferred outcome space, outperforming BEACON and the competing baselines.

Taken together, these examples show that UG-BEACON can incorporate user guidance in a simple and effective way, making the framework useful not only for unconstrained NS, but also for targeted discovery tasks in which some outcome regions are more important than others.

### 5.6 Practical considerations for sharp outcome spaces

A practical limitation of BEACON, and more broadly of GP-guided sequential design methods, is that discovery depends on the surrogate model being able to represent the relevant input-outcome structure. Thompson sampling encourages exploration of uncertain regions, but it cannot guarantee discovery of an arbitrarily small or highly localized behavior region if the initial data provide no signal that such a region exists. This issue is relevant in materials discovery problems with, *e.g.*, phase changes or other significant/sharp changes in mechanism.

In the most difficult case, the response may appear nearly flat over most of the input space but contain abrupt, weakly

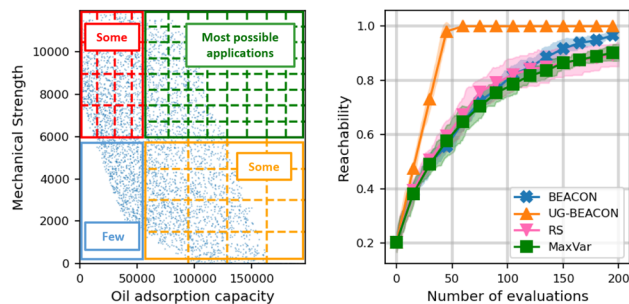


Fig. 8 Oil sorbent material discovery case study. Left: user-defined partition of the two-dimensional outcome space, with finer discretization in the most desirable performance regions. Right: reachability *versus* iteration for UG-BEACON and the benchmark methods. UG-BEACON attains full preferred-space coverage substantially faster than the competing approaches.

structured jumps. A standard stationary GP can easily over-smooth such behavior, and posterior samples from a misspecified model may not assign enough probability to the rare region to guide sampling there. More flexible surrogates can help when the sharp structure is learnable from the available data. Examples include non-stationary kernels or input warping for spatially varying length scales,<sup>40</sup> adaptive hyperparameter treatments that reduce the risk of over-smoothing,<sup>41</sup> deep kernel learning models that learn representations jointly with the GP,<sup>42</sup> and methods designed to account for GP model misspecification.<sup>43</sup> These approaches improve the modeling side of the problem, but they cannot overcome the absence of learnable structure. If a rare behavior is effectively uncorrelated with nearby sampled points and no available side information indicates where it may occur, then no GP-guided acquisition can reliably target it from a small number of evaluations; discovery is possible only through sufficiently broad exploration or chance sampling.

For prospective applications where missing such regions would be costly, BEACON can be combined with more deliberate exploration safeguards. These include broader space-filling initialization, maintaining an input-space diversity component, or adding scheduled exploration when reachability begins to plateau. Such modifications are directly compatible with the framework because they change the surrogate model or the candidate-generation strategy while preserving the core idea of scoring novelty in outcome space. Another potential strategy is to iteratively shrink the search region around the most novel previously discovered outcomes, thereby progressively increasing the resolution of the surrogate model in the sharp regions likely to contain highly novel outcomes.<sup>44</sup> Developing these hybrid strategies is an important direction for future work, particularly for discovery problems where broad coverage must be balanced against reliable identification of rare high-performing regions.

## 6 Conclusions

In this work, we presented BEACON, a Bayesian novelty search algorithm for discovering diverse behaviors of expensive, noisy



black-box systems. By combining multi-output Gaussian process surrogates with a Thompson sampling-based novelty acquisition function, BEACON directs evaluations toward previously unexplored regions of outcome space while still accounting for model uncertainty and observation noise. This means that BEACON is particularly well suited to low-data discovery settings, where evaluations are costly and a broad coverage of outcomes is more valuable than optimizing a single predefined target.

Across synthetic benchmarks and real-world case studies in materials and molecular discovery, BEACON consistently achieved stronger reachability than competing methods under limited evaluation budgets. These results suggest that Bayesian surrogate modeling can make novelty-driven exploration practical in domains where traditional evolutionary NS approaches are too sample-inefficient to be useful.

There are several important directions for future work. From a methods standpoint, reducing computational overhead and developing stronger theoretical guarantees remain worthwhile challenges. From an applications standpoint, a particularly promising next step is to integrate BEACON-like strategies into closed-loop autonomous experimentation platforms, where the goal is not only to optimize a known target but also to uncover previously unseen phenomena or property regimes.<sup>12</sup> We view this as an especially compelling opportunity for truly autonomous discovery in chemistry and materials science.

## Author contributions

Wei-Ting Tang: conceptualization, methodology, software, formal analysis, investigation, visualization, writing – original draft. Ankush Chakrabarty: methodology, software, formal analysis, writing – review and editing. Joel A. Paulson: conceptualization, supervision, funding acquisition, project administration, writing – review and editing.

## Conflicts of interest

There are no conflicts to declare

## Data availability

The code for BEACON is publicly available at <https://github.com/PaulsonLab/BEACON> and archived on Zenodo (<https://doi.org/10.5281/zenodo.20435036>). The repository includes the main scripts used to run BEACON in continuous and discrete settings for both single- and multi-outcome problems, together with the Thompson sampling implementation used throughout the study. It also contains problem-specific code and supporting folders for the synthetic, materials, molecular, and MNIST case studies, as well as plotting utilities. These resources are intended to support reproduction of the empirical results reported in this paper and to provide a starting point for applying BEACON to related novelty search (NS) problems. Datasets for molecular and material problems studied in this work are available at <https://doi.org/10.6084/m9.figshare.32493084>.

Supplementary information (SI): additional methodological and computational details supporting the main text. Section 1 presents a derivation of a direct reachability acquisition function and compares that to BEACON. Section 2 contains all the core implementation details for BEACON and the baseline methods. Section 3 provides additional details on the suite of benchmark and application problems. Section 4 discusses various ablation studies on the BEACON strategy. Finally, Section 5 discusses how to scale BEACON to problems with large evaluation budgets using sparse GP approximations. See DOI: <https://doi.org/10.1039/d6dd00123h>.

## Acknowledgements

This work was supported by the National Science Foundation (NSF) CAREER Award 2621493.

## Notes and references

- 1 J. Grizou, L. J. Points, A. Sharma and L. Cronin, *Sci. Adv.*, 2020, **6**, eaay4237.
- 2 K. Terayama, M. Sumita, R. Tamura, D. T. Payne, M. K. Chahal, S. Ishihara and K. Tsuda, *Chem. Sci.*, 2020, **11**, 5959–5968.
- 3 E. Stach, B. DeCost, A. G. Kusne, J. Hatrick-Simpers, K. A. Brown, K. G. Reyes, J. Schrier, S. Billinge, T. Buonassisi, I. Foster, C. P. Gomes, J. M. Gregoire, A. Mehta, J. Montoya, E. Olivetti, C. Park, E. Rotenberg, S. K. Saikin, S. Smullin, V. Stanev and B. Maruyama, *Matter*, 2021, **4**, 2702–2726.
- 4 K. Wang and A. W. Dowling, *Curr. Opin. Chem. Eng.*, 2022, **36**, 100728.
- 5 L. M. Rios and N. V. Sahinidis, *J. Global Optim.*, 2013, **56**, 1247–1293.
- 6 B. Shahriari, K. Swersky, Z. Wang, R. P. Adams and N. De Freitas, *Proc. IEEE*, 2015, **104**, 148–175.
- 7 S. J. Russell and P. Norvig, *Artificial intelligence: A modern approach*, Pearson, 2016.
- 8 P. I. Frazier, *arXiv*, 2018, preprint arXiv:1807.02811, DOI: [10.48550/arXiv.1807.02811](https://doi.org/10.48550/arXiv.1807.02811).
- 9 J. Lehman and K. O. Stanley, *Evol. Comput.*, 2011, **19**, 189–223.
- 10 J. Lehman and K. O. Stanley, *Genetic Programming Theory and Practice IX*, 2011, 37–56.
- 11 E. C. Jackson and M. Daley, *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 2019, pp. 173–174.
- 12 R. Bulanadi, J. Chowdhury, H. Funakubo, M. Ziatdinov, R. Vasudevan, A. Biswas and Y. Liu, *ACS Nanosci. Au*, 2026, **6**(1), 86–94.
- 13 J. Gomes, P. Mariano and A. L. Christensen, *Knowl. Base Syst.*, 2015, **82**, 1–20.
- 14 J.-B. Mouret and S. Doncieux, *IEEE Trans. Evol. Comput.*, 2012, **16**, 1–15.
- 15 J.-B. Mouret and J. Clune, *Artif. Life*, 2015, **22**, 245–268.
- 16 C. K. Williams and C. E. Rasmussen, *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, MA, 2006, vol. 2.



- 17 A. Kudva, W.-T. Tang and J. A. Paulson, *Comput. Chem. Eng.*, 2024, **181**, 108515.
- 18 H. Liu, J. Cai and Y.-S. Ong, *Knowl. Base Syst.*, 2018, **144**, 102–121.
- 19 W. J. Maddox, M. Balandat, A. G. Wilson and E. Bakshy, *Adv. Neural Inf. Process. Syst.*, 2021, **34**, 19274–19287.
- 20 N. Srinivas, A. Krause, S. M. Kakade and M. Seeger, *arXiv*, 2009, preprint, arXiv: 0912.3995, DOI: [10.48550/arXiv.0912.3995](https://doi.org/10.48550/arXiv.0912.3995).
- 21 S. Ament, S. Daulton, D. Eriksson, M. Balandat and E. Bakshy, *Adv. Neural Inf. Process. Syst.*, 2023, **36**, 20577–20612.
- 22 W. R. Thompson, *Biometrika*, 1933, **25**, 285–294.
- 23 K. Kandasamy, A. Krishnamurthy, J. Schneider and B. Póczos, *International Conference on Artificial Intelligence and Statistics*, 2018, pp. 133–142.
- 24 J. Wilson, V. Borovitskiy, A. Terenin, P. Mostowsky and M. Deisenroth, *International Conference on Machine Learning*, 2020, pp. 10292–10302.
- 25 R. H. Byrd, P. Lu, J. Nocedal and C. Zhu, *SIAM J. Sci. Comput.*, 1995, **16**, 1190–1208.
- 26 S. Prillo and J. Eisenschlos, *International Conference on Machine Learning*, 2020, pp. 7793–7802.
- 27 D. Eriksson and M. Jankowiak, *Uncertainty in Artificial Intelligence*, 2021, pp. 493–503.
- 28 R.-R. Griffiths, L. Klarner, H. Moss, A. Ravuri, S. Truong, Y. Du, S. Stanton, G. Tom, B. Rankovic, A. Jamasb, A. Deshwal, J. Schwartz, A. Tripp, G. Kell, S. Frieder, A. Bourached, A. Chan, J. Moss, C. Guo, J. P. Dürholt, S. Chaurasia, J. W. Park, F. Strieth-Kalthoff, A. Lee, B. Cheng, A. Aspuru-Guzik, P. Schwaller and J. Tang, *Adv. Neural Inf. Process. Syst.*, 2023, 76923–76946.
- 29 S. Doncieux, G. Paolo, A. Laflaquière and A. Coninx, *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, 2020, pp. 85–93.
- 30 S. Lee, B. Kim, H. Cho, H. Lee, S. Y. Lee, E. S. Cho and J. Kim, *ACS Appl. Mater. Interfaces*, 2021, **13**, 23647–23654.
- 31 S. M. Moosavi, A. Nandy, K. M. Jablonka, D. Ongari, J. P. Janet, P. G. Boyd, Y. Lee, B. Smit and H. J. Kulik, *Nat. Commun.*, 2020, **11**, 1–10.
- 32 S. Ghude and C. Chowdhury, *Chem. – Eur. J.*, 2023, **29**, e202301840.
- 33 H. Daglar and S. Keskin, *ACS Appl. Mater. Interfaces*, 2022, **14**, 32134–32148.
- 34 S. Boobier, D. R. Hose, A. J. Blacker and B. N. Nguyen, *Nat. Commun.*, 2020, **11**, 5753.
- 35 J. S. Delaney, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 1000–1005.
- 36 Z.-M. Win, A. M. Cheong and W. S. Hopkins, *J. Chem. Inf. Model.*, 2023, **63**, 1906–1913.
- 37 D. R. Jones, M. Schonlau and W. J. Welch, *J. Global Optim.*, 1998, **13**, 455–492.
- 38 L. Deng, *IEEE Signal Process. Mag.*, 2012, **29**, 141–142.
- 39 C. Doersch, *arXiv*, 2016, preprint, arXiv:1606.05908, DOI: [10.48550/arXiv.1606.05908](https://doi.org/10.48550/arXiv.1606.05908).
- 40 J. Snoek, K. Swersky, R. Zemel and R. Adams, *International Conference on Machine Learning*, 2014, pp. 1674–1682.
- 41 F. Berkenkamp, A. P. Schoellig and A. Krause, *J. Mach. Learn. Res.*, 2019, **20**, 1–24.
- 42 A. G. Wilson, Z. Hu, R. R. Salakhutdinov and E. P. Xing, *Adv. Neural Inf. Process. Syst.*, 2016, **29**, 2594–2602.
- 43 I. Bogunovic and A. Krause, *Adv. Neural Inf. Process. Syst.*, 2021, **34**, 3004–3015.
- 44 A. E. Siemenn, Z. Ren, Q. Li and T. Buonassisi, *npj Comput. Mater.*, 2023, **9**, 79.

