Multi-Fidelity Active Learning with GFlowNets

Anonymous authors
Paper under double-blind review

Abstract

In the last decades, the capacity to generate large amounts of data in science and engineering applications has been growing steadily. Meanwhile, the progress in machine learning has turned it into a suitable tool to process and utilise the available data. Nonetheless, many relevant scientific and engineering problems present challenges where current machine learning methods cannot yet efficiently leverage the available data and resources. For example, in scientific discovery, we are often faced with the problem of exploring very large, structured and high-dimensional spaces, and where querying a high fidelity, black-box objective function is very expensive. Progress in machine learning methods that can efficiently tackle such problems would help accelerate currently crucial areas such as drug and materials discovery. In this paper, we propose a multi-fidelity active learning algorithm with GFlowNets as a sampler, to efficiently discover diverse, high-scoring candidates where multiple approximations of the black-box function are available at lower fidelity and cost. Our evaluation on molecular discovery tasks show that multi-fidelity active learning with GFlowNets can discover high-scoring candidates at a fraction of the budget of its single-fidelity counterpart while maintaining diversity, unlike RL-based alternatives. These results open new avenues for multi-fidelity active learning to accelerate scientific discovery and engineering design.

1 Introduction

To tackle the most pressing challenges for humanity, such as the climate crisis and the threat of pandemics or antibiotic resistance, there is a growing need for new scientific discoveries. By way of illustration, materials discovery can play an important role in improving the efficiency of energy production and storage; and reducing the costs and duration of drug discovery cycles has the potential to effectively and rapidly mitigate the consequences of new diseases. In recent years, researchers in materials science, biochemistry and other fields have increasingly adopted machine learning as a tool since it holds the promise to drastically accelerate scientific discovery (Butler et al., 2018; Zitnick et al., 2020; Bashir et al., 2021; Das et al., 2021).

Although machine learning has already made a positive impact in scientific discovery applications (Stokes et al., 2020; Jumper et al., 2021), unleashing its full potential will require improving the current algorithms Agrawal & Choudhary (2016). For example, typical tasks in potentially impactful applications in materials and drug discovery require exploring combinatorially large, structured and high-dimensional spaces (Bohacek et al., 1996; Polishchuk et al., 2013), where only small, noisy data sets are available. Furthermore, obtaining new annotations computationally or experimentally is often very expensive. Such scenarios present serious challenges even for the most advanced current machine learning methods.

In the search for a useful discovery, we typically define a quantitative proxy for usefulness, which we can view as a black-box function. One promising avenue for improvement is developing methods that more efficiently leverage the availability of multiple approximations of the target black-box function at lower fidelity but much lower cost than the highest fidelity oracle (Chen et al., 2021; Fare et al., 2022). For example, a standard tool to characterise the properties of materials and molecules is quantum mechanics simulations such as Density Functional Theory (DFT) (Parr, 1980; Sholl & Steckel, 2022). However, DFT is computationally too expensive for high-throughput exploration of large search spaces. Thus, large-scale exploration can only be achieved through cheaper but less accurate oracles. Nonetheless, solely relying on low-fidelity approximations is clearly suboptimal. Ideally, such "needle-in-a-haystack" problems would be

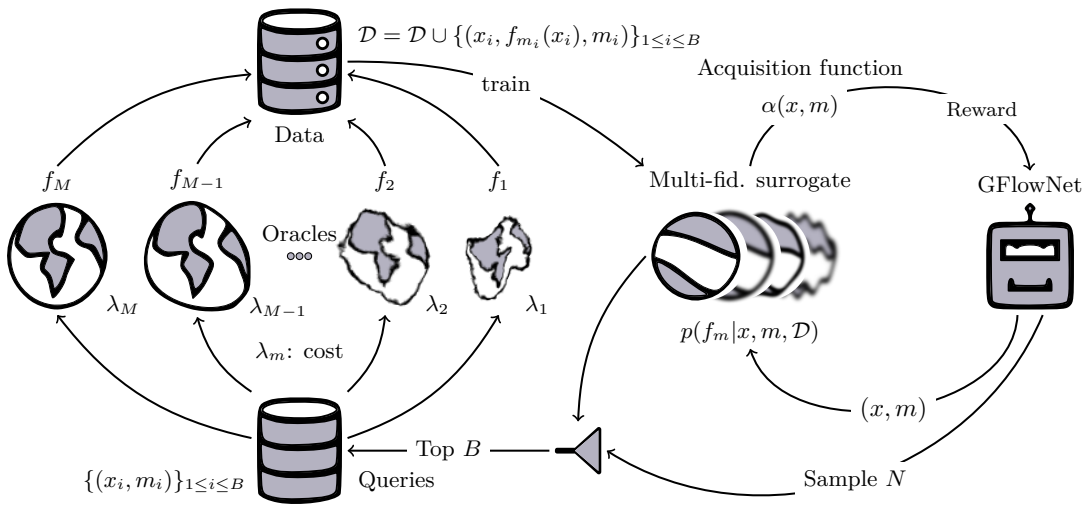


Figure 1: Illustration of multi-fidelity active learning with GFlowNets (Algorithm 1). Given a set of M oracles $f_1, \ldots, f_M$ (center left) with varying fidelities and costs $\lambda < \ldots < \lambda_M$, respectively, we can construct a data set $\mathcal{D}$ (top left) with annotations from the oracles. With this data, we fit a multi-fidelity surrogate (center), modelling the posterior $p(f_m(x)|x, m, \mathcal{D})$. Using the surrogate, we compute a multi-fidelity acquisition function—max-value entropy search in our experiments— which is used as the reward to train a GFlowNet (right). GFlowNet samples both an object x and the fidelity m proportionally to this reward. Once GFlowNet is trained, we sample N tuples (x, m) and select the top B according to the acquisition function (bottom left). Finally, we annotate each new candidate with the selected oracle, add them to the data set and repeat the process.

best tackled by methods that can efficiently and adaptively distribute the available computational budget between the multiple oracles depending on the already acquired information.

Another challenge is that even the highest fidelity oracles are often underspecified with respect to the actual, relevant, downstream applications. This underspecification problem can be mitigated by finding multiple candidate solutions (Jain et al., 2023a). However, most current machine learning methods used in scientific discovery problems, such as Bayesian optimisation (BO, Song et al., 2018; Garnett, 2023) and reinforcement learning (RL, Angermueller et al., 2020), are designed for global optimisation of the target function. Therefore, it is imperative to develop methods that not only find the global optimum, but also discover sets of diverse, high-scoring candidates.

Recently, generative flow networks (GFlowNets, Bengio et al., 2021a) have demonstrated their ability to find diverse candidates through discrete probabilistic modelling, with particularly promising results when used in an active learning loop (Jain et al., 2022). In this paper, we propose a multi-fidelity active learning algorithm enhanced with these capabilities of GFlowNets. Our contributions can be summarized as follows:

- We introduce a multi-fidelity active learning algorithm designed for combinatorially large, structured and high-dimensional spaces.
- We propose an extension of GFlowNets for this multi-fidelity setting, to sample both candidates and oracle indices, proportionally to a given acquisition function.
- We conduct a comprehensive empirical evaluation across four scientific discovery tasks and demonstrate that multi-fidelity active learning with GFlowNets:
  - Discovers high-scoring samples with reduced computational costs compared to its single-fidelity counterpart.

- Identifies multiple modes of the target function, unlike methods relying on reinforcement learning or standard Bayesian optimization, thereby facilitating diverse sampling.

2 Related Work

Our work can be framed within the broad field of active learning (AL), a class of machine learning methods whose goal is to learn an efficient data sampling scheme to accelerate training (Settles, 2009). For the bulk of the literature in AL, the goal is to train an accurate model $h(x)$ of an unknown target function $f(x)$, as in classical supervised learning. However, in certain scientific discovery problems, which motivate our work, a desirable goal is often instead to discover multiple, diverse candidates x with high values of $f(x)$, as discussed in Section 1.

Our work is also closely connected to Bayesian optimisation (BO, Garnett, 2023; Snoek et al., 2015), which aims at optimising a black-box objective function $f(x)$ that is expensive to evaluate. In contrast to the problems we address in this paper, standard BO typically considers continuous domains and works best in relatively low-dimensional spaces (Frazier, 2018). Nonetheless, in recent years, approaches for BO with structured data (Deshwal & Doppa, 2021; Papenmeier et al., 2023) and high-dimensional domains (Grosnit et al., 2021) have been proposed in the literature. The main difference between BO and the problem we tackle in this paper is that we are interested in finding multiple, diverse samples with high value of f and not only the optimum. Recent work by Maus et al. (2022) has proposed a variant of traditional BO to find diverse solutions.

This goal, as well as the discrete nature of the search space, is shared with active search (Garnett et al., 2012), a variant of active learning in which the task is to efficiently find multiple samples of a valuable (binary) class from a discrete domain $\mathcal{X}$. This objective was already considered in the early 2000s by Warmuth et al. (2001) for drug discovery , and more formally analysed in later work (Jiang et al., 2017; 2019). Another recent research area in stochastic optimisation that considers diversity is so-called Quality-Diversity (Chatzilygeroudis et al., 2021), which typically uses evolutionary algorithms that perform search in a latent space. These and other problems such as multi-armed bandits (Robbins, 1952) and the general framework of experimental design (Chaloner & Verdinelli, 1995) all share the objective of optimising or exploring an expensive black-box function. Formal connections between some of these areas have been established in the literature (Srinivas et al., 2010; Foster, 2021; Jain et al., 2023a; Fiore et al., 2023).

Multi-fidelity methods have been proposed in most of these areas of research. An early survey on multi-fidelity methods for Bayesian optimisation was compiled by Peherstorfer et al. (2018), and research on the subject has continued since with the proposal of specific acquisition functions (Takeno et al., 2020) and the use of deep neural networks to improve the modelling (Li et al., 2020). Recently, works on multi-fidelity active search have also appeared in the literature (Nguyen et al., 2021), but interestingly, the literature on multi-fidelity active learning (Li et al., 2022a) is scarcer. Finally, while multi-fidelity methods have started to be applied in scientific discovery problems (Chen et al., 2021; Fare et al., 2022) the literature is still scarce probably because most approaches cannot tackle the specifics of scientific discovery, such as the need for diverse samples. Here, we aim at addressing this need with the use of GFlowNets (Bengio et al., 2021a; Jain et al., 2023b) for multi-fidelity active learning.

3 Method

In this section, we first briefly introduce the necessary background on GFlowNets and active learning. Then, we describe the proposed algorithm for multi-fidelity active learning with GFlowNets.

3.1 Background

GFlowNets Generative Flow Networks (GFlowNets; Bengio et al., 2021a;b) are amortised samplers designed for sampling from discrete high-dimensional distributions. Given a space of compositional objects $\mathcal{X}$ and a non-negative reward function $R(x)$, GFlowNets are designed to learn a stochastic policy π that generates $x \in \mathcal{X}$ with a probability proportional to the reward, that is $\pi(x) \propto R(x)$. This distinctive property

induces sampling of diverse, high-reward objects, which is a desirable property for scientific discovery, among other applications (Jain et al., 2023a).

The objects $x \in \mathcal{X}$ are constructed sequentially by sampling transitions $s_t \rightarrow s_{t+1} \in \mathbb{A}$ between partially constructed objects (states) $s \in \mathcal{S}$, which includes a unique empty state s_0. The stochastic forward policy is typically parameterised by a neural network $P_F(s_{t+1}|s_t;\theta)$, where θ denotes the learnable parameters, which models the distribution over transitions $s_t \rightarrow s_{t+1}$ from the current state s_t to the next state s_{t+1}. The backward transitions are parameterised too and denoted $P_B(s_t|s_{t+1};\theta)$. Objects x are generated by the sequential application of P_F, forming trajectories $\tau = (s_0 \rightarrow s_1 \ldots \rightarrow x)$. To learn the parameters θ such that $\pi(x) \propto R(x)$ we use the trajectory balance learning objective (Malkin et al., 2022)

$$\mathcal{L}_{TB}(\tau;\theta) = \left(\log \frac{Z_\theta \prod_{t=0}^{n} P_F(s_{t+1}|s_t;\theta)}{R(x)\prod_{t=1}^{n} P_B(s_t|s_{t+1};\theta)}\right)^2, \tag{1}$$

where Z_θ is an approximation of the partition function $\sum_{x\in\mathcal{X}} R(x)$ that is learnt. The GFlowNet learning objective supports training from off-policy trajectories, so during training the trajectories are typically sampled from a mixture of the current policy with a uniform random policy. The reward is also tempered to make the policy focus on the modes.

Active Learning In its simplest formulation, the active learning problem that we consider is as follows: we start with an initial data set $\mathcal{D} = \{(x_i, f(x_i))\}$ of samples $x \in \mathcal{X}$ and their evaluations by an expensive, black-box objective function (oracle) $f : \mathcal{X} \rightarrow \mathbb{R}$, which we use to train a surrogate model $h(x)$. A GFlowNet can then be trained to learn a generative policy $\pi_\theta(x)$ using $h(x)$ as reward function, that is $R(x) = h(x)$. Optionally, we can instead train a probabilistic surrogate $p(f|\mathcal{D})$ and use as reward the output of an acquisition function $\alpha(x, p(f|\mathcal{D}))$ that considers the epistemic uncertainty of the surrogate model, as typically done in Bayesian optimisation. This is the approach by Jain et al. (2022) with GFlowNet-AL. An important difference between traditional BO and active learning with GFlowNets is that the latter samples from the acquisition function instead of optimising it (Jain et al., 2023a). Finally, we use the policy $\pi(x)$ to generate a batch of samples to be evaluated by the oracle f, we add them to our data set and repeat the process a number of active learning rounds.

While much of the active learning literature (Settles, 2009) has focused on so-called *pool-based* active learning, where the learner selects samples from a pool of unlabelled data, we here consider the scenario of *de novo query synthesis*, where samples are selected from the entire object space $\mathcal{X}$. This scenario is particularly suited for scientific discovery (King et al., 2004; Xue et al., 2016; Yuan et al., 2018; Kusne et al., 2020). The ultimate goal pursued in active learning applications is also heterogeneous. Often, the goal is the same as in classical supervised machine learning: to train an accurate (surrogate) model $h(x)$ of the unknown target function $f(x)$. For some problems in scientific discovery, we are usually not interested in the accuracy across the entire input space $\mathcal{X}$, but rather in discovering new, diverse objects with high values of f. We have reviewed the literature that is connected to our work in Section 2.

3.2 Multi-Fidelity Active Learning

We now consider the following active learning problem with multiple oracles of different fidelities. Our ultimate goal is to generate a batch of K samples $x \in \mathcal{X}$ according to the following desiderata:

- The samples obtain a high value when evaluated by the objective function $f : \mathcal{X} \rightarrow \mathbb{R}^+$.
- The samples should be diverse, covering distinct high-valued regions of f.

Furthermore, we are constrained by a computational budget Λ that limits our capacity to evaluate f. While f is extremely expensive to evaluate, we have access to a discrete set of approximate functions (oracles) $\{f_m\}_{1\leq m\leq M} : \mathcal{X} \rightarrow \mathbb{R}^+$, where m represents the fidelity index and each oracle has an associated cost λ_m—we assume, without loss of generality, that the larger m, the higher the fidelity and that $\lambda_1 < \lambda_2 < \ldots < \lambda_M$. We also assume $f_M = f$ because, even though there may exist more accurate oracles, we do not have access

to them. This scenario resembles many practically relevant problems in scientific discovery and motivates our approach: because the objective function f_M is not a perfect proxy of the true usefulness of objects x, we seek diversity; and because f_M may be expensive to evaluate, we make use of approximate models.

In multi-fidelity active learning—as well as in multi-fidelity Bayesian optimisation—the iterative sampling scheme consists of not only selecting the next object x (or batch of objects) to evaluate, but also the level of fidelity m, such that the procedure is cost-effective.

Briefly, our algorithm, MF-GFN follows these iterative steps: we use the currently available data to train a probabilistic multi-fidelity surrogate model $h(x, m)$. We can use the surrogate to compute the value of annotating a candidate x with the oracle f_m via an acquisition function $\alpha(x, m)$. Next, we train a GFlowNet with the acquisition function as a reward. Once trained we sample N tuples (x, m) and select the top B. Finally, we annotate each candidate x with the selected oracle m and start over. Figure 1 contains a visual illustration of MF-GFN and more detailed descriptions are provided in Algorithm 1 and in Appendices A and B. Below, we discuss further the surrogate model and the acquisition function, and in Section 3.3 we introduce the multi-fidelity GFlowNet.

Surrogate Model Given a dataset $\mathcal{D}$, a candidate x and an oracle index m, we want to model the posterior distribution over the output of the oracle, $p(f_m(x)|x, m, \mathcal{D})$. A natural modelling choice is Gaussian Processes (GP) (Rasmussen & Williams, 2005), commonly used in Bayesian optimisation. However, in order to better model structured, high-dimensional data, we use deep kernel learning (Wilson et al., 2016): First, a non-linear embedding of the inputs x is learnt by a deep neural network. Then, a multi-fidelity GP kernel is applied by combining a Matern kernel applied to the latent representations from the network along with a linear downsampling kernel over the fidelity index m (Wu et al., 2019).

Acqusition Function Multi-fidelity methods proposed in the Bayesian optimization literature have adapted information theory-based acquisition functions (Li et al., 2022a; Wu et al., 2023; Li et al., 2022b). In our work, we use the multi-fidelity version (Takeno et al., 2020) of max-value entropy search (MES) Wang & Jegelka (2017). MES captures the mutual information between the value of candidate x and the maximum value attained by the objective function, $f^\star$. The multi-fidelity variant is designed to select the candidate x and the fidelity m that maximise the mutual information between $f_M^\star$ and the oracle at fidelity m, f_m, weighted by the cost of the oracle (see Appendix B for the details). MES has been shown to be more efficient than plain entropy search (ES) (Wang & Jegelka, 2017) and it even makes random sampling a strong baseline. Additionally, for efficiency, we adopt the GIBBON approximation of MF-MES, which has demonstrated good performance in the context of multi-fidelity optimization (Moss et al., 2021).

3.3 Multi-Fidelity GFlowNets

A multi-fidelity acquisition function can be regarded as a cost-adjusted utility function. Therefore, in order to carry out a cost-aware search, we seek to sample diverse objects with high value of the acquisition function. To this purpose, we propose to use a GFlowNet as a generative model by training it to sample the fidelity m in addition to the candidate x itself. Formally, given a GFlowNet with state and transition spaces $\mathcal{S}$ and $\mathbb{A}$, we augment the state space with a new dimension for the fidelity $\mathcal{M}' = \{0, 1, 2, \ldots, M\}$ (including $m = 0$, which corresponds to unset fidelity), such that the augmented, multi-fidelity space is $\mathcal{S}_{\mathcal{M}'} = \mathcal{S} \cup \mathcal{M}'$. The set of allowed transitions $\mathbb{A}_M$ is augmented such that a fidelity $m > 0$ of a trajectory must be selected once, and only once, from any intermediate state.

Intuitively, allowing the selection of the fidelity at any step in the trajectory should give flexibility for better generalisation. At the end, complete trajectories are the concatenation of an object x and the fidelity m, that is $(x, m) \in \mathcal{X}_{\mathcal{M}} = \mathcal{X} \cup \mathcal{M}$. In summary, the proposed approach learns a policy that samples jointly objects in a possibly very large, structured and high-dimensional space, together with the level of fidelity, that maximise a given multi-fidelity acquisition function as reward.

Algorithm 1: MF-GFN: Multi-fidelity active learning with GFlowNets. A graphical summary of this algorithm is shown in Fig. 1.

Input: $\{(f_m, \lambda_m)\}$: M oracles and their corresponding costs;
$\mathcal{D}_0 = \{(x_i, f_m(x_i), m_i)\}$: Initial data set;
$h(x, m)$: Multi-fidelity Gaussian Process surrogate model;
$\alpha(x, m)$: Multi-fidelity acquisition function;
$R(\alpha(x, m), \beta)$: reward function to train the GFlowNet;
B: Batch size of oracles queries;
Λ: Maximum available budget;
K: Number of top-scoring candidates to be evaluated at termination;
Result: Top-$K(\mathcal{D})$, Diversity
Initialization: $\Lambda_j = 0$, $\mathcal{D} = \mathcal{D}_0$
while $\Lambda_j < \Lambda$ **do**

- Fit h on data set $\mathcal{D}$;
- Train GFlowNet with reward $R(\alpha(x, m), \beta)$ to obtain policy $\pi_\theta(x)$;
- Sample $N \gg B$ tuples $(x_i, m_i) \sim \pi_\theta$;
- Score each tuple using $\alpha(x, m)$ and select the top B tuples with the highest scores;
- Evaluate each tuple with the corresponding oracle to form batch $\mathcal{B} = \{(x_1, f_m(x_1), m_1), \ldots, (x_B, f_m(x_B), m_B)\}$;
- Update data set $\mathcal{D} = \mathcal{D} \cup \mathcal{B}$ and budget $\Lambda_j = \Lambda_j + \sum_{i=1}^{i=B} \lambda_{m_i}$;

end

4 Empirical Evaluation

In this section, we present empirical evaluations of multi-fidelity active learning with GFlowNets. Through our experiments, we aim to answer the following questions:

- Can our multi-fidelity active learning approach find high-scoring, diverse samples at lower cost than with a single-fidelity oracle?
- Does MF-GFN, which samples objects and fidelities (x, m), provide any advantage over sampling only x and selecting m randomly?

4.1 Metrics

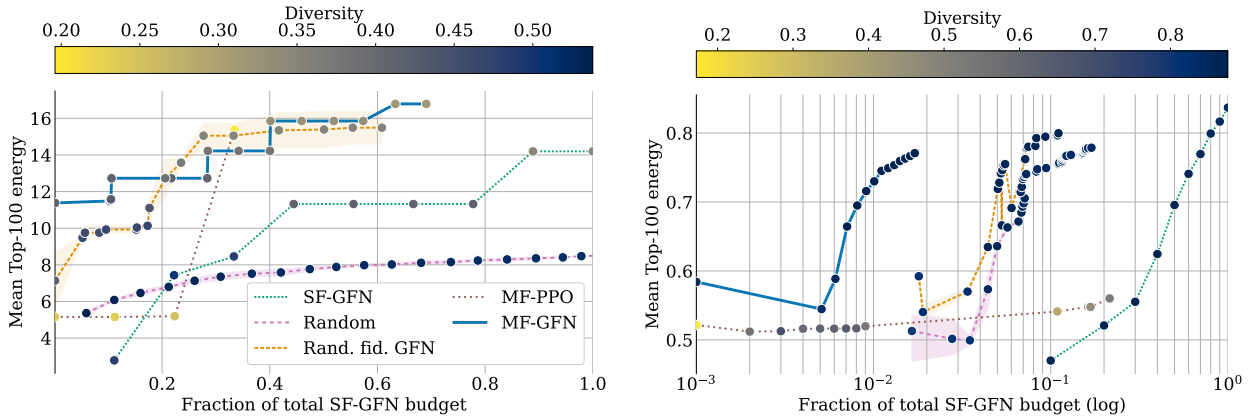
As discussed in Section 3.2, our goal is to sample diverse objects with high scores according to a reward function. Following Gao et al. (2022) and Jain et al. (2022), we here consider a pair of metrics that capture both the score and the diversity of the final batch of candidates:

- **Mean top-K score**: mean score, per the highest fidelity oracle f_M, of the top-K samples.
- **Top-K diversity**: mean pairwise distance within the top-K samples.

For additional details, see Appendix D. Since here we are interested in the cost effectiveness of the active learning process, we evaluate the above metrics as a function of the cost accumulated in querying the oracles. It is important to note that multi-fidelity approaches are not aimed at achieving *better* mean top-K scores than a single-fidelity active learning counterpart, but rather *the same* mean top-K scores but *with a smaller budget*.

4.2 Baselines

In order to evaluate our approach, and to shed light on the questions stated above, we consider the following baselines:



(a) DNA aptamers task

(b) Anti-microbial peptides (AMP) task

Figure 2: Results on the DNA aptamers and AMP tasks. The curves indicate the mean energy f_M within the top-100 samples computed at the end of each active learning round and plotted as a function of the budget used. The colour of the round markers indicates the diversity within the batch (darker colour indicating more diversity), computed as the average sequence identity distance (see Appendix D). In both the DNA and AMP tasks, MF-GFN outperforms all baselines in terms of cost efficiency, while obtaining great diversity in the final batch of top-K candidates.

GFlowNet with highest fidelity (SF-GFN) GFlowNet-based active learning (GFlowNet-AL) as in Jain et al. (2022) with the highest fidelity oracle, to establish a benchmark for performance without considering the cost-accuracy trade-offs.

GFlowNet with random fidelities (Random fid. GFN) Variant of SF-GFN where the candidates are generated with the GFlowNet but the multi-fidelity acquisition function is evaluated with random fidelities, to investigate the benefit of deciding the fidelity with GFlowNets.

Random candidates and fidelities (Random) Quasi-random approach where both candidates and fidelities are randomly sampled. We query the top (x, m) pairs according to the acquisition function.

Multi-fidelity PPO (MF-PPO) Instantiation of multi-fidelity Bayesian optimisation where the acquisition function is optimised using proximal policy optimisation (PPO, Schulman et al., 2017). Unlike with the other baselines, we include an initialisation of $n/3$ steps where n is the maximum number of steps allowed. We do this to help exploration and diversity, since without this PPO tends to collapse to generation of very similar candidates.

4.3 Benchmark Tasks

As a proof of concept, we perform experiments on two low-dimensional synthetic functions: Branin and Hartmann, widely used in the multi-fidelity Bayesian optimisation literature (Perdikaris et al., 2017; Song et al., 2018; Kandasamy et al., 2019; Li et al., 2020; Folch et al., 2023). These tasks show that MF-GFN is able to obtain results comparable to other multi-fidelity BO methods. We provide these results in Appendix C.4. To obtain a solid assessment of the performance of MF-GFN on large, structured and high-dimensional problems, we evaluate it on more complex tasks of practical scientific relevance. We present results on a variety of discovery domains: DNA aptamers (Section 4.3.1), anti-microbial peptides (Section 4.3.2) and small molecules (Section 4.3.3).

4.3.1 DNA Aptamers

DNA aptamers are single-stranded nucleotide sequences of nucleobases A, C, T and G, with multiple applications in polymer design due to their specificity and affinity as sensors in crowded biochemical environments (Zhou et al., 2017; Corey et al., 2022; Yesselman et al., 2019; Kilgour et al., 2021). The objective is to maximize the (negative) free energy of the secondary structure of DNA sequences. This free energy can be seen as a proxy of the stability of the sequences. Diversity is computed as one minus the mean pairwise sequence identity among a set of DNA sequences.

Setting In our experiments, we consider fixed-length sequences of 30 bases and design a GFlowNet environment where the action space $\mathbb{A}$ consists of the choice of base to append to the sequence, starting from an empty sequence. This yields a design space of size $|\mathcal{X}| = 4^{30}$ (ignoring the selection of fidelity in MF-GFN). Further details about the task are discussed in Appendix C.5.1.

Oracles NUPACK (Zadeh et al., 2011), a nucleic acid structure analysis software, is used as the highest fidelity oracle, f_M. As a low fidelity oracle, we trained a transformer model on 1 million randomly sampled sequences annotated with f_M, and assigned it a cost $100\times$ smaller than the highest-fidelity oracle. The cost difference is selected to simulate practical scenarios where wet lab experiments take hours for evaluation, while cheap online simulations take a few minutes.

Results As presented in Fig. 2a, MF-GFN reaches the best mean top-K energy achieved by its single-fidelity counterpart with just about 25 % of the budget. It is also more efficient than GFlowNet with random fidelities and MF-PPO. Crucially, we also see that MF-GFN maintains a high level of diversity (0.32), even after converging to the top-K scores. On the contrary, MF-PPO (0.20) is not able to discover diverse samples, as is expected based on prior work (Jain et al., 2022).

4.3.2 Antimicrobial Peptides

Antimicrobial peptides are short protein sequences which possess antimicrobial properties. As proteins, these are sequences of amino-acids—a vocabulary of 20 along with a special stop token. The aim is to identify sequences with a high antimicrobial activity, as measured by a classifier trained on DBAASP (Pirtskhalava et al., 2021). The diversity calculation mirrors that of DNA.

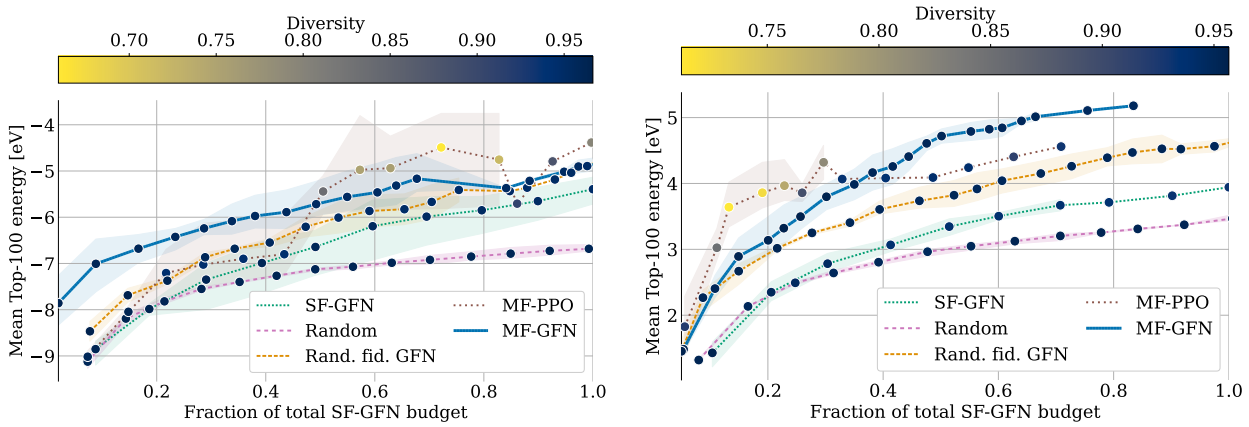
Setting We consider variable-length protein sequences with up to 50 residues. Analogous to DNA, if we ignore the fidelity, this yields a design space of size $|\mathcal{X}| = 20^{50}$.

Oracles We construct a three-oracle setup by training deep learning models with different capacities on exclusive subsets of data points. We simulated a setup wherein the two lower fidelity oracles are trained on specifc subgroups of the peptides. Details in Appendix C.5.2. Similar to the DNA experiment, the lower-fidelity oracles had (equivalent) costs $100\times$ less than the highest fidelity oracle.

Results Fig. 2b indicates that in this task MF-GFN obtains even greater advantage over all other baselines in terms of cost-efficiency. It reaches the same maximum mean top-K score as the random baselines with $10\times$ less budget and almost $100\times$ less budget than SF-GFN. In this task, MF-PPO did not achieve comparable results. Crucially, the diversity of the final batch found by MF-GFN stayed high (0.87), satisfying this important criterion in the motivation of this method.

4.3.3 Small Molecules

Molecules are clouds of interacting electrons (and nuclei) described by a set of quantum mechanical properties. These properties dictate their chemical behaviours and applications. To demonstrate the capability of MF-GFN in the setting of quantum chemistry, we consider two proof-of-concept tasks in molecular electronic potentials: maximisation of the (negative) adiabatic ionisation potential (IP) and of the adiabatic electron affinity (EA). These electronic potentials dictate the molecular redox chemistry, and are crucial in organic semiconductors, photoredox catalysis and organometallic synthesis. In this task, the diversity measure is



(a) Molecules ionisation potential (IP) task

(b) Molecules electron affinity (EA) task

Figure 3: Results on the molecular discovery tasks: (a) ionisation potential (IP), (b) electron affinity (EA). These visualisations are analogous to those in Fig. 2. The diversity of molecules is computed as the average pairwise Tanimoto distance (see Appendix D). Results generally show MF-GFN's faster convergence in discovering diverse molecules with desirable properties.

computed as the average pairwise Tanimoto distance among the top-K scoring molecules (Bajusz et al., 2015).

Setting We designed the GFlowNet state space by considering variable length sequences of SELFIES tokens (Krenn et al., 2020) to represent molecules, with a vocabulary size of 26. The maximum length was 64, resulting in a design space of $|\mathcal{X}| = 26^{64}$.

Oracles Numerous approximations of these quantum mechanical properties have been developed with different methods at different fidelities, with the famous example of Jacob's ladder in density functional theory (Perdew & Schmidt, 2001). We employed three oracles that correlate with experimental results as approximations of the scoring function, by using various levels of geometry optimisation to obtain approximations to the adiabatic geometries. The calculation of IP or EA was carried out with semi-empirical quantum chemistry method XTB (Neugebauer et al., 2020). These three oracles had costs of 1, 3 and 7 (respectively), proportional to their computational running demands. See Appendix C.5.3 for further details.

Results The realistic configuration and practical relevance of these tasks allow us to draw stronger conclusions about the usefulness of multi-fidelity active learning with GFlowNets in scientific discovery applications. As in the other tasks evaluated, we here also found MF-GFN to achieve better cost efficiency at finding high-score top-K molecules (Fig. 3), especially for ionization potentials (Fig. 3a). By clustering the generated molecules, we find that MF-GFN captures as many modes as random generation, far exceeding that of MF-PPO. Indeed, while MF-PPO is able to quickly optimise the target function in the task of electron affinity (Fig. 3b), all generated molecules were from a few clusters (low diversity), which is of much less utility for chemists.

4.4 Ablation studies and additional results

Besides the main experiments presented above, we carried out additional experiments to gain further insights about MF-GFN and study the influence of its various components. We provide detailed results in Appendix E and summarise the main conclusions here:

- Analysing the results in terms of the top-K *diverse* samples confirms that the GFlowNet-based approaches are able to jointly optimize scores and diversity, while RL approaches trade diversity for high scores.

- As is expected, the advantage of MF-GFN over its single-fidelity counterpart decreases as the cost of the lower fidelity oracles increases. Nonetheless, even with a cost ratio of $1:2$ in the DNA task, MF-GFN still outperforms all other methods.
- The same conclusions hold for various values of the final batch size, $K \in \{50, 100, 200\}$.
- We expect MF-GFN to query cheap oracles to prune the input space and costly oracles for high-reward candidates. We validate this through a visualization of a synthetic task.

5 Conclusions, Limitations and Future Work

In this paper, we have presented MF-GFN, a multi-fidelity active learning algorithm that leverages GFlowNets to achieve exploration with diversity for scientific discovery applications. MF-GFN samples candidates as well as the fidelity at which the candidate is to be evaluated, when multiple oracles are available with varying fidelities and costs. We evaluated MF-GFN on benchmark tasks of practical relevance, such as DNA aptamer generation, antimicrobial peptide and small molecule design. Through comparisons with previously proposed methods as well as with variants of our method designed to understand the contributions of different components, we conclude that multi-fidelity active learning with GFlowNets not only outperforms its single-fidelity active learning counterpart in terms of cost effectiveness and diversity of sampled candidates, but it also offers an advantage over other multi-fidelity methods due to its ability to learn a stochastic policy to jointly sample objects and the fidelity of the oracle to be used to evaluate them.

Limitations and Future Work Aside from the molecular modelling tasks, our empirical evaluations in this paper involved simulated oracles with manually selected costs. Future work should evaluate MF-GFN with more practical oracles and costs that reflect their computational or financial demands. Furthermore, a promising avenue that we do not study in this paper is the application of MF-GFN in more complex, structured design spaces, such as hybrid (discrete and continuous) domains (Lahlou et al., 2023), as well as multi-fidelity, multi-objective problems (Jain et al., 2023b).

Statement of Broader Impact

Our work is motivated by pressing challenges to sustainability and public health, and we envision applications of our approach to drug discovery and materials discovery. However, as with all work on these topics, there is a potential risk of dual use of the technology by nefarious actors (Urbina et al., 2022). The authors strongly oppose any uses or derivations of this work intended to cause harm to humans or the environment.

Reproducibility Statement

We have made an effort to include the most relevant details of our proposed algorithm in the main body of the paper. For example, a detailed procedure of the steps of the algorithm is presented in Algorithm 1. Besides this, we have included additional details about the algorithm in Appendices A and B. We have also provided the most relevant information about the experiments in Section 4, for instance including a description of the data representation and the oracles for each of the benchmark tasks. The rest of the details about the experiments are provided in Appendix C for the sake of better clarity, transparency and reproducibility. Finally, our submission includes the original code of our algorithm and experiments and, since it has been developed as open source, the link to the code will be included in the manuscript at the end of the review process.

References

RDKit: Open-source cheminformatics. `https://www.rdkit.org`, 2023. URL `https://zenodo.org/record/8254217`.

Ankit Agrawal and Alok Choudhary. Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science. *APL Materials*, 4(5):053208, 2016.

Christof Angermueller, David Dohan, David Belanger, Ramya Deshpande, Kevin Murphy, and Lucy Colwell. Model-based reinforcement learning for biological sequence design. In *International Conference on Learning Representations*, 2020. URL `https://openreview.net/forum?id=HklxbgBKvr`.

Dávid Bajusz, Anita Rácz, and Károly Héberger. Why is tanimoto index an appropriate choice for fingerprint-based similarity calculations? *Journal of Cheminformatics*, 7, 05 2015. doi: 10.1186/s13321-015-0069-3.

Maximilian Balandat, Brian Karrer, Daniel R. Jiang, Samuel Daulton, Benjamin Letham, Andrew Gordon Wilson, and Eytan Bakshy. BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 33, 2020. URL `http://arxiv.org/abs/1910.06403`.

Christoph Bannwarth, Sebastian Ehlert, and Stefan Grimme. GFN2-xTB—an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *Journal of Chemical Theory and Computation*, 15(3):1652–1671, 2019.

Ali Bashir, Qin Yang, Jinpeng Wang, Stephan Hoyer, Wenchuan Chou, Cory McLean, Geoff Davis, Qiang Gong, Zan Armstrong, Junghoon Jang, et al. Machine learning guided aptamer refinement and discovery. *Nature Communications*, 12(1):2366, 2021.

Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, and Yoshua Bengio. Flow network based generative models for non-iterative diverse candidate generation. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 34, 2021a.

Yoshua Bengio, Salem Lahlou, Tristan Deleu, Edward J. Hu, Mo Tiwari, and Emmanuel Bengio. GFlowNet foundations. *arXiv preprint arXiv:2111.09266*, 2021b.

Regine S Bohacek, Colin McMartin, and Wayne C Guida. The art and practice of structure-based drug design: a molecular modeling perspective. *Medicinal Research Reviews*, 16(1):3–50, 1996.

Keith T Butler, Daniel W Davies, Hugh Cartwright, Olexandr Isayev, and Aron Walsh. Machine learning for molecular and materials science. *Nature*, 559(7715):547–555, 2018.

Kathryn Chaloner and Isabella Verdinelli. Bayesian experimental design: A review. *Statistical Science*, pp. 273–304, 1995.

Konstantinos Chatzilygeroudis, Antoine Cully, Vassilis Vassiliades, and Jean-Baptiste Mouret. Quality-diversity optimization: a novel branch of stochastic optimization. In *Black Box Optimization, Machine Learning, and No-Free Lunch Theorems*, pp. 109–135. Springer, 2021.

Chi Chen, Yunxing Zuo, Weike Ye, Xiangguo Li, and Shyue Ping Ong. Learning properties of ordered and disordered materials from multi-fidelity data. *Nature Computational Science*, 1(1):46–53, 2021.

Laming Chen, Guoxin Zhang, and Hanning Zhou. Fast greedy map inference for determinantal point process to improve recommendation diversity. *arXiv preprint arXiv:1709.05135*, 2018.

David R Corey, Masad J Damha, and Muthiah Manoharan. Challenges and opportunities for nucleic acid therapeutics. *Nucleic Acid Therapeutics*, 32(1):8–13, 2022.

Payel Das, Tom Sercu, Kahini Wadhawan, Inkit Padhi, Sebastian Gehrmann, Flaviu Cipcigan, Vijil Chenthamarakshan, Hendrik Strobelt, Cicero Dos Santos, Pin-Yu Chen, et al. Accelerated antimicrobial discovery via deep generative models and molecular dynamics simulations. *Nature Biomedical Engineering*, 5(6):613–623, 2021.

Aryan Deshwal and Janardhan Rao Doppa. Combining latent space and structured kernels for bayesian optimization over combinatorial spaces. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 34, 2021.

Clyde Fare, Peter Fenner, Matthew Benatan, Alessandro Varsi, and Edward O Pyzer-Knapp. A multi-fidelity machine learning approach to high throughput materials screening. *npj computational materials*, 8(1):257, 2022.

Francesco Di Fiore, Michela Nardelli, and Laura Mainini. Active learning and Bayesian optimization: a unified perspective to learn with a goal. *arXiv preprint arXiv:2303.01560*, 2023.

Jose Pablo Folch, Robert M Lee, Behrang Shafei, David Walz, Calvin Tsay, Mark van der Wilk, and Ruth Misener. Combining multi-fidelity modelling and asynchronous batch Bayesian optimization. *Computers & Chemical Engineering*, 172:108194, 2023.

Adam Evan Foster. *Variational, Monte Carlo and policy-based approaches to Bayesian experimental design*. PhD thesis, University of Oxford, 2021.

Peter I. Frazier. A tutorial on bayesian optimization. *arXiv preprint arXiv:1807.02811*, 2018.

Wenhao Gao, Tianfan Fu, Jimeng Sun, and Connor W. Coley. Sample efficiency matters: A benchmark for practical molecular optimization, 2022.

Jacob R. Gardner, Geoff Pleiss, David Bindel, Kilian Q. Weinberger, and Andrew Gordon Wilson. Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration, 2021.

Roman Garnett. *Bayesian optimization*. Cambridge University Press, 2023.

Roman Garnett, Yamuna Krishnamurthy, Xuehan Xiong, Jeff Schneider, and Richard Mann. Bayesian optimal active search and surveying. *arXiv preprint arXiv:1206.6406*, 2012.

Antoine Grosnit, Rasul Tutunov, Alexandre Max Maraval, Ryan-Rhys Griffiths, Alexander I. Cowen-Rivers, Lin Yang, Lin Zhu, Wenlong Lyu, Zhitang Chen, Jun Wang, Jan Peters, and Haitham Bou-Ammar. High-dimensional Bayesian optimisation with variational autoencoders and deep metric learning. *arXiv preprint arXiv:2106.03609*, 2021.

Thomas A Halgren. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, 17(5-6):490–519, 1996.

Moksh Jain, Emmanuel Bengio, Alex Hernandez-Garcia, Jarrid Rector-Brooks, Bonaventure FP Dossou, Chanakya Ajit Ekbote, Jie Fu, Tianyu Zhang, Michael Kilgour, Dinghuai Zhang, et al. Biological sequence design with GFlowNets. In *International Conference on Machine Learning (ICML)*, volume 162. PMLR, 2022.

Moksh Jain, Tristan Deleu, Jason Hartford, Cheng-Hao Liu, Alex Hernandez-Garcia, and Yoshua Bengio. GFlowNets for AI-driven scientific discovery. *Digital Discovery*, 2023a.

Moksh Jain, Sharath Chandra Raparthy, Alex Hernandez-Garcia, Jarrid Rector-Brooks, Yoshua Bengio, Santiago Miret, and Emmanuel Bengio. Multi-objective GFlowNets. In *International Conference on Machine Learning (ICML)*, 2023b.

Shali Jiang, Gustavo Malkomes, Geoff Converse, Alyssa Shofner, Benjamin Moseley, and Roman Garnett. Efficient nonmyopic active search. In *International Conference on Machine Learning (ICML)*, volume 70. PMLR, 2017.

Shali Jiang, Roman Garnett, and Benjamin Moseley. Cost effective active search. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 32, 2019.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, 2021.

Kirthevasan Kandasamy, Gautam Dasarathy, Junier B. Oliva, Jeff Schneider, and Barnabas Poczos. Multi-fidelity gaussian process bandit optimisation. *Journal of Artificial Intelligence Research (JAIR)*, 66:151–196, 2019.

Michael Kilgour, Tao Liu, Brandon D Walker, Pengyu Ren, and Lena Simine. E2EDNA: Simulation protocol for DNA aptamers with ligands. *Journal of Chemical Information and Modeling*, 61(9):4139–4144, 2021.

Ross D King, Kenneth E Whelan, Ffion M Jones, Philip GK Reiser, Christopher H Bryant, Stephen H Muggleton, Douglas B Kell, and Stephen G Oliver. Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature*, 427(6971):247–252, 2004.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

Mario Krenn, Florian Häse, AkshatKumar Nigam, Pascal Friederich, and Alan Aspuru-Guzik. Self-referencing embedded strings (SELFIES): A 100% robust molecular string representation. *Machine Learning: Science and Technology*, 1(4):045024, 2020.

Patrick Kunzmann and Kay Hamacher. Biotite: a unifying open source computational biology framework in Python. *BMC Bioinformatics*, 19:1–8, 2018.

A Gilad Kusne, Heshan Yu, Changming Wu, Huairuo Zhang, Jason Hattrick-Simpers, Brian DeCost, Suchismita Sarker, Corey Oses, Cormac Toher, Stefano Curtarolo, et al. On-the-fly closed-loop materials discovery via Bayesian active learning. *Nature Communications*, 11(1):5966, 2020.

Salem Lahlou, Tristan Deleu, Pablo Lemos, Dinghuai Zhang, Alexandra Volokhova, Alex Hernández-García, Léna Néhale Ezzine, Yoshua Bengio, and Nikolay Malkin. A theory of continuous generative flow networks. In *International Conference on Machine Learning (ICML)*, 2023.

Shibo Li, Wei Xing, Mike Kirby, and Shandian Zhe. Multi-fidelity bayesian optimization via deep neural networks. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 33, 2020.

Shibo Li, Robert M Kirby, and Shandian Zhe. Deep multi-fidelity active learning of high-dimensional outputs. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 151, pp. 1694–1711. PMLR, 2022a.

Shibo Li, Jeff M. Phillips, Xin Yu, Robert M. Kirby, and Shandian Zhe. Batch multi-fidelity active learning with budget constraints, 2022b.

Wesley J. Maddox, Samuel Stanton, and Andrew Gordon Wilson. Conditioning sparse variational gaussian processes for online decision-making. *CoRR*, abs/2110.15172, 2021. URL `https://arxiv.org/abs/2110.15172`.

Nikolay Malkin, Moksh Jain, Emmanuel Bengio, Chen Sun, and Yoshua Bengio. Trajectory balance: Improved credit assignment in GFlowNets. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 35, 2022.

Natalie Maus, Kaiwen Wu, David Eriksson, and Jacob Gardner. Discovering many diverse solutions with bayesian optimization. *arXiv preprint arXiv:2210.10953*, 2022.

Henry B. Moss, David S. Leslie, Javier Gonzalez, and Paul Rayson. GIBBON: General-purpose information-based bayesian optimisation. *Journal of Machine Learning Research (JMLR)*, 22(235):1–49, 2021.

Hagen Neugebauer, Fabian Bohle, Markus Bursch, Andreas Hansen, and Stefan Grimme. Benchmark study of electrochemical redox potentials calculated with semiempirical and dft methods. *The Journal of Physical Chemistry A*, 124(35):7166–7176, 2020.

Quan Nguyen, Arghavan Modiri, and Roman Garnett. Nonmyopic multifidelity active search. In *International Conference on Machine Learning (ICML)*, volume 139. PMLR, 2021.

Leonard Papenmeier, Luigi Nardi, and Matthias Poloczek. Bounce: Reliable high-dimensional bayesian optimization for combinatorial and mixed spaces. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 36, pp. 1764–1793, 2023.

Robert G Parr. Density functional theory of atoms and molecules. In *Horizons of Quantum Chemistry: Proceedings of the Third International Congress of Quantum Chemistry Held at Kyoto, Japan, October 29-November 3, 1979*, pp. 5–15. Springer, 1980.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 32, 2019.

Benjamin Peherstorfer, Karen Willcox, and Max Gunzburger. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Review*, 60(3):550–591, 2018.

John P Perdew and Karla Schmidt. Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conference Proceedings*, 577(1):1–20, 2001.

Paris Perdikaris, M. Raissi, Andreas C. Damianou, ND Lawrence, and George Em Karniadakis. Nonlinear information fusion algorithms for data-efficient multi-fidelity modelling. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 473(2198), 2017.

Malak Pirtskhalava, Anthony A Amstrong, Maia Grigolava, Mindia Chubinidze, Evgenia Alimbarashvili, Boris Vishnepolsky, Andrei Gabrielian, Alex Rosenthal, Darrell E Hurt, and Michael Tartakovsky. DBAASP v3: database of antimicrobial/cytotoxic activity and structure of peptides as a resource for development of new therapeutics. *Nucleic Acids Research*, 49(D1):D288–D297, 2021.

Pavel G Polishchuk, Timur I Madzhidov, and Alexandre Varnek. Estimation of the size of drug-like chemical space based on GDB-17 data. *Journal of Computer-Aided Molecular Design*, 27:675–679, 2013.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 11 2005. ISBN 9780262256834. doi: 10.7551/mitpress/3206.001.0001. URL `https://doi.org/10.7551/mitpress/3206.001.0001`.

Herbert E. Robbins. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58:527–535, 1952.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

Burr Settles. Active learning literature survey. *Independent Technical Report*, 2009.

David S Sholl and Janice A Steckel. *Density functional theory: a practical introduction*. John Wiley & Sons, 2022.

Jasper Snoek, Oren Rippel, Kevin Swersky, Ryan Kiros, Nadathur Satish, Narayanan Sundaram, Mostofa Patwary, Mr Prabhat, and Ryan Adams. Scalable Bayesian optimization using deep neural networks. In *International Conference on Machine Learning (ICML)*. PMLR, 2015.

András Sobester, Alexander Forrester, and Andy Keane. *Engineering Design via Surrogate Modelling. Appendix: Example Problems*, pp. 195–203. John Wiley & Sons, Ltd, 2008. ISBN 9780470770801. doi: https://doi.org/10.1002/9780470770801.app1. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470770801.app1`.

Jialin Song, Yuxin Chen, and Yisong Yue. A general framework for multi-fidelity Bayesian optimization with Gaussian processes. In *International Conference on Artificial Intelligence and Statistics*, volume 89. PMLR, 2018.

Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *International Conference on Machine Learning (ICML)*, 2010.

Samuel Stanton, Wesley Maddox, Nate Gruver, Phillip Maffettone, Emily Delaney, Peyton Greenside, and Andrew Gordon Wilson. Accelerating Bayesian optimization for biological sequence design with denoising autoencoders. In *International Conference on Machine Learning (ICML)*, volume 162. PMLR, 2022.

Jonathan M Stokes, Kevin Yang, Kyle Swanson, Wengong Jin, Andres Cubillos-Ruiz, Nina M Donghia, Craig R MacNair, Shawn French, Lindsey A Carfrae, Zohar Bloom-Ackermann, et al. A deep learning approach to antibiotic discovery. *Cell*, 180(4):688–702, 2020.

Shion Takeno, Hitoshi Fukuoka, Yuhki Tsukada, Toshiyuki Koyama, Motoki Shiga, Ichiro Takeuchi, and Masayuki Karasuyama. Multi-fidelity Bayesian optimization with max-value entropy search and its parallelization. In *International Conference on Machine Learning (ICML)*, volume 119. PMLR, 2020.

Fabio Urbina, Filippa Lentzos, Cédric Invernizzi, and Sean Ekins. Dual use of artificial-intelligence-powered drug discovery. *Nature Machine Intelligence*, 4(3):189–191, 2022.

Zi Wang and Stefanie Jegelka. Max-value entropy search for efficient Bayesian optimization. In *International Conference on Machine Learning (ICML)*, volume 70. PMLR, 2017.

Manfred KK Warmuth, Gunnar Rätsch, Michael Mathieson, Jun Liao, and Christian Lemmen. Active learning in the drug discovery process. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 14, 2001.

Andrew Gordon Wilson, Zhiting Hu, Ruslan Salakhutdinov, and Eric P Xing. Deep kernel learning. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 370–378. PMLR, 2016.

Dongxia Wu, Ruijia Niu, Matteo Chinazzi, Yian Ma, and Rose Yu. Disentangled multi-fidelity deep bayesian active learning, 2023.

Jian Wu, Saul Toscano-Palmerin, Peter I. Frazier, and Andrew Gordon Wilson. Practical multi-fidelity bayesian optimization for hyperparameter tuning. In *Uncertainty in Artificial Intelligence Conference (UAI)*, volume 115, pp. 788–798. PMLR, 2019.

Dezhen Xue, Prasanna V Balachandran, John Hogden, James Theiler, Deqing Xue, and Turab Lookman. Accelerated search for materials with targeted properties by adaptive design. *Nature Communications*, 7 (1):1–9, 2016.

Joseph D Yesselman, Daniel Eiler, Erik D Carlson, Michael R Gotrik, Anne E d'Aquino, Alexandra N Ooms, Wipapat Kladwang, Paul D Carlson, Xuesong Shi, David A Costantino, et al. Computational design of three-dimensional RNA structure and function. *Nature Nanotechnology*, 14(9):866–873, 2019.

Ruihao Yuan, Zhen Liu, Prasanna V Balachandran, Deqing Xue, Yumei Zhou, Xiangdong Ding, Jun Sun, Dezhen Xue, and Turab Lookman. Accelerated discovery of large electrostrains in BaTiO3-based piezoelectrics using active learning. *Advanced Materials*, 30(7):1702884, 2018.

Joseph N Zadeh, Conrad D Steenberg, Justin S Bois, Brian R Wolfe, Marshall B Pierce, Asif R Khan, Robert M Dirks, and Niles A Pierce. NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry*, 32(1):170–173, 2011.

Wenhu Zhou, Runjhun Saran, and Juewen Liu. Metal sensing by DNA. *Chemical Reviews*, 117(12):8272–8325, 2017.

C Lawrence Zitnick, Lowik Chanussot, Abhishek Das, Siddharth Goyal, Javier Heras-Domingo, Caleb Ho, Weihua Hu, Thibaut Lavril, Aini Palizhati, Morgane Riviere, et al. An introduction to electrocatalyst design using machine learning for renewable energy storage. *arXiv preprint arXiv:2010.09435*, 2020.

A MF-GFN Algorithm

Our algorithm, MF-GFN, detailed in Algorithm 1, proceeds as follows: An active learning round j starts with a data set of annotated samples $\mathcal{D}_j = \{(x_i, f_m(x_i), m_i)\}_{1 \leq m \leq M}$. The data set is used to fit a probabilistic *multi-fidelity surrogate* model $h(x, m)$ of the posterior $p(f_m(x)|x, m, \mathcal{D})$. The output of the surrogate model is then used to compute the value of a *multi-fidelity acquisition function* $\alpha(x, m)$. In our experiments, we use the multi-fidelity version (Takeno et al., 2020) of max-value entropy search (MES) (Wang & Jegelka, 2017), which is an information-theoretic acquisition function widely used in Bayesian optimisation.

An active learning round terminates by generating N objects from the sampler (here the GFlowNet policy π) and forming a batch with the best B objects, according to α. Note that $N \gg B$, since sampling from a GFlowNet is relatively inexpensive. The selected objects are annotated by the corresponding oracles and incorporated into the data set, such that $\mathcal{D}_{j+1} = \mathcal{D}_j \cup \{(x_1, f_m(x_1), m_1), \ldots (x_B, f_m(x_B), m_B)\}$.

B Surrogate Models and Acquisition Function

In this appendix, we provide additional details about the surrogate models used in our active learning experiments, as well as about the acquisition function.

B.1 Gaussian Processes

In our multi-fidelity active learning experiments, we model the posterior distribution over the outputs of the oracles f_m, $p(f_m(x)|x, m, \mathcal{D})$, assuming that observations are perturbed by noise from a normal distribution $\mathcal{N}(0, \sigma^2)$. The assumption of normally distributed noise with constant variance is widely used in the Bayesian Optimization literature. Consider a set of n points $z_{(1:n)} = (x_1, m_1), (x_2, m_2), \ldots, (x_n, m_n)$ with observed values $y_{(1:n)} = y_1, y_2, \ldots, y_n$. We can then use Gaussian Processes such that $f|z_{(1:n)}, y_{(1:n)} \sim GP(\mu_n, K_n)$ with mean μ_n and covariance function or kernel K_n evaluated at point $z = (x, m)$ as

$$\mu_n(x) = \mu(x) + K(x, x_{1:n})(K(x_{1:n}, x_{1:n}) + \sigma^2 I)^{-1}(y_{1:n} - \mu(x_{1:n}))$$
$$K_n(x_1, x_2) = K((x_1, x_2) - K(x_1, x_{1:n})(K(x_{1:n}, x1 : n) + \sigma^2 I)^{-1} K(x_{1:n}, x_2).$$

We adapt the multi-fidelity kernel as proposed in (Wu et al., 2019). The authors implement a Downsampling Kernel for the data fidelity parameter, in cases where it is relevant, along with an Exponential Decay Kernel for the iteration fidelity parameter, when applicable. As our experimental approach treats fidelity as akin to a data point, the implementation of the Downsampling Kernel has been incorporated.

Hence, the kernel function of the GP is

$$K(z_1, z_2) = K_1(x_1, x_2) \times K_2(m_1, m_2),$$

where $K_1(\cdot, \cdot)$ is a square-exponential kernel and $K_2(\cdot, \cdot)$, the downsampling kernel is given by

$$K_2(m_1, m_2) = c + (1 - m_1)^{1+\delta}(1 - m_2)^{(1+\delta)},$$

where $c, \delta > 0$ are hyper-parameters.

B.2 Deep Kernel Learning

While for the synthetic (simpler) tasks we use exact GPs, it is well know that GPs are less effective with high-dimensional data, because of the reliance on the Euclidean distance. Furthermore, standard GPs are not directly applicable to discrete, structured data. Therefore, for the benchmark tasks we implement deep kernel learning (DKL; Wilson et al., 2016). In DKL, the inputs are transformed by

$$k(x_i, x_j|\theta) \to k(g(x_i, w), g(x_j, w)|\theta, w),$$

where the non-linear mapping $g(x, w)$ is a low-dimensional continuous embedding, learnt via a deep neural network—a transformer in our tasks. To scale the GP to large datasets, we implement the stochastic variational GP based on the greedy inducing point method (Chen et al., 2018). We adopt the deep kernel learning experimental setup from (Stanton et al., 2022).

B.3 Acquisition Function

Max Value Entropy Search (MES) (Wang & Jegelka, 2017) is an information-theoretic acquisition function. The standard, single-fidelity MES seeks to query the objective function at locations that reduce our current uncertainty in the maximum value of f^*. It aims to maximise the mutual information between the value of the objective function f when choosing point x and the maximum of the objective function, $f^\star$. This contrasts with previously proposed entropy search (ES) criterion, which instead considers the `arg max` of the objective function. The MES criterion is defined as follows:

$$\alpha = I(f^\star; y|\mathcal{D}_j) = H(y|\mathcal{D}_j) - \mathbb{E}_{f^\star}[H(y|\mathcal{D}_j, f^\star)|\mathcal{D}_j],$$

where y is the outcome of experiment x and $\mathcal{D}_j$ is the data set at the jth active learning iteration. and $H(Y) = \mathbb{E}_Y[-\log(p(Y))]$ is the differential entropy of random variable Y.

The information gain is defined as the reduction in entropy of y provided by knowing the maximal value $f^\star$

$$IG(y, m|\mathcal{D}_n) = H(y|D_n) - H(y|f^\star < m, D_n)$$

It follows that the MES acquisition function can be expressed in terms of IG:

$$\alpha = \mathbb{E}_{m\sim f^\star}[IG_n(y, m|\mathcal{D}_n)],$$

where $y \sim \mathcal{N}(\mu_\mathcal{A}, \sigma_\mathcal{A})$, $f(x) \sim \mathcal{N}(\mu_\mathcal{B}, \sigma_\mathcal{B})$ and the difference between y and $f(x)$ is just independent Gaussian noise.

Replacing the maximisation of an intractable quantity with the maximisation of a lower bound is a well-established strategy. Instead of attempting to evaluate the intractable quantity, IG, we evaluate its lower bound IG^{Approx}.

Thus, the acquisition function becomes

$$\alpha = \frac{1}{\mathcal{M}} \sum_{m\in\mathcal{M}} IG^{Approx}(y, m|\mathcal{D}_n)$$

$$IG^{Approx} = \frac{1}{2} log|R| - \frac{1}{2\mathcal{M}} \sum_{m\in\mathcal{M}} log(1 - \rho^2 \frac{\phi(\gamma(m))}{\Phi(\gamma(m))}[\gamma(m) + \frac{\phi(\gamma(m))}{\Phi(\gamma(m))}]),$$

where ϕ and Φ are the standard normal cumulative distribution and probability density functions (as arising from the expression for the differential entropy of a truncated Gaussian), $\gamma(m) = \frac{m-\mu_n(x)}{\sigma_n(x)}$ and R is the correlation matrix with elements $R_{i,j} = \frac{\Sigma^y_{i,j}}{\Sigma^y_{i,i}\Sigma^y_{j,j}}$.

This construction is called the General Information-Based Bayesian OptimisatioN (GIBBON) acquisition function (Moss et al., 2021).

Multi-Fidelity Formulation Let the maximum of the highest fidelity function f_M (when M different fidelities are available to querying) be $f^\star_M$. We obtain a pair (x, m) which maximally gains information of the optimal value $f^\star$ of the highest fidelity per unit cost. Formally, the multi-fidelity max value entropy search acquisition function that use in our algorithm is the following:

$$\alpha(x, m) = \frac{1}{\lambda_m} I(f^\star_M; f_m|\mathcal{D}_j), \tag{2}$$

where λ_m is the cost of the oracle at fidelity m.

C Experimental Details

This appendix presents the details about the experiments presented in the main Section 4. First, we provide general details about all tasks and then present details specific to each task in separate sections.

C.1 Initial data set and budget

We define a budget (Λ_0) for the initial data set. Let λ_m be the cost of evaluation with oracle f_m, and n_{SF}, n_{MF} be the total number of initial training points in the single- and multi-fidelity experiments respectively. Also let n_m be the number of training points evaluated against f_m in the multi-fidelity experiment such that $n_{MF} = \sum_{m=1}^{m=M} n_m$, then

$$\Lambda_0 = n_{SF} \times \lambda_M = \sum_{m=1}^{m=M} n_m \times \lambda_m$$

The initial data set is split into train-validation in the ratio of 9:1 for all tasks. Task specific information is summarized in Table 1.

For each task, we assign a total active learning budget $\Lambda = \gamma \times \lambda_M$ (Table 2). γ was selected based on the rate of convergence of the algorithms to the modes. Note that during an active learning round, only the oracle evaluations of the sampled batch contribute to Λ. The cost of sampling from a trained GFlowNet is nearly negligible compared to the oracle evaluations. This is why we can afford to sample a large number of samples ($N = 5 \times B$) to then select the best B, according to the acquisition function (Algorithm 1).

C.2 DKL Implementation Details

We are describe our implementation of DKL, which is inspired by (Stanton et al., 2022).

Neural Network Architecture For all experiments, the same base architecture was used, featuring transformer encoder layers with position masking for padding tokens. Standard pre-activation residual blocks were implemented, comprising two convolutional layers, layer normalization, and swish activations. The encoder embeds input sequences with standard vocabulary and sinusoidal position embeddings. The encoder is trained with the Masked Language Modeling (MLM) objective which is calculated by randomly masking input tokens and subsequently computing the empirical cross-entropy between the original sequence and the predictive distribution generated by the MLM head for the masked positions.

Optimizer Hyperparameters The running estimates of the first two moments in the Adam optimizer (Kingma & Ba, 2015) were disabled by setting $\beta_1 = 0.0$ and $\beta_2 = 0.01$.

Kernel Hyperparameters In order to force the encoder to learn features appropriate for the initial lengthscale, we place a tight Gaussian prior $\sigma = 0.01$ around the intial lengthscale value. The reinitialization procedure for inducing point locations and variational parameters outlined in Maddox et al. (2021) was followed.

C.3 Policy Implementation Details

Neural Network Architecture For all tasks, the architecture of the forward policy (P_F) model is a multi-layer perceptron with 2 hidden layers and 2048 units per layer. The backward policy (P_B) model was set to share all but the last layer parameters with P_F. We use LeakyReLU as our activation function as in Bengio et al. (2021a). All models are trained with the Adam optimiser Kingma & Ba (2015).

Reward Function As detailed in 1, the GFlowNet is trained to generate samples with a higher value of MES (and its multi-fidelity variant) in single- and multi-fidelity experiments respectively. In order to increase the relative reward of higher values of the acquisition function, we transform the MES value $\alpha(x, m)$ with the reward function $R(\alpha(x, m), \beta)$. On an additional note, MES exhibits increased sparsity as more samples are discovered. Hence, in order to facilitate optimization, a linear reduction of the parameter β

(with a scaling factor denoted by ρ) is implemented at each successive active learning round so as to scale up the rewards. Given an active learning round j,

$$R(\alpha(x,m)) = \frac{\alpha(x,m) \times \rho^{j-1}}{\beta}$$

Note that within an active learning round (j), the GFN samples from this fixed reward function and thus the policy need not be conditioned on j. Details for all tasks are summarized in Table 2.

Our models are implemented in `pytorch` (Paszke et al., 2019), and rely on `botorch` (Balandat et al., 2020) and `GPytorch` (Gardner et al., 2021).

Table 1: Oracle costs (indexed by increasing level of fidelity) and initial data set details

Task	Oracle Costs			Initial Data Set				
				Λ_0	n_{SF}	n_{MF}		
	λ_1	λ_2	λ_M			n_1	n_2	n_M
Branin	0.01	0.1	1	4	4	20	20	2
Hartmann 6D	0.125	0.25	1	25	25	80	40	5
DNA	–	0.2	20	1600	80	–	3000	50
AMP	0.5	0.5	50	2500	50	2000	2000	10
Molecules	1	3	7	1050	150	700	68	16

Table 2: Active-learning and policy reward function and hyperparameters of multi-fidelity experiments

Task	Surrogate Model	Active-learning		Policy reward function	
		γ	B	β	ρ
Branin	Exact GP	300	30	1	1
Hartmann 6D	Exact GP	100	10	1e-2	1
DNA	DKL	256	512	1e-5	2
Antimicrobial Peptides	DKL	20	32	1e-5	1
Molecules	DKL	180	128	1e-6	1.5

C.4 Synthetic Tasks

C.4.1 Branin

We consider an active learning problem in a two-dimensional space where the target function f_M is the Branin function, as modified in (Sobester et al., 2008) and implemented in `botorch` (Balandat et al., 2020). In the domain $[-5, 10] \times [0, 15]$, the Branin function has three modes and is evaluated using the following expression:

$$f(x) = (x_2 - \frac{-1.25x_1^2}{\pi^2} + \frac{5x_1}{\pi} - 6)^2 + (10 - \frac{5}{4\pi})\cos(x_1) + 10.$$

This corresponds to the modification introduced in (Sobester et al., 2008). As lower fidelity functions, we used the expressions from (Perdikaris et al., 2017), which involve non-linear transformations of the true function as well as shifts and non-uniform scalings. The functions, indexed by increasing level of fidelity, are the following:

$$f_1(x) = f_2(1.2(x+2)) - 3x_2 + 1$$
$$f_2(x) = 10\sqrt{f(x-2)} + 2(x_1 - 0.5) - 3(3x_2 - 1) - 1$$

We simulate three levels of fidelity, including the true function. The lower-fidelity oracles, the costs of the oracles $(0.01, 0.1, 1.0)$ as well as the number of points queried in the initial training set were adopted from (Li et al., 2020).

In order to consider a discrete design space, we map the domain to a discrete 100×100 grid. We model this grid with a GFlowNet as in (Bengio et al., 2021a; Malkin et al., 2022): starting from the origin $(0, 0)$, for any state $s = (x_1, x_2)$, the action space consists of the choice between the exit action or the dimension to increment by 1, provided the next state is in the limits of the grid.

We use the `botorch` implementation of an exact multi-fidelity Gaussian process as described in B.1 for regression. The active learning batch size $\mathcal{B}$ is 30 in the Branin task.

Fig. 4a illustrates the results for this task. We observe that MF-GFN is able to reach the minimum of the Branin function with a smaller budget than the single-fidelity counterpart and the baselines.

C.4.2 Hartmann 6D

We consider the 6-dimensional Hartmann function as objective f_M on a hyper-grid domain. It is typically evaluated on the hyper-cube $x_i \in [0, 1]^6$ and consists of six local maxima. The true Hartmann function is given by

$$f(x) = \sum_{i=1}^{4} \alpha_i exp(-\sum_{j=1}^{3} A_{ij}(x_j - P_{ij})^2),$$

where $\alpha = [1.0, 1.2, 3.0, 3.2]$ and $A, P \in R^{4\times 6}$ are the following fixed matrices:

$$A = \begin{bmatrix} 10 & 3 & 17 & 3.5 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 1 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 1 \end{bmatrix}$$

$$P = 10^{-4} \times \begin{bmatrix} 3689 & 1170 & 267 \\ 4699 & 4387 & 7470 \\ 1091 & 8732 & 5547 \\ 381 & 5743 & 8828 \end{bmatrix}$$

To simulate the lower fidelities, we modify α to $\alpha(m)$ where $\alpha(m) = \alpha + (M - m)\delta$ where $\delta = [0.01, -0.01, -0.1, 0.1]$ and $M = 3$. The domain is $X = [0, 1]^6$. This implementation was adopted from Kandasamy et al. (2019). As with Branin, we consider three oracles, adopting the lower-fidelity oracles and the set of costs $(0.125, 0.25, 1.0)$ from (Song et al., 2018).

We discretize the domain into a six-dimensional hyper-grid of length 10, yielding 10^6 possible candidate points. For the surrogate, we use the same exact multi-fidelity GP implementation as of Branin. The active learning batch size B is 10.

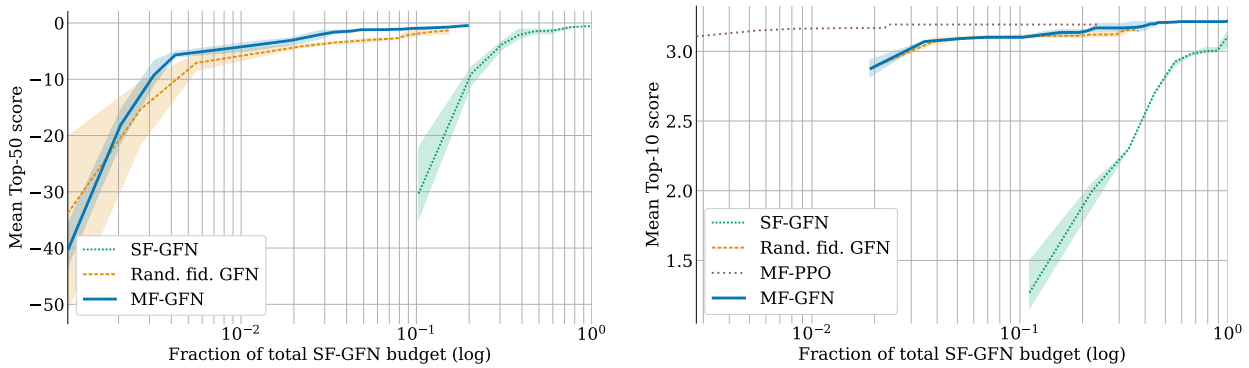
The results for the task are illustrated in Fig. 4b, which indicate that multi-fidelity active learning with GFlowNets (MF-GFN) offers an advantage over single-fidelity active learning (SF-GFN) as well as some of the other baselines in this higher-dimensional synthetic problem as well. The better performance on MF-PPO can be attributed to the fact that while the GFN initiates its exploration from the origin point, the PPO commences from a random starting point within a bounded range, allowing at most three units of displacement (maximum possible displacement is 10 units) along each of the six axes. We hypothesise that this aids the PPO algorithm in expediting the discovery of modes within the optimization process. While MF-PPO performs better in this task, as shown in the benchmark experiments, it tends to collapse to single modes of the function in complex high-dimensional scenarios.

.

C.5 Benchmark Tasks

C.5.1 DNA

We conduct experiments using a two-oracle setup (f_M, f_1) with costs $\lambda_M = 20$ and $\lambda_1 = 0.2$ for the high and low fidelity oracles, respectively. As f_M, we used the free energy of the secondary structure of DNA



(a) Branin task

(b) Hartmann task

Figure 4: Results on the synthetic tasks—Branin and Hartmann functions. The curves indicate the mean score f_M within the top-50 and top-10 samples (for Branin and Hartmann, respectively) computed at the end of each active learning round and plotted as a function of the budget used. The random baseline is omitted from this plot to facilitate the visualisation since the results were significantly worse in these tasks. We observe that MF-GFN clearly outperforms the single-fidelity counterpart (SF-GFN) and slightly improves upon the GFlowNet baseline that samples random fidelities. On Hartmann, MF-PPO initially outperforms the other methods.

Table 3: Oracles for the antimicrobial peptides task

Oracle	Training points	Model	Layers	Hidden units	Training Epochs
f_1	3447	MLP	2	512	51
f_2	3348	MLP	2	512	51
f_M	6795	MLP	2	1024	101

sequences obtained via the software NUPACK (Zadeh et al., 2011), setting the temperature at 310 K. f_1 is a transformer (with 8 encoder layers, 1024 hidden units per layer and 16 heads) trained on 1 million random sequences annotated by f_M. To evaluate the performance of f_1 (with respect to f_M), we construct a test set by sampling sequences from a uniform distribution of the free energy. On this test set, the explained variance of the f_1 is calculated to be 0.8. For the probabilistic surrogate model, we implement deep kernel learning, the hyper-parameters of which are provided in Table 4. The active learning batch size B is 512.

C.5.2 Antimicrobial Peptides

We use data from DBAASP (Pirtskhalava et al., 2021) containing antimicrobial activity labels, which is split into three sets - D_1 for training the oracle, D_2 as the initial data set in the active learning loop and D_3 as the test set (Jain et al., 2022).

This is a three-oracle setup (f_M, f_2, f_1) where each oracle is a different neural network model. The configurations of the oracle models are presented in Table 3. Biologically, each antimicrobial peptide can be classified into an antimicrobial group. f_M is trained on the entire dataset D_1. However, for f_1 and f_2, we divide D_1 into two equally-sized subpart such the set of antimicrobial groups present in one subpart is mutually exclusive of the other. This simulated a setup wherein each lower fidelity oracle specialised in different sub-regions of the entire sample space. We set costs $\lambda_M = 50$ and $\lambda_1 = \lambda_2 = 0.5$ as f_1 and f_2 have similar configurations. Further, as for DNA, the explained variance of f_1 and f_2 (with respect to f_M) on a uniform test set, D_3 was 0.1435 and 0.099 respectively. For the surrogate, we implement deep kernel learning with hyperparameters in Table 4. The active learning batch size B is 32.

Table 4: Deep Kernel hyperparameters for the DNA and Antimicrobial tasks

	Hyperparameter	Value
Architecture	Num. of Layers	8
	Num. of Heads	8
	Latent dimension	64
	GP likelihood variance init.	0.25
	GP length scale prior	$\mathcal{N}(0.7, 0.01)$
	Num. of inducing points (SVGP head)	64
Optimisation	Batch Size	128
	Learning Rate	1e-3
	Adam EMA parameters $(\beta 1, \beta 2)$	(0., 1e-2)
	Max. number of Epochs	512
	Early stopping patience (number of epochs)	15
	Early stopping holdout ratio	0.1

C.5.3 Small Molecules

This is a three oracle setup (f_M, f_2, f_1) with costs representing the actual compute time. We implement the oracles using `RDKit` 2023.03 (rdk, 2023) and the semi-empirical quantum chemistry package `xTB`. We use `GFN2-xTB` (Bannwarth et al., 2019) method for the single point calculation of ionization potential and electron affinity with empirical correction terms.

In f_1, we consider one conformer obtained by `RDKit` with its geometry optimised via force-field `MMFF94`(Halgren, 1996). This geometry is used to calculate (vertical) IP/EA. In f_2, we consider two conformers obtained by `RDKit`, and take the lowest energy conformer after optimisation by `MMFF94`, and further optimise it via `GFN2-xTB` to obtain the ground state geometry; this remains a vertical IP/EA calculation. In f_M, we consider four conformers obtained by `RDKit`, and take the lowest energy conformer after optimisation by `MMFF94`, and further optimise it via `GFN2-xTB`; the corresponding ion is then optimised by `GFN2-xTB`, and the adiabatic energy difference is obtained via total electronic energy. The fidelities are based on the fact that vertical IP/EA approximates that of adiabatic ones (to varying degrees, depending on the molecule). On a uniform test set of 1400 molecules, the explained variance of f_1 and f_2 (with respect to f_M) is $0.1359, 0.279$ and $0.79, 0.86$ for the EA and IP tasks respectively.

The surrogate model is a deep kernel. Further details about the hyperparameters are provided in Table 5. The active learning batch size B is 128. In the environment for GFN, we consider a set of SELFIES vocabularies containing aliphatic and aromatic carbon, boron, nitrogen, oxygen, fluorine, sulfur, phosphorous, chlorine, and bromine, subject to standard valency rules.

We note that this is proof-of-concept and hence we do not conduct a full search of conformers, and nor do we use Density Functional Theory calculations, but we note that the highest fidelity oracle has a good correlation with experiments (Neugebauer et al., 2020). We do not consider synthesisability in this study and we note it may negatively impact GFN as unphysical molecules could produce false results for the semi-empirical oracle.

D Metrics

In this section, we provide additional details about the metrics used for the evaluation of the proposed MF-GFN as well as the baselines.

Mean Top-K Score We adapt this metric from Bengio et al. (2021a). At the end of an active learning round, we sample N (x, m) candidates and then select the top-K candidates ($K \ll N$) according to the acquisition function value. In the experiments, we score these top-K candidates with the corresponding oracle. In the figures, we report the mean score according to the highest-fidelity oracle.

Table 5: Deep Kernel hyperparameters for the molecular tasks

	Hyperparameter	Value
Architecture	Number of Layers	8
	Number of Heads	8
	Latent dimension	32
	GP likelihood variance init.	0.25
	GP length scale prior	$\mathcal{N}(0.7, 0.01)$
	Number of inducing points (SVGP head)	64
Optimisation	Batch Size	128
	Learning Rate	1e-3
	Adam EMA parameters $(\beta 1, \beta 2)$	(0., 1e-2)
	Max. number of Epochs	512
	Early stopping patience (number of epochs)	15
	Early stopping holdout ratio	0.1

Mean Diverse Top-K Score This is a version of the previous metric by which we restricts the selection of the K candidates to examples that are diverse between each other. We use similarity measures (*vide infra*) such that we sample the top-K candidates where each candidate is at most similar to each other by a certain threshold. For antimicrobial peptides, the sequence identity threshold is 0.35; for DNA aptamers, the sequence identity threshold is 0.60; for molecules, the Tanimoto similarity distance threshold is 0.35.

Diversity In order to measure the diversity of a set of candidates, we use use one minus the similarity index with the following details for each of the tasks:

- **DNA aptamers**: The similarity measure is calculated by the mean of pairwise sequence identity between a set of DNA sequences. We utilize global alignment with Needleman-Wunsch algorithm and standard nucleotide substitution matrix, as calculated by `biotite` package (Kunzmann & Hamacher, 2018).

- **Antimicrobial peptides**: The similarity measure is calculated by the mean of pairwise sequence identity between a set of peptide sequences. We utilize global alignment with Needleman-Wunsch algorithm and BLOSUM62 substitution matrix, as calculated by `biotite` package (Kunzmann & Hamacher, 2018).

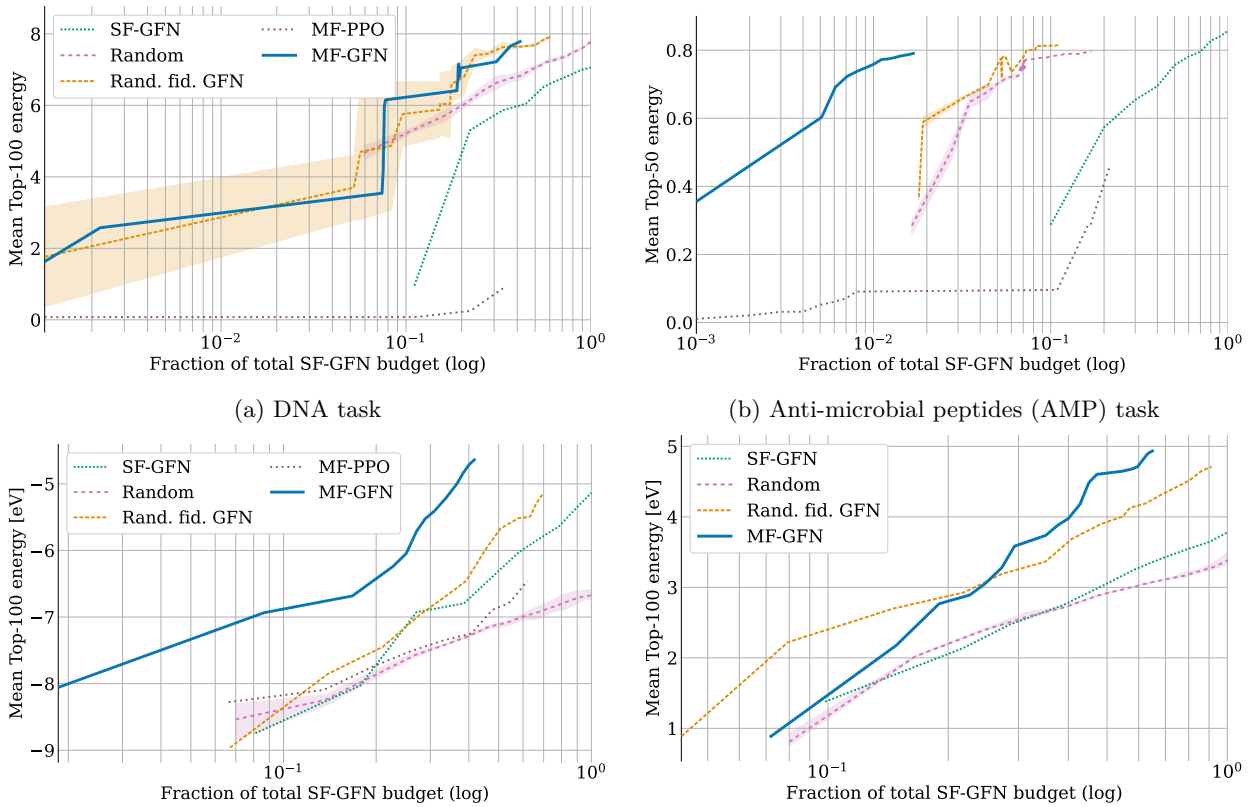
- **Molecules**: The similarity measure is calculated by the mean of pairwise Tanimoto similarity between a set of molecules. Tanimoto metrics are calculated from Morgan Fingerprints (radius of two, size of 2048 bits) as implemented in `RDKit` package (rdk, 2023).

E Additional Results

E.1 Energy of Diverse Top-K

In this section we complement the results presented in Section 4 with the mean diverse top-K scores, as defined in Appendix D. This metric combines that mean top-K score and the measure of diversity. Figure 5 shows the results on the DNA, AMP and the molecular tasks.

The results with this metric allow us to further confirm that multi-fidelity active learning with GFlowNets is able to discover sets of diverse candidates with high mean scores, as is sought in many scientific discovery applications. In contrast, methods that do not encourage diversity such as RL-based algorithms (MF-PPO) obtain comparatively much lower results with this metric.



(a) DNA task

(b) Anti-microbial peptides (AMP) task

(c) Molecules ionisation potential (IP) task

(d) Molecules electron affinity (EA) task

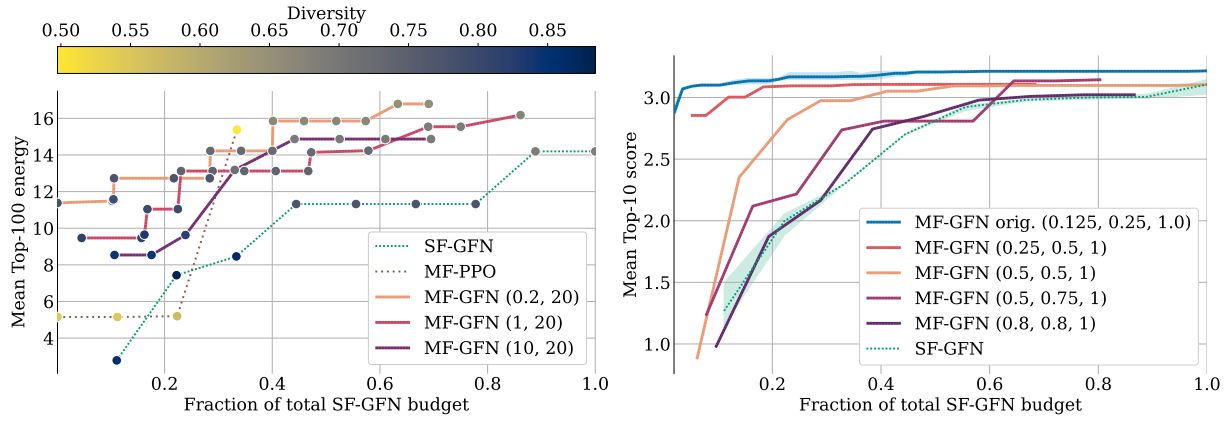
Figure 5: Mean scores (energy) of diverse top-K candidates on the DNA (top left), AMP (top right) and molecular (bottom) tasks. The mean energy is computed across the top-K examples at each active learning round that also satisfy the criteria of diversity. Consistent with the diversity metrics observed in Fig. 2, we here see that GFlowNet-based methods, and especially MF-GFN, obtain good results according to this metric, while MF-PPO achieves comparatively much lower mean energy.

E.2 Understanding the Impact of Oracle Costs

As discussed in 1, a multi-fidelity acquisition function like the one we use—defined in Eq. (2)—is a cost-adjusted utility function. Consequently, the cost of each oracle plays a crucial role in the utility of acquiring each candidate. In our tasks with small molecules (Section 4.3.3), for instance, we used oracles with costs proportional to their computational demands and observed that multi-fidelity active learning largely outperforms single-fidelity active learning. However, depending on the costs of the oracles, the advantage of multi-fidelity methods can diminish significantly.

In order to analyse the impact of the oracle costs on the performance of MF-GFN, we run several experiments on the DNA task (Section 4.3.1), which consists of two oracles, with a variety of oracle costs. In particular, besides the costs used in the experiments presented in Section 4.3.1, with costs $(0.2, 20)$ for the lowest and highest fidelity oracles, we run experiments with costs $(1, 20)$ and $(10, 20)$.

The results, presented in Fig. 6a, indeed confirm that the advantage of MF-GFN over SF-GFN decreases as the cost of the lowest-fidelity oracle becomes closer to the cost of the highest-fidelity oracle. However, it is remarkable that even with a ratio of costs as small as $1:2$, MF-GFN still outperforms not only SF-GFN but also MF-PPO in terms of cost effectiveness, without diversity being negatively impacted. It is important to note that in practical scenarios of scientific discovery, the cost of lower fidelity oracles is typically orders of magnitude smaller than the cost of the most accurate oracles, since the latter correspond to wet-lab experiments or expensive computer simulations.



(a) DNA (b) Hartmann

Figure 6: Analysis of the impact of the oracle costs on the performance of MF-GFN on the DNA task and the synthetic Hartmann task. On the DNA task, we observe that the advantage over SF-GFN and MF-PPO (0.2, 20) decreases as the cost of the lower fidelity oracle becomes closer to the cost of the highest fidelity oracle. Nonetheless, even with a cost ratio of 1 : 2 MF-GFN displays remarkable performance with respect to other methods. Similar conclusions can be drawn from the analysis on the Hartmann task.

E.3 Batch Size Ablation

We evaluate the impact of the batch size on the performance of MF-GFN and its comparison with the baselines for the molecule IP task with different batch sizes. We notice that the reward curve becomes steeper with higher batch sizes.

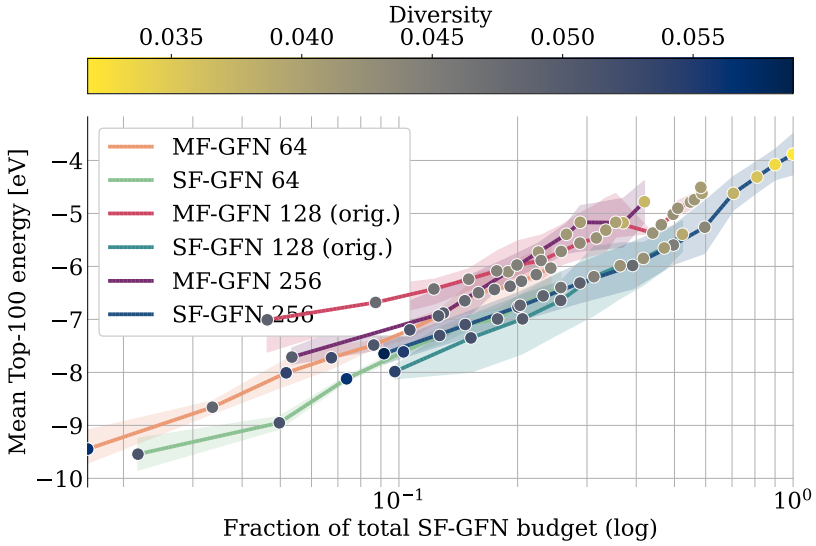


Figure 7: Molecules IP: Impact of acquisition size (64/128/256)

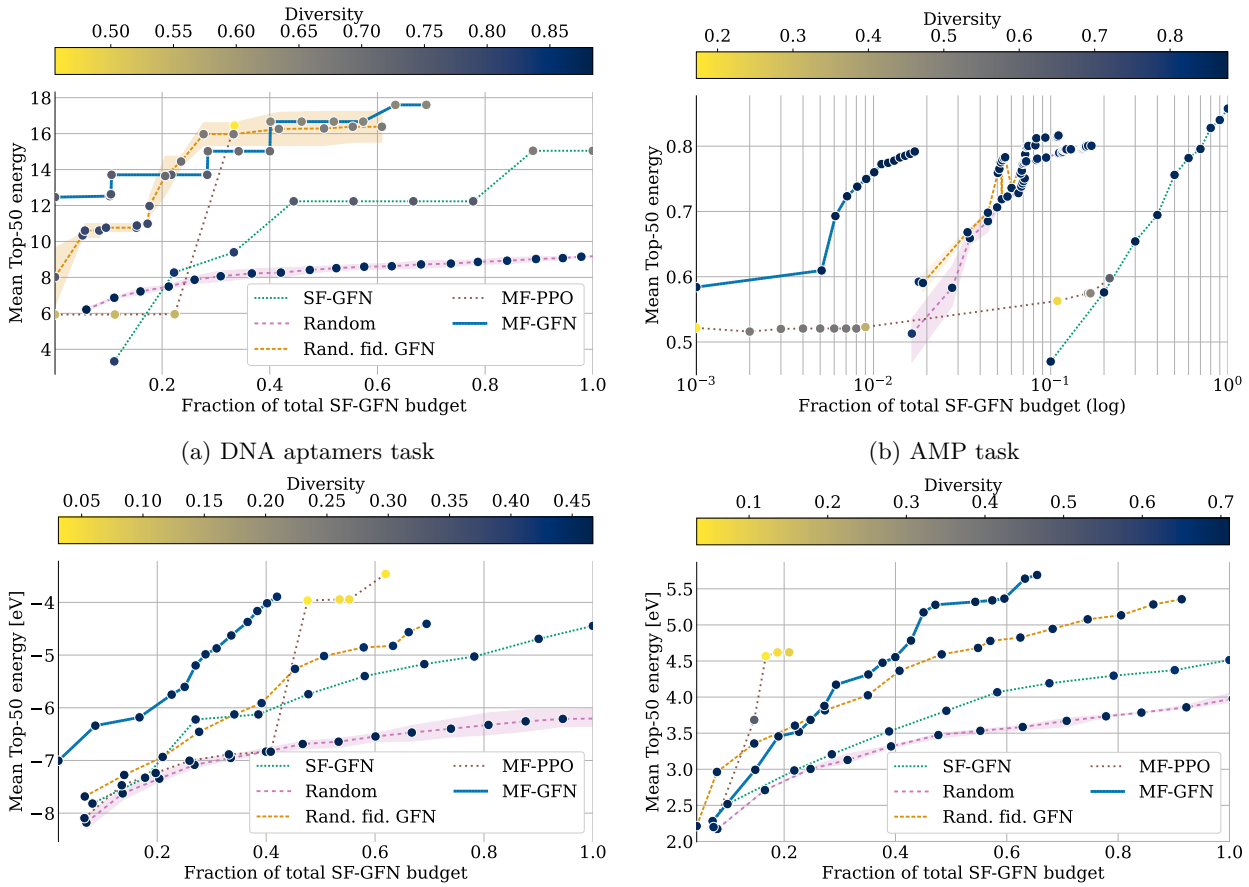
E.4 Impact of the choice of the final batch size

For the set of results presented in the main paper, we computed the mean top-K energy and diversity on the final batch of size $K = 100$. While the choice of K is not arbitrary as it is related to the acquisition size in the active learning loop and in turn to reasonable numbers in the domains of application, it is interesting

to study whether our conclusions are robust to other choices of K. In Fig. 8, we provide the equivalent set of results for all the task with $K = 50$ and in Fig. 9 with $K = 200$, half and double the size, respectively.

In view of these results, we can conclude that the results are robust to the choice of this parameter, since we can derive the same conclusions for all values of $K \in \{50, 100, 200\}$:

- MF-GFN obtains the best trade-off between mean energies and diversity of all the evaluated methods.
- All other GFlowNet-based methods are able to discover diverse samples.
- The multi-fidelity method with PPO is able to discover high-scoring samples, but with a strong lack of diversity.



(a) DNA aptamers task

(b) AMP task

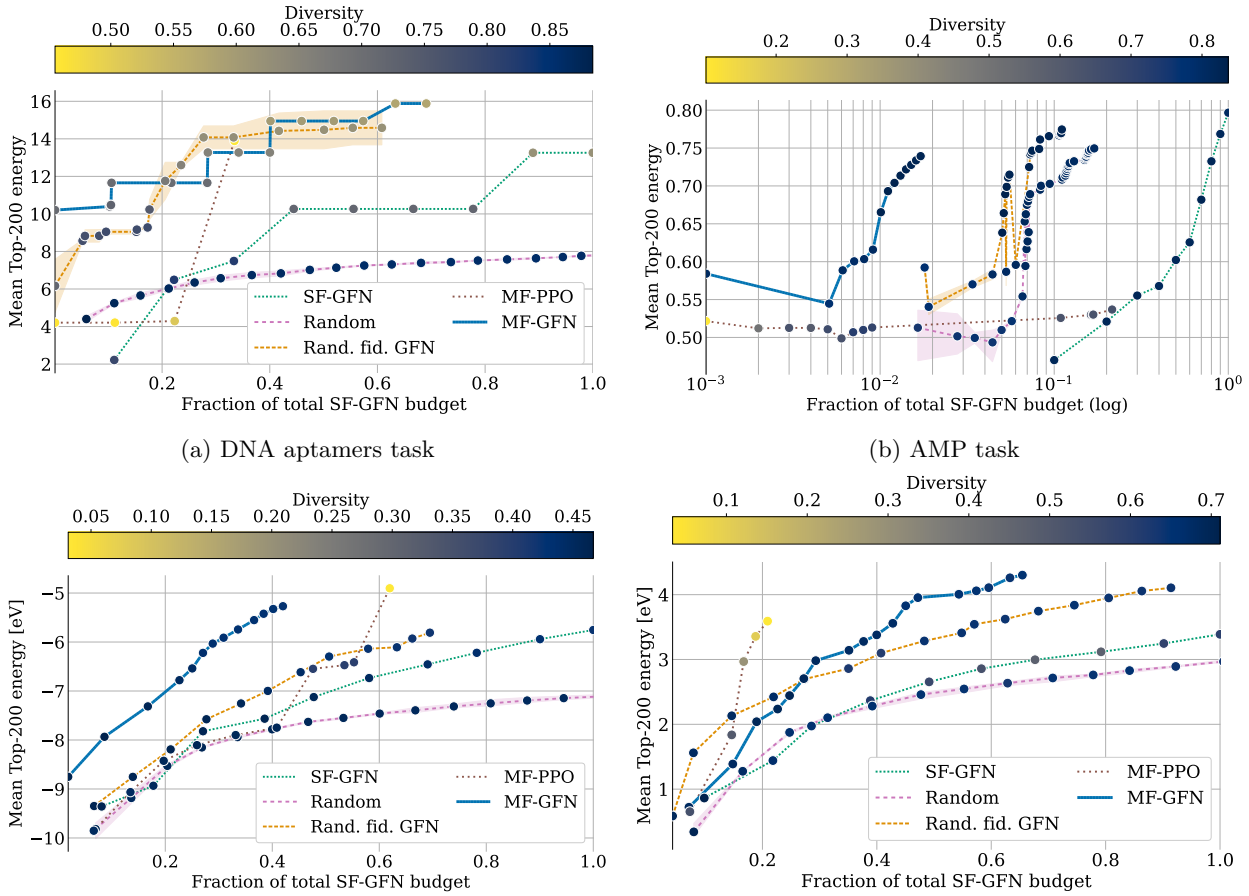
(c) Molecules IP task

(d) Molecules EA task

Figure 8: Results as in the original figures, but with $K = 50$, instead of $K = 100$.

E.5 Visualisation of Sampled Candidates

Given that MF-GFN conducts a cost-aware search with the help of the multi-fidelity acquisition function, our expectation is that the algorithm will selectively query the less costly oracles for input space exploration and will query the more expensive oracles on high-reward candidates. To substantiate this hypothesis, we provide a two-dimensional visualization (Figure 10) of the sampled candidates after expending the allocated budget in the synthetic Branin task. C.4.1.



(a) DNA aptamers task

(b) AMP task

(c) Molecules IP task

(d) Molecules EA task

Figure 9: Results as in the original figures, but with $K = 200$, instead of $K = 100$.

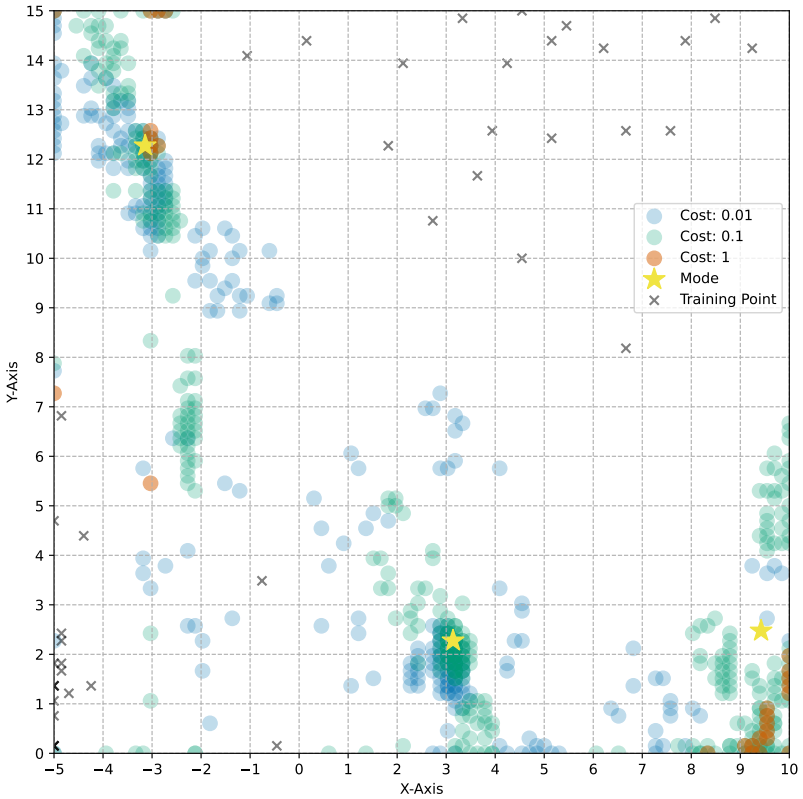


Figure 10: We present a visualization of the sampled candidates (x, m) in the synthetic Branin task (Appendix C.4.1). The domain of Branin is defined in $[-5, 10] \times [0, 15]$. Each round marker, identified by grid-specific coordinates, represents a sampled candidate, x. The markers are color-coded based on the oracle the candidate is to be evaluated with, m. Our observation reveals that the lower fidelity oracles (with costs of 0.01 and 0.1) are primarily used for exploration across the input domain, while evaluations using the high-fidelity oracle (cost=1) are predominantly concentrated near the modes (denoted by the star marker). Furthermore, it's important to note that the training points were intentionally chosen to exclude any modes of the Branin function.