
Active Boltzmann Sampling: Uncertainty-Driven Exploration of High-Dimensional Energy Landscapes via Gaussian Process Surrogates

Alex Guangyuan Wang
Mila, McGill University
Montreal, QC, Canada
guangyuan.wang@mail.mcgill.ca

Abstract

Sampling from equilibrium distributions of high-dimensional, multimodal potential energy surfaces remains a central challenge in computational physics and biochemical discovery. Standard methods, such as Markov Chain Monte Carlo (MCMC) and Molecular Dynamics, suffer from severe mode-trapping in the presence of high energy barriers, while recent deep generative models typically require large offline datasets that are themselves expensive to generate. In this work, we propose *Active Boltzmann Sampling* (ABS), a Bayesian active learning framework that efficiently explores phase space by treating the potential energy landscape as an expensive black-box function modeled by a Gaussian Process (GP). Unlike traditional Bayesian Optimization which seeks to minimize regret (finding the global minimum), ABS aims to minimize distributional divergence. We introduce a novel stochastic acquisition strategy rooted in Thompson Sampling that leverages epistemic uncertainty to tunnel through high-energy barriers and theoretically recovers the Boltzmann distribution in the asymptotic limit. We validate our approach with sine-cosine and double well energies and the challenging 39-dimensional Lennard-Jones 13 (LJ13) cluster. We show that ABS achieves competitive distributional fidelity in the low data regime and recovers key metastable states with much fewer energy evaluations than state-of-the-art annealing baselines, as quantified by the Wasserstein-2 distance. Our code is open-sourced at <https://github.com/al2wang/bayesian/blob/main/main.py>.

1 Introduction

The ability to generate independent, identically distributed (i.i.d.) samples from the Boltzmann distribution, $\pi(\mathbf{x}) \propto \exp(-U(\mathbf{x})/k_B T)$, is the cornerstone of statistical mechanics. It underpins our ability to calculate thermodynamic observables, such as free energy, heat capacity, and reaction rates that are essential for applications ranging from drug discovery and protein folding to the design of novel materials. However, for many systems of interest, the potential energy function $U(\mathbf{x})$ is computationally expensive and defines a landscape fraught with deep metastable minima separated by high energy barriers.

Standard sampling techniques, such as Metropolis-Hastings MCMC or Langevin Dynamics, rely on local gradient information to traverse the landscape. In rough, high-dimensional spaces, these methods exhibit *broken ergodicity*: they remain trapped in a single local basin for timescales that exceed feasible computation, failing to visit other statistically significant modes. While advanced sampling techniques like parallel tempering attempt to mitigate this, they often require heuristic tuning of reaction coordinates or temperature schedules, which can be difficult to define a priori for unknown systems. More recently, deep generative models, particularly Normalizing Flows (e.g.,

Boltzmann Generators [Noé et al., 2019] and iDEM [Akhound-Sadegh et al., 2024]), have shown promise in learning mappings from simple latent distributions to complex target densities. However, these amortized inference methods typically face the problem that they require a representative dataset of samples from the high-energy regions to train the flow, but generating such a dataset is precisely the problem we wish to solve. Purely energy-based training of flows via reverse KL divergence is prone to *mode collapse*, where the model captures only a subset of the modes Midgley et al. [2023].

We approach this challenge through the lens of *Active Learning*. By modeling the potential energy surface (PES) with a Gaussian Process (GP), we can leverage Bayesian uncertainty quantification to drive exploration. However, a critical gap exists in the current literature: standard Bayesian Optimization (BO) algorithms [Srinivas et al., 2010, Snoek et al., 2012] are designed for *optimization* (i.e., finding the global minimum $\mathbf{x}^* = \arg \min U(\mathbf{x})$) rather than *sampling*. Acquisition functions like Expected Improvement (EI) are aggressively greedy; once the global minimum is located, they cease to explore the suboptimal basins that contribute most to the thermodynamic entropy.

In this paper, we bridge the gap between Bayesian optimization and sampling. We propose a methodology that uses the epistemic uncertainty of a GP not merely to reduce error, but to actively construct a stochastic surrogate landscape that facilitates barrier crossing. Our contributions are as follows: First, we formulate the *Active Boltzmann Sampling* (ABS) algorithm, a closed-loop framework that uses Gaussian Processes to iteratively construct a high-fidelity surrogate of the Boltzmann distribution. Second, we introduce a *stochastic acquisition strategy* based on Boltzmann-weighted Thompson Sampling. We provide an intuitive justification for why this approach balances exploration (via uncertainty-driven barrier crossing) and exploitation (via mean-driven settling) more effectively for sampling tasks than deterministic BO acquisition functions. Third, we incorporate *Automatic Relevance Determination* (ARD) via Matérn-5/2 kernels to handle the highly anisotropic nature of atomic interactions, where energy sensitivity varies drastically across spatial dimensions. Empirically, we evaluate our method on the benchmark Lennard-Jones 13 (LJ13) system ($d = 39$). Using the Wasserstein-2 distance on physical invariants, we demonstrate that our method achieves decent mode coverage and distributional fidelity compared to MCMC, Progressive Inference-Time Annealing (PITA) [Akhound-Sadegh et al., 2025], and BGflow [Noé et al., 2019] baselines.

2 Background and Related Work

Formally, for the problem of sampling from equilibrium distributions, we seek to generate samples from the Boltzmann distribution $\pi(\mathbf{x}) \propto \exp(-U(\mathbf{x})/k_B T)$, where $U : \mathbb{R}^d \rightarrow \mathbb{R}$ is the potential energy. In this section, we review the theoretical foundations of Gaussian Processes (GP), contrast our active sampling approach with Bayesian Optimization (BO), and detail the baseline methodologies used for evaluation.

2.1 Gaussian Processes: Theoretical Foundations

Our methodology relies on Gaussian Processes (GPs) as a probabilistic surrogate model for the underlying potential energy surface $U(\mathbf{x})$. We adhere to the standard definitions and theorems provided by Rasmussen and Williams [2006].

Definition 1 (Gaussian Process). *A Gaussian Process is a collection of random variables, any finite number of which have a joint multivariate Gaussian distribution. A GP is completely specified by its mean function $m(\mathbf{x})$ and covariance function $k(\mathbf{x}, \mathbf{x}')$:*

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (1)$$

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})] \quad (2)$$

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))]. \quad (3)$$

In the context of regression, we assume observed energies y_i at configurations \mathbf{x}_i are corrupted by additive Gaussian noise $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$, such that $y_i = f(\mathbf{x}_i) + \epsilon$.

Remark 2 (RKHS and Regularity). *Associated with the covariance kernel k is a Reproducing Kernel Hilbert Space (RKHS), denoted $\mathcal{H}_k(\mathcal{X})$, which is the closure of the linear span of functions $\{k(\cdot, \mathbf{x}) : \mathbf{x} \in \mathcal{X}\}$ equipped with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}_k}$ satisfying the reproducing property $\langle f, k(\cdot, \mathbf{x}) \rangle_{\mathcal{H}_k} = f(\mathbf{x})$. For the Matérn family of kernels used in this work, the RKHS norm $\|f\|_{\mathcal{H}_k}$ serves as a measure of the smoothness of the potential energy surface. Specifically, assuming*

$U(\mathbf{x}) \in \mathcal{H}_k(\mathcal{X})$ with $\|U\|_{\mathcal{H}_k} \leq B$, we can bound the approximation error. The epistemic uncertainty $\sigma_n(\mathbf{x})$ corresponds to the worst-case error in this function class:

$$\sigma_n(\mathbf{x}) = \sup_{f \in \mathcal{H}_k, \|f\|_{\mathcal{H}_k} \leq 1} |f(\mathbf{x}) - \mu_n(\mathbf{x})|. \quad (4)$$

This geometric interpretation ensures that our surrogate model is well-posed for Lipschitz-continuous energy landscapes typical in molecular dynamics.

Definition 3 (Posterior Inference). Given a dataset of observations $\mathcal{D} = \{(\mathbf{X}, \mathbf{y})\}$, where $\mathbf{X} \in \mathbb{R}^{n \times d}$ is the design matrix and $\mathbf{y} \in \mathbb{R}^n$ is the vector of target energies, the joint distribution of the observed targets and the function values \mathbf{f}_* at test locations \mathbf{X}_* is given by

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N} \left(\mathbf{0}, \begin{bmatrix} K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I & K(\mathbf{X}, \mathbf{X}_*) \\ K(\mathbf{X}_*, \mathbf{X}) & K(\mathbf{X}_*, \mathbf{X}_*) \end{bmatrix} \right) \quad (5)$$

where $K(\cdot, \cdot)$ denotes the Gram matrix computed via the kernel function.

Definition 4 (Conditional Distribution). The posterior distribution $p(\mathbf{f}_* | \mathbf{X}, \mathbf{y}, \mathbf{X}_*)$ is a multivariate Gaussian $\mathcal{N}(\boldsymbol{\mu}_*, \boldsymbol{\Sigma}_*)$ with moments

$$\boldsymbol{\mu}_* = K(\mathbf{X}_*, \mathbf{X}) [K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I]^{-1} \mathbf{y} \quad (6)$$

$$\boldsymbol{\Sigma}_* = K(\mathbf{X}_*, \mathbf{X}_*) - K(\mathbf{X}_*, \mathbf{X}) [K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 I]^{-1} K(\mathbf{X}, \mathbf{X}_*). \quad (7)$$

The computational bottleneck lies in the inversion of the matrix $(K + \sigma_n^2 I)$, which incurs a complexity of $\mathcal{O}(n^3)$. This cubic scaling motivates our choice of iterative active learning, where sample efficiency is paramount over raw throughput.

2.2 Baselines and State-of-the-Art Sampling Methods

To rigorously evaluate our GP-based methodology, we compare against three distinct classes of algorithms: classical MCMC, flow-based generative models, and annealing strategies.

Markov Chain Monte Carlo (MCMC) MCMC remains the gold standard for asymptotically exact sampling. We employ the Metropolis-Hastings algorithm [Metropolis et al., 1953, Hastings, 1970], which constructs a Markov chain via a proposal distribution $q(\mathbf{x}' | \mathbf{x})$. A proposed move is accepted with probability

$$A(\mathbf{x} \rightarrow \mathbf{x}') = \min \left\{ 1, \frac{\pi(\mathbf{x}') q(\mathbf{x} | \mathbf{x}')}{\pi(\mathbf{x}) q(\mathbf{x}' | \mathbf{x})} \right\} = \min \left\{ 1, \exp \left(-\frac{U(\mathbf{x}') - U(\mathbf{x})}{k_B T} \right) \right\}. \quad (8)$$

While MCMC guarantees convergence to $\pi(\mathbf{x})$ as $t \rightarrow \infty$, it suffers severely from *mode trapping* in rough energy landscapes. In systems like Lennard-Jones clusters, high energy barriers separate metastable states. This results in mixing times that exceed feasible computational horizons.

Boltzmann Generators (BGflow) Boltzmann Generators [Noé et al., 2019] represent the state-of-the-art in deep generative sampling. They use Normalizing Flows to learn a bijective mapping $F_\phi : \mathcal{Z} \rightarrow \mathcal{X}$ between a simple latent distribution $\mu(\mathbf{z})$ (e.g., isotropic Gaussian) and the complex target distribution. The change of variables formula gives the density of the generated samples

$$p_X(\mathbf{x}) = \mu(\mathbf{z}) |\det \mathbf{J}_{F^{-1}}(\mathbf{x})| \quad (9)$$

where $\mathbf{z} = F_\phi^{-1}(\mathbf{x})$ and \mathbf{J} is the Jacobian. BGflow is trained by minimizing the combination of the reverse KL divergence and a maximum likelihood term

$$\mathcal{L}_{BG} = \mathbb{E}_{\mathbf{z} \sim \mu} \left[\frac{U(F_\phi(\mathbf{z}))}{k_B T} - \log |\det \mathbf{J}_{F_\phi}(\mathbf{z})| \right] - \mathbb{E}_{\mathbf{x} \sim \mathcal{D}_{\text{data}}} [\log p_X(\mathbf{x})]. \quad (10)$$

While powerful, BGflow requires a significant volume of training data to learn the transformation topology and often struggles with broken flows in regions of high curvature.

Physics-Informed Temperature Annealing (PITA) As a baseline for heuristic sampling improvements, we consider the PITA method [Akhound-Sadegh et al. \[2025\]](#). PITA integrates physical constraints into a simulated annealing framework. The core principle involves exploring the landscape at an artificially high temperature T_{high} to cross energetic barriers, followed by a cooling schedule to settle into local minima:

$$T(t) = T_{\text{high}} \cdot \alpha^t, \quad 0 < \alpha < 1. \quad (11)$$

PITA augments this with gradient-based Langevin dynamics. However, annealing methods are fundamentally one-way processes; once the system cools into a basin, it cannot easily escape if it is the wrong basin, leading to kinetic trapping similar to MCMC.

2.3 Bayesian Optimization vs. Active Sampling

Our work sits at the intersection of Bayesian Optimization (BO) and Active Learning (AL). It is crucial to distinguish our objective from standard BO.

Standard Bayesian Optimization The goal of traditional BO is global optimization: finding the global minimizer $\mathbf{x}^* = \arg \min_{\mathbf{x}} U(\mathbf{x})$. Standard acquisition functions like Expected Improvement (EI) [\[Jones et al., 1998\]](#) are designed to greedily reduce the regret $R_T := U(\mathbf{x}^*) - \min_{t \leq T} U(\mathbf{x}_t)$. EI selects points that are expected to improve upon the best value observed so far. It evaluates a candidate \mathbf{x} by the expected decrease in objective value relative to the current minimum as

$$\alpha_{EI}(\mathbf{x}) = \mathbb{E}_{p(y|\mathbf{x}, \mathcal{D})} [\max(0, f_{\min} - y)]. \quad (12)$$

While highly effective for locating a single optimum, acquisition functions such as EI tend to focus sampling increasingly near the current best solution, leading to a collapse in sampling diversity [\[Jain et al., 2023\]](#). In settings concerned with sampling rather than optimization, such as molecular thermodynamics, identifying only the global minimum is insufficient; instead, one must characterize the full ensemble of thermally accessible states and their relative probabilities, which are governed by entropy as well as energy.

Active Sampling for Distribution Learning Our approach differs fundamentally in that we treat the sampling process as a distribution matching problem. Rather than regret minimization, we aim to minimize the divergence between the surrogate-induced Boltzmann distribution $\hat{\pi}$ and the true distribution π . Recent work by [Zhang et al. \[2024\]](#) highlights the connection between active learning and GFlowNets, proposing acquisition functions that target *epistemic uncertainty* not to minimize f , but to clarify the integral of the partition function Z . By using a Thompson Sampling-based acquisition (as detailed in Section 3), we align our strategy with the theoretical framework of maximum-entropy reinforcement learning

$$\mathbf{x}_{\text{new}} \sim \text{Softmax} \left(-\frac{\hat{U}(\mathbf{x})}{T} \right), \quad (13)$$

where \hat{U} is a realization from the GP posterior. This implicitly balances exploration (via the variance in \hat{U}) and exploitation (via the mean) in a manner that asymptotically recovers the Boltzmann distribution, unlike EI which asymptotically focuses solely on the global minimum.

2.4 Connections to Classical Importance Sampling

Classical importance sampling [\[Burda et al., 2016\]](#) estimates expectations $\mathbb{E}_{\pi}[O(\mathbf{x})]$ using samples from a proposal distribution $q(\mathbf{x})$:

$$\mathbb{E}_{\pi}[O(\mathbf{x})] \approx \frac{1}{N} \sum_{i=1}^N O(\mathbf{x}_i) \frac{\pi(\mathbf{x}_i)}{q(\mathbf{x}_i)}, \quad \mathbf{x}_i \sim q(\mathbf{x}). \quad (14)$$

The variance of this estimator is minimized when $q(\mathbf{x}) \propto |O(\mathbf{x})|\pi(\mathbf{x})$. In our active GP framework, the GP posterior serves as an adaptive proposal distribution $q_t(\mathbf{x})$. Unlike static importance sampling, our $q_t(\mathbf{x})$ improves over time. As the GP variance $\sigma_n^2(\mathbf{x})$ decreases in relevant regions, the mismatch between the surrogate Boltzmann distribution and the true target diminishes, reducing the variance of thermodynamic averages computed from the generated ensemble.

3 Methodology

Our proposed framework integrates Bayesian non-parametric modeling with uncertainty-guided active sampling to efficiently explore high-dimensional energy landscapes. We formulate the problem of sampling as a sequential decision-making process where a GP surrogate model approximates the potential energy surface. In this section, we detail the theoretical underpinnings of our approach, including the specific kernel definitions, the derivation of the posterior predictive distribution, and our novel candidate-based acquisition strategy rooted in Boltzmann-weighted Thompson Sampling.

3.1 Problem Formulation: The Energy Landscape

We consider a physical system described by a state vector $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^d$, where d represents the dimensionality of the configuration space. For a system of N_p particles in 3D space, $d = 3N_p$. The system is governed by a potential energy function $U : \mathcal{X} \rightarrow \mathbb{R}$. Our objective is to generate samples from the equilibrium Boltzmann distribution $\pi(\mathbf{x})$ at a thermodynamic temperature T :

$$\pi(\mathbf{x}) = \frac{1}{Z} \exp\left(-\frac{U(\mathbf{x})}{k_B T}\right), \quad \text{where} \quad Z = \int_{\mathcal{X}} \exp\left(-\frac{U(\mathbf{x}')}{k_B T}\right) d\mathbf{x}' \quad (15)$$

is the intractable partition function and k_B is the Boltzmann constant. In our experiments, for example, we specifically target the Lennard-Jones (LJ) cluster systems, where the potential energy $U_{LJ}(\mathbf{r})$ for a configuration of particles $\mathbf{r} = \{\mathbf{r}_1, \dots, \mathbf{r}_{N_p}\}$ is defined by pairwise interactions:

$$U_{LJ}(\mathbf{r}) = 4\epsilon \sum_{i=1}^{N_p} \sum_{j>i}^{N_p} \left[\left(\frac{\sigma_{rm}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{rm}}{r_{ij}}\right)^6 \right] + \frac{1}{2} k_{osc} \sum_{i=1}^{N_p} \|\mathbf{r}_i\|^2. \quad (16)$$

Here, $r_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|_2$ is the Euclidean distance between particles i and j , ϵ is the depth of the potential well, and σ_{rm} is the finite distance at which the inter-particle potential is zero. We impose a harmonic trap (oscillator) with stiffness k_{osc} to ensure the compactness of the cluster during sampling. The fundamental challenge lies in the evaluation cost of $U(\mathbf{x})$ and the extreme multimodality of the landscape, which renders standard MCMC methods inefficient due to mode collapse.

3.2 Gaussian Process Surrogate Modeling

To minimize the number of expensive evaluations of the true potential $U(\mathbf{x})$, we construct a probabilistic surrogate model using Gaussian Processes.

Anisotropic Matern Covariance Structure Given the potential roughness of the energy landscape and the need to model distinct lengthscales across dimensions (though isotropic approximations are often used for rotational invariance), we employ the Matérn class of kernels. Specifically, we use the Matérn-5/2 kernel [Matérn, 1960], which provides a balance between the smoothness of the Radial Basis Function (RBF) and the non-differentiability of the Exponential kernel. The covariance between two states \mathbf{x} and \mathbf{x}' is defined as

$$k_{\nu=5/2}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \left(1 + \sqrt{5} d_M(\mathbf{x}, \mathbf{x}') + \frac{5}{3} d_M^2(\mathbf{x}, \mathbf{x}') \right) \exp\left(-\sqrt{5} d_M(\mathbf{x}, \mathbf{x}')\right) \quad (17)$$

where σ_f^2 is the signal variance, and d_M is the weighted Euclidean distance induced by the automatic relevance determination (ARD) lengthscales $\ell = [\ell_1, \dots, \ell_d]^\top$:

$$d_M(\mathbf{x}, \mathbf{x}') = \sqrt{(\mathbf{x} - \mathbf{x}')^\top \mathbf{\Lambda}^{-2} (\mathbf{x} - \mathbf{x}')}, \quad \mathbf{\Lambda} = \text{diag}(\ell_1, \dots, \ell_d). \quad (18)$$

This choice of kernel allows the model to learn that potential energy changes more rapidly along certain interaction axes (e.g., when two particles are in close proximity) compared to others. The hyperparameters $\theta = \{\sigma_f, \ell_1, \dots, \ell_d\}$ are optimized by maximizing the Exact Marginal Log-Likelihood (MLL) of the observed data $\mathcal{D}_n = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$, where $y_i = U(\mathbf{x}_i)$.

Posterior Inference and Computational Complexity Assuming a homoscedastic Gaussian noise model where observations $y = f(\mathbf{x}) + \epsilon$ with $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$, the posterior distribution at a test point

\mathbf{x}_* is analytically tractable. Let K be the $n \times n$ covariance matrix with entries $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$, $\mathbf{k}_* = [k(\mathbf{x}_1, \mathbf{x}_*), \dots, k(\mathbf{x}_n, \mathbf{x}_*)]^\top$, and $k_{**} = k(\mathbf{x}_*, \mathbf{x}_*)$. The predictive posterior is Gaussian $\mathcal{N}(\mu_n(\mathbf{x}_*), \sigma_n^2(\mathbf{x}_*))$ with moments given by Definition 4 as

$$\mu_n(\mathbf{x}_*) = m(\mathbf{x}_*) + \mathbf{k}_*^\top (K + \sigma_n^2 I)^{-1} (\mathbf{y} - m(X)) \quad (19)$$

$$\sigma_n^2(\mathbf{x}_*) = k_{**} - \mathbf{k}_*^\top (K + \sigma_n^2 I)^{-1} \mathbf{k}_* \quad (20)$$

Hyperparameter Optimization via Type-II Maximum Likelihood The performance of the GP surrogate is governed by the kernel hyperparameters $\boldsymbol{\theta} = \{\sigma_f^2, \ell_1, \dots, \ell_d, \sigma_n^2\}$. While manual tuning or grid search is feasible in low dimensions, it becomes intractable for the high-dimensional energy landscapes considered here ($d = 39$ for LJ-13). Furthermore, static hyperparameters fail to adapt to the varying local curvature of the potential energy surface as the active learner explores new basins of attraction. Instead, we employ a data-driven approach known as *Type-II Maximum Likelihood Estimation* (or Empirical Bayes). We select $\boldsymbol{\theta}$ by maximizing the marginal likelihood (evidence), which is obtained by integrating out the latent function values \mathbf{f} as follows,

$$p(\mathbf{y}|\mathbf{X}, \boldsymbol{\theta}) = \int p(\mathbf{y}|\mathbf{f}, \sigma_n^2) p(\mathbf{f}|\mathbf{X}, \boldsymbol{\theta}) d\mathbf{f}. \quad (21)$$

Since both the likelihood $p(\mathbf{y}|\mathbf{f})$ and the prior $p(\mathbf{f}|\mathbf{X})$ are Gaussian, this integral is analytically tractable. The resulting marginal log-likelihood (MLL) is then given by

$$\mathcal{L}(\boldsymbol{\theta}) := \log p(\mathbf{y}|\mathbf{X}, \boldsymbol{\theta}) = \underbrace{-\frac{1}{2} \mathbf{y}^\top \mathbf{K}_\theta^{-1} \mathbf{y}}_{\mathcal{T}_{\text{fit}}} - \underbrace{\frac{1}{2} \log |\mathbf{K}_\theta|}_{\mathcal{T}_{\text{complexity}}} - \frac{N}{2} \log 2\pi, \quad (22)$$

where $\mathbf{K}_\theta = K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}$. This objective function embodies a rigorous trade-off, often referred to as the Bayesian Occam’s Razor [MacKay, 2003], which automatically regulates model complexity without the need for a separate validation set. Note that (22) is composed of two competing terms that justify the superiority of this approach over manual selection: The \mathcal{T}_{fit} term represents the Mahalanobis distance between the observations and the prior mean. Maximizing this term encourages the kernel to adjust lengthscales such that the data is *expected* under the prior. A model with extremely short lengthscales that approximates white noise can minimize reconstruction error perfectly, but fails to generalize. The $\mathcal{T}_{\text{complexity}}$ term penalizes the volume of the function space accessible to the model. A model with short lengthscales is highly flexible and can fit any dataset, resulting in a large determinant $|\mathbf{K}_\theta|$ and a correspondingly heavy penalty. Conversely, a model with very long lengthscales is rigid (low complexity) but will yield a poor data fit \mathcal{T}_{fit} .

In our high-dimensional setting ($d = 39$), manual hyperparameter tuning is impossible due to the combinatorial explosion of the search space. More critically, uniform manual lengthscales assume isotropy, which is physically invalid for molecular systems where degrees of freedom (e.g., bond vibrations vs. van der Waals interactions) exhibit vastly different energy gradients. By maximizing $\mathcal{L}(\boldsymbol{\theta})$ via gradients $\nabla_{\boldsymbol{\theta}} \mathcal{L}$, our method performs Automatic Relevance Determination (ARD) [Wipf and Nagarajan, 2007]. Dimensions k that are irrelevant to the energy changes will automatically be assigned large lengthscales ($\ell_k \rightarrow \infty$), effectively removing them from the inference. This allows the GP to dynamically focus its limited data budget on the lower-dimensional manifold where the relevant physics occurs, a property that static manual tuning cannot achieve.

3.3 Active Sampling via Posterior-Guided Boltzmann Exploration

Standard Bayesian Optimization (BO) [Xu et al., 2024, Hvarfner et al., 2024] seeks to maximize a function using acquisition functions like Expected Improvement (EI). However, our goal for this project is not optimization, but *sampling* from the Boltzmann distribution defined by the latent function. We propose an active learning strategy that utilizes the epistemic uncertainty of the GP to drive exploration while adhering to the statistical mechanics of the target system.

Our proposed Active Boltzmann Sampling framework operates as a closed-loop sequential decision process. The core mechanism integrates a Bayesian non-parametric surrogate with a stochastic acquisition strategy designed to explore the equilibrium distribution of the target potential $U(\mathbf{x})$. The procedure is formalized in Algorithm 1 and detailed step-by-step below.

Data Preprocessing and Numerical Stability GP inference is notoriously sensitive to the scaling of input features and target values, particularly when modeling high-energy physics potentials where $U(\mathbf{x})$ can span tens of orders of magnitude. To ensure the convergence of the L-BFGS-B optimizer and the stability of Cholesky decompositions, we employ a strict double-precision standardization pipeline at each iteration t . Let $\mathcal{D}_t = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N_t}$ be the dataset at iteration t . We define the domain mappings $\Psi : \mathcal{X} \rightarrow [0, 1]^d$ and $\Phi_t : \mathbb{R} \rightarrow \mathbb{R}$ as follows:

$$\tilde{\mathbf{x}}_i = \Psi(\mathbf{x}_i) = (\mathbf{x}_i - \mathbf{b}_{\text{low}}) \oslash (\mathbf{b}_{\text{high}} - \mathbf{b}_{\text{low}}), \quad \tilde{y}_{i,t} = \Phi_t(y_i) = \frac{y_i - \mu_{\mathcal{D}_t}}{\sigma_{\mathcal{D}_t} + \epsilon} \quad (23)$$

where \oslash denotes element-wise division, and $\mu_{\mathcal{D}_t}, \sigma_{\mathcal{D}_t}$ are the running empirical mean and standard deviation of the observed energies. The kernel hyperparameters are optimized in this latent standardized space $\mathcal{Z} = [0, 1]^d \times \mathbb{R}$, which effectively preconditions the optimization landscape.

Candidate Generation Due to the non-convexity of the acquisition surface in high dimensions, continuous optimization of acquisition functions is prone to local optima. Instead, we employ a discrete candidate-based approach. At iteration t , we generate a set of candidate points $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_M\}$ drawn uniformly from the domain bounds with $\mathbf{c}_j \sim \mathcal{U}(\mathbf{b}_{\text{low}}, \mathbf{b}_{\text{high}})$, $j = 1, \dots, M$, where M is the candidate pool size.

Thompson Sampling on the Energy Surface To balance exploration (visiting high-uncertainty regions) and exploitation (visiting low-energy regions), we employ a strategy analogous to Thompson Sampling [Russo et al., 2018]. Rather than using the deterministic mean $\mu_n(\mathbf{x})$ as a surrogate for energy, we draw a realization from the posterior predictive distribution for each candidate. Let the random variable $E(\mathbf{x})$ represent the probabilistic belief of the energy at state \mathbf{x} . We define the *stochastic energy surrogate* $\tilde{U}_t(\mathbf{x})$ as $\tilde{U}_t(\mathbf{x}) \sim \mathcal{N}(\mu_n(\mathbf{x}), \beta^2 \sigma_n^2(\mathbf{x}))$, where β is a scaling factor controlling the aggressiveness of uncertainty incorporation. In our posterior sampling mode, we perform the reparameterization trick,

$$\tilde{U}_t(\mathbf{c}_j) = \mu_n(\mathbf{c}_j) + \epsilon_j \cdot \sigma_n(\mathbf{c}_j), \quad \epsilon_j \sim \mathcal{N}(0, 1). \quad (24)$$

In regions where the GP is uncertain (high σ_n), the sampled energy \tilde{u}_t has a higher probability of being significantly lower than the mean prediction. Since the Boltzmann probability decays exponentially with energy, these low energy samples that are driven by high variance become highly probable candidates for selection. This naturally incentivizes the algorithm to investigate unknown regions, as uncertainty is implicitly converted into potential thermodynamic probability.

Probabilistic Selection We define the selection probability for the next sample \mathbf{x}_{t+1} over the discrete candidate set \mathcal{C} using the Boltzmann weights derived from the stochastic surrogate. The probability mass function $P(\mathbf{c}_j)$ is given by

$$P(\mathbf{c}_j | \mathcal{D}_t) = \frac{\exp(-\tilde{U}_t(\mathbf{c}_j)/T)}{\sum_{k=1}^M \exp(-\tilde{U}_t(\mathbf{c}_k)/T)}, \quad (25)$$

where T is the sampling temperature. Note that this T matches the physical temperature of the target distribution. The next query point is then sampled categorically according to $\mathbf{x}_{t+1} \sim \text{Categorical}(\mathcal{C}; \mathbf{P})$. To ensure numerical stability when computing these probabilities, we use the Log-Sum-Exp trick. Let $L_j = -\tilde{U}_t(\mathbf{c}_j)/T$. The log-probability is computed as

$$\log P(\mathbf{c}_j) = L_j - \left(\max_k (L_k) + \log \sum_{k=1}^M \exp(L_k - \max_m (L_m)) \right). \quad (26)$$

This prevents floating-point overflow for low-energy configurations and underflow for high-energy barriers. Our approach naturally balances exploration and exploitation. Exploitation occurs because low μ_n lowers \tilde{U} . Exploration occurs because high σ_n increases the variance of \tilde{U} , allowing it to occasionally effectively tunnel through high-energy barriers in the surrogate landscape.

Algorithm 1 Active Boltzmann Sampling via Gaussian Processes

Require: Domain $\Omega \subset \mathbb{R}^d$ (bounds $[\mathbf{l}, \mathbf{u}]$), oracle potential energy $U : \Omega \rightarrow \mathbb{R}$.
Require: Temperature T , horizon T_{\max} , batch size M (candidates), initial size N_{init} .
Require: Kernel prior (Matérn-5/2 with ARD, $k_{\nu=2.5}$).
 1: Sample $\mathbf{X}_0 \sim \mathcal{U}(\Omega)^{N_{\text{init}}}$ and valuate $\mathbf{y}_0 = U(\mathbf{X}_0)$. Set dataset $\mathcal{D}_0 \leftarrow (\mathbf{X}_0, \mathbf{y}_0)$.
 2: **for** $t = 0$ to $T_{\max} - 1$ **do**
 3: // DATA STANDARDIZATION
 4: Calculate moments: $\mu_y \leftarrow \text{mean}(\mathbf{y}_t)$, $\sigma_y \leftarrow \text{std}(\mathbf{y}_t) + \epsilon$.
 5: Normalize inputs $\tilde{\mathbf{X}} \leftarrow (\mathbf{X}_t - \mathbf{l}) \oslash (\mathbf{u} - \mathbf{l})$ and standardize targets $\tilde{\mathbf{y}} \leftarrow (\mathbf{y}_t - \mu_y)/\sigma_y$.
 6: // GP FITTING
 7: Construct GP via $f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k_\theta(\mathbf{x}, \mathbf{x}'))$.
 8: Optimize θ^* by maximizing MLL $\theta^* = \text{argmax}_\theta \log p(\tilde{\mathbf{y}}|\tilde{\mathbf{X}}, \theta)$.
 9: // CANDIDATE GENERATION
 10: Sample candidate pool $\mathbf{C} = \{\mathbf{c}_j\}_{j=1}^M \sim \mathcal{U}(\Omega)$ and normalize: $\tilde{\mathbf{C}} \leftarrow (\mathbf{C} - \mathbf{l}) \oslash (\mathbf{u} - \mathbf{l})$.
 11: // POSTERIOR PREDICTION
 12: Compute posterior moments for standardized candidates $\tilde{\mathbf{c}}_j$: $\hat{\mu}_j, \hat{\sigma}_j^2 \leftarrow \text{Predict}(\tilde{\mathbf{c}}_j|\tilde{\mathbf{X}}, \tilde{\mathbf{y}}, \theta^*)$.
 13: Transform back to physical scale: $\mu_j \leftarrow \hat{\mu}_j \cdot \sigma_y + \mu_y$, $\sigma_j \leftarrow \sqrt{\hat{\sigma}_j^2} \cdot \sigma_y$.
 14: // STOCHASTIC ENERGY SURROGATE
 15: Draw Thompson samples from the posterior: $\tilde{U}_j = \mu_j + \xi_j \cdot \sigma_j$, $\xi_j \sim \mathcal{N}(0, 1)$.
 16: // BOLTZMAN SELECTION
 17: Compute logits $\ell_j = -\tilde{U}_j/T$ and probabilities via Softmax: $P(\mathbf{c}_j) = \frac{\exp(\ell_j)}{\sum_{k=1}^M \exp(\ell_k)}$.
 18: Sample index $k^* \sim \text{Categorical}(P(\mathbf{c}_1), \dots, P(\mathbf{c}_M))$.
 19: Update next query point $\mathbf{x}_{\text{new}} \leftarrow \mathbf{c}_{k^*}$.
 20: Observe true energy and update history: $y_{\text{new}} \leftarrow U(\mathbf{x}_{\text{new}})$, $\mathcal{D}_{t+1} \leftarrow \mathcal{D}_t \cup \{(\mathbf{x}_{\text{new}}, y_{\text{new}})\}$.
 21: **end for**

Measure-theoretic Posterior Sampling Strategy Let λ be the Lebesgue measure on Ω . We aim to construct a sequence of sampling measures $\{\nu_t\}_{t=1}^T$ that converges to the target Boltzmann measure Π , where $d\Pi(\mathbf{x}) = Z^{-1} e^{-U(\mathbf{x})/k_B T} d\lambda(\mathbf{x})$. In the absence of the true potential U , we construct a *random measure* $\tilde{\Pi}_t$ induced by the GP posterior process $\mathcal{GP}(\mu_t, k_t)$. The acquisition policy is defined via the Radon-Nikodym derivative of the stochastic surrogate measure with respect to the base uniform measure

$$\frac{d\tilde{\Pi}_t}{d\lambda}(\mathbf{x}) \propto \exp\left(-\frac{\tilde{U}_t(\mathbf{x})}{k_B T}\right), \quad \tilde{U}_t(\cdot) \sim \mathcal{GP}(\mu_t(\cdot), k_t(\cdot, \cdot)). \quad (27)$$

The candidate-based selection is effectively a Monte Carlo approximation of this stochastic integral. By drawing \mathbf{x}_{t+1} proportional to this density, we maximize the expected reduction in the Kullback-Leibler divergence between the surrogate posterior and the true equilibrium distribution, implicitly solving the equivalent of (25):

$$\mathbf{x}_{t+1} \sim \mathbb{E}_{\tilde{U}_t} \left[\frac{\exp(-\tilde{U}_t(\mathbf{x})/T)}{\int_{\mathcal{X}} \exp(-\tilde{U}_t(\mathbf{x}')/T) d\mathbf{x}'} \right]. \quad (28)$$

3.4 Alternative Sampling Modes

While our primary method relies on the posterior sampling described above, our framework supports deterministic modifications to analyze the impact of uncertainty.

Greedy Sampling We set $\sigma_n(\mathbf{x}) \equiv 0$, effectively using $\tilde{U}(\mathbf{x}) = \mu_n(\mathbf{x})$. This mode relies purely on the mean prediction, typically resulting in getting trapped in the nearest local minimum.

Pessimistic Exploration (LCB-style) We define the surrogate energy as the Lower Confidence Bound of the energy (which corresponds to an Upper Confidence Bound on the Boltzmann probability)

$$\tilde{U}_{LCB}(\mathbf{x}) = \mu_n(\mathbf{x}) - \beta \sigma_n(\mathbf{x}). \quad (29)$$

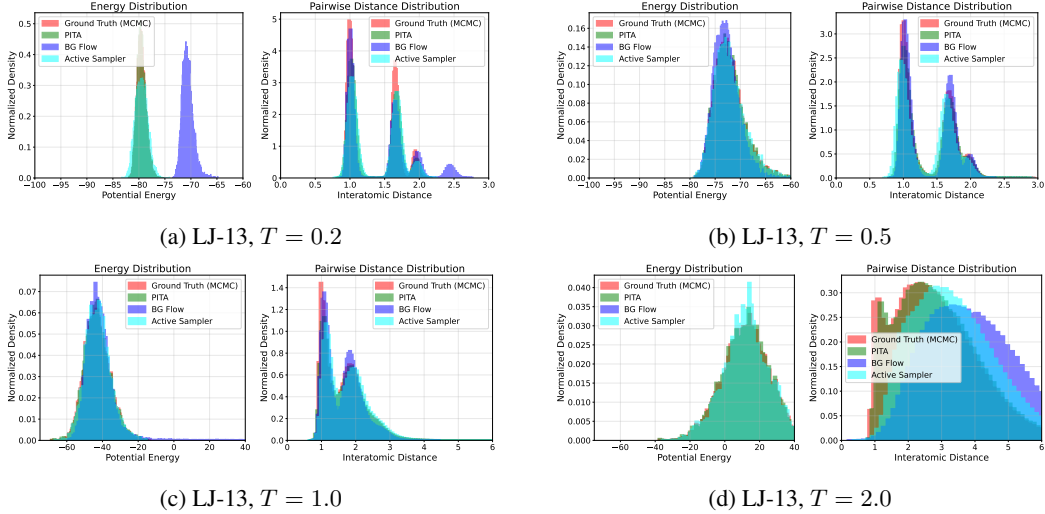


Figure 1: LJ-13 sampling task overlaid plots with $T \in \{0.2, 0.5, 1.0, 2.0\}$ and no annealing. We compare the distribution of the interatomic distances and energy of the particles in the MCMC dataset (ground-truth, in red), best run samples generated using a PITA model (in green), a BGflow model (in blue), and our active sampler (in cyan).

Unlike the stochastic posterior sampling, this is a deterministic acquisition function. It aggressively favors high-uncertainty regions regardless of the mean prediction, often leading to over-exploration of high-energy non-physical states.

Our empirical results demonstrate that the stochastic posterior sampling mode strikes the optimal balance required for mixing between metastable states in the LJ-13 cluster.

3.5 Evaluation Metrics and Ground Truth Comparison

To validate the fidelity of the generated ensemble, we employ ground truth samples generated via exhaustive MCMC or Parallel Tempering. We compare the learned distribution against the ground truth using the 2-Wasserstein distance and visual inspection of marginalized distributions. For the LJ-13 system, direct density estimation in 39D is infeasible. Therefore, we evaluate physical invariants, specifically the distribution of pairwise inter-atomic distances given by

$$D = \{\|\mathbf{r}_i - \mathbf{r}_j\|_2 | 1 \leq i < j \leq N_p\} \quad (30)$$

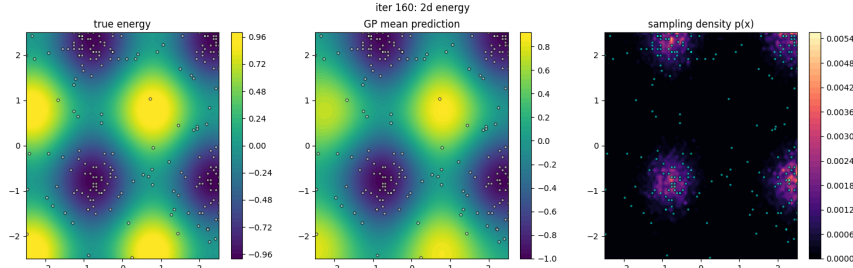
and the scalar potential energy distribution $P(E)$. The overlap between the histograms of these invariants generated by our GP active learner and the ground truth MCMC data serves as our primary metric for success.

4 Experiments

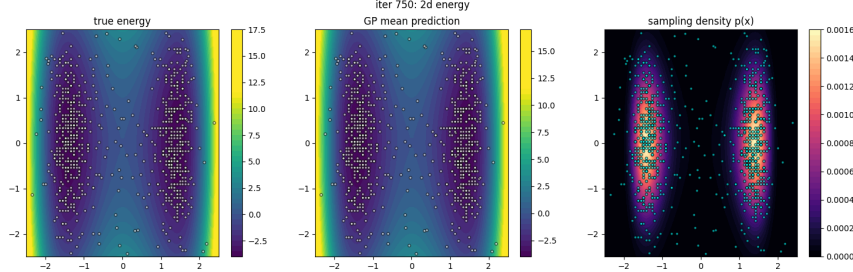
Evaluation method. We evaluate our active sampler on a toy sine-cosine energy, the double well energy, and on molecular conformation sampling tasks such as the Lennard-Jones system of 13 particles (LJ-13). For LJ-13 specifically, the target distribution is defined by the Lennard-Jones potential augmented with a harmonic oscillator term, following Akhound-Sadegh et al. [2025]. For metrics, we use sample-based metrics such as the interatomic 2-Wasserstein distance $\mathcal{D}\text{-}\mathcal{W}_2$ and the potential energy 2-Wasserstein distance $\mathcal{E}\text{-}\mathcal{W}_2$ between the distributions induced by the ground-truth and generated samples, to assess mode coverage and precision, respectively (see Appendix A). It provides a geometrically meaningful quantification of the “work” required to transport the probability mass from the generated ensemble to the ground truth. As an alternative, one could also observe the changes of the reverse KL divergence between the target and sampling densities.

Algorithm	Distance- $\mathcal{W}_2 \downarrow$	Energy- $\mathcal{W}_2 \downarrow$	#Energy Calls \downarrow
PITA	0.0751 ± 0.0298	3.2765 ± 1.5251	5,000,000
BGflow	0.1513 ± 0.0472	16.4310 ± 11.1438	5,000,000
ABS	0.1087 ± 0.1695	13.8924 ± 9.0782	5,000

Table 1: LJ-13 sampling task with $T = 1.0$ and no annealing. Metrics are calculated using only samples at the end of training, with respect to the MCMC ground truth (i.e., with Distance- $\mathcal{W}_2 = \text{Energy-}\mathcal{W}_2 = 0$) over 10k samples and standard deviations averaged over 3 seeds.

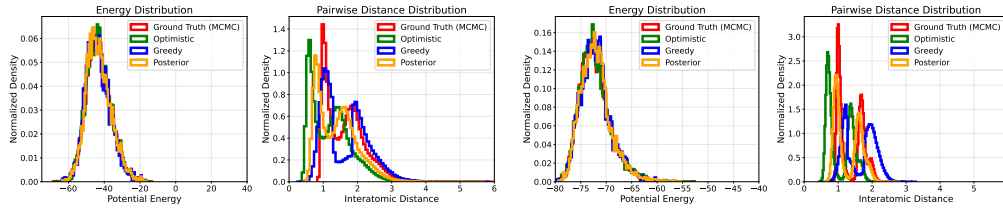


(a) 2D sine-cosine densities.



(b) Double well densities.

Figure 2: Sampling densities matching in the energy and probability spaces at convergence, $T = 0.3$.



(a) LJ-13, different sampling modes at $T = 1.0$

(b) LJ-13, different sampling modes at $T = 0.5$

Figure 3: Effect of different sampling modes on ABS.

Baselines. We compare with three different baselines: BGflow [Noé et al., 2019], PITA [Akhound-Sadeh et al., 2025], and an MCMC chain as the ground truth. We assume access to a short MCMC chain run at high temperature throughout the PITA experiments.

Architecture. In training PITA, we follow the design choices proposed by Akhound-Sadeh et al. [2025] and use EGNN [Satorras et al., 2021] as the model backbone for LJ-13. Our training follows a sequential temperature schedule, proceeding from high to low temperatures. After training at a given temperature for a fixed number of epochs, we generate samples at the next lower temperature and continue training at that temperature. For LJ-13, we train a single diffusion model conditioning it on β and using the data across all previously seen temperatures.

For the BGflow baseline, we train a Boltzmann Generator [Noé et al., 2019] by minimizing the forward KL divergence between the flow-induced distribution and the Boltzmann distribution. The model consists of a RealNVP-style architecture with six coupling blocks, with each block using an affine coupling transformation parameterized by multilayer perceptrons and operates on a split of the 39D particle coordinate space. A standard multivariate Gaussian is used as the prior. Samples from the prior are transformed through the flow to generate molecular configurations, and the log-determinant of the Jacobian is tracked explicitly.

Our active sampler is based on a GP surrogate trained on previously evaluated energies and updated sequentially. At each iteration, we fit a GP with an ARD Matérn kernel following Xu et al. [2024] to the accumulated data and construct a sampling distribution proportional to $\exp(-\tilde{E}(x)/T)$, where $\tilde{E}(x)$ is derived from the GP posterior mean and uncertainty according to the chosen sampling mode. In low dimensions (1D and 2D), we evaluate the GP posterior on a dense uniform grid over the domain and sample the next query point directly from this discretized distribution. In higher dimensions, where grid-based evaluation becomes intractable, including LJ-13 (39D), we instead draw a fixed set of uniformly random candidate points and sample the next query from the induced categorical distribution over these candidates. This allows the method to scale to high-dimensional molecular systems while preserving uncertainty-aware exploration.

Further training details and hyperparameters are provided in Appendix B.

5 Results and Discussions

Effect of kernel regularity and candidate density. The inductive bias of the Gaussian Process is primarily governed by the smoothness class of the covariance kernel. Our ablation studies indicate that the Matérn-5/2 kernel ($\nu = 2.5$) yields superior sampling fidelity compared to the Radial Basis Function (RBF) and the Matérn-3/2 kernel ($\nu = 1.5$). The RBF kernel implies infinite differentiability (C^∞), a strong smoothness assumption that leads to over-regularization of the sharp, stiff repulsive walls characteristic of Lennard-Jones potentials. In contrast, the Matérn-5/2 kernel enforces only C^2 continuity, which better aligns physically with the existence of continuous inter-atomic forces while allowing for the rapid local variations necessary to distinguish adjacent metastable minima.

For the Matérn-5/2 kernel in d dimensions specifically, its information gain is measured by the quantity $\gamma_T = \mathcal{O}(T^{\frac{d(d+1)}{2\nu+d(d+1)}} \log T)$. Since our energy $\mathbf{x}_{t+1} \sim \exp(-\tilde{U}_t(\mathbf{x})/\tau)$ has full support over Ω (due to the unboundedness of the Gaussian noise in the Thompson sample), the sequence of observation locations $\{\mathbf{x}_t\}$ is dense in Ω almost surely as $T \rightarrow \infty$. Consequently, the posterior variance $\sigma_T^2(\mathbf{x}) \rightarrow 0$ pointwise. This ensures that the surrogate Boltzmann distribution $\hat{\pi}_T$ converges to the true distribution π in the weak topology of measures.

Moreover, for LJ-13, the size of the candidate pool M acts as a discretization parameter for the continuous acquisition integral. The candidate-based Boltzmann weighting can be viewed as an importance sampling estimator of the partition function

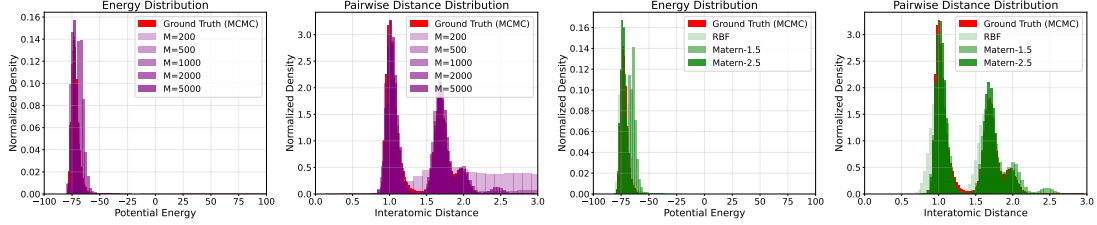
$$Z_T \approx \left(\frac{\text{Vol}(\Omega)}{M} \right) \sum_{j=1}^M \exp \left(\frac{-\tilde{U}(\mathbf{c}_j)}{T} \right). \quad (31)$$

We observe in Figure 4a that low candidate density ($M < 10^3$) results in high variance estimators of Z_T , causing the sampler to miss narrow basins of attraction (mode collapse). Increasing M provides asymptotic consistency, ensuring that the discrete categorical distribution converges to the true posterior-induced Boltzmann distribution, albeit at a linear computational cost during the posterior prediction step.

Empirically, we observe in Figure 4b diminishing returns beyond a moderate candidate budget ($M = 2000$), suggesting that kernel expressivity dominates performance once the candidate set sufficiently covers the domain. Together, these results highlight a bias-variance trade-off in kernel smoothness and a sampling-resolution trade-off governed by the candidate pool size.

For the 1D toy sine-cosine problem, the candidate pool size M acts as the discretization resolution for the continuous Boltzmann policy. With sparse candidate sets ($M < 50$), the importance sampling estimator of the partition function exhibits high variance, introducing significant stochasticity that prevents the sampler from resolving the precise location of energy minima. Increasing M ensures

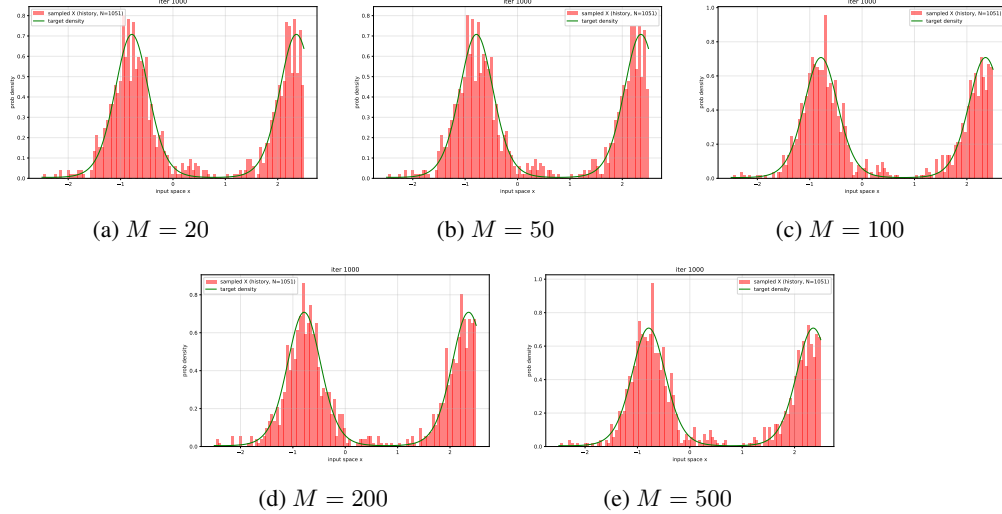
asymptotic consistency, again allowing the discrete categorical distribution to rapidly converge to the true continuous Boltzmann distribution. Unlike the high-dimensional case, the coverage of the 1D domain saturates quickly; we observe diminishing returns beyond a small budget ($M = 100$), where the discretization error becomes negligible compared to the surrogate’s epistemic uncertainty.



(a) LJ-13 with varying candidates, Matérn-2.5 kernel

(b) LJ-13 with different kernels, 5,000 candidates

Figure 4: LJ-13 sampling task overlaid plots with $T = 0.8$ fixed and no annealing, tested over different number of candidates $M \in \{200, 500, 1000, 2000, 5000\}$ and three kernel types $K \in \{k_{\text{rbf}}, k_{\nu=1.5}, k_{\nu=2.5}\}$. We compare the distribution of the interatomic distances and energy of the particles in the MCMC dataset (ground-truth, in red) and samples generated using our active sampler.



(a) $M = 20$

(b) $M = 50$

(c) $M = 100$

(d) $M = 200$

(e) $M = 500$

Figure 5: 1D sine-cosine sampling task with $T = 0.3$ fixed and no annealing, tested over different number of candidates $M \in \{20, 50, 100, 200, 500\}$ at iteration $1e3$. We compare the sampling distribution using our active sampler and the oracle target energy.

Effect of different sampling modes. Mathematically, the optimistic mode defined as $E_{\text{acq}}(x) = \mu(x) + \beta\sigma(x)$ corresponds to the Upper Confidence Bound (UCB) in the context of energy minimization. This is a risk-averse policy that penalizes epistemic uncertainty, effectively treating unexplored regions as high-energy states.

The superior performance of the optimistic mode over the greedy mode arises from the *safety against smoothing* effect. In high-dimensional physical systems like LJ-13, the valid low-energy configuration space is a thin manifold surrounded by high-energy singularities (particle clashes). GP with standard kernels tend to over-smooth these sharp barriers in regions where data is sparse (high σ). Consequently, the greedy sampler, following only the smooth mean μ , erroneously assigns significant probability mass to nonphysical transition regions, leading to poor distributional fidelity as seen in Figures 3a and 3b.

By contrast, the optimistic mode adds a variance penalty $+\beta\sigma$. This effectively reconstructs the energy barriers: it assumes that any region with high uncertainty is likely a high-energy wall. This prior belief aligns perfectly with the physics of Lennard-Jones potentials, where most of the

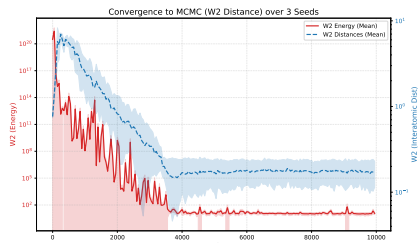


Figure 6: LJ-13, change of \mathcal{W}_2 distances at $T = 1.0$

Figure 7: Convergence of energy and interatomic distance \mathcal{W}_2 between ABS and MCMC baseline. This would only make sense if the sampling algorithm is online.

(T_{\max}, B)	(1250, 8)	(625, 16)	(312, 32)	(156, 64)	(78, 128)
BGflow Energy- \mathcal{W}_2	5.13×10^{17}	1.80×10^5	7.59×10^4	3.41×10^5	2.98×10^7

Table 2: LJ-13 sampling task with $T = 1.0$ and no annealing. The BGflow energy \mathcal{W}_2 distance is calculated with respect to the MCMC ground truth over 10k energy calls.

configuration space is indeed invalid. This constrains the sampler to the trust regions. The posterior mode, where $E_{\text{acq}}(x) \sim \mathcal{N}(\mu(x), \sigma^2(x))$, effectively implements randomized probability matching. It dynamically balances the exploration-exploitation tradeoff without the static bias of fixed- β strategies. The performance similarity between optimistic and posterior modes confirms that robustly accounting for epistemic uncertainty, either deterministically (UCB) or stochastically (Thompson Sampling), is strictly necessary to resolve the disjoint metastable states of the LJ13 cluster, whereas pure mean-based exploitation (greedy) fails to respect the landscape’s boundaries.

Sample efficiency and fair comparison with deep generative baselines. A critical advantage of our active sampling framework is its sample efficiency compared to state-of-the-art deep generative models. Recent benchmarks on generative sampling for molecular conformations, such as those utilizing score-based diffusion models or continuous normalizing flows [Midgley et al., 2023], indicate that these methods typically require training datasets in the order of 10^5 to 10^7 samples to accurately estimate the score function $\nabla_{\mathbf{x}} \log p(\mathbf{x})$ or the transport map [He et al., 2025]. This “data hunger” arises from the need to constrain the vast parameter space of the neural networks (which often contain millions of weights in the coupling layers or denoising networks) to solve the high-dimensional density matching problem. To illustrate this disparity, we evaluated the performance of the BGflow baseline restricted to the same computational budget as our active sampler (fix $N = 10,000$ energy evaluations while varying training steps and batch sizes such that their product is N ; see Table 2). Under this constraint, BGflow failed to converge to a valid equilibrium distribution and exhibited severe mode collapse, where the generated samples clustered into a single, non-physical basin, exhibiting the same mode collapse artifacts highlighted by He et al. [2025] in simulation-free training regimes. The expressivity of the neural density estimator is a liability in the low-data regime; with only 10,000 points, the energy-based KL-divergence loss is insufficient to regularize the flow’s Jacobian. This leads to overfitting on the few observed modes.

In contrast, our GP-based approach leverages the strong inductive bias of the Matérn kernel. By modeling the potential energy surface non-parametrically, the active sampler does not need to *learn how to sample* from scratch; instead, it efficiently learns the boundaries of the low-energy manifold and achieves high distributional fidelity with orders of magnitude fewer calls to the oracle potential.

6 Future Work and Directions

Our current results open several avenues for hybridizing active learning with amortized inference. A promising direction is to use the converged GP surrogate, \mathcal{M}_{GP} , as a differentiable, cheap-to-evaluate proxy for the true energy function to train highthroughput samplers like BGflow or PITA. In this scheme, the active sampler would first expend a fixed budget (e.g., $N = 10,000$ calls) to explore the

landscape and construct a high-fidelity surrogate. Subsequently, a flow model could be trained offline by minimizing the reverse KL divergence with respect to the *surrogate* energy $\tilde{U}_{\text{GP}}(\mathbf{x})$, thereby generating unlimited samples without incurring additional computational cost from the expensive ground-truth oracle. Rigorous evaluation of this method would require strictly matching the number of true energy evaluations-ensuring that the baselines do not implicitly access the oracle during their training phase beyond the initial data collection. It would be relevant to report the computational expense of all methods by controlling the total number of energy evaluation calls. It may also be interesting to compare with Fully Bayesian GP and/or Fully Bayesian Single Task GP’s fit.

Additionally, our current candidate generation strategy, $\mathbf{c}_j \sim \mathcal{U}(\Omega)$, relies on uniform sampling, which suffers from the curse of dimensionality as the volume of the low-energy manifold vanishes in \mathbb{R}^{39} . This limitation is rooted in the *concentration of measure* phenomenon. As $d \rightarrow \infty$, the volume of the ϵ -shell at the boundary of the hypercube dominates the total volume. Consequently, uniform candidates $\mathbf{c} \sim \mathcal{U}(\Omega)$ increasingly fail to probe the internal basins of attraction where low-energy states reside, leading to *candidate starvation* in the landscape’s interior. Future work should investigate adaptive importance sampling distributions, $q_t(\mathbf{c})$, that bias candidate generation towards the trust regions identified by the GP. One can potentially use normalizing flows to propose candidates rather than fully model the target.

Finally, the theoretical connection between our sampling modes and multi-armed bandit algorithms warrants deeper exploration. While our optimistic mode $(\mu + \beta\sigma)$ parallels the UCB algorithm used in regret minimization, the objective of sampling is distribution matching rather than optimization. Future work could formulate the choice of sampling hyperparameters as a sequential decision process, potentially leveraging GFlowNet [Bengio et al., 2021, Jain et al., 2023, Lahlou et al., 2023] objectives to optimally balance the exploration of high-uncertainty regions with the exploitation of known modes to ensure the empirical distribution converges asymptotically to the true Boltzmann measure.

Acknowledgments and Disclosure of Funding

This report was written as part of a COMP 400 Honours Project in Computer Science in Fall 2025. The author is indebted to Dr. Mohammad Pedramfar and Prof. Siamak Ravanbakhsh for their time, patience, and extensive guidance.

References

- Tara Akhound-Sadegh, Jarrod Rector-Brooks, Avishek Joey Bose, Sarthak Mittal, Pablo Lemos, Cheng-Hao Liu, Marcin Sendera, Siamak Ravanbakhsh, Gauthier Gidel, Yoshua Bengio, et al. Iterated denoising energy matching for sampling from boltzmann densities. *arXiv preprint arXiv:2402.06121*, 2024.
- Tara Akhound-Sadegh, Jungyoon Lee, Avishek Joey Bose, Valentin De Bortoli, Arnaud Doucet, Michael M Bronstein, Dominique Beaini, Siamak Ravanbakhsh, Kirill Neklyudov, and Alexander Tong. Progressive inference-time annealing of diffusion models for sampling from boltzmann densities. *arXiv preprint arXiv:2506.16471*, 2025.
- Yoshua Bengio, Emmanuel Bengio, Joseph Cloutier, Xavier Groleau, Marc-Alexandre Côté, and Pascal Vincent. Flow network based generative models for non-iterative diverse candidate generation. In *Advances in Neural Information Processing Systems*, 2021.
- Yuri Burda, Roger Grosse, and Ruslan Salakhutdinov. Importance weighted autoencoders. In *Proceedings of the 33rd International Conference on Machine Learning*, pages 1510–1519, 2016.
- W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- Jiajun He, Yuanqi Du, Francisco Vargas, Dinghuai Zhang, Shreyas Padhy, RuiKang OuYang, Carla Gomes, and José Miguel Hernández-Lobato. No trick, no treat: Pursuits and challenges towards simulation-free training of neural samplers. *arXiv preprint arXiv:2502.06685*, 2025.
- Carl Hvarfner, Erik Orm Hellsten, and Luigi Nardi. Vanilla bayesian optimization performs great in high dimensions. *arXiv preprint arXiv:2402.02229*, 2024.
- Moksh Jain, Yoshua Bengio, and Alex Hernandez-Garcia. Gflownets for ai-driven scientific discovery. In *Digital Discovery*, 2023. Available at <https://openreview.net/pdf?id=kX8h23UG6v>.
- Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, 1998.
- Salem Lahlou, Moksh Jain, Tristan Deleu, and Yoshua Bengio. A theory of generative flow networks. In *Proceedings of the 40th International Conference on Machine Learning*, 2023.
- David J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, Cambridge, UK, 2003. ISBN 978-0521642989.
- Bertil Matérn. *Spatial Variation: Stochastic Models and Their Application to Some Problems in Forest Surveys and Other Sampling Investigations*. Statens Skogsforskningsinstitut, 1960.
- N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- Laurence I Midgley, Vincent Stimper, Gregor NC Simidjievski, Bernhard Schölkopf, and José Miguel Hernández-Lobato. Flow annealed importance sampling bootstrap. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=X8suS1k5_N.
- Frank Noé, Simon Olsson, Jonas Köhler, and Hao Wu. Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. *Science*, 365(6457):eaaw1147, 2019.
- Carl Edward Rasmussen and Christopher KI Williams. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA, 2006.
- Daniel J Russo, Benjamin Van Roy, Abbas Kazerouni, Ian Osband, Zheng Wen, et al. A tutorial on thompson sampling. *Foundations and Trends® in Machine Learning*, 11(1):1–96, 2018.
- Victor Garcia Satorras, Emiel Hoogetboom, and Max Welling. $e(n)$ equivariant graph neural networks. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 9323–9332. PMLR, 2021. URL <https://proceedings.mlr.press/v139/satorras21a.html>.

- Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. Practical bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems*, volume 25, 2012.
- Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the 27th International Conference on Machine Learning*, 2010.
- David P. Wipf and Srikantan S. Nagarajan. A new view of automatic relevance determination. In John C. Platt, Daphne Koller, Yoram Singer, and Sam T. Roweis, editors, *Advances in Neural Information Processing Systems 20*, 2007.
- Zhitong Xu, Haitao Wang, Jeff M Phillips, and Shandian Zhe. Standard gaussian process is all you need for high-dimensional bayesian optimization. *arXiv preprint arXiv:2402.02746*, 2024.
- Cheng Zhang, Thang Bui, and Yingzhen Li. Bayesian active learning for sampling. *arXiv preprint arXiv:2402.02229*, 2024.

Appendix

A Quantitative Evaluation: The Wasserstein-2 Distance

Measure-theoretic formulation. Let $\mathcal{P}_2(\Omega)$ denote the space of probability measures on the metric space (Ω, d) with finite second moments. Formally, for two measures $\mu, \nu \in \mathcal{P}_2(\Omega)$, the squared 2-Wasserstein distance is defined as the solution to the Kantorovich optimal transport problem:

$$\mathcal{W}_2^2(\mu, \nu) := \inf_{\gamma \in \Pi(\mu, \nu)} \int_{\Omega \times \Omega} d(x, y)^2 d\gamma(x, y), \quad (32)$$

where $\Pi(\mu, \nu)$ is the set of all joint probability measures (couplings) on $\Omega \times \Omega$ having marginals μ and ν , respectively. That is, for all measurable sets $A \subseteq \Omega$, $\gamma(A \times \Omega) = \mu(A)$ and $\gamma(\Omega \times A) = \nu(A)$.

Empirical estimation for physical invariants. Direct computation of \mathcal{W}_2 in the high-dimensional configuration space \mathbb{R}^{3N} is computationally intractable due to the curse of dimensionality. Therefore, we evaluate the fidelity of the sampled ensemble by projecting onto physically meaningful 1D invariants: the scalar potential energy $E(\mathbf{x})$ and the set of pairwise interatomic distances $\{r_{ij}(\mathbf{x})\}$.

For one-dimensional distributions, the optimal transport problem admits a closed-form analytical solution. Let μ and ν be the continuous distributions of a 1D invariant. The distance is given by the L_2 distance between their inverse cumulative distribution functions F_μ^{-1} and F_ν^{-1} :

$$\mathcal{W}_2^2(\mu, \nu) = \int_0^1 |F_\mu^{-1}(z) - F_\nu^{-1}(z)|^2 dz. \quad (33)$$

In the discrete empirical setting, consider a set of generated samples $X = \{x_i\}_{i=1}^N$ and ground truth samples $Y = \{y_j\}_{j=1}^M$. We denote their sorted order statistics as $x_{(1)} \leq \dots \leq x_{(N)}$ and $y_{(1)} \leq \dots \leq y_{(M)}$. If $N = M$, the empirical Wasserstein distance is exactly the Euclidean distance between the sorted vectors $\hat{\mathcal{W}}_2^2(X, Y) = \frac{1}{N} \sum_{i=1}^N |x_{(i)} - y_{(i)}|^2$. In the general case where $N \neq M$, we estimate the integral via linear interpolation of the quantiles. We define the empirical quantile functions $\hat{F}_X^{-1}(z)$ and $\hat{F}_Y^{-1}(z)$ and approximate the integral numerically over the uniform grid $z \in [0, 1]$. This formulation provides a robust, parameter-free metric for assessing whether our active sampling algorithm has correctly captured both the mode locations (metastable states) and the relative probability mass (entropy) of the target Boltzmann distribution.

B Further Training Details and Hyperparameters

In this section, we provide detailed specifications for the training of our Active Boltzmann Sampler (ABS) and the baseline models (PITA and BGflow) to ensure reproducibility. All experiments were conducted using PyTorch with CUDA acceleration.

B.1 Active Boltzmann Sampler (Ours)

The Active Sampler leverages the BoTorch and GPyTorch libraries. To ensure numerical stability given the steep gradients of the Lennard-Jones potential, all Gaussian Process operations are performed in double precision (float64).

Gaussian Process Surrogate We use a Single Task Gaussian Process `SingleTaskGP` with an Exact Marginal Log Likelihood (MLL) objective. We employ a Scale Kernel wrapping a Matérn-5/2 kernel ($\nu = 2.5$). Crucially, we enable Automatic Relevance Determination (ARD) by setting `ard_num_dims=39`, allowing the model to learn separate lengthscales for each dimension of the configuration space. We place a Gamma prior on the lengthscales to encourage regularization. Hyperparameters are optimized at every iteration by maximizing the MLL using the L-BFGS-B optimizer via `botorch.fit.fit_gpytorch_mll`.

Data Standardization To facilitate kernel learning, inputs and outputs are rigorously normalized at each step. The inputs \mathbf{x} are linearly scaled from the domain bounds $[-2.5, 2.5]$ to the unit hypercube $[0, 1]^d$. Energy values y are standardized to zero mean and unit variance based on the current observed pool statistics: $\tilde{y} = (y - \mu_y)/\sigma_y$.

Acquisition and Sampling For high-dimensional experiments (LJ-13), we generate a candidate pool with $N_{\text{cand}} = 5,000$ uniform random candidates at each step. The process begins with $N_{\text{init}} = 50$ samples drawn from a Sobol sequence or uniform distribution to initialize the GP. The total budget is set to $T_{\text{max}} = 5,000$ oracle calls for the comparisons in Table 1.

B.2 Baseline: Physics-Informed Temperature Annealing (PITA)

We use the official implementation and configuration provided by Akhoun-Sadegh et al. [2025] at <https://github.com/taraak/pita>. The model is trained using a sequential temperature annealing schedule to facilitate mode mixing. Its backbone is an $E(n)$ Equivariant Graph Neural Network (EGNN) conditioned on the current temperature. This ensures the model respects the rotational and translational symmetries inherent to the molecular system. The model is trained sequentially from high to low temperatures, assuming access to an MCMC chain at high temperature at all times. The specific schedule used for LJ-13 is temperatures $T \in \{4.0, 3.0, 2.0, 1.5, 1.2, 1.0\}$, epochs per temperature $\in \{150, 200, 150, 200, 300\}$, and batch size 512. The total loss is a weighted sum of score matching and energy constraints $\mathcal{L} = w_1 \mathcal{L}_{\text{score}} + w_2 \mathcal{L}_{\text{energy_score}} + w_3 \mathcal{L}_{\text{energy_match}}$, where we set $w_1 = w_2 = w_3 = 1.0$. We use the elucidating noise schedule with $\sigma_{\text{min}} = 0.05$.

B.3 Baseline: Boltzmann Generators (BGflow)

We directly implement the BGflow baseline using the bgflow library at https://github.com/al2wang/bayesian/blob/main/bgflow_train.py. The model is trained via energy-based training by minimizing the reverse KL divergence: $\mathcal{L} = \mathbb{E}_{z \sim p_z} [U(T(z))/(k_B T) - \log |\det J_T(z)|]$. For the BGflow architecture, we employ a RealNVP-based normalizing flow consisting of 6 affine coupling blocks. For Coupling Layers, each block uses a SplitFlow (splitting dimensions into 19 and 20), a SwapFlow, and a CouplingFlow. For Transformers, the affine transformations (shift and scale) are parameterized by dense neural networks (DenseNet) with two hidden layers of 128 units and ReLU activations. The sampling prior is a standard isotropic Gaussian $\mathcal{N}(0, \mathbf{I})$ in 39 dimensions.

We use Adam with a learning rate of 1×10^{-3} and gradient clipping and a batch size of 128. For the sample-efficiency comparison, we limit the training to match the evaluation budget of the active sampler. For the converged baseline, we train for 10,000 steps.

C Extended Analysis and Failure Modes

C.1 Wall-Clock Time vs. Sample Efficiency Trade-off

While our method minimizes the number of oracle calls to the potential energy function, it incurs a growing computational cost due to the cubic scaling of the Cholesky decomposition required for exact GP inference. For the LJ-13 system using a fast analytical potential (milliseconds per call), the wall-clock time of training the GP on $N = 10,000$ points (approximately 6 hours on an RTX8000) exceeds that of executing 5×10^6 PITA energy evaluation calls. However, our framework is designed for the regime where $U(\mathbf{x})$ takes minutes or hours to evaluate (e.g., density functional theory). In such regimes, the computational overhead of the GP is amortized, and the sample efficiency becomes the dominant factor instead.

C.2 The Curse of Dimensionality in Candidate Generation

A key limitation observed in the 39D experiments is the diminishing coverage of the candidate pool. The effective density of $M = 5,000$ candidates in \mathbb{R}^{39} is exponentially small. The probability that a uniform candidate falls into a specific basin of attraction with volume ratio $V_{\text{basin}}/V_{\Omega} \approx 10^{-9}$ is negligible. Our method relies on the GP *expanding* the basins via the lengthscales ℓ . However, if the initial random initialization does not land near a basin, the gradients of the GP mean are zero (reverting to the prior mean), and the acquisition function becomes a pure uncertainty search (equivalent to random search). This necessitates the use of N_{init} samples from a space-filling design rather than pure random sampling to guarantee at least one sample lands in a informative region.