STATE-ADAPTIVE MULTI-SCALE SEARCH FOR SCALABLE BAYESIAN OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Bayesian optimization (BO) is attractive for expensive black-box problems, but exact Gaussian-process (GP) inference scales as $O(n^3)$ in the number of observations and high-dimensional search introduces a separate statistical and optimization challenge. These pressures are acute in robotics and other costly tuning settings: reusable hyperparameter histories may initially be small because each trial is expensive, yet a sequential campaign can rapidly accumulate hundreds or thousands of observations. We introduce **Adaptive Multi-Scale Dirichlet-Guided Variational Gaussian Process (AMS-DVGP)**, a scalable sparse-GP BO method that combines variational inference with state-adaptive spatial allocation. Beyond the usual incumbent-centered exploitation and uncertainty-driven exploration signals, a Dirichlet-process mixture summarizes promising occupied modes and provides an additional data-driven search center. Across six synthetic benchmarks spanning 2–500 dimensions and 500–5000 total evaluations, AMS-DVGP achieves the best mean endpoint result among the three directly compared methods on every task, reducing final regret relative to adaptive-depth FocalBO by 48.2% on average. It also obtains the best final objective on 11-D HalfCheetah and the 32-D FrankaBell manipulation task. Algorithm analysis shows that the controller uses all four search actions and that its high-dimensional gains persist throughout the optimization trajectory rather than appearing only at the endpoint. Recorded wall-clock cost across the eight retained tasks is about 60.4% lower than adaptive-depth FocalBO, while an acquisition study shows robust performance under TS, UCB, and EI.

1 INTRODUCTION

Bayesian optimization (BO) addresses expensive black-box problems of the form $x^\star \in \arg\max_{x\in\mathcal{X}} f(x)$, where evaluating f may require simulation, physical experimentation, or training a costly model (Jones et al., 1998; Snoek et al., 2012; Shahriari et al., 2016). BO alternates between a probabilistic surrogate and an acquisition rule that selects new evaluations. Gaussian processes (GPs) are especially useful because they provide predictions and calibrated uncertainty, but exact GP inference requires $O(n^2)$ memory and $O(n^3)$ computation in the number of observations n (Rasmussen and Williams, 2006). Sparse variational GPs (SVGPs) replace the expensive dependence on all observations with a smaller inducing set and stochastic variational training (Titsias, 2009; Hensman et al., 2013).

The word *scalable* in BO is used for several related bottlenecks. For GP computation, the principal scaling variable is the observation count n; FocalBO explicitly targets large offline and online datasets with sparse GPs (Wei et al., 2024). High input dimension d is a distinct difficulty because the search space and surrogate estimation become harder even when n is moderate, motivating trust-region, embedding, sparse-subspace, and adaptive-subspace methods (Eriksson et al., 2019; Wang et al., 2016; Eriksson and Jankowiak, 2021; Papenmeier et al., 2022). We therefore use *scalable BO* operationally across both axes while keeping their roles separate: sparse variational inference addresses observation-scale computation, whereas our 300-D and 500-D experiments test whether the spatial search mechanism remains effective as dimension grows.

This distinction matters in costly applications. Robot and embodied-AI tuning campaigns can begin without a large reusable hyperparameter archive because real or high-fidelity trials are time-consuming and expensive. Nevertheless, sequential optimization can quickly build a large observation history, so

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

the method must remain computationally tractable while extracting as much value as possible from each new batch. In other words, scalable BO needs both *computational efficiency* as n grows and *optimization efficiency* in how finite evaluations are allocated.

FocalBO (Wei et al., 2024), our principal scalable sparse-GP baseline, tackles this problem by focalizing variational training and hierarchically optimizing acquisition functions over progressively smaller incumbent-centered regions. Its adaptive-depth mechanism controls how deeply to focus. We take a complementary step: rather than adapting only the depth of an incumbent-centered hierarchy, AMS-DVGP adapts the *search action itself*. The controller decides whether to search around the incumbent, probe multiple spatial scales, contract around another promising visited mode, or escape globally. This introduces a third spatial signal beyond the canonical exploitation/exploration cues of incumbent quality and predictive uncertainty: a Dirichlet-process mixture summarizes occupied modes in the visited space and supplies alternative data-driven centers.

We propose **Adaptive Multi-Scale Dirichlet-Guided Variational Gaussian Process (AMS-DVGP)**. A progress detector selects among information-guided regional search, multi-scale probing, cluster contraction, and global escape; a DP-GMM proposes promising occupied modes; and controller-aware likelihood weighting emphasizes uncertain observations near currently active regions. The acquisition function remains modular, and the same regional controller supports TS, UCB, and EI.

Our evaluation deliberately spans different scalability regimes. We include two low-dimensional synthetic benchmarks, Bohachevsky 2-D and standard Ackley 5-D; four larger-scale synthetic benchmarks, Griewank 100-D, Ackley 300-D on the wider $[-32.768, 32.768]^{300}$ domain, Rastrigin 500-D with 3000 total evaluations, and Rosenbrock 500-D with a 2000-point initial design and 5000 total evaluations; and two embodied-control tasks, 11-D HalfCheetah and a custom 32-D FrankaBell manipulation benchmark. Under matched budgets and matched initial designs, AMS-DVGP achieves the best mean endpoint result on all eight retained tasks relative to adaptive-depth FocalBO and SVGP.

Contributions.

1. We introduce a **state-adaptive multi-scale search policy** for scalable sparse-GP BO that dynamically controls center, scale, locality, and batch allocation rather than only focal depth.
2. We add **Dirichlet-guided occupied-mode proposals** as an additional spatial signal beyond incumbent-centered exploitation and uncertainty-guided exploration, and couple the active regions to variational likelihood weighting.
3. We provide **broad scalable-BO evidence across observation count and dimension**: 2–500 dimensions, 500–5000 total evaluations, six synthetic tasks plus two embodied-control tasks, acquisition robustness, recorded wall-clock cost, and an algorithm analysis showing both internal strategy switching and lower trajectory-averaged regret than adaptive-depth FocalBO.

2 Related Work

2.1 Bayesian optimization and sparse variational GPs

Given observations $\mathcal{D}_t = \{(x_i, y_i)\}_{i=1}^t$ with $y_i = f(x_i) + \epsilon_i$ and $\epsilon_i \sim \mathcal{N}(0, \sigma_n^2)$, a GP prior yields predictive mean $\mu_t(x)$ and variance $\sigma_t^2(x)$, and BO selects new queries through an acquisition function $a_t(x)$. Expected improvement (EI), probability of improvement (PI), upper confidence bound (UCB), and Thompson sampling (TS) are standard examples (Kushner, 1964; Thompson, 1933; Srinivas et al., 2010; Vakili et al., 2021).

For large or sequential datasets, an SVGP introduces inducing variables $u = f(Z)$ at $m \ll n$ inducing inputs Z and uses $q(u) = \mathcal{N}(m_u, S_u)$ with $q(f, u) = p(f \mid u)q(u)$. Its standard evidence lower bound is

$$\mathcal{L}_{\mathrm{SVGP}} = \sum_{i=1}^{t} \mathbb{E}_{q(f_i)}[\log p(y_i \mid f_i)] - \mathrm{KL}[q(u) \| p(u)]. \tag{1}$$

With fixed m, sparse variational training avoids the exact GP's cubic dependence on all n observations and supports stochastic mini-batch optimization (Titsias, 2009; Hensman et al., 2013; Maddox et al., 2021). This is the computational foundation of the large-observation scalability considered here.

2.2 SCALABLE AND HIGH-DIMENSIONAL BO

Large n and large d create different bottlenecks. Sparse-GP methods target the computational cost of growing observation histories (Titsias, 2009; Hensman et al., 2013; Vakili et al., 2021), while high-dimensional BO methods constrain or learn where search should occur, including TuRBO, random embeddings, sparse axis-aligned priors, and adaptive subspaces (Eriksson et al., 2019; Wang et al., 2016; Eriksson and Jankowiak, 2021; Papenmeier et al., 2022). FocalBO combines sparse-GP inference with focalized local modeling and demonstrates scalable BO with large offline/online datasets and high-dimensional control (Wei et al., 2024). Recent work explores still other regimes, including gradient-informed Bayesian neural surrogates, computation-aware GPs, and tabular foundation-model BO (Makrygiorgos et al., 2025; Sinaga et al., 2025; Yu et al., 2026). Because these methods address different assumptions and surrogate families, our empirical claims are scoped to the directly compared sparse-GP baselines rather than universal dominance over all scalable BO paradigms.

2.3 WHAT FOCALBO ADAPTS, AND WHAT WE ADAPT

FocalBO (Wei et al., 2024) allocates sparse-GP capacity through focalized training and a hierarchy of incumbent-centered search regions; its central adaptive quantity is the focal optimization depth. AMS-DVGP adapts a broader search action: whether to search locally or globally, which center to use, what scale to probe, and how to allocate the batch. Incumbent value and predictive uncertainty are canonical BO signals for exploitation and exploration; AMS-DVGP adds a Dirichlet-guided occupied-mode center so search budget can also be directed toward promising visited structure away from the current best point. The main comparison fixes TS across methods, while Figure 5 tests AMS-DVGP itself with TS, UCB, and EI.

3 METHOD

3.1 METHOD OVERVIEW

Each BO iteration first summarizes recent optimization progress, then uses the DP-GMM state to propose promising visited regions. The controller constructs one or more candidate regions, allocates the batch across those regions, applies the chosen acquisition rule within the resulting candidate domains, evaluates the selected points, and finally updates both the sparse GP surrogate and the maintained clustering/controller state. Thus, acquisition selection and candidate-domain adaptation are separated: TS is used for the main baseline comparison, but EI and UCB can be dispatched through the same regional controller.

Algorithm 1 AMS-DVGP STEP

In $\mathcal{D}_t, q_t, \mathcal{A}, B, \mathcal{M}_t, \Theta$
$x_t^\star \leftarrow \arg\max_i y_i,\ s_t \leftarrow \mathrm{Regime}(\mathcal{D}_t; \Theta)$
$c_t^\star \leftarrow \mathrm{BestCluster}(\mathcal{M}_t),\ u_t^\star \leftarrow \arg\max_x \sigma_t^2(x)$
if $\mathrm{Escape}(\mathcal{M}_t; \Theta)$: $X^+ \leftarrow \mathrm{Global}(q_t, B; \Theta)$
else if s_t = fast and $\mathrm{ClusterGate}(c_t^\star; \Theta)$: $R_t \leftarrow \mathrm{Contract}(c_t^\star; \Theta)$
else if s_t = stagnation and $a_{t-1} \neq$ MS: $R_t \leftarrow \mathrm{MultiScale}(x_t^\star; \Theta)$
else: $R_t \leftarrow \mathrm{InfoGuide}(x_t^\star, u_t^\star, c_t^\star; \Theta)$
$(R_t, B_t) \leftarrow \mathrm{Allocate}(R_t, B; \Theta)$
$X^+ \leftarrow \mathrm{RegionalAcq}(\mathcal{A}, q_t, R_t, B_t)$, $Y^+ \leftarrow f(X^+)$
$\mathcal{D}_{t+1} \leftarrow \mathcal{D}_t \cup (X^+, Y^+)$
$\mathcal{M}_{t+1} \leftarrow \mathrm{Refresh}(\mathcal{M}_t, \mathcal{D}_t, X^+; \Theta)$
$q_{t+1} \leftarrow \mathrm{FitAMS}(\mathcal{D}_{t+1}, \mathcal{M}_{t+1}; \Theta)$
Out $\mathcal{D}_{t+1}, q_{t+1}, \mathcal{M}_{t+1}$

3.2 PROGRESS-REGIME SUMMARY

Let $b_t = \max_{i \le t} y_i$ be the incumbent best objective value after t evaluations. The implementation monitors recent incumbent improvements and maps them into coarse progress regimes such as `fast`, `steady`, `slow`, and `stagnation`. These labels do *not* each invoke a separate bespoke operator. Instead, the executed controller distinguishes three practical cases: (1) under **fast progress**, it may contract around a high-potential visited cluster; (2) after **repeated stagnation**, it invokes multi-scale probing or periodic global escape; and (3) in **all other cases**, it uses information-guided regional search. This is intentionally simpler, and more faithful to the current code path, than claiming a fully distinct strategy for every named state.

3.3 CONTROLLER ACTIONS

The controller exposes four region-generation operators to the acquisition rule: **(1) information-guided search**, combining the incumbent, a promising visited-cluster center, and a posterior-uncertainty maximum; **(2) multi-scale search**, distributing the batch across fine, medium, and coarse neighborhoods according to a budget vector $\boldsymbol{\beta} = (\beta_f, \beta_m, \beta_c)$ with $\sum_j \beta_j = 1$; **(3) cluster contraction**, refining around the best-ranked cluster when its potential exceeds a threshold τ_c; and **(4) global escape**, mixing domain-wide uniform and uncertainty-guided candidates with exploration fraction η_g after a stagnation patience h_{stag} is reached. The progress regime selects the operator, while the acquisition function ranks points within the resulting candidate domains when applicable. These quantities are implementation hyperparameters: the method description uses symbols, while the executed values are reported in Appendix A. We do not claim that the particular numerical settings are theoretically optimal; Algorithm 1 collects the fixed controller and training settings in Θ.

3.4 DIRICHLET-GUIDED PROPOSALS FROM VISITED CLUSTERS

To summarize the geometry of visited inputs $X_t = \{x_i\}_{i=1}^t$, we fit a truncated Dirichlet-process Gaussian mixture, following the classical DP mixture construction and its variational approximations (Ferguson, 1973; Blei and Jordan, 2006). At a high level, $G \sim \mathrm{DP}(\alpha, G_0)$, with the practical implementation realized through a Bayesian Gaussian mixture with `weight_concentration_prior_type=dirichlet_process`. This does *not* prove that the method discovers true function basins; rather, it groups *visited input regions* and lets the data decide how many components are effectively occupied.

Let z_i denote the assigned component for point x_i. We rank occupied clusters by their observed objective quality $Q_k = |C_k|^{-1} \sum_{i:z_i=k} y_i$, where $C_k = \{i : z_i = k\}$. The best-ranked occupied cluster provides an alternative search center, not a proof of global structure. Hence we describe the mechanism as *ranking visited spatial clusters by observed objective quality*.

3.5 CONTROLLER-WEIGHTED VARIATIONAL TRAINING

The method also reweights the variational likelihood terms. For each observation, posterior uncertainty is summarized as $u_i = \mathrm{Var}_{q_t}[f(x_i)]$. If $c_1, \ldots, c_M$ are recent controller region centers, its region affinity is $a_i = \max_{m \le M} k(x_i, c_m)$. These two signals are blended using an uncertainty coefficient $\lambda_u \in [0, 1]$,

$$r_i = \lambda_u \widehat{u}_i + (1 - \lambda_u)\widehat{a}_i, \tag{2}$$

where $\widehat{u}_i$ and $\widehat{a}_i$ denote normalized versions of uncertainty and affinity. The final weights are produced by a softmax with concentration multiplier $\kappa_w > 0$, followed by normalization,

$$\widetilde{w}_i = \mathrm{softmax}_i(\kappa_w r_i), \qquad w_i = \frac{\widetilde{w}_i}{\max_j \widetilde{w}_j}. \tag{3}$$

Because these weights are constructed with detached statistics, we write $\mathrm{sg}(w_i)$ for stop-gradient weights. The effective training objective is therefore best described as a *controller-weighted variational training objective*:

$$\mathcal{L}_{\text{AMS}} = \frac{\sum_i \mathrm{sg}(w_i)\, \mathbb{E}_{q(f_i)}[\log p(y_i \mid f_i)]}{\sum_i \mathrm{sg}(w_i) + \varepsilon} - \mathrm{KL}[q(u) \| p(u)] + \mathcal{R}_{\text{aux}}. \tag{4}$$

Here $\varepsilon > 0$ is a numerical stabilizer and $\mathcal{R}_{\text{aux}}$ collects the additional code-level regularization terms. We intentionally avoid overselling this as a new theoretically characterized GP model; it is an executed training objective used by the complete method.

4 EXPERIMENTS

4.1 EXPERIMENTAL SETUP

We compare AMS-DVGP with adaptive-depth FocalBO (Wei et al., 2024) and a standard SVGP baseline. For FocalBO, we use the adaptive focal-depth configuration corresponding to the published

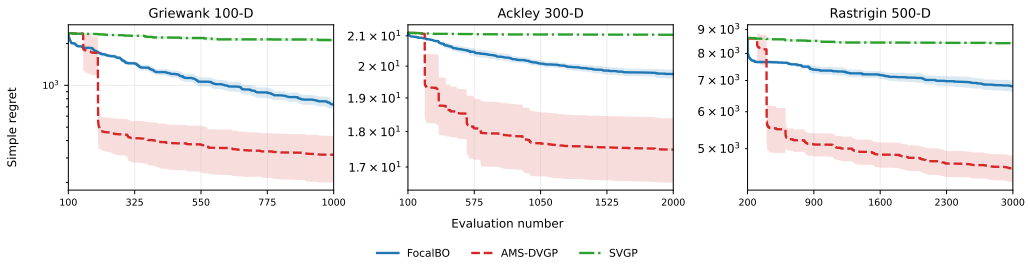


Figure 1: Representative scalable synthetic benchmarks with exact evaluation-number axes: Griewank 100-D, Ackley 300-D, and Rastrigin 500-D. Curves show means across ten matched runs and shaded regions show 95% confidence intervals. The three-panel layout emphasizes the larger-space regimes used in the main scalable-BO narrative.

method rather than the earlier fixed-depth implementation setting. All three methods use TS in the main comparison and share matched initial designs seed-by-seed. Synthetic objectives use BoTorch/standard benchmark implementations (Surjanovic and Bingham, 2013; Balandat et al., 2020); inputs are normalized to $[0, 1]^d$ before GP modeling. Exact domains, initial designs, total evaluation budgets, batch sizes, and inducing-set sizes are listed in Appendix A.1. All reported endpoints and trajectory analyses use ten matched runs, and paired Wilcoxon signed-rank tests (Wilcoxon, 1945) are used throughout.

We report simple regret $r_t = f^\star - \max_{i \le t} f(x_i)$ for the six synthetic minimization benchmarks under the code's sign convention, and best observed objective for the two MuJoCo tasks. The two Ackley settings are intentionally complementary rather than a one-factor scaling ablation: Ackley 5-D uses the standard code domain $[-5, 10]^5$ with 500 total evaluations, whereas Ackley 300-D uses the wider $[-32.768, 32.768]^{300}$ domain and 2000 evaluations to test whether the same method remains effective under a much larger search space and observation history.

4.2 Primary quantitative results

Table 1 summarizes the eight retained settings. AMS-DVGP has the best mean endpoint result in every retained task among the three compared methods; significance tests are reported in Appendix B.

Across the six synthetic tasks, the relative reduction in final mean regret versus FocalBO is 27.1% on Bohachevsky 2-D, 72.9% on Ackley 5-D, 56.6% on Griewank 100-D, 11.5% on Ackley 300-D, 33.4% on Rastrigin 500-D, and 87.5% on Rosenbrock 500-D, averaging 48.2%. Every synthetic comparison against SVGP is significant ($p \le 0.0039$). Against adaptive-depth FocalBO, all synthetic comparisons except Bohachevsky are significant ($p \le 0.0273$), making Bohachevsky the lone low-dimensional case where the mean improvement is not statistically separated. The low-dimensional Bohachevsky and Ackley curves are deferred to Appendix C, while Figure 1 focuses the main text on the larger-scale scalable-BO benchmarks.

Table 1: Primary results with adaptive-depth FocalBO. Lower is better for synthetic regret; higher is better for MuJoCo objectives.

Task	FocalBO	AMS-DVGP	SVGP
Bohachevsky 2-D	0.704 ± 0.354	$\mathbf{0.513 \pm 0.102}$	0.791 ± 0.260
Ackley 5-D	1.582 ± 1.011	$\mathbf{0.428 \pm 0.168}$	2.984 ± 0.459
Griewank 100-D	728.9 ± 80.4	$\mathbf{316.3 \pm 187.9}$	2119.7 ± 101.9
Ackley 300-D	19.745 ± 0.232	$\mathbf{17.484 \pm 1.460}$	21.035 ± 0.014
Rastrigin 500-D	6803.2 ± 315.0	$\mathbf{4528.6 \pm 483.6}$	8417.1 ± 73.6
Rosenbrock 500-D	$(1.092 \pm 0.098) \times 10^7$	$\mathbf{(1.367 \pm 0.040) \times 10^6}$	$(4.414 \pm 0.178) \times 10^7$
HalfCheetah	1076.7 ± 305.2	$\mathbf{1919.1 \pm 79.8}$	1427.4 ± 115.8
FrankaBell	36.395 ± 1.845	$\mathbf{39.040 \pm 1.826}$	32.773 ± 1.341

4.3 COMPUTATIONAL COST

The adaptive controller adds DP-GMM fitting and regional bookkeeping, but it avoids the repeated traversal of increasingly deep focal hierarchies used by adaptive-depth FocalBO. Across the eight retained tasks, the mean of the recorded per-task wall-clock times is 2694 ± 1038 seconds for AMS-DVGP, 6798 ± 2578 seconds for FocalBO, and 1393 ± 526 seconds for SVGP (mean $\pm$ standard deviation across task-level means). Thus, AMS-DVGP uses about 60.4% less recorded wall-clock time than adaptive-depth FocalBO on this suite, while SVGP remains the lightest method. We use *computational cost* for this wall-clock measure and reserve *trajectory-wide optimization efficiency* for the regret-based analysis in Section 4.6.

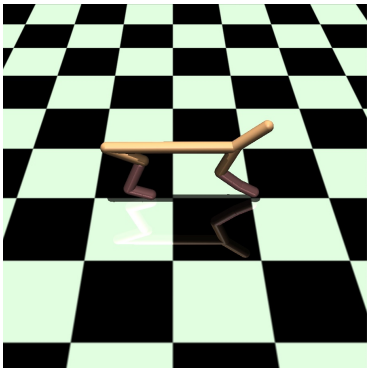
4.4 HALFCHEETAH COMPOSITE PARAMETER SEARCH

We evaluate the Gymnasium `HalfCheetah-v4` environment (Towers et al., 2024), simulated with MuJoCo (Todorov et al., 2012), over three environment/reward parameters and eight controller parameters. The exact ranges are listed in Appendix D. The controller uses grouped sinusoidal actions; for front joints $i \in \{0, 1, 2\}$ and rear joints $j \in \{3, 4, 5\}$, $a_i(t) = A_1 \sin(2\pi f_1 t + \phi_1 + 0.5i)$ and $a_j(t) = A_2 \sin(2\pi f_2 t + \phi_2 + 0.5(j-3))$. Each candidate is evaluated over three episodes with a 500-step cap. The black-box score is

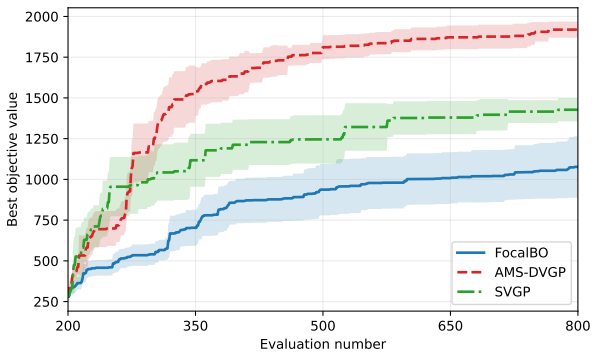
$$J = \overline{R} + 100s + \min(\overline{L}/100, 5), \tag{5}$$

where $\overline{R}$ is mean episodic return, s is the fraction of episodes with positive total return, and $\overline{L}$ is mean episode length. The objective does not explicitly reward "standing upright": the dominant Gymnasium return rewards forward locomotion while penalizing control effort, and the two auxiliary terms favor positive-return episodes and longer rollouts. Because reward-related parameters are optimized jointly with controller parameters, this is a composite parameter-search problem rather than a pure fixed-reward controller benchmark.

AMS-DVGP attains 1919.1 ± 79.8, versus 1076.7 ± 305.2 for adaptive-depth FocalBO and 1427.4 ± 115.8 for SVGP ($p = 0.0020$ against each baseline).



(a) Representative rollout frame.

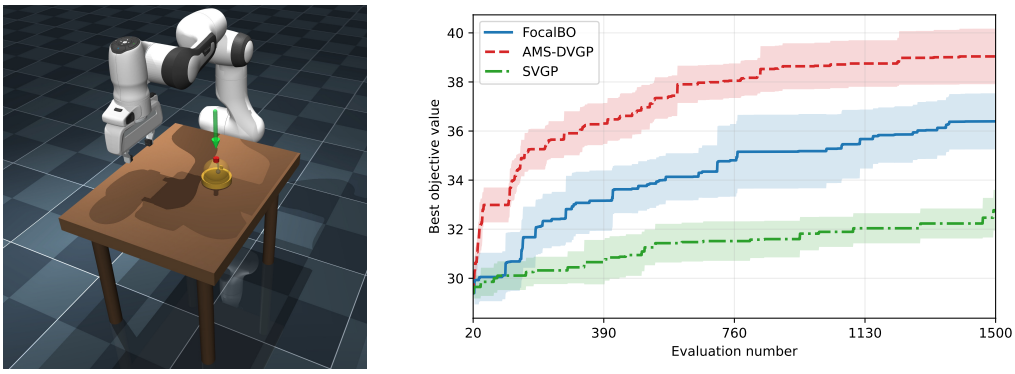


(b) Optimization curve.

Figure 2: Representative HalfCheetah frame and optimization performance.

4.5 FRANKABELL: 32-D MANIPULATION PARAMETER SEARCH

We additionally introduce **FrankaBell**, our name for a custom 32-D bell-pressing benchmark simulated in MuJoCo (Todorov et al., 2012). The Franka Emika Panda robot model and mesh assets are taken from MuJoCo Menagerie (Zakka et al., 2022); the bell scene, controller, search space, and objective are defined for this study. The search includes seven proportional-gain multipliers, seven derivative-gain ratios, seven joint-damping parameters, planar approach offsets, hover and press depths, wrist orientation, four phase durations, and gripper opening. Appendix E gives the exact parameter groups and objective. We use 20 initial observations, a nominal 1500-evaluation budget, batch size 50, and 20 inducing variables.



(a) Task setup.

(b) Optimization curve.

Figure 3: FrankaBell manipulation experiment. AMS-DVGP reaches 39.040 ± 1.826, versus 36.395 ± 1.845 for adaptive-depth FocalBO and 32.773 ± 1.341 for SVGP.

The advantage is significant against both baselines (paired Wilcoxon $p = 0.0059$ vs. FocalBO and $p = 0.0020$ vs. SVGP), providing a higher-dimensional embodied-control validation in addition to HalfCheetah.

4.6 Algorithm analysis: adaptive behavior and trajectory-wide efficiency

We first ask whether the controller actually exercises its multiple search operators rather than collapsing to one policy. Figure 4 summarizes representative diagnostic traces for one seed each on Ackley 5-D, Rastrigin 500-D, and Rosenbrock 500-D. Cluster-gate fallback is merged into the action that is actually executed, information-guided search. Across all three traces, multi-scale search and cluster contraction account for most controller iterations, while information-guided search and global escape are invoked selectively. Global escape is intentionally sparse because it is triggered only after repeated stagnation. These shares are descriptive mechanism diagnostics, not estimates of component-wise causal contribution.

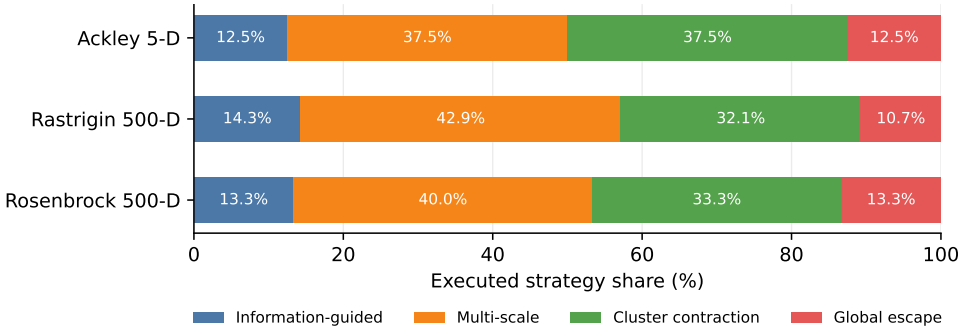


Figure 4: Representative AMS-DVGP strategy allocation. Percentages are shares of executed controller iterations in one diagnostic run per task; fallback events are counted under the action actually executed.

We next test whether the endpoint gains reflect sustained search quality throughout the high-dimensional BO horizon. For each run, define the trajectory-averaged simple regret

$$\bar{r}_{\text{traj}} = \frac{1}{T - N_0 + 1} \sum_{t=N_0}^{T} r_t, \tag{6}$$

where N_0 is the initial-design size and r_t is best-so-far simple regret after evaluation t. Lower values indicate that a method maintains stronger incumbents throughout the sequential campaign, not merely

Table 2: Trajectory-averaged simple regret on the high-dimensional tasks. Lower is better.

Task	FocalBO	AMS-DVGP	SVGP	AMS vs. Focal reduction
Griewank 100-D	1183	**559**	2207	52.8%
Ackley 300-D	20.19	**18.08**	21.05	10.5%
Rastrigin 500-D	7221	**5146**	8472	28.7%
Rosenbrock 500-D	1.67×10^7	$\mathbf{4.57 \times 10^6}$	4.58×10^7	72.6%

at the endpoint. Table 2 shows that AMS-DVGP reduces this quantity relative to adaptive-depth FocalBO by 52.8%, 10.5%, 28.7%, and 72.6% on Griewank 100-D, Ackley 300-D, Rastrigin 500-D, and Rosenbrock 500-D, respectively; all paired comparisons are significant ($p \leq 0.0039$).

The corresponding reductions versus SVGP are 74.7%, 14.1%, 39.3%, and 90.0%. Together, the strategy-share diagnostic and trajectory-wide metric support the intended interpretation: AMS-DVGP dynamically uses multiple search modes, and its high-dimensional advantage is not confined to a late-stage improvement.

5 Discussion

Scalability across observation count and dimension. The exact-GP $O(n^3)$ bottleneck is fundamentally an observation-count problem, while high dimension creates a separate search and statistical bottleneck. AMS-DVGP addresses the first through a sparse variational GP and the second through adaptive spatial allocation; the experiments deliberately cover both. The low-dimensional Bohachevsky and Ackley tasks show that the controller remains useful with only 20 initial observations, whereas Griewank 100-D, Ackley 300-D, Rastrigin 500-D, and Rosenbrock 500-D combine larger spaces with longer histories. Rosenbrock is especially demanding because it starts from 2000 initial observations and continues to 5000 total evaluations. We therefore view the method as scalable sparse-GP BO with demonstrated robustness to high dimension, not as a claim that large n and large d are the same phenomenon.

Why Dirichlet-guided proposals are useful. Classical BO acquisitions balance exploitation of high predicted value with exploration driven by predictive uncertainty (Shahriari et al., 2016; Srinivas et al., 2010; Thompson, 1933). An incumbent-centered regional method operationalizes the exploitation side directly, and explicit uncertainty maxima provide a natural exploration anchor. AMS-DVGP adds a third data-driven spatial cue: a DP mixture summarizes promising *occupied* modes in the visited space, so budget can be assigned to a region that is neither the current incumbent neighborhood nor simply the most uncertain location. The role of the DP is therefore not to claim discovery of true basins, but to diversify region proposals using the structure already revealed by observations.

Acquisition robustness. The main comparison fixes TS across methods, but the controller is not structurally tied to TS. AMS-DVGP remains effective across representative acquisition functions (TS, UCB, and EI), indicating that the method does not rely on a single acquisition choice. Final regret is 0.428 ± 0.168 for TS, 0.661 ± 0.540 for UCB, and 0.741 ± 0.303 for EI. Paired Wilcoxon tests show no significant difference for TS vs. UCB ($p = 0.2754$) or EI vs. UCB ($p = 0.6250$), whereas TS is significantly better than EI ($p = 0.0273$). We therefore interpret Figure 5 as acquisition robustness rather than full acquisition invariance.

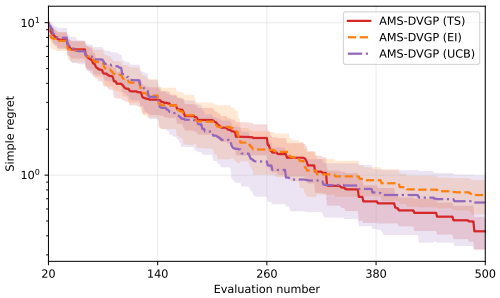


Figure 5: AMS-DVGP acquisition robustness on Ackley. Curves show means with 95% confidence intervals over ten runs.

Position within scalable BO. On the eight retained benchmarks, AMS-DVGP has the best mean endpoint result among the directly compared scalable sparse-GP methods, adaptive-depth FocalBO and SVGP. Against adaptive-depth FocalBO,

the paired improvement is significant in seven of the eight tasks, with Bohachevsky the lone non-significant case; against SVGP, all eight paired comparisons are significant. This is strong evidence within the matched comparison suite, but it is not a universal "best scalable BO" claim: recent methods target different assumptions, surrogate families, and high-dimensional regimes (Makrygiorgos et al., 2025; Sinaga et al., 2025; Yu et al., 2026). The contribution is instead a resource-allocation mechanism that complements sparse inference and remains effective across both large-observation and high-dimensional settings.

6 Conclusion

AMS-DVGP combines sparse variational GP inference with Dirichlet-guided occupied-mode proposals and state-adaptive multi-scale search for scalable Bayesian optimization. Across six synthetic tasks spanning 2–500 dimensions and 500–5000 evaluations, it achieves the best mean endpoint result among the three directly compared methods, with a 48.2% average relative regret reduction over adaptive-depth FocalBO. It also obtains the strongest final objective on HalfCheetah and FrankaBell, uses about 60.4% less recorded wall-clock time than adaptive-depth FocalBO across the retained suite, and remains effective under TS, UCB, and EI. Algorithm analysis further shows that the controller actively uses multiple search modes and maintains lower trajectory-averaged regret throughout high-dimensional optimization, supporting adaptive spatial allocation as a practical complement to sparse-GP approximation.

AI Use Statement

Generative AI tools were used to assist with literature search, code inspection, experimental-result checking and interpretation, manuscript organization, drafting and language editing, and formatting. All AI-assisted content was reviewed by the authors, who take responsibility for the final equations, claims, experiments, statistics, citations, and manuscript text.

Reproducibility Statement

The accompanying supplementary package documents the implementation, exact experimental configurations, plots, validation notes, and provenance of third-party simulation assets. The archival release is intended to include exact commit hashes, environment specifications, random seeds, raw result files, and plotting scripts.

References

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*, volume 33, pages 21524–21538, 2020.

David M. Blei and Michael I. Jordan. Variational inference for dirichlet process mixtures. *Bayesian Analysis*, 1(1):121–144, 2006. doi: 10.1214/06-BA104.

David Eriksson and Martin Jankowiak. High-dimensional bayesian optimization with sparse axis-aligned subspaces. In *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, volume 161 of *Proceedings of Machine Learning Research*, pages 493–503, 2021.

David Eriksson, Michael Pearce, Jacob R. Gardner, Ryan Turner, and Matthias Poloczek. Scalable global optimization via local bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Thomas S. Ferguson. A bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1(2):209–230, 1973. doi: 10.1214/aos/1176342360.

James Hensman, Nicolò Fusi, and Neil D. Lawrence. Gaussian processes for big data. In *Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence*, pages 282–290, 2013.

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. doi: 10.1023/A:1008306431147.

Harold J. Kushner. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *Journal of Basic Engineering*, 86(1):97–106, 1964. doi: 10.1115/1.3653121.

Wesley J. Maddox, Samuel Stanton, and Andrew Gordon Wilson. Conditioning sparse variational gaussian processes for online decision-making. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Georgios Makrygiorgos, Joshua Hang Sai Ip, and Ali Mesbah. Towards scalable bayesian optimization via gradient-informed bayesian neural networks. *IFAC-PapersOnLine*, 59(6):157–162, 2025. doi: 10.1016/j.ifacol.2025.07.138.

Leonard Papenmeier, Luigi Nardi, and Matthias Poloczek. Increasing the scope as you learn: Adaptive bayesian optimization in nested subspaces. In *Advances in Neural Information Processing Systems*, volume 35, 2022. doi: 10.52202/068431-0842.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando de Freitas. Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2016. doi: 10.1109/JPROC.2015.2494218.

Marshal Sinaga, Julien Martinelli, and Samuel Kaski. Robust and computation-aware gaussian processes. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

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 W. Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the 27th International Conference on Machine Learning*, pages 1015–1022, 2010.

Sonja Surjanovic and Derek Bingham. Virtual library of simulation experiments: Test functions and datasets. Simon Fraser University, 2013. Optimization test-function library; accessed 2026-08-29.

William R. Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3–4):285–294, 1933. doi: 10.1093/biomet/25.3-4.285.

Michalis Titsias. Variational learning of inducing variables in sparse gaussian processes. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*, volume 5 of *Proceedings of Machine Learning Research*, pages 567–574. PMLR, 2009.

Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5026–5033. IEEE, 2012. doi: 10.1109/IROS.2012.6386109.

Mark Towers, Ariel Kwiatkowski, Jordan Terry, John U. Balis, Gianluca De Cola, Tristan Deleu, Manuel Goulão, Andreas Kallinteris, Markus Krimmel, Arjun KG, Rodrigo Perez-Vicente, Andrea Pierré, Sander Schulhoff, Jun Jet Tai, Hannah Tan, and Omar G. Younis. Gymnasium: A standard interface for reinforcement learning environments, 2024. URL `https://arxiv.org/abs/2407.17032`.

Sattar Vakili, Henry Moss, Artem Artemev, Vincent Dutordoir, and Victor Picheny. Scalable thompson sampling using sparse gaussian process models. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Ziyu Wang, Frank Hutter, Masrour Zoghi, David Matheson, and Nando de Freitas. Bayesian optimization in a billion dimensions via random embeddings. *Journal of Artificial Intelligence Research*, 55:361–387, 2016. doi: 10.1613/jair.4806.

Yunyue Wei, Vincent Zhuang, Saraswati Soedarmadji, and Yanan Sui. Scalable bayesian optimization via focalized sparse gaussian processes. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-3828.

Frank Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945. doi: 10.2307/3001968.

Rosen Yu, Cyril Picard, and Faez Ahmed. Git-bo: High-dimensional bayesian optimization with tabular foundation models. In *International Conference on Learning Representations*, 2026.

Kevin Zakka, Yuval Tassa, and MuJoCo Menagerie Contributors. MuJoCo Menagerie: A collection of high-quality simulation models for MuJoCo, 2022. URL `https://github.com/google-deepmind/mujoco_menagerie`.

A Implementation and Protocol Details

A.1 Experimental configurations

Table 3 lists the eight retained experiments. The synthetic domains are the physical objective domains before the normalized $[0,1]^d$ GP interface.

Table 3: Experimental configurations used for the retained results.

Task	d	Init	Eval	Batch	m	Domain
Bohachevsky	2	20	500	20	20	$[-100, 100]^2$
Ackley	5	20	500	20	20	$[-5, 10]^5$
Griewank	100	100	1000	50	50	$[-600, 600]^{100}$
Ackley (wide)	300	100	2000	50	50	$[-32.768, 32.768]^{300}$
Rastrigin	500	200	3000	100	200	$[-5.12, 5.12]^{500}$
Rosenbrock	500	2000	5000	100	200	$[-5, 10]^{500}$
HalfCheetah	11	200	800	20	50	normalized parameter box
FrankaBell	32	20	1500	50	20	normalized parameter box

A.2 Executed controller constants

Table 4: Key controller constants used in the current AMS-DVGP implementation.

Component	Executed setting
Progress thresholds	$\tau_{\text{fast}} = 0.01$, $\tau_{\text{steady}} = 0.003$, $\tau_{\text{slow}} = 0.001$; stagnation band $\lvert\Delta_t\rvert \leq 0.001$
Repeated stagnation rule	$h_{\text{stag}} = 4$ accumulated stagnation events
DP-GMM	full covariance; up to 200 EM/VI iterations; Dirichlet-process prior via BayesianGaussianMixture
Multi-scale allocation	$\boldsymbol{\beta} = (0.4, 0.4, 0.2)$ across fine / medium / coarse scales
Information-guided priorities	$\boldsymbol{\pi} = (1.0, 0.8, 0.6)$ for incumbent / cluster / uncertainty regions
Global escape mix	$\eta_g = 0.7$ global-uniform, $1 - \eta_g = 0.3$ uncertainty-guided
Cluster contraction	$\tau_c = 0.7$; local radius $r_c = 0.05$
Regional selector	named acquisition dispatch (TS/EI/UCB/PI supported); TS in main comparisons
Likelihood weighting	$\lambda_u = 0.6$, $1 - \lambda_u = 0.4$, $\kappa_w = 2$, $\varepsilon = 10^{-7}$
Auxiliary regularizer	0.7 KL-alignment $+0.3$ smoothness; adaptive regularization weight from the executed code path

These values are fixed implementation settings used consistently in the reported experiments. They are reported for reproducibility rather than presented as theoretically derived or individually optimized choices; the current study does not include a dedicated sensitivity analysis for these controller hyperparameters.

B Statistical Tests

Table 5: Paired Wilcoxon p-values for the eight retained endpoint comparisons.

Task	p(AMS vs FocalBO)	p(AMS vs SVGP)
Bohachevsky 2-D	0.1934	0.0039
Ackley 5-D	0.0273	0.0020
Griewank 100-D	0.0039	0.0020
Ackley 300-D	0.0039	0.0020
Rastrigin 500-D	0.0020	0.0020
Rosenbrock 500-D	0.0020	0.0020
HalfCheetah	0.0020	0.0020
FrankaBell	0.0059	0.0020

Recorded wall-clock cost

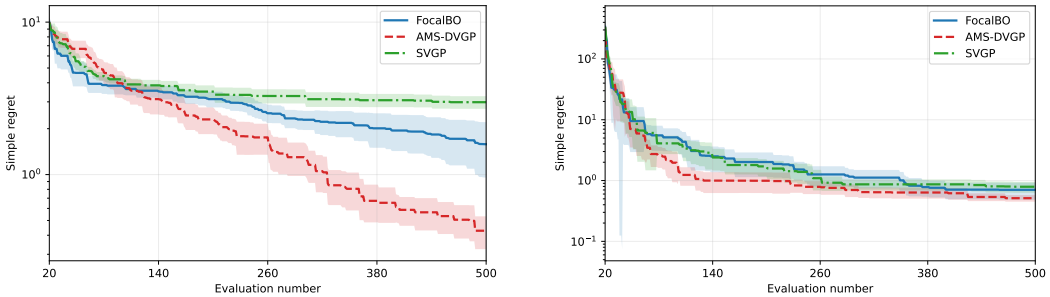
For completeness, Table 6 reports the cross-task summary used in the main-text computational-cost discussion. Each entry is the mean $\pm$ standard deviation across the eight retained task-level mean runtimes; it is a wall-clock summary rather than an asymptotic complexity estimate.

Table 6: Recorded wall-clock cost aggregated across the eight retained tasks.

Method	Mean task-level wall-clock (s)
FocalBO	6798 ± 2578
AMS-DVGP	2694 ± 1038
SVGP	1393 ± 526

C Additional Synthetic Curves

Figure 6 reports the two low-dimensional synthetic curves that are no longer shown in the main text. Ackley 5-D remains a useful small-budget reference, while Bohachevsky 2-D is included for completeness because its endpoint statistics are reported in Table 1 and Table 5.



(a) Ackley 5-D. (b) Bohachevsky 2-D.

Figure 6: Additional low-dimensional synthetic results with exact evaluation-number axes. Curves show means across ten matched runs and shaded regions show 95% confidence intervals.

D HalfCheetah Search Space and Objective

Table 7: HalfCheetah 11-D composite parameter-search space.

Group	Parameter	Range
Environment / reward	forward reward weight	$[0.1, 3.0]$
	control cost weight	$[0.001, 1.0]$
	reset noise scale	$[0.01, 0.5]$
Controller	action scale	$[0.3, 1.5]$
	action-noise s.d.	$[0, 0.3]$
	front frequency f_1	$[0.5, 5.0]$
	rear frequency f_2	$[0.5, 5.0]$
	front amplitude A_1	$[0.1, 1.0]$
	rear amplitude A_2	$[0.1, 1.0]$
	front phase ϕ_1	$[0, 6.28]$
	rear phase ϕ_2	$[0, 6.28]$

Each candidate is evaluated in `HalfCheetah-v4` for three episodes capped at 500 steps. The six joint actions are generated by two sinusoidal groups (front and rear) and clipped to $[-1, 1]$. The composite objective used throughout the reported experiment is

$$J_{\mathrm{HC}} = \overline{R} + 100s + \min(\overline{L}/100, 5), \tag{7}$$

where $\overline{R}$ is mean episodic return, s is the fraction of episodes with positive return, and $\overline{L}$ is mean episode length. The first three optimized variables modify Gymnasium's environment/reward configuration; the remaining eight parameterize the sinusoidal controller.

E FrankaBell Search Space and Objective

FrankaBell is a custom bell-pressing benchmark built around the Franka Emika Panda model from MuJoCo Menagerie (Zakka et al., 2022). The Menagerie Panda model and meshes are released under Apache-2.0; our benchmark adds the bell scene, task logic, controller parameterization, and objective. The normalized BO interface maps $[0, 1]^{32}$ to the physical ranges in Table 8.

Table 8: FrankaBell 32-D search space used in the reported 1500-evaluation experiment.

Group	Parameters	Physical range
Joint proportional gains	$7\times$ k_p scale	$[0.18, 2.20]$
Derivative ratios	$7\times$ k_d/k_p	$[0.03, 0.32]$
Joint damping	$7\times$ damping	$[0.20, 6.00]$
Approach offsets	x, y	$[-0.24, 0.24]$ m
Vertical motion	hover height	$[0.030, 0.350]$ m
	press depth	$[0.000, 0.070]$ m
Wrist orientation	yaw	$[-1.50, 1.50]$ rad
	pitch	$[-1.00, 1.00]$ rad
Phase timing	approach	$[0.18, 2.50]$ s
	press	$[0.08, 1.80]$ s
	hold	$[0.04, 1.20]$ s
	retract	$[0.10, 1.60]$ s
Gripper	opening command	$[0, 255]$

The objective used in the reported run is

$$J_{\mathrm{bell}} = 20I_{\mathrm{strike}} + 12E_{\mathrm{ring}} + 8P_{\mathrm{press}} + 4I_{\mathrm{touch}} - 12\min(e_{xy}, 0.50) - 1.5 \times 10^{-4}E_{\mathrm{act}}, \tag{8}$$

where I_{strike} indicates clapper–plate contact, E_{ring} is the clipped maximum clapper velocity normalized by the simulator scale, P_{press} is the clipped maximum button displacement normalized by the button travel range, I_{touch} indicates button contact, e_{xy} is mean planar end-effector alignment error

after the approach phase, and E_{act} is accumulated squared actuator force integrated over simulation time. Numerically unstable rollouts receive a score of -50. The reported BO comparison uses 20 initial observations, nominal budget 1500, batch size 50, and 20 inducing variables; all three methods use TS. Because the loop advances in full batches, stored traces can overshoot the nominal target, but all reported FrankaBell statistics and plots are evaluated at the 1500th observation.

F LIMITATIONS

Component attribution. The paper validates the complete AMS-DVGP method and includes acquisition robustness and mechanism diagnostics, but it does not isolate the causal contribution of every controller component through exhaustive removal ablations.

Scaling characterization. The retained suite reaches 500 dimensions and 5000 total evaluations, but it is not a controlled complexity experiment that sweeps n at fixed d (or d at fixed n) to estimate an empirical scaling law. Sparse variational inference supplies the large-n computational mechanism, while the high-dimensional experiments demonstrate performance robustness rather than a new asymptotic complexity bound.