BAYESIAN OPTIMIZATION ON AMORTIZED SEARCH SPACE

Anonymous authors
Paper under double-blind review

ABSTRACT

In real-world black-box optimization problems, a differentiable mechanistic simulator is often available that approximately predicts the behavior of the system as a function of the parameters being optimized. Here, we show how an amortized inverse map of such a simulator can be used in Bayesian optimization to restrict the search space while using a Gaussian process kernel that is aware of the original problem geometry. This allows for efficiently leveraging the low intrinsic dimensionality of the optimization problem. We derive regret bounds that highlight the benefits and limitations of this method and illustrate it with an application to drug dosing personalization in Parkinson's disease. We use, as a simulator, a pharmacokinetic/pharmacodynamic (PK/PD) model of a drug, a mechanistic model calibrated on population data with quantified parameter uncertainty. A dosing neural network is trained offline against the model, amortizing the map from patient PK/PD parameters to the regimen predicted to be optimal, and online Bayesian optimization then searches over this learned, low-dimensional representation rather than over raw dosages. Numerical simulations demonstrate the improved performance of our method in this application, as well as under simulator misspecification.

1 INTRODUCTION

We consider the problem of optimizing a system with objective function U and parameters $\mathbf{x}$ in a search space (or design space) $\mathcal{D} \subseteq \mathbb{R}^d$:

$$x^\star = \arg\max_{\mathbf{x}\in\mathcal{D}} U(\mathbf{x}), \tag{1}$$

U is an expensive black-box function, with no available derivative information. Bayesian optimization (BO) is a sample-efficient approach for optimizing black-box functions that are expensive to evaluate, and has been applied across domains as varied as hyperparameter tuning, robotics, and drug discovery (Brochu et al., 2010b; Shahriari et al., 2016; Frazier, 2018). At each iteration, a probabilistic surrogate, typically a Gaussian process (GP) (Rasmussen & Williams, 2006), models the unknown objective from past observations, and an acquisition function balances exploration and exploitation to select the next candidate to evaluate. Under budgets of only a few tens of experiments, learning an informative surrogate can be difficult even at moderate dimensionality. Although this difficulty has traditionally been attributed to the curse of dimensionality (Wang et al., 2013; Kandasamy et al., 2015), recent evidence identifies prior-induced model complexity as an important and potentially avoidable cause (Hvarfner et al., 2024a).

We consider the setting where a differentiable mechanistic simulator approximating the behavior of the system being optimized is available. We propose to use the simulator to learn a restriction of the search space. Offline, we train a neural network that amortizes the solution of the simulator-defined optimization problem into a single forward pass; its image $\mathcal{M} \subseteq \mathcal{D}$ is the subset of simulator-optimal candidates. Online, BO searches over the network inputs, using a Gaussian process kernel pulled back from $\mathcal{D}$ to preserve the geometry of the original problem. Because the simulator is used offline only, the online optimization loop can be conducted under any form of feedback, including scalar evaluations, preferences, or other types of observations that may not directly correspond to the simulator output.

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.

Personalized drug dosing, our motivating application, is a concrete instance of this setting. Here the simulator is a pharmacokinetic-pharmacodynamic (PK/PD) model, describing how a dosing regimen translates into drug concentration across body compartments (pharmacokinetics) and into a clinical effect (pharmacodynamics). These models are estimated on clinical-trial data, with inter-individual variability captured by random effects on the parameters. The PD model comes with a known target level of efficacy, so the dosing problem can be solved inside the simulator for any patient parameters. In this setting, the image of the amortized network is a low-dimensional subspace of pharmacologically coherent regimens, so the online search is restricted to regimens that are sensible by construction, which matters both for patient safety and for keeping the number of trials small. Figure S1 illustrates our methodology applied to this use-case.

Our contributions can be summarized as follows:

- We introduce BOASS (Bayesian Optimization on Amortized Search Space), a method where a model is trained to amortize the solution of the simulator-defined optimization problem, and its image defines a low-dimensional set of simulator-optimal candidates over which BO runs online. Within BOASS, we define the pullback kernel for the GP surrogate to make the model aware of amortized map geometry;
- We derive regret bounds that highlight the benefits and limitations of this method, decomposing the contribution of the simulator misspecification, the amortization error, and the online optimization error in the regret;
- We evaluate BOASS on synthetic experiments and on a dose personalization problem of Levodopa-Carbidopa for Parkinson's disease;

The remainder of the paper is organized as follows: Section 2 discusses related works, Section 3 presents BOASS in a general setting, Section 4 contains the regret analysis, and Section 5-6 presents experimental results. We conclude with a discussion in Section 7

2 Related Work

Bayesian optimization, high dimensions and learned search spaces. A recurring strategy for high-dimensional or structured domains is to run BO on a lower-dimensional representation of the raw design space. REMBO (Wang et al., 2016) optimizes inside a random linear embedding, assuming that the objective varies along a low-dimensional subspace only; SAASBO (Eriksson & Jankowiak, 2021) places a sparsity-inducing prior on the kernel lengthscales, which restricts it to axis-aligned structure; TuRBO (Eriksson et al., 2019) keeps the full space and confines the search to local trust regions. These methods scale to hundreds of dimensions, but the representation is either drawn at random or inferred from the observations, requiring a large iteration budget and several restarts, which is not applicable in contexts such as precision medicine. Alternatively, latent-space BO learns a generative model of the inputs, typically a variational autoencoder, and optimizes in its latent space (Gómez-Bombarelli et al., 2018; Tripp et al., 2020; Maus et al., 2022; Stanton et al., 2022); amortized optimization instead trains a network offline to map a problem instance to its solution, replacing per-instance optimization with a forward pass (Amos, 2023). Our approach combines these two ideas: the dosing network is an amortized optimizer mapping patient parameters to the model-optimal regimen, and its image defines the low-dimensional space over which online BO runs, obtained offline, before any patient evaluation is spent.

Leveraging simulators in Bayesian optimization. Several BO methods have been proposed to exploit mechanistic simulators. One line of work builds the GP kernel on a latent representation of simulated trajectories, so that configurations with similar simulated behaviour are treated as similar by the surrogate (Rai et al., 2019; Antonova et al., 2020); another uses the simulator to bias the acquisition function, the surrogate mean, or the presumed location of the optimum with an informative prior (Hvarfner et al., 2022; 2024b; Boltz et al., 2025). Multi-fidelity BO (Kandasamy et al., 2016; Moss et al., 2021) is the most standard way to combine a simulator with real-world evaluations, treating the simulator as a cheap, low-fidelity approximation of the expensive, high-fidelity real-world objective. This assumption does not necessarily hold in our setting: for example, the PK/PD model predicts the plasma concentration $\mathbf{c}(t)$ or the treatment effect $r(t)$, whereas the feedback we can realistically collect from a patient (a symptom score, a wearable-derived signal, or a pairwise preference), is not a noisier or cheaper version of that same quantity but a genuinely different observation. We instead

use the simulator purely offline to restrict the search space (Section 3), and let the online BO loop operate directly on whatever feedback is actually available. Closest to our approach, Granley et al. (2023) personalize the encoding of visual stimuli into electrical stimulation in visual prostheses, by first training a neural network encoder against a static perception model acting as a decoder, and then fine-tuning a low-dimensional set of patient-specific parameters online with Preferential Bayesian optimization (Brochu et al., 2010a; Fauvel & Chalk, 2022). We build on the same idea, with three main differences: we frame the method in a general BO setting, benchmarking it on multiple scenarios and providing regret bounds. Moreover, we introduce a pullback kernel which allows leveraging the original problem space geometry. Finally in our use case, the encoder now amortizes a dosing optimization problem, and the decoder is an ODE-based model rather than a static perception model: the offline objective equation 7 is defined on the simulated trajectory, so training backpropagates through the ODE solution (Section 6.2).

3 BOASS METHOD

Let $\Phi \subseteq \mathbb{R}^p$ be the parameter space of the simulator. We assume that the user has a prior knowledge of the simulator parameters in the form of a distribution $p(\phi)$ over it. For a design $\mathbf{x} \in \mathcal{D}$ and parameters ϕ, the simulator returns an output $\mathbf{s}(\mathbf{x}; \phi) : \mathcal{D} \times \Phi \to \mathbb{R}^k$ that predicts the behavior of the system and is differentiable in $\mathbf{x}$ and ϕ. By assumption, the desired system output $\mathbf{s}^\star_\phi$ at parameter ϕ is known. This turns the simulator into an objective $f_{\text{sim}}(\cdot\,; \phi) : \mathcal{D} \to \mathbb{R}$ that scores a design by the proximity of its simulated output to the target, $f_{\text{sim}}(\mathbf{x}; \phi) = d\left(\mathbf{s}(\mathbf{x}; \phi) - \mathbf{s}^\star_\phi\right)$, where d is some relevant distance on $\mathbb{R}^k$. The BOASS methodology decomposes in the following two steps.

Amortization step. We denote the simulator-optimal design by $\mathcal{E}^\star(\phi) = \arg\min_{\mathbf{x} \in \mathcal{D}} f_{\text{sim}}(\mathbf{x}; \phi)$, and we assume it exists for every $\phi \in \Phi$. Computing $\mathcal{E}^\star(\phi)$ anew for each ϕ is expensive, so we amortize it offline into a neural network $\mathcal{E}_w : \Phi \to \mathcal{D}$ with weights w, trained to minimize the simulator objective in expectation:

$$\min_w \; \mathbb{E}_{\phi \sim p(\phi)} \left[f_{\text{sim}}\big(\mathcal{E}_w(\phi); \phi\big) \right]. \tag{2}$$

Its image $\mathcal{M} = \mathcal{E}_w(\Phi) = \{\mathcal{E}_w(\phi) : \phi \in \Phi\} \subseteq \mathcal{D}$ is the *amortized search space*, the designs that are approximately simulator-optimal for some ϕ.

Online optimization step. To solve the online optimization problem equation 1, we optimize the target objective over Φ instead of $\mathcal{D}$,

$$\phi^\star = \arg\max_{\phi \in \Phi} \; U\big(\mathcal{E}_w(\phi)\big), \tag{3}$$

and return the design $\mathcal{E}_w(\phi^\star) \in \mathcal{M}$. At iteration n, we observe a feedback $y_n = U\big(\mathcal{E}_w(\phi_n)\big) + \epsilon_n$ with $\epsilon_n \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$. A Gaussian process surrogate with kernel k on Φ is fit to the history $\mathcal{H}_n = \{(\phi_i, y_i)\}_{i \le n}$ of candidates and feedback, and an acquisition function α proposes the next candidate $\phi_{n+1} = \arg\max_{\phi \in \Phi} \; \alpha\big(\phi; \mathcal{H}_n\big)$, whose design $\mathcal{E}_w(\phi_{n+1})$ is evaluated to return the feedback y_{n+1}. The kernel k of the surrogate GP can be defined directly on the search coordinates, $k_\Phi(\phi, \phi')$, or on the regimens they produce, $k_\mathcal{D}\big(\mathcal{E}_w(\phi), \mathcal{E}_w(\phi')\big)$, i.e., the pullback of a kernel on $\mathcal{D}$ through the deterministic map $\mathcal{E}_w$. The latter is a valid positive-definite kernel for any positive-definite $k_\mathcal{D}$, being its pullback through a deterministic map.

BOASS restricts the search space rather than assuming a particular observation model, so nothing in the construction of $\mathcal{M}$ depends on the Gaussian-noise feedback assumed above. The proposed method directly applies to other BO settings, including preference feedback, which we use in Section 6.4, as well as multi-objective optimization and other non-Gaussian observation models. Restricting the search to $\mathcal{M}$ reduces the optimization dimension from $D = \dim(\mathcal{D})$ to the intrinsic dimensionality of $\mathcal{M}$, which is often much lower, at the cost of a bias whenever $\mathbf{x}^\star \notin \mathcal{M}$. We analyze this trade-off in terms of the regret in Section 4.

4 REGRET ANALYSIS UNDER BOASS

Our amortized search-space restriction trades approximation bias for a simpler online optimization problem. $\widehat{\mathbf{x}}_n = \operatorname{argmax}_{i \le n} U(\mathbf{x}_i)$ is the solution to the optimization problem reported after n iterations. By noting $\mathbf{x}^\star$ the maximizer of U over $\mathcal{D}$, and $\mathbf{x}^\star_{\mathcal{M}}$ its maximizer over $\mathcal{M}$, the simple regret after n evaluations decomposes as $r_n = \Delta_{\mathcal{M}} + r_n^{\mathcal{M}}$, with search-space bias $\Delta_{\mathcal{M}} = U(\mathbf{x}^\star) - U(\mathbf{x}^\star_{\mathcal{M}})$, and online optimization error $r_n^{\mathcal{M}} = U(\mathbf{x}^\star_{\mathcal{M}}) - U(\widehat{\mathbf{x}}_n)$. The following proposition provides a regret bound on the online optimization error when using the pullback kernel under iid Gaussian noise, with a GP-UCB acquisition function, and with reasonable hypotheses for $\mathcal{M}$.

Proposition 1 *Let $\Phi \subset \mathbb{R}^p$ be compact and convex, let $\mathcal{E}_w : \Phi \to \mathcal{D}$, and let $\mathcal{M} = \mathcal{E}_w(\Phi)$. Define the pullback kernel $k_{\mathrm{pb}}(\phi, \phi') = k_{\mathcal{D}}(E_w(\phi), E_w(\phi'))$. Suppose that:*

1. *$\mathcal{M} \subset \mathbb{R}^D$ is a smooth, compact, ρ-dimensional embedded submanifold without boundary;*
2. *$k_{\mathcal{D}}$ is a Matérn-ν kernel restricted to $\mathcal{M}$;*
3. *$g = U \circ \mathcal{E}_w$ is distributed according to the GP prior with kernel k_{pb} and satisfies the derivative-tail condition of Srinivas et al. (2010).*
4. *Observations have the form $y_t = g(\phi_t) + \epsilon_t$, with independent $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$;*
5. *The Mercer eigenfunctions of $k_{\mathcal{D}}$ restricted to $\mathcal{M}$ satisfy the uniform boundedness condition used in the information-gain bound of Vakili et al. (2021).*

Then, for a GP-UCB acquisition function, simple regret of the best evaluated point satisfies $r_n^{\mathcal{M}} = \widetilde{O}\left(n^{-\frac{\nu}{2\nu+\rho}}\right)$, compared with $\widetilde{O}(n^{-\nu/(2\nu+D)})$ on the full D-dimensional input space.

The proof is detailed in Appendix B.4. Proposition 1 makes explicit the central trade-off of amortized search-space restriction: the simulator and its amortization can introduce bias, while restricting online BO to a subspace of lower intrinsic dimension can accelerate optimization.

The search-space bias itself separates into simulator misspecification and amortization error.

Proposition 2 *Let $\mathcal{E}^\star(\phi)$ denote the exact optimizer of the simulator-defined optimization problem. $\mathcal{M}^\star = \mathcal{E}^\star(\Phi)$ is the simulator-induced search space. Let $\mathbf{x}^\star \in \arg\max_{\mathbf{x} \in \mathcal{D}} U(\mathbf{x})$ and $\mathbf{x}^\star_{\mathcal{M}} \in \arg\max_{\mathbf{x} \in \mathcal{M}} U(\mathbf{x})$ be respectively the globally optimal system parameters and the best system parameters contained in the learned search space. Define the search-space bias as $\Delta_{\mathcal{M}} = U(\mathbf{x}^\star) - U(\mathbf{x}^\star_{\mathcal{M}})$. Under the hypotheses of Proposition 1, if U is L_U-Lipschitz and $\sup_{\phi \in \Phi} \delta\left(\mathcal{E}_w(\phi), \mathcal{E}^\star(\phi)\right) \le \varepsilon_{\mathrm{am}}$, then :*

$$\Delta_{\mathcal{M}} \le r_n \le \underbrace{\Delta_{\mathrm{model}}}_{\text{simulator misspecification}} + \underbrace{L_U \varepsilon_{\mathrm{am}}}_{\text{amortization error}} + \underbrace{\widetilde{O}\left(n^{-\frac{\nu}{2\nu+\rho}}\right)}_{\text{online optimization error}} . \tag{4}$$

The proof is detailed in Appendix B.3.

Kernel k_Φ defined directly on the p-dimensional simulator parameters has the generic Matérn rate $\widetilde{O}(n^{p/(2\nu+p)})$, with $\rho \le p$. If $\mathcal{E}_w$ is an embedding, then $p = \rho$ and the two kernels have the same generic polynomial exponent. However, the pullback kernel is geometry-aware. In particular, two points may be far away in the simulator parameter space, $\|\phi - \phi'\| \gg 0$, but produce almost identical regimens, $\|\mathcal{E}_w(\phi) - \mathcal{E}_w(\phi')\| \approx 0$. While a Matérn kernel defined on Φ may regard these as nearly independent, the pullback kernel captures their similarity : $k_{\mathcal{D}}(\mathcal{E}_w(\phi), \mathcal{E}_w(\phi')) \approx k_{\mathcal{D}}(\mathcal{E}_w(\phi), \mathcal{E}_w(\phi))$.

5 SYNTHETIC EXPERIMENTS

We illustrate the trade-off of Section 4 on a problem where the dimension D of the design space and the search-space bias can be set independently. We embed the Branin function in $\mathcal{D} = [-1, 1]^D$ through the exact map $\mathcal{E}^\star(\phi) = (\phi, m(\phi))$, $\phi \in \Phi = [-1, 1]^2$, with a custom $m : \Phi \to \mathbb{R}^{D-2}$ (see equation 50, Appendix K). We minimize $f : \mathcal{D} \to \mathbb{R}$ defined as $f_\delta(\mathbf{x}) = \widetilde{B}(\mathbf{x}_{1:2}) + \|\mathbf{x}_{3:D} - m(\mathbf{x}_{1:2}) - \delta\, \mathbf{e}_1\|_2^2$, with $\widetilde{B}$ the Branin function rescaled to Φ and shifted to minimum 0. Since $\min_{\mathcal{D}} f_\delta = 0$ and $f_\delta(\mathcal{E}^\star(\phi)) = \widetilde{B}(\phi) + \delta^2$, the search-space bias is $\Delta_{\mathcal{M}} = \delta^2$.

Table 1: BO benchmark for the Branin function optimization, across search-space biases and ambient dimension. Simple regret at iteration 60, in units of 10^{-3}: mean $\pm$ s.e.m. over 25 seeds. $\Delta_{\mathcal{M}} = \delta^2$ (in units of 10^{-3}) is the smallest regret attainable by a method restricted to Φ, and is 0 for the methods that search $\mathcal{D}$ directly. Within each block, **bold** marks, for each D column, the lowest mean together with every method whose $\hat{m} \pm$ s.e.m. interval overlaps that of the lowest.

δ	Method	$\Delta_{\mathcal{M}}$	$D = 6$	$D = 8$	$D = 16$	$D = 32$
0	BOASS-k_Φ	0	$\mathbf{6.3 \pm 2.5}$	$\mathbf{6.3 \pm 2.5}$	$\mathbf{6.3 \pm 2.5}$	$\mathbf{6.3 \pm 2.5}$
	BO-on-$\mathcal{D}$	-	37.2 ± 7.1	74.7 ± 18.0	149.1 ± 26.0	687.5 ± 120.0
	SAASBO	-	$\mathbf{9.0 \pm 2.0}$	21.6 ± 3.7	177.0 ± 36.0	489.7 ± 72.0
	TuRBO-1	-	$\mathbf{7.7 \pm 1.6}$	31.2 ± 7.6	275.2 ± 48.0	609.8 ± 99.0
	Random search on Φ	0	12.1 ± 2.3	12.1 ± 2.3	12.1 ± 2.3	12.1 ± 2.3
0.25	BOASS-k_Φ	62.5	68.8 ± 2.5	68.8 ± 2.5	$\mathbf{68.8 \pm 2.5}$	$\mathbf{68.8 \pm 2.5}$
	BO-on-$\mathcal{D}$	-	41.5 ± 11.0	81.0 ± 20.0	154.7 ± 22.0	761.8 ± 110.0
	SAASBO	-	$\mathbf{11.4 \pm 2.3}$	$\mathbf{19.0 \pm 3.0}$	108.3 ± 17.0	534.0 ± 59.0
	TuRBO-1	-	$\mathbf{8.3 \pm 1.6}$	28.5 ± 6.2	292.2 ± 54.0	636.7 ± 110.0
	Random search on Φ	62.5	74.6 ± 2.3	74.6 ± 2.3	74.6 ± 2.3	74.6 ± 2.3
0.5	BOASS-k_Φ	250	256.3 ± 2.5	256.3 ± 2.5	256.3 ± 2.5	$\mathbf{256.3 \pm 2.5}$
	BO-on-$\mathcal{D}$	-	48.1 ± 8.2	69.7 ± 14.0	$\mathbf{184.1 \pm 31.0}$	853.8 ± 92.0
	SAASBO	-	$\mathbf{8.8 \pm 2.0}$	$\mathbf{22.2 \pm 4.2}$	$\mathbf{140.1 \pm 24.0}$	637.2 ± 57.0
	TuRBO-1	-	$\mathbf{7.3 \pm 1.2}$	40.3 ± 8.0	308.3 ± 51.0	764.7 ± 110.0
	Random search on Φ	250	262.1 ± 2.3	262.1 ± 2.3	262.1 ± 2.3	262.1 ± 2.3

We focus on BOASS with the kernel k_Φ (BOASS-k_Φ), and compare it with random search on Φ and with three methods searching $\mathcal{D}$: standard BO (BO-on-$\mathcal{D}$), SAASBO (Eriksson & Jankowiak, 2021) and TuRBO-1 (Eriksson et al., 2019). All methods are initialised in $\mathcal{M} = \mathcal{E}(\Phi)$.

Table 1 reports the simple regret after 60 evaluations, which for BOASS-k_Φ does not depend on D by construction. Without bias, BOASS-k_Φ has the lowest regret for all D. With a bias, BOASS-k_Φ cannot go below $\Delta_{\mathcal{M}} = \delta^2$ and the methods searching $\mathcal{D}$ gradually do better at smaller D as the bias increases. Restricting the search to $\mathcal{M}$ is thus beneficial as long as the bias is below the regret that a search over $\mathcal{D}$ reaches within the budget, and this regret grows with D. We report results with additional values of δ in Appendix K, together with further details and extra experiments on the amortization error and the comparison between the kernels $k_{\mathcal{D}}$ and k_Φ.

6 Use-case: Personalized Levodopa dosing for Parkinson's disease

As a concrete case study, we consider personalised dosing of Levodopa-Carbidopa , the most common treatment for Parkinson's disease (Dhall & Kreitzman, 2016; Houghton et al., 2019). As Parkinson's progresses, the therapeutic window (the safe and effective range of a drug's dose) narrows (Pezzoli & Zini, 2010), and advanced patients increasingly struggle to find a dosage that controls symptoms without triggering side effects, see Appendix C for further details. Traditionally, clinicians optimize dosing through a time-consuming trial-and-error titration process, guided by protocols and experience. We tackle here the problem of personalisation with Bayesian optimisation.

Problem setup. We fix the maximum number of daily drug intakes $M = 7$ and denote by $\mathbf{d} = [d_1, \dots, d_M]^\intercal$ and $\mathbf{t} = [t_1, \dots, t_M]^\intercal$ the dose amounts and the intake times. For a given patient, we aim at solving $(\mathbf{d}^\star, \mathbf{t}^\star) = \arg\max_{(\mathbf{d},\mathbf{t})\in\mathcal{D}} U(\mathbf{d}, \mathbf{t})$, where U is the black-box patient utility, a proxy for treatment efficacy. The observations collected from the patient, which we call *feedback*, can take different forms, such as symptoms rating or pairwise preference between two regimens $(\mathbf{d}_k, \mathbf{t}_k) \succ (\mathbf{d}_j, \mathbf{t}_j)$. We represent the space of admissible dosing regimens as $\mathcal{D} = \{(\mathbf{d}, \mathbf{t}) \mid 0 \le d_i \le d_{\max};\ 0 \le t_1,\ t_i < t_{i+1},\ t_M \le T\}$. Reflecting common clinical constraints for this treatment, we require intake times to be at least $\Delta = 60$ min apart and to lie within the intake window $[t_0, t_f]$, and dose amounts to lie in $[0, d_{\max}]$, with $d_{\max} = 250$ mg. Microtablets are available in increments of $\Delta_{\text{tab}} = 5$ mg, so we round each dose to the nearest multiple of Δ_{tab} before administering the regimen to the patient. We optimize dosing regimen sequentially, assuming washout of the effect of the previous regimen between one day and the next.

6.1 MECHANISTIC PK/PD MODELS FOR DOSING

As a differentiable mechanistic simulator, we use a PK/PD model, which describes how a dosing regimen translates into drug concentration over time (pharmacokinetics, PK) and, in turn, into clinical effect (pharmacodynamics, PD). Let $\phi \in \Phi \subset \mathbb{R}^p$ denote a vector of patient-specific PK/PD parameters (absorption and elimination rates, clearance, volumes of distribution, etc.) in the parameter space Φ. The PK model describes the concentration trajectories $\mathbf{c}(t;\phi)$ for $t > 0$ across the K compartments (e.g., plasma, effect site, ...) as the solution of a system of ODEs:

$$\begin{cases} c_1(0) = g(d;\phi), \quad c_i(0) = c_{i0}, \quad i = 2,\dots,K \\ \dot{\mathbf{c}}(t) = f_{\text{PK}}\big(\mathbf{c}(t);\phi\big), \end{cases} \tag{5}$$

where $c_1(t)$ is the concentration in the compartment where the dose is administered, $c_i(t)$ are the concentrations in the other compartments. The initial condition $g(d;\phi)$ maps the dose amount to an initial concentration in the dosing compartment. The PD output $r(t;\phi) = f_{\text{PD}}\big(c_e(t;\phi);\phi\big)$ maps the concentration in the effect compartment e, denoted by by $c_e(t)$, to a clinical quantity such as a symptom or treatment-response score. These models are estimated on clinical-trial data, and inter-individual variability in the parameters is modeled through random effects, whose distribution define $p(\phi)$ (Appendix G). For our experiments, we combine the PK model of Senek et al. (2018) with the PD model of Westin et al. (2011), whose output is the treatment response scale (TRS, ranging from $r = -3$ for parkinsonism to $r = +3$ for dyskinesia, with $r^\star = 0$ the optimal response).

Let $c^\star_\phi$ be the lowest drug concentration that achieves the target clinical state of the patient $r^\star$ under the patient's pharmacodynamics, $c^\star_\phi = \min\{c > 0 : f_{\text{PD}}\big(c;\phi\big) = r^\star\}$. We assume that $c^\star_\phi$ always exists, meaning that the patient can be efficiently treated with the drug. We define the model-optimal regimen as:

$$(\mathbf{d}^\star_\phi, \mathbf{t}^\star_\phi) \;=\; \underset{(\mathbf{d},\mathbf{t})\in\mathcal{D}}{\arg\min} \int_0^T |c_e(s;\phi \mid (\mathbf{d},\mathbf{t})) - c^\star_\phi|^2 \, ds, \tag{6}$$

where $c_e(s;\phi \mid (\mathbf{d},\mathbf{t}))$ is the concentration in e at time s induced by the regimen $(\mathbf{d},\mathbf{t})$. In practice, it is computed by integrating the ODE system piecewise between successive intakes, updating the initial conditions at each dosing time, see Appendix G.3 for the computation details.

6.2 OFFLINE TRAINING OF THE DOSING NETWORK

We start by training a *dosing neural network* that maps a patient's PK/PD parameters ϕ to the regimen the PK/PD model predicts to be optimal for that patient (panel A Figure S1). We define the training loss at ϕ as the average squared deviation of the induced concentration from the target $c^\star_\phi$ over the treatment horizon $[0,T]$, and approximate it by trapezoidal quadrature over a fine grid of $J+1$ time points $[0 = \tau_0,\dots,\tau_J = T]$:

$$\mathcal{L}(w;\phi) \;=\; \frac{1}{T}\int_0^T |c_e\left(s;\phi \mid \mathcal{E}_w(\phi)\right) - c^\star_\phi|^2 \, ds \approx \frac{1}{T}\sum_{j=1}^{J} \frac{\epsilon(\tau_j;\phi) + \epsilon(\tau_{j-1};\phi)}{2}(\tau_j - \tau_{j-1}), \tag{7}$$

with $\epsilon(\tau_j;\phi) = |c_e\left(\tau_j;\phi \mid \mathcal{E}_w(\phi)\right) - c^\star_\phi|^2$. We train $\mathcal{E}_w$ on a training set of virtual patients $\{\phi_i\}_{i=1}^N$ sampled from the population distribution $p(\phi)$, minimizing the Monte Carlo estimate of the average loss over the population $\mathcal{E}_{\phi\sim p(\phi)}\left[\mathcal{L}(w;\phi)\right] \;\approx\; \frac{1}{N}\sum_{i=1}^N \mathcal{L}(w;\phi_i)$. Since the data is generated synthetically, we can train the network with arbitrarily large N. Panel A in Figure S1 gives a schematic of the training procedure.

Training the network requires backpropagating the gradient of the loss equation 7 through the ODE solution. The network output enters the initial conditions of each integration interval (see equation 44, Appendix G.3). Following the recommendations of Kidger (2021; 2022), we take the discretize-then-optimize approach, which backpropagates through the discretized ODE solution produced by the numerical solver.

The network takes the PK/PD parameters ϕ as input and returns $2M$ values: the dose amounts $\{d_i\}_{i=1}^M$, the lags between consecutive intakes $\{t_i - t_{i-1} - \Delta\}_{i=2}^M$, and the slack $t_f - t_M$ left between the last intake and the end of the intake window. It is a feedforward network, trained on 20,000 samples from $p(\phi)$ (80/20 train/validation split). We provide more details on the architecture and the training procedure in Appendix D.

Table 2: The four evaluation settings differ in the query submitted to the patient, a pair of regimens in the *preference feedback* setting and a single regimen in the other three, and in the utility that generates the feedback, either PK/PD or misspecified. In the *preference feedback* setting, $\mathbf{U}$ and $\mathbf{U}'$ denote the utilities of the two regimens in the query.

	preference	concentration	misspecified	fully misspecified
Query	$\{(\mathbf{d}, \mathbf{t}), (\mathbf{d}', \mathbf{t}')\}$	$(\mathbf{d}, \mathbf{t})$	$(\mathbf{d}, \mathbf{t})$	$(\mathbf{d}, \mathbf{t})$
Feedback	$y \sim B\left[\Psi\left((\mathbf{U} + \epsilon - \mathbf{U}' - \epsilon')/\nu\right)\right]$	$y = \mathbf{U} + \epsilon$	$y = \mathbf{U} + \epsilon$	$y = \mathbf{U} + \epsilon$
Noise	$\epsilon, \epsilon' \overset{\text{iid}}{\sim} \mathcal{N}(0, 0.05)$	$\epsilon \sim \mathcal{N}(0, 0.1)$	$\epsilon \sim \mathcal{N}(0, 0.2)$	$\epsilon \sim \mathcal{N}(0, 0.2)$
Utility $\mathbf{U}$	$\mathbf{U}_{\text{pkpd}}$	$\mathbf{U}_{\text{pkpd}}$	$\mathbf{U}_{\text{oracle}}$	$\mathbf{U}_{\text{oracle}}$
Oracle regimen	–	–	perturbed $\mathcal{E}_w(\phi)$	random in $\mathcal{D}$
Maximizer of $\mathbf{U}$	in $\mathcal{M}$	in $\mathcal{M}$	outside $\mathcal{M}$	outside $\mathcal{M}$

6.3 Online Bayesian optimization

In this use case, the dose regimens have $D = 13$ free coordinates, and $p = \dim(\Phi) = 14$ PK/PD parameters. However, $\mathcal{M}$ is low dimensional due to structural non-identifiability of the PK/PD parameters from the optimal regimen. We provide provide a detailed analysis in Appendix H, together with a study of the Jacobian of $\mathcal{E}_w$ which gives us an estimate the intrinsic dimension of $\mathcal{M}$ between 3 and 5, with a bound $\dim(\mathcal{M}) \leq 7$. This makes this use case a relevant example of optimization over a low-dimensional amortized search space.

For the optimization, we add two dosage constraints of clinical relevance: a minimum daily dose of $D_{\min} = 320$ mg, and a variation of the total daily dose of at most 40% with respect to the previous regimen. The first constraint restricts $\mathcal{D}$ but not $\mathcal{M}$, as it is already satisfied by the regimens the PK/PD model deems optimal. The second one limits the search to total-dose neighborhoods of the previous regimen, as a form of clinical prudence relevant to realistic applications. We enforce both constraints for all methods by restricting the optimization of the acquisition function at each step (Appendix F).

6.4 Experiment settings and baselines

We evaluate the online optimization on 100 held-out patients, under two well-specified settings and two misspecified settings, summarized in Table 2. The settings differ in the query submitted to the patient, in the feedback returned, and in the latent utility generating it. In the well-specified settings, the utility $\mathbf{U}_{\text{pkpd}}$ is the negative concentration error induced by the regimen under the chosen PK/PD model, and in the misspecified settings is $\mathbf{U}_{\text{oracle}}$ the negative distance to an oracle regimen $(\mathbf{d}^{\text{or}}, \mathbf{t}^{\text{or}})$:

$$\begin{aligned} \mathbf{U}_{\text{pkpd}}(\mathbf{d}, \mathbf{t}, \phi) &= -\frac{1}{T} \int_0^T \left| c_e\left(s; \phi \mid (\mathbf{d}, \mathbf{t})\right) - c_\phi^\star \right|^2 ds, \\ \mathbf{U}_{\text{oracle}}(\mathbf{d}, \mathbf{t}) &= -\delta\left[(\mathbf{d}, \mathbf{t}), (\mathbf{d}^{\text{or}}, \mathbf{t}^{\text{or}})\right], \end{aligned} \tag{8}$$

where ϕ is the patient's true PK/PD parameter vector, and $\delta^2\left[(\mathbf{d}, \mathbf{t}), (\mathbf{d}', \mathbf{t}')\right] = \sum_{q=1}^{Q} \left(\sum_{i=1}^{M} d_i \mathbf{1}_{t_i \leq \lambda_q} - \sum_{i=1}^{M} d_i' \mathbf{1}_{t_i' \leq \lambda_q} \right)^2$ measures how much two cumulative dose profiles differ on an equispaced grid of time points $0 = \lambda_0 < \lambda_1 < \cdots < \lambda_Q = T$, with $Q = 20$. The four settings are the following:

1. **concentration feedback**: the feedback is a noisy observation of $\mathbf{U}_{\text{pkpd}}$;
2. **preference feedback**: the feedback is a pairwise comparison between two regimens, drawn from a probit model on $\mathbf{U}_{\text{pkpd}}$. At iteration n, the comparison is between the regimens tested at iterations n and $n-1$. That is, $y_n \sim B(p_n)$ with $p_n = P((\mathbf{d}_n, \mathbf{t}_n) \succ (\mathbf{d}_{n-1}, \mathbf{t}_{n-1})) = \Psi\left((U(\mathbf{d}_n, \mathbf{t}_n) - U(\mathbf{d}_{n-1}, \mathbf{t}_{n-1}))/\nu\right)$, where Ψ the normal c.d.f. , and $\nu = 1$;
3. **misspecified feedback**: the feedback is a noisy observation of $\mathbf{U}_{\text{oracle}}$, with a biased oracle regimen obtained by perturbing the doses of $\mathcal{E}_w(\phi)$, the regimen the network returns on the patient's true ϕ (Appendix F.3). By construction, the oracle dosage is outside, but close, to $\mathcal{M}$.
4. **fully misspecified feedback**: the feedback is as in the previous setting, but the oracle regimen is now sampled at random in $\{(\mathbf{d}, \mathbf{t}) \in \mathcal{D} : \sum_{i=1}^{M} d_i \geq D_{\min}\}$.

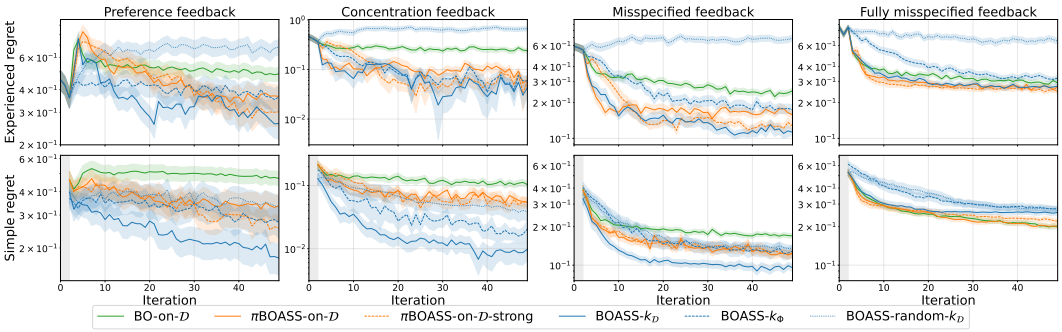

Figure 1: Results from the Levodopa-Carbidopa experiments (evaluation settings of Table 2). Top row: experienced regret to the oracle regimen. Bottom row: simple regret for the posterior mean maximizer at each iteration computed after the initialisation, represented by the grey band. Averages are over 100 simulated patients. Shaded bands: 10th-90th percentile of the estimated mean, assuming Gaussian regret.

The optimization loop starts with the output of the dosing network on the population-average PK/PD parameters, and two random samples in $\mathcal{M}$ lying in the 40%-total-dose neighborhood of that regimen. The optimization then runs for a total of 50 iterations, using log noisy EI as acquisition function (Ament et al., 2023) in all settings except *preference feedback*, where we use EUBO (Lin et al., 2022).

We compare the two kernel variants of our method to several baselines.Unlike in Section 5, we do not include TuRBO and SAASBO. TuRBO centres its trust region on the best observed regimen, while the daily dose variation constraint is relative to the previous regimen, so the trust region may contain no feasible points. The BoTorch[1] implementation of SAASBO does not support preference feedback, and we exclude it to keep the same set of methods in all settings. We consider the following methods:

- **BOASS-$k_{\mathcal{D}}$**, **BOASS-k_{Φ}**: BO over Φ with, respectively, the pullback kernel $k_{\mathcal{D}}\big(\mathcal{E}_w(\phi), \mathcal{E}_w(\phi')\big)$ and the kernel $k_{\Phi}(\phi, \phi')$ on the PK/PD parameters;
- **BOASS-random-$k_{\mathcal{D}}$**: BO over Φ with the pullback kernel $k_{\mathcal{D}}$ and random acquisition;
- **BO-on-$\mathcal{D}$**: BO over $\mathcal{D}$ with the kernel $k_{\mathcal{D}}$;
- **πBOASS-on-$\mathcal{D}$** and **πBOASS-on-$\mathcal{D}$-strong**: as a soft counterpart to the amortized search space, we optimize over $\mathcal{D}$ with a πBO prior over the location of the optimum (Hvarfner et al., 2022), informed by the dosing network. For each patient, the prior π is a Gaussian mixture model fitted to samples $\mathcal{E}(\phi_s)$ of dosing regimens, with $\phi_s \sim p(\phi)$. At iteration n, the acquisition function is $\alpha(\mathbf{x})\,\pi(\mathbf{x})^{\gamma_n}$, with $\gamma_n = \beta/n$. We use $\beta = 10$, following Hvarfner et al. (2022), for the strong prior (**-strong**), and $\beta = 0.808$, set by our own calibration procedure, see Appendix I for details. This variant can be seen as a robustified version of BOASS, which we expect would be less sensitive to search space bias;
- **Oracle parameters**: the parameters $\mathbf{x}$ returned by directly optimizing $\mathbf{x}$ on the patient-specific simulator with parameters ϕ, using L-BFGS-B optimizer. In the two misspecified settings, this is the biased oracle regimen.

For all methods, the search space, either $\mathcal{D}$ or Φ, is rescaled to the unit box (Appendix F), and $k_{\mathcal{D}}$ and k_{Φ} are Matérn 5/2 kernels with automatic relevance determination, over the regimen coordinates and over the PK/PD parameters. $k_{\mathcal{D}}$ is defined on the non-discretized output of the dosing network, although the patient feedback depends on the discretized doses. Differentiability is required for optimizing the acquisition function and fitting the GP hyperparameters. Because the discretization interval is small relative to the range of clinically relevant dose variation, we expect the resulting kernel approximation to be minor. All reported utilities and regrets are evaluated after rounding. GP hyperparameters are re-estimated at each iteration by maximizing the marginal likelihood.

6.5 RESULTS

The dosing amortized optimization yields results nearly indistinguishable from direct optimization using L-BFGS-B, at a fraction of the computational cost (see analysis in Appendix E.1).

In our use-case, two forms of regret are relevant, both measured at each iteration as a utility gap to the oracle regimen: the *experienced regret* refers to the regimen prescribed to the patient, the *simple regret* to the estimated optimal regimen, i.e. the posterior mean maximizer. The posterior mean is maximized under the constraint that the daily dose is greater than the minimum daily dose D_{min}. Figure 1 reports the two regrets over the 100 held-out patients. In the two well-specified settings, *concentration feedback* and *preference feedback*, optimizing over the simulator-induced search space (methods *BOASS-**) brings a clear advantage on both regrets over all baselines. The *preference feedback* setting gives the worst experienced regret trajectories for every method, which is expected since the surrogate is harder to learn from pairwise comparisons than from direct observations of the utility.

The two misspecified settings quantify the cost of restricting the search to $\mathcal{M}$. In the *misspecified feedback* setting, where the optimum lies outside $\mathcal{M}$ by construction, BOASS variants are still the best performing. Here the optimum restricted to $\mathcal{M}$ is close enough to the true one that the cost of the restriction is negligible compared to the gain in sample efficiency obtained by searching a lower-dimensional space. The *fully misspecified feedback* setting introduces a stronger bias: BOASS- $k_{\mathcal{D}}$ and BOASS-$k_{\oplus}$ converge to a suboptimal regimen and are outperformed by the baselines optimizing over the full space $\mathcal{D}$. The experienced regret of BOASS, however, matches that of the best baselines. πBOASS is more robust to strong misspecification, as its search and the posterior mean maximizer are not restricted to $\mathcal{M}$. Between the two kernel options, *BOASS-$k_{\mathcal{D}}$* is consistently better than *BOASS-k_{Φ}*. The mapping is not injective, and the pullback kernel $k_{\mathcal{D}}$ accounts for this property whereas the kernel on Φ does not.

7 DISCUSSION

In this paper, we introduced BOASS, a BO approach that leverages mechanistic simulators to effectively reduce the search to a lower dimensional subspace, while preserving the initial problem geometry for learning GP surrogates. We show through theoretical analysis, synthetic experiments and a concrete case study of individualized drug dosing, that BOASS and πBOASS improve sample-efficiency of BO, outperforming all baselines. Performance gains can be explained by the fact that the maximum information gain is dependent on the amortized search space intrinsic dimension. Sharper regret bounds could be obtained when the simulator-optimal parameters locally vary along only fewer independent directions than the number of mechanistic model input parameters. We leave the theoretical analysis of these sharper bounds for future work.

However, because the search is confined to the subspace, a strongly misspecified model can limit the search to a subset excluding the true optimum. The method thus trades-off search space bias with search space dimensionality. We showed it remains relatively robust to misspecification of the simulator, the advantage of searching a lower-dimensional space over a finite time horizon being able to compensate for the potential bias introduced. We also introduce a variant, πBOASS, that is more robust to such misspecification.

BOASS is especially relevant in a setting of moderate dimensionality with very small evaluation budget, for which standard high-dimensional BO methods are not suited. Moreover, the use of such an amortized search space could be generalized beyond BO, e.g. for evolutionary optimization, in multi-arm bandits or in Reinforcement Learning.

In the case of Parkinson's disease, most previous work on algorithmic dosing and prescription is retrospective or in silico. However, a performance gap may occur when going from simulation to real-life. A logical next step is a clinical trial of our dosing schedule recommendation system with physician oversight, with a control group following a dosing schedule designed by physician's standard practice in real-life patients.

[1] We use BoTorch (Balandat et al., 2020) for implementing all methods.

8 AI USE STATEMENT

LLM-based coding assistants were used for the development and review of the code. Additionally, commercially available LLMs were used for manuscript proofreading, for refining the design of the synthetic experiments, for complementing our literature review, and for providing feedback on the manuscript. LLMs were used for drafting some sections of the appendix, before careful editing by the authors. The authors reviewed all AI-assisted work and take full responsibility for the final content of this work.

9 ETHICS STATEMENT

This work focuses on leveraging differentiable mechanistic simulators for more efficient global optimization of black-box objective functions. In particular, it presents an application of the proposed method for personalizing medication dosing, with the potential benefit of reducing the burden and duration of empirical dose titration. The present study uses only simulated patients generated from previously published aggregate PK/PD models. It involves no human participants, identifiable patient information, or individual-level clinical records, and therefore did not require informed consent or institutional review board approval.

The proposed method is evaluated exclusively in silico and is not validated for clinical use. Errors in the mechanistic model, noisy or biased patient feedback, distribution shifts, and optimization outside the model's domain of validity could all lead to inappropriate recommendations, potentially resulting in ineffective treatment or adverse effects.

In particular, the population PK/PD models used in this work were estimated from particular study cohorts and may not represent all patient groups equally. The safety constraints considered in this study are not comprehensive, and should therefore not be interpreted as clinical safety guarantees.

Any prospective use would require independent evaluation on representative patient populations, approval by the relevant ethics and regulatory bodies, and a carefully controlled clinical study with informed consent. Recommendations should remain subject to qualified clinician review rather than being administered autonomously.

10 REPRODUCIBILITY STATEMENT

All empirical results are based on synthetic experiments or simulated patients. No proprietary data, individual-level clinical records, or human participants are used. Virtual patients are generated from the published aggregate pharmacokinetic and pharmacodynamic models cited in the paper. The manuscript specifies the model equations and parameter distributions, BOASS algorithm, neural-network architecture and training procedure, data-generation and pruning rules, search-space specification, numerical ODE solver and tolerances, acquisition-function optimization, Gaussian process fitting, baselines, evaluation protocols, sample sizes, and uncertainty summaries. These details are provided in Sections 3–5 and Appendices A and C–H. The implementation uses PyTorch, JAX, diffrax, BoTorch, and GPyTorch.

We have documented in details the instantiation of the Levodopa-Carbidopa use-case in Appendix A-D-E-F-G-I — architecture, training procedure, solver settings, search space, baselines, and evaluation protocol, and for the synthetic experiments in Appendix K. The code to reproduce all experiments is submitted with the manuscript and will be released in a public repository upon acceptance, specifying the exact package versions and random seeds.

REFERENCES

T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019.

S. Ament, S. Daulton, D. Eriksson, M. Balandat, and E. Bakshy. Unexpected improvements to expected improvement for bayesian optimization. *Advances in neural information processing systems*, 36:20577–20612, 2023.

B. Amos. Tutorial on amortized optimization. *Foundations and Trends in Machine Learning*, 16(5): 592–732, 2023.

R. Antonova, A. Rai, T. Li, and D. Kragic. Bayesian optimization in variational latent spaces with dynamic compression. In *Proceedings of the Conference on Robot Learning*, volume 100, pp. 456–465. PMLR, 2020.

M. Balandat, B. Karrer, D. Jiang, S. Daulton, B. Letham, A. G. Wilson, and E. Bakshy. BoTorch: A framework for efficient Monte-Carlo Bayesian optimization. *Advances in neural information processing systems*, 33:21524–21538, 2020.

T. Boltz, J. L. Martinez, C. Xu, K. R. Baker, Z. Zhu, J. Morgan, R. Roussel, D. Ratner, B. Mustapha, and A. L. Edelen. Leveraging prior mean models for faster Bayesian optimization of particle accelerators. *Scientific Reports*, 15(1):12232, 2025.

J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, Y. Katariya, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang. JAX: Composable transformations of Python+NumPy programs, 2018.

E. Brochu, T. Brochu, and N. de Freitas. A Bayesian interactive optimization approach to procedural animation design. In *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '10, pp. 103–112, Goslar, DEU, 2010a. Eurographics Association.

E. Brochu, V. M. Cora, and N. de Freitas. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning, 2010b.

P. L. S. Chan, J. G. Nutt, and N. H. G. Holford. Importance of Within Subject Variation in Levodopa Pharmacokinetics: A 4Year Cohort Study in Parkinson's Disease. *Journal of Pharmacokinetics and Pharmacodynamics*, 32(3):307–331, 2005.

E. C. M. de Lange and M. Hammarlund-Udenaes. Understanding the blood-brain barrier and beyond: Challenges and opportunities for novel CNS therapeutics. *Clinical Pharmacology & Therapeutics*, 111(4):758–773, 2022.

R. Dhall and D. L. Kreitzman. Advances in levodopa therapy for Parkinson disease: Review of RYTARY (carbidopa and levodopa) clinical efficacy and safety. *Neurology*, 86(14 Suppl 1):S13–24, 2016.

Y. Dorn, C. Chen, and N. Xie. Manifold-Aware Information Gain and Lower Bounds for Gaussian-Process Bandits on Riemannian Quotient Spaces, 2026.

D. Eriksson and M. Jankowiak. High-dimensional Bayesian optimization with sparse axis-aligned subspaces. In C. de Campos and M. H. Maathuis (eds.), *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, volume 161 of *Proceedings of Machine Learning Research*, pp. 493–503. PMLR, 2021.

D. Eriksson, M. Pearce, J. Gardner, R. D. Turner, and M. Poloczek. Scalable global optimization via local bayesian optimization. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

T. Fauvel and M. Chalk. Human-in-the-loop optimization of visual prosthetic stimulation. *Journal of Neural Engineering*, 19(3):036038, 2022.

J. P. Folch, C. Tsay, R. M. Lee, B. Shafei, W. Ormaniec, A. Krause, M. van der Wilk, R. Misener, and M. Mutnỳ. Transition constrained Bayesian optimization via Markov decision processes. *Advances in Neural Information Processing Systems*, 37:88194–88235, 2024.

P. I. Frazier. A Tutorial on Bayesian Optimization, 2018.

J. Gardner, G. Pleiss, K. Q. Weinberger, D. Bindel, and A. G. Wilson. Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration. *Advances in neural information processing systems*, 31, 2018.

R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*, 4(2):268–276, 2018.

J. Granley, T. Fauvel, M. Chalk, and M. Beyeler. Human-in-the-loop optimization for deep stimulus encoding in visual prostheses. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, pp. 79376–79398, Red Hook, NY, USA, 2023. Curran Associates Inc.

M. Hammarlund-Udenaes. Active-site concentrations of chemicals—are they a better predictor of effect than plasma/organ/tissue concentrations? *Basic & Clinical Pharmacology & Toxicology*, 106(3):215–220, 2010.

R. A. Hauser, F. Deckers, and P. Lehert. Parkinson's disease home diary: Further validation and implications for clinical trials. *Movement Disorders: Official Journal of the Movement Disorder Society*, 19(12):1409–1413, 2004.

R. A. Hauser, D. E. Kremens, L. W. Elmer, D. L. Kreitzman, R. R. Walsh, R. Johnson, R. Howard, J. T. Nguyen, and R. Patni. Prevalence of Dyskinesia and OFF by 30-Minute Intervals Through the Day and Assessment of Daily Episodes of Dyskinesia and OFF: Novel Analyses of Diary Data from Gocovri Pivotal Trials. *Journal of Parkinson's Disease*, 9(3):591–600, 2019.

R. Houghton, F. Boess, L. Verselis, Y. Ding, R. Freitas, N. Constantinovici, and R. Ong. Treatment Patterns in Patients with Incident Parkinson's Disease in the United States. *Journal of Parkinson's Disease*, 9(4):749–759, 2019.

C. Hvarfner, D. Stoll, A. Souza, M. Lindauer, F. Hutter, and L. Nardi. πBO: Augmenting acquisition functions with user beliefs for bayesian optimization. In *International Conference on Learning Representations*, 2022.

C. Hvarfner, E. O. Hellsten, and L. Nardi. Vanilla Bayesian Optimization Performs Great in High Dimensions. In *Proceedings of the 41st International Conference on Machine Learning*, pp. 20793–20817. PMLR, 2024a.

C. Hvarfner, F. Hutter, and L. Nardi. A general framework for user-guided Bayesian optimization. In *International Conference on Learning Representations*, 2024b.

J. Jankovic and E. K. Tan. Parkinson's disease: Etiopathogenesis and treatment. *Journal of Neurology, Neurosurgery & Psychiatry*, 91(8):795–808, 2020.

D. Janz, A. Akhavan, and A. B. Tsybakov. Maximum effective dimension and information gain, 2026.

K. Kandasamy, J. Schneider, and B. Poczos. High Dimensional Bayesian Optimisation and Bandits via Additive Models. In *Proceedings of the 32nd International Conference on Machine Learning*, pp. 295–304. PMLR, 2015.

K. Kandasamy, G. Dasarathy, J. B. Oliva, J. Schneider, and B. Póczos. Gaussian process bandit optimisation with multi-fidelity evaluations. In *Advances in Neural Information Processing Systems*, volume 29, pp. 992–1000, 2016.

P. Kidger. *On Neural Differential Equations*. PhD thesis, University of Oxford, 2021.

P. Kidger. Diffrax documentation: Adjoints, 2022.

Z. J. Lin, R. Astudillo, P. Frazier, and E. Bakshy. Preference exploration for efficient Bayesian optimization with multiple outcomes. In *International Conference on Artificial Intelligence and Statistics*, pp. 4235–4258. PMLR, 2022.

I. Loshchilov and F. Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019.

N. Maus, H. Jones, J. Moore, M. J. Kusner, J. Bradshaw, and J. Gardner. Local latent space Bayesian optimization over structured inputs. In *Advances in Neural Information Processing Systems*, volume 35, pp. 34505–34518, 2022.

H. B. Moss, D. S. Leslie, J. Gonzalez, and P. Rayson. GIBBON: General-purpose information-based Bayesian optimisation. *Journal of Machine Learning Research*, 22(235):1–49, 2021.

D. Nyholm, A. I. Nilsson Remahl, N. Dizdar, R. Constantinescu, B. Holmberg, R. Jansson, S. M. Aquilonius, and H. Askmark. Duodenal levodopa infusion monotherapy vs oral polypharmacy in advanced Parkinson disease. *Neurology*, 64(2):216–223, 2005.

A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

J. A. Paulson, F. Sorouifar, and A. Chakrabarty. Efficient multi-step lookahead Bayesian optimization with local search constraints. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 123–129. IEEE, 2022.

G. Pezzoli and M. Zini. Levodopa in Parkinson's disease: From the past to the future. *Expert Opinion on Pharmacotherapy*, 11(4):627–635, 2010.

A. Rai, R. Antonova, F. Meier, and C. G. Atkeson. Using simulation to improve sample-efficiency of Bayesian optimization for bipedal robots. *Journal of Machine Learning Research*, 20(49):1–24, 2019.

C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. MIT Press, 2006.

P. Rosa. {L^2}-posterior contraction rates for Gaussian process and random series priors in Bayesian nonparametric regression models, 2025.

P. Rosa, V. Borovitskiy, A. Terenin, and J. Rousseau. Posterior Contraction Rates for Matérn Gaussian Processes on Riemannian Manifolds. In *Neural Information Processing System*, 2023.

M. Senek, D. Nyholm, and E. I. Nielsen. Population pharmacokinetics of levodopa/carbidopa microtablets in healthy subjects and Parkinson's disease patients. *European journal of clinical pharmacology*, 74(10):1299–1307, 2018.

B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE*, 104(1):148–175, 2016.

N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the 27th International Conference on Machine Learning (ICML)*, 2010.

S. Stanton, W. Maddox, N. Gruver, P. Maffettone, E. Delaney, P. Greenside, and A. G. Wilson. Accelerating Bayesian optimization for biological sequence design with denoising autoencoders. In *International Conference on Machine Learning*, pp. 20459–20478. PMLR, 2022.

N. Titova and K. R. Chaudhuri. Personalized medicine in Parkinson's disease: Time to be precise. *Movement Disorders*, 32(8):1147–1154, 2017.

A. Tripp, E. Daxberger, and J. M. Hernández-Lobato. Sample-efficient optimization in the latent space of deep generative models via weighted retraining. In *Advances in Neural Information Processing Systems*, volume 33, pp. 11259–11272, 2020.

S. Vakili, K. Khezeli, and V. Picheny. On information gain and regret bounds in Gaussian process bandits. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2021.

Z. Wang, M. Zoghi, F. Hutter, D. Matheson, and N. De Freitas. Bayesian optimization in high dimensions via random embeddings. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, IJCAI '13, pp. 1778–1784, Beijing, China, 2013. AAAI Press.

Z. Wang, F. Hutter, M. Zoghi, D. Matheson, and N. de Freitas. Bayesian optimization in a billion dimensions via random embeddings. *Journal of Artificial Intelligence Research*, 55:361–387, 2016.

J. Westin, D. Nyholm, S. Pålhagen, T. Willows, T. Groth, M. Dougherty, and M. O. Karlsson. A pharmacokinetic-pharmacodynamic model for duodenal levodopa infusion. *Clinical neuropharmacology*, 34(2):61–65, 2011.

Appendices

A Method summary **16**

B Regret bound for Bayesian optimization on Amortized Search Space **16**

B.1 Regret decomposition 16

B.2 Reduced information gain 17

B.3 Error introduced by amortization 18

B.4 Intrinsic-dimension-dependent information gain 19

B.5 Pullback kernel vs kernel defined in the simulator parameters space 20

C Challenges in Levodopa treatment for Parkinson's disease **20**

D Dosing network architecture and training details for Levodopa use-case **21**

E Training data **22**

E.1 Amortization analysis for Levodopa-Carbidopa dosing use-case 22

F Bayesian optimization details **23**

F.1 Search space and constraints 23

F.2 Implementation details 24

F.3 Evaluation settings 24

G PK/PD model for Levodopa-Carbidopa in Parkinson's disease **25**

G.1 Pharmacokinetic (PK) model 25

G.2 Pharmacodynamic (PD) model 27

G.3 PKPD solution under multiple intakes 27

H Intrinsic dimension of the amortized search space **27**

I Bayesian optimization with πBOASS **28**

J Supplementary results **30**

K Synthetic experiments **30**

K.1 Branin embedded in an ambient space 30

K.1.1 Comparison with high-dimensional BO methods and bias analysis 32

K.1.2 Learned embedding and amortization error 32

K.2 Objective drawn from a Gaussian process 36

Algorithm 1 Bayesian optimization on amortized search-space for precision dosing

Require: population distribution $p(\phi)$; differentiable simulator $(f_{\text{PK}}, h, f_{\text{PD}})$; target effect $r^\star$; training size N; step size η; GP kernel k; acquisition function α; online budget T_{BO}

Offline: train the dosing network
sample virtual patients $\{\phi_i\}_{i=1}^N \sim p(\phi)$
for $i = 1, \dots, N$ **do**
  $c^\star_{\phi_i} \leftarrow \min\{c > 0 : f_{\text{PD}}(c; \phi_i) = r^\star\}$ {target concentration}
end for
initialize network weights w
repeat
  sample a minibatch $B \subseteq \{1, \dots, N\}$
  for $i \in B$ **do**
    $(\mathbf{d}, \mathbf{t}) \leftarrow \mathcal{E}_w(\phi_i)$; integrate the PK/PD ODEs equation 44 and evaluate the loss $\mathcal{L}(w; \phi_i)$ equation 7
  end for
  $w \leftarrow w - \eta \nabla_w \frac{1}{|B|} \sum_{i \in B} \mathcal{L}(w; \phi_i)$ {backpropagate through the ODE solver}
until convergence
the trained network defines the search space $\mathcal{M} = \{\mathcal{E}_w(\phi) : \phi \in \Phi\}$
Online: personalize the regimen for a patient
initialize the history $\mathcal{H} \leftarrow \emptyset$ {optionally seed with the population mean $\bar{\phi}$}
for $t = 0, 1, \dots, T_{\text{BO}} - 1$ **do**
  fit the GP surrogate with kernel k to $\mathcal{H}$
  $\phi_{t+1} \leftarrow \arg\max_{\phi \in \Phi} \alpha(\phi; \mathcal{H})$ {acquisition step}
  administer the regimen $(\mathbf{d}_{t+1}, \mathbf{t}_{t+1}) = \mathcal{E}_w(\phi_{t+1})$ to the patient
  observe feedback y_{t+1} {symptom score, wearable-derived signal, preference, ...}
  $\mathcal{H} \leftarrow \mathcal{H} \cup \{(\phi_{t+1}, y_{t+1})\}$
end for

Ensure: recommended regimen $\mathcal{E}_w(\phi^\star)$, with $\phi^\star$ the best candidate in $\mathcal{H}$

A Method summary

Algorithm 1 summarizes the two stages of the BOASS method in the precision drug dosing context. Offline, the dosing network $\mathcal{E}_w$ is trained against the PK/PD model (Section 6.2). Online, Bayesian optimization searches over the network inputs ϕ, using $\mathcal{E}_w$ to turn each candidate into a regimen administered to the patient (Section 6.3).

The method is illustrated in the context of individualized patient-in-the-loop drug dosing in Figure S1.

B Regret bound for Bayesian optimization on Amortized Search Space

B.1 Regret decomposition

Restricting Bayesian optimization to the image of the simulator-inverting network reduces the set of candidate parameters that must be explored, but may exclude the parameters that maximize the true utility. We make this trade-off explicit below.

Let $\mathbf{x} \in \mathcal{D}$ denote the system's parameters being optimized, and let $U(\mathbf{x})$ be the unknown patient-specific utility. Let $\mathcal{M} = \{\mathcal{E}_w(\phi) : \phi \in \Phi\} \subseteq \mathcal{D}$ denote the search space induced by the simulator inversion neural network. This is the set over which the online BO stage operates.

Let $\mathbf{x}^\star \in \arg\max_{\mathbf{x} \in \mathcal{D}} U(\mathbf{x})$ and $\mathbf{x}^\star_{\mathcal{M}} \in \arg\max_{\mathbf{x} \in \mathcal{M}} U(\mathbf{x})$ be respectively the globally optimal system parameters and the best system parameters contained in the learned search space. We define the search-space bias as :

$$\Delta_{\mathcal{M}} = U(\mathbf{x}^\star) - U(\mathbf{x}^\star_{\mathcal{M}}) \geq 0, \tag{9}$$

with equality when $\mathbf{x}^\star \in \mathcal{M}$.

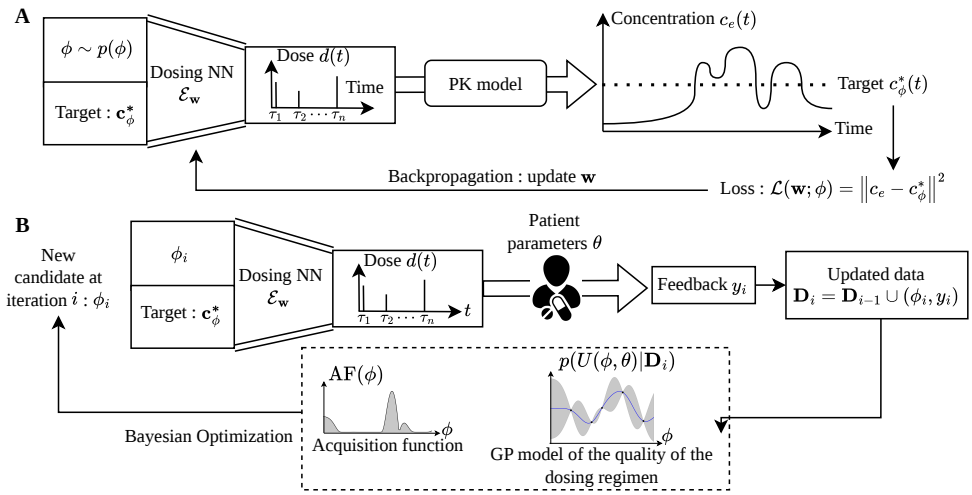

Figure S1: Illustration of the BOASS method in an application to individualized drug dosing. **(A)** The dosing network $\mathcal{E}_w$ is trained offline against the PK/PD model; **(B)** Online, BO searches over patient parameters ϕ.

Suppose that after n online evaluations, BO has tested regimens $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathcal{M}$. In the noise-free scenario, the reported solution is typically $\max_{i \le n} U(\mathbf{x}_i)$. Alternatively, in practice, one may report instead a maximum of the posterior mean $\widehat{\mathbf{x}}_n \in \arg\max_{\mathbf{x} \in \mathcal{M}} \mu_n(\mathbf{x})$, where μ_n is the posterior mean of the GP surrogate. The corresponding regret is a distinct quantity and is not directly controlled by the standard cumulative-regret argument used below. Our theoretical results therefore concern the best evaluated point.

Its global simple regret admits the decomposition :

$$\begin{aligned} r_n &= U(\mathbf{x}^\star) - \max_{i \le n} U(\mathbf{x}_i) \\ &= \underbrace{U(\mathbf{x}^\star) - U(\mathbf{x}^\star_{\mathcal{M}})}_{\text{search-space bias}} + \underbrace{U(\mathbf{x}^\star_{\mathcal{M}}) - U(\widehat{\mathbf{x}}_n)}_{\text{optimization error on } \mathcal{M}} \\ &= \Delta_{\mathcal{M}} + r_n^{\mathcal{M}}. \end{aligned} \tag{10}$$

Equation equation 10 captures the central motivation for the method: amortized search-space restriction may introduce an irreducible bias $\Delta_{\mathcal{M}}$, but it can simultaneously make the remaining optimization problem substantially easier.

B.2 Reduced information gain

Srinivas et al. (2010) established regret bounds for the GP-UCB acquisition function with iid Gaussian noise, and showed that regret is controlled by the maximum information gain. For a domain A and kernel k, the maximum information gain is defined as:

$$\gamma_n(A;k) = \max_{\mathbf{x}_{1:n} \in A^n} I(y_{1:n}; U_{1:n}), \tag{11}$$

where $U_{1:n} = (U(\mathbf{x}_1), \ldots, U(\mathbf{x}_n))$, and $y_{1:n}$ denotes noisy observations.

They showed that for a general compact and convex domain, for $f \sim GP(0, k(\mathbf{x}, \mathbf{x}'))$ (under some derivative-tail condition) : $R_n = \widetilde{O}(\sqrt{n\beta_n\gamma_n})$.

$\mathcal{M}$ is not guaranteed to be convex. Therefore, we instead define the objective in the simulator-induced coordinates as $g(\phi) = U(\mathcal{E}_w(\phi))$ and use the pullback kernel $k_{\text{pb}}(\phi, \phi') = k_{\mathcal{D}}(\mathcal{E}_w(\phi), \mathcal{E}_w(\phi'))$. We assume for our theoretical analysis that $\Phi \subset [0, r]^p$ is compact and convex, where $p = \dim(\Phi)$, and that observations are of the form $y_t = g(\phi_t) + \epsilon_t$, with independent Gaussian noise $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$.

We further assume that the pullback GP satisfies the derivative-tail condition of Srinivas et al. (2010): for some constants $a, b > 0$,

$$\Pr\left(\sup_{\phi\in\Phi}\left|\frac{\partial g(\phi)}{\partial \phi_j}\right| > L\right) \le a\exp\left(-\left(\frac{L}{b}\right)^2\right), \qquad j = 1,\dots,p. \tag{12}$$

Under these assumptions, Theorem 2 of Srinivas et al. (2010) implies that, for an appropriate exploration sequence β_n, GP-UCB gives a regret bound of $\widetilde{O}\left(\sqrt{n\beta_n\gamma_n(\Phi;k_{\mathrm{pb}})}\right)$.

(Here and below, the $\widetilde{O}$ notation is a variant of O, where log factors are suppressed).

Since $\mathcal{M} = \mathcal{E}_w(\Phi)$ and $k_{\mathrm{pb}}(\phi,\phi') = k_{\mathcal{D}}(\mathcal{E}_w(\phi),\mathcal{E}_w(\phi'))$, the two domains generate exactly the same set of kernel Gram matrices. Hence,

$$\gamma_n(\Phi;k_{\mathrm{pb}}) = \gamma_n(\mathcal{M};k_{\mathcal{D}}). \tag{13}$$

Therefore cumulative regret satisfies :

$$R_n^{\mathcal{M}} = \widetilde{O}\left(\sqrt{n\beta_n\gamma_n(\mathcal{M};k_{\mathcal{D}})}\right), \tag{14}$$

and hence the simple regret of the best evaluated point satisfies :

$$r_n^{\mathcal{M}} \le \frac{R_n^{\mathcal{M}}}{n} = \widetilde{O}\left(\sqrt{\frac{\beta_n\gamma_n(\mathcal{M};k_{\mathcal{D}})}{n}}\right). \tag{15}$$

Because $\mathcal{M} \subseteq \mathcal{D}$, the set of admissible n-point designs in $\mathcal{M}$ is a subset of those in $\mathcal{D}$. Therefore,

$$\gamma_n(\mathcal{M};k_{\mathcal{D}}) \le \gamma_n(\mathcal{D};k_{\mathcal{D}}). \tag{16}$$

Combining Equations equation 10 and equation 15 yields :

$$r_n \le \Delta_{\mathcal{M}} + \widetilde{O}\left(\sqrt{\frac{\beta_n\gamma_n(\mathcal{M};k_{\mathcal{D}})}{n}}\right). \tag{17}$$

Equation equation 17 makes the bias-complexity trade-off explicit: the learned search space is advantageous at a finite evaluation budget whenever the reduction in optimization error outweighs the approximation error introduced by excluding parts of $\mathcal{D}$.

Note that when BO is run over Φ with the pullback kernel, its statistical complexity is that of the simulator-induced set of system parameters rather than that of the raw parameterization Φ. This also explains why the pullback kernel is insensitive to redundant coordinates: distinct ϕ, ϕ' satisfying $\mathcal{E}_w(\phi) = \mathcal{E}_w(\phi')$ are identical from the perspective of the GP.

B.3 Error introduced by amortization

The previous argument treats $\mathcal{M}$ as given. In practice, it is itself learned by amortizing the simulator-defined optimization problem in Equation 3.

Let $\mathcal{E}^\star(\phi)$ denote an exact solution of the simulator optimization problem, and define the corresponding ideal simulator-induced search space as $\mathcal{M}^\star = \{\mathcal{E}^\star(\phi) : \phi \in \Phi\}$.

The dosing network $\mathcal{E}_w$ approximates $\mathcal{E}^\star$, producing $\mathcal{M} = \mathcal{E}_w(\Phi)$. Suppose that U is L_U-Lipschitz with respect to a regimen distance δ, i.e. $|U(\mathbf{x}) - U(\mathbf{x}')| \le L_U\delta(\mathbf{x},\mathbf{x}')$, and that the amortization error is uniformly bounded as :

$$\sup_{\phi\in\Phi}\delta\left(\mathcal{E}_w(\phi),\mathcal{E}^\star(\phi)\right) \le \varepsilon_{\mathrm{am}}. \tag{18}$$

Define the error due solely to simulator misspecification as :

$$\Delta_{\mathrm{model}} = U(\mathbf{x}^\star) - \max_{\mathbf{x}\in\mathcal{M}^\star} U(\mathbf{x}). \tag{19}$$

Then

$$\Delta_{\mathcal{M}} \leq \Delta_{\text{model}} + L_U \varepsilon_{\text{am}}. \tag{20}$$

Indeed, if $\phi^\star$ produces the best regimen in $\mathcal{M}^\star$, then Lipschitz continuity implies $U(\mathcal{E}_w(\phi^\star)) \geq U(\mathcal{E}^\star(\phi^\star)) - L_U \varepsilon_{\text{am}}$, while the best point in $\mathcal{M}$ can only have higher utility than $\mathcal{E}_w(\phi^\star)$.

Combining Equations equation 17 and equation 20 gives the overall decomposition :

$$r_n \leq \underbrace{\Delta_{\text{model}}}_{\substack{\text{simulator} \\ \text{misspecification}}} + \underbrace{L_U \varepsilon_{\text{am}}}_{\substack{\text{amortization} \\ \text{error}}} + \underbrace{\widetilde{O}\left(\sqrt{\frac{\beta_n \gamma_n(\mathcal{M}; k_{\mathcal{D}})}{n}}\right)}_{\substack{\text{online} \\ \text{optimization error}}}. \tag{21}$$

Equation equation 21 separates the three sources of error in our method. A more accurate mechanistic model reduces Δ_{model}, a better-trained dosing network reduces ε_{am}, and a more structured, lower-complexity search space reduces the number of online evaluations required by Bayesian optimization.

B.4 Intrinsic-dimension-dependent information gain

The set-inclusion argument in Equation equation 16 guarantees that restricting the search to $\mathcal{M}$ cannot increase the maximum information gain. A sharper characterization can be obtained from the spectral decay of the kernel on the restricted search space.

Let $\{\lambda_j\}_{j \geq 1}$ denote the Mercer eigenvalues of $k_{\mathcal{D}}$ restricted to $\mathcal{M}$ with respect to the Riemannian volume measure on $\mathcal{M}$, ordered such that $\lambda_1 \geq \lambda_2 \geq \cdots$.

Assume additionally that the Mercer eigenfunctions satisfy the boundedness condition required by Vakili et al. (2021). If, for some $\beta > 1$, the eigenvalues satisfy the polynomial decay $\lambda_j \leq C j^{-\beta}$, then

$$\gamma_n(\mathcal{M}; k_{\mathcal{D}}) = O\left(n^{1/\beta} (\log n)^{1-1/\beta}\right). \tag{22}$$

Thus, the information-gain rate is controlled by the spectral complexity of the kernel on the simulator-induced search space. In particular, if restriction to $\mathcal{M}$ induces faster eigenvalue decay than on $\mathcal{D}$, Equation equation 22 gives a sharper rate.

A useful sufficient condition for such faster decay can be obtained for Matérn kernels when the simulator-induced search space has additional geometric regularity. Vakili et al. (2021) use the fact that for a Matérn kernel with smoothness parameter $\nu > \frac{1}{2}$, $\lambda_m = O\left(m^{-\frac{2\nu+D}{D}}\right)$, and for the Squared Exponential kernel, $\lambda_m = O\left(\exp(-m^{\frac{1}{D}})\right)$. Substituting into Equation equation 22 leads to bounds on the maximum information gain $O\left(n^{\frac{D}{2\nu+D}} (\log n)^{\frac{2\nu}{2\nu+D}}\right)$ and $O\left(\log^{D+1}(n)\right)$, respectively.

Now, suppose that $\mathcal{M} \subset \mathcal{D} \subset \mathbb{R}^D$ is a compact, smooth ρ-dimensional embedded submanifold without boundary, and that $k_{\mathcal{D}}$ is a Matérn kernel with smoothness parameter ν. Proposition 27 of Rosa et al. (2023) establishes that the RKHS $\mathcal{H}_k$ of the restricted extrinsic Matérn kernel is norm-equivalent to the Sobolev space $H^{\nu+\rho/2}(\mathcal{M})$. The proof of Corollary 1.2 of Rosa (2025) then implies that the eigenvalues satisfy:

$$\lambda_j \asymp j^{-(1+2\nu/\rho)} = j^{-(2\nu+\rho)/\rho}. \tag{23}$$

In particular, $\lambda_j = O(j^{-(2\nu+\rho)/\rho})$.

Combined with the information-gain results of Vakili et al. (2021), this gives an information-gain rate depending on ρ rather than on the dimension D of the full design space. Indeed, substituting $\beta = (2\nu + \rho)/\rho$ into Equation equation 22 gives

$$\gamma_n(\mathcal{M}; k_{\mathcal{D}}) = O\left(n^{\frac{\rho}{2\nu+\rho}} (\log n)^{\frac{2\nu}{2\nu+\rho}}\right). \tag{24}$$

Under the GP-UCB assumptions stated above, this results in $r_n^{\mathcal{M}} = \widetilde{O}\left(n^{-\frac{\nu}{2\nu+\rho}}\right)$ and $r_n^{\mathcal{D}} = \widetilde{O}(n^{-\nu/(2\nu+D)})$. Since $\rho < D$, $\frac{\rho}{2\nu+\rho} < \frac{D}{2\nu+D}$, restricting optimization to a ρ-dimensional simulator-induced search space improves the polynomial dependence of the maximum information gain on the

evaluation budget n, and therefore improves the simple regret exponent, which is more favorable on $\mathcal{M}$.

Finally, combining the optimization rate with the simulator and amortization errors gives

$$r_n \leq \underbrace{\Delta_{\text{model}}}_{\text{simulator misspecification}} + \underbrace{L_U \varepsilon_{\text{am}}}_{\text{amortization error}} + \underbrace{\widetilde{O}\left(n^{-\frac{\nu}{2\nu+\rho}}\right)}_{\text{online optimization error}} . \tag{25}$$

Equation equation 25 makes explicit the trade-off induced by restricting the search space. A more accurate mechanistic model reduces Δ_{model}, better amortization reduces ε_{am}, while a lower-dimensional simulator-induced solution set improves the rate at which the remaining online optimization error decreases.

Moreover, it follows immediately from 10 that

$$\Delta_{\mathcal{M}} \leq r_n. \tag{26}$$

The Mercer-eigenvalue information-gain result of Vakili et al. (2021) has been widely used. However, it assumes uniformly bounded eigenfunctions. Janz et al. (2026) recently noted that this assumption has not been verified even for the Matérn or squared-exponential kernels and recognize it as a genuine gap in the information-gain literature. We therefore emphasize that our result is conditional to this assumption.

B.5 Pullback kernel vs kernel defined in the simulator parameters space

The pullback kernel defines similarity between two search coordinates $\phi, \phi' \in \Phi$ through the regimens they produce, $k_{\text{pb}}(\phi, \phi') = k_{\mathcal{D}}\left(\mathcal{E}_w(\phi), \mathcal{E}_w(\phi')\right)$. For an isotropic Matérn kernel $k_{\mathcal{D}}$, similarity is therefore determined by the regimen-space distance $\|\mathcal{E}_w(\phi) - \mathcal{E}_w(\phi')\|$.

For comparison, suppose that BO instead uses a Matérn-ν kernel k_Φ directly on the search coordinates Φ, which measures similarity using Euclidean distance $\|\phi - \phi'\|$.

Under the corresponding Matérn information-gain assumptions,

$$\gamma_n(\Phi; k_\Phi) = \widetilde{O}\left(n^{\frac{p}{2\nu+p}}\right), \tag{27}$$

with $p = \dim(\Phi)$.

In general, $p = \dim(\Phi) \geq \dim(\widehat{\mathcal{M}}) = \rho$, so the pullback kernel can have a more favorable polynomial information gain rate.

Note that if $\mathcal{E}_w : \Phi \to \widehat{\mathcal{M}}$ is an embedding, $\dim(\Phi) = \dim(\widehat{\mathcal{M}}) = \rho$.

In general, the two kernels nevertheless encode different geometries.

The pullback kernel inherits the geometry induced by the map $\mathcal{E}_w$, and therefore accounts for directions in Φ that have different effects on the produced regimen. In particular, two points may be far away in the simulator parameter space, $\|\phi - \phi'\| \gg 0$, but produce almost identical regimens, $\|\mathcal{E}_w(\phi) - \mathcal{E}_w(\phi')\| \approx 0$. While a Matérn kernel defined on Φ may regard these as nearly independent, the pullback kernel captures their similarity : $k_{\text{pb}}(\phi, \phi') \approx k_{\mathcal{D}}(x, x)$, where $x = \mathcal{E}_w(\phi)$.

This distinction is consistent with recent results showing that manifold geometry can affect GP-bandit regret beyond the dimension-dependent rate. Dorn et al. (2026) derive regret bounds for Matérn GP bandits on compact Riemannian manifolds in which geometric quantities such as the Riemannian volume, injectivity radius, and curvature explicitly enter the bound (Dorn et al., 2026). Their setting differs from ours as they focused on intrinsic Matérn kernels, but it supports the observation that two kernels with the same dimensional asymptotic exponent may nevertheless exhibit different regret due the geometry encoded by the kernel.

C Challenges in Levodopa treatment for Parkinson's disease

Four aspects make the dosing especially challenging in Parkinson's disease. First, the response to a regimen is highly patient-specific due to inter-individual physiological differences influencing drug

absorption, metabolism, and distribution (Hammarlund-Udenaes, 2010; de Lange & Hammarlund-Udenaes, 2022; Chan et al., 2005; Titova & Chaudhuri, 2017), and drug response (Jankovic & Tan, 2020). Second, the optimization objective is only weakly observed: relevant real-time biomarkers of treatment response are rarely available, and clinicians often rely instead on patient- and clinician-reported outcome measures such as symptom diaries (Hauser et al., 2004; 2019), which are subjective, sparse, and noisy. Third, even if a mechanistic simulator of the drug is often available in the form of a pharmacokinetic/pharmacodynamic (PK/PD) model, it predicts a quantity, drug concentration in the effective body compartment or a pharmacological effect, that cannot be readily measured in routine care. Fourth, the optimization problem is typically moderate to high dimensional, with potentially complex dosing schedules.

D Dosing network architecture and training details for Levodopa use-case

Architecture. The network takes as input the PK/PD parameters ϕ of Appendix G together with the target concentration $c^\star_\phi$, itself a function of the PD parameters. All inputs are standardized with the mean and standard deviation of the sampled population. It is a feedforward network with 6 fully connected layers of width 256 and ReLU activations, followed by a linear layer returning $z \in \mathbb{R}^{2M}$.

Output layer. The vector z is turned into a regimen by the deterministic layer

$$d_i \;=\; d_{\max}\,\frac{1+\tanh(z_i)}{2}, \qquad (\eta_1,\dots,\eta_M) \;=\; W\,\mathrm{softmax}(z_{M+1},\dots,z_{2M}), \tag{28}$$

for $i=1,\dots,M$, where $W=(t_f-t_0)-(M-1)\Delta$ is the free time to allocate between intakes. The administration times follow as $t_1=t_0$ and $t_i=t_{i-1}+\Delta+\eta_{i-1}$ for $i=2,\dots,M$. The $\tanh$ keeps each dose in $[0,d_{\max}]$, and the softmax makes the lags positive and constrains them to sum to W, so that consecutive intakes are at least Δ apart and all of them fall within $[t_0,t_f]$. The last lag η_M enters no administration time: it is the slack between the last intake and the end of the window, and it lets the network leave part of the window unused. Layer equation 28 is the decoding map equation 31 evaluated at $u_i=(1+\tanh(z_i))/2$ and $v_i=\exp(z_{M+i})$, up to the positive rescaling of $\mathbf{v}$ that leaves equation 31 unchanged. The image $\mathcal{M}=\mathcal{E}_w(\Phi)$ is therefore contained in $\mathcal{D}$ by construction.

Smooth dose discretization. Microtablets come in multiples of $\Delta_{\text{tab}}=5$ mg, so the admissible doses are the $K=d_{\max}/\Delta_{\text{tab}}+1=51$ values $\{0,\Delta_{\text{tab}},\dots,d_{\max}\}$. Rounding d_i to the nearest of them is piecewise constant, hence has zero gradient almost everywhere and cannot be trained through. During training we replace it by the smooth surrogate

$$\rho_\kappa(d) \;=\; \frac{\Delta_{\text{tab}}}{2}\sum_{k=0}^{K-2}\left[1+\tanh\left(\kappa\left(d-k\,\Delta_{\text{tab}}-\frac{\Delta_{\text{tab}}}{2}\right)\right)\right], \tag{29}$$

a sum of $K-1$ smooth steps of height Δ_{tab}, each centred at the midpoint between two consecutive admissible doses. The steepness κ interpolates between the two regimes of interest: for $\kappa\,\Delta_{\text{tab}}\ll 1$ the map is close to the identity away from the endpoints of the dose range and leaves the continuous doses essentially untouched; as $\kappa\to\infty$, it converges to exact rounding, with vanishing gradients. We anneal it during training: discretization is disabled for the first half of the epochs, then κ increases linearly from 0.1 to 1 over the remaining ones. The validation loss and the regimens actually administered always use exact rounding.

We note that although in our application, the dosing resolution is 5mg, this is much smaller than the typical range of dose variation so that we could still use a GP defined on a continuous dosing space.

ODE integration. The PK/PD system is integrated with the Tsit5 solver of *diffrax* (Kidger, 2022), restarting at each intake with the updated initial conditions of equation 44, using an adaptive step size (relative tolerance 10^{-7}, absolute tolerance 10^{-9}) and 200 saved points per inter-intake segment. These points form the grid of the trapezoidal rule in equation 7. The horizon $[0,T]$ starts 30 min before the first intake and ends 60 min after t_f, so that the loss also covers the elimination phase following the last intake. The loss is computed on the concentration scale, as in equation 7.

Table S1: Per-patient induced-concentration error and compute time to obtain one patient's dosage ($N = 500$ patients), for the amortized, continuous-direct (per-patient optimization with L-BFGS-B), and discrete-direct solutions (per-patient optimization with L-BFGS-B rounded to the nearest 5 mg increment). Error is the trapezoid-integrated squared error between the induced effect-site concentration and the patient's target. Time is the mean wall-clock time for a single-patient call (not batched). †discrete-direct's doses are continuous-direct's solution rounded to the dose grid; the rounding itself is negligible and was not separately timed, so its reported time is continuous-direct's.

	Time-normalized squared conc. error ((mg/L)^2)			Mean time / patient
Method	q_{10}	Median	q_{90}	
Amortized (dosing net)	0.0558	0.2163	1.2013	0.174 ms
Continuous-direct (optimized)	0.0543	0.2142	1.1993	52.971 s
Discrete-direct (5 mg grid)	0.0544	0.2143	1.1993	52.971 s†

Optimization. We train for 25 epochs with AdamW (Loshchilov & Hutter, 2019), batch size 64, learning rate $2 \cdot 10^{-3}$ decayed by a factor 0.2633 every 4 epochs, weight decay $4.13 \cdot 10^{-5}$, and gradient-norm clipping at 10. The hidden layers use dropout with probability 0.07. We retain the checkpoint with the lowest validation loss. The training hyperparameters were selected with Optuna (Akiba et al., 2019).

E Training data

PK/PD studies typically report only the marginal of the estimated distribution of each parameter, not their joint. Sampling the parameters independently from these marginals can produce combinations that are physiologically implausible, which inflates $\mathcal{M}$ relative to what the true joint distribution would give. When possible, it is therefore advisable to prune implausible parameter samples from the training set using clinical knowledge. A simple approach that we found effective in the dosing context is to discard the samples whose peak effect-compartment concentration is too high under a low dose d_l, or too low under a high dose d_h, relative to the target concentration, that is

$$\text{training set} = \{\phi \sim p(\phi) : C_{\max}(d_l; \phi) < c^\star_\phi < C_{\max}(d_h; \phi)\} \tag{30}$$

with $C_{\max}(d; \phi) = \max_{t \in [0, \tau^*]} c_e(t; \phi \mid d)$, where the values d_l, d_h and τ^* are chosen based on the clinical context.

For the dosing network, we sample the training data and prune the parameters as in equation 30, with $\tau^* = 120$ min, $d_l = 80$ mg and $d_h = 600$ mg, values chosen for the clinical context of Levodopa treatment in advanced Parkinson's patients.

E.1 Amortization analysis for Levodopa-Carbidopa dosing use-case

We compare the concentration-tracking error of the regimen obtained with a forward pass of the dosing network against a direct, per-patient optimization of the same concentration error using L-BFGS-B (4 restarts: 1 warm-start at the dosing network forward pass, 3 random Sobol starts), evaluated on 500 held-out virtual patients. For the direct optimization, we report both the continuous-dose solution ("Continuous-direct") and its rounding to the nearest 5 mg increment ("Discrete-direct", $\Delta = 5$ mg). Table S1 shows that the three approaches achieve essentially indistinguishable concentration errors. The table also reports the average time needed to obtain each solution: for the dosing network this is a single forward pass, excluding the one-time offline training cost; for direct optimization it is roughly five orders of magnitude slower, since each optimization step requires solving the ODE and backpropagating through it to compute gradients.

Figure S2 compares the induced concentration trajectories and prescribed regimens of the amortized and continuous-direct solutions for two example patients. The left panel shows the results for a typical patient, for which the two solutions remain nearly indistinguishable. The right panel shows the worst case among the 500 held-out virtual patients: the amortized solution prescribes slightly higher maintenance doses than the numerically optimized one, pushing its induced concentration slightly further above the target. Even in this worst case, the discrepancy stays small, and the amortized regimen still offers a fast, high-quality solution to the PK/PD dosing problem.

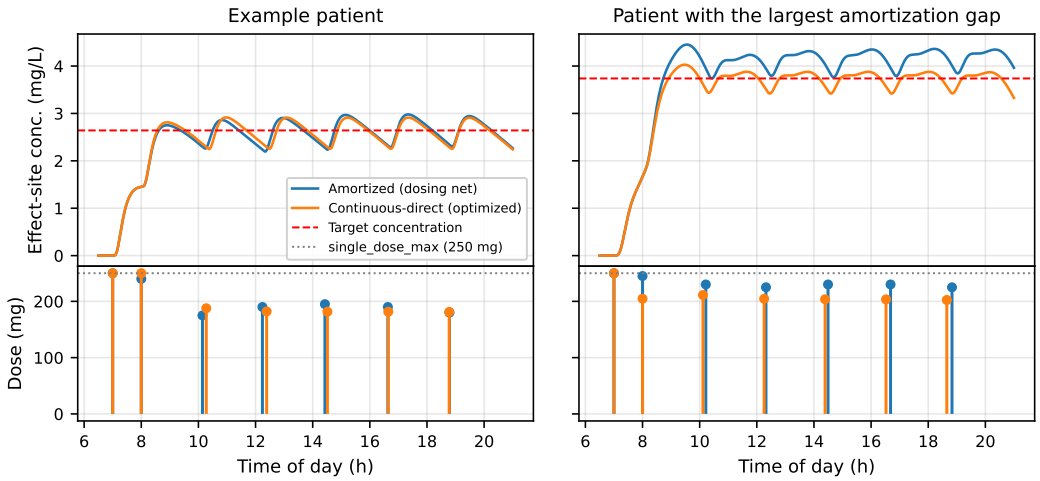

Figure S2: Two examples of optimal dosing and induced PK trajectory. On the right, we select the virtual patient with the largest discrepancy in concentration error between the amortized and the continuous-direct dosing solutions.

F Bayesian optimization details

F.1 Search space and constraints

Dosing space. We define the dosing space $\mathcal{D}$ of Section 6 through a map $[0,1]^{2M} \to \mathcal{D}$, built so that every point of the unit box corresponds to an admissible regimen. Let $W = (t_f - t_0) - (M-1)\,\Delta$ be the free time to allocate between intakes where Δ is the minimum time between two intakes. For $\mathbf{x} = (\mathbf{u}, \mathbf{v}) \in [0,1]^M \times [0,1]^M$ and $i = 1, \dots, M$, the corresponding dose d_i and intake time t_i are

$$\begin{cases} d_i = u_i\, d_{\max}, \\ t_i = t_0 + (i-1)\,\Delta + \sum_{k=1}^{i-1} \frac{v_k}{\sum_{j=1}^{M} v_j}\, W. \end{cases} \tag{31}$$

Search space rescaling. For all methods, the BO search space is rescaled to the unit box. When the search space is $\mathcal{D}$, the rescaling is the right inverse of equation 31 that normalizes $\mathbf{v}$ to sum to one; this fixes the one redundant degree of freedom of equation 31, whose decoded regimen is invariant to a positive rescaling of $\mathbf{v}$. When it is Φ, each parameter is rescaled affinely in the scale in which its random effect is Gaussian (logarithmic for the log-normal parameters, and reflected for those, such as the PD baseline, defined on the negative scale), centred on that patient's covariate-adjusted typical value, so that $x = 0$ and $x = 1$ are the 0.025 and 0.975 quantiles of the patient's own marginal of $p(\phi)$. For our method with the pullback kernel $k_{\mathcal{D}}$, the kernel is evaluated on the regimens $\mathcal{E}_w(\phi)$, mapped to the unit box with the same right inverse of equation 31, so that both search spaces present the kernel with identical coordinates.

Constraint handling. Dosing regimens are usually subject to practical constraints, for instance a minimum and maximum intake dose, a maximum total daily dose, and a discrete set of available tablet strengths. These constraints can often be enforced directly in the output layer of $\mathcal{E}_w$, rather than handled during the online Bayesian optimization step. The constraint on the variation of the total daily dose is a form of transition-constrained BO, which can be treated as a Markov decision process (see, e.g., Paulson et al., 2022; Folch et al., 2024). Here we take a simpler approach, restricting the optimization of the acquisition function at each step: in practice, we assign an arbitrary low value to the points violating the constraints. This simplification does not affect the purpose of the comparison, as we use the same approach for all methods. Discarding infeasible points in this way makes the acquisition landscape more challenging to optimize, as it introduces flat regions and sharp transitions; we counterbalance it with a large number of restarts in the acquisition function optimization.

Constrained acquisition function. Let $\mathbf{x}$ be a point of the unit-box search space (of $\mathcal{D}$ or of Φ), $(\mathbf{d}(\mathbf{x}), \mathbf{t}(\mathbf{x}))$ the regimen it encodes, and $D(\mathbf{x}) = \sum_{i=1}^{M} d_i(\mathbf{x})$ its total daily dose. Writing D_t for the total daily dose administered at iteration t, the feasible set at that iteration is

$$\mathcal{C}_t \;=\; \{\mathbf{x} : D(\mathbf{x}) \geq D_{\min}\} \;\cap\; \{\mathbf{x} : |D(\mathbf{x}) - D_t| \leq \rho\, D_t\}, \tag{32}$$

with $D_{\min} = 320$ mg and $\rho = 0.4$ as in Section 6, and the acquisition function we actually optimize is

$$\tilde{\alpha}(\mathbf{x}\,;\,\mathcal{H}_t) \;=\; \begin{cases} \alpha(\mathbf{x}\,;\,\mathcal{H}_t), & \mathbf{x} \in \mathcal{C}_t, \\ \alpha_\emptyset, & \text{otherwise.} \end{cases} \tag{33}$$

We set $\alpha_\emptyset = 0$ for EUBO, which is its infimum, and $\alpha_\emptyset = -10^9$ for log expected improvement, far below any value it attains. The total dose $D(\mathbf{x})$ is evaluated on the undiscretized regimen, before rounding to the tablet grid, so that $\mathbf{x} \mapsto D(\mathbf{x})$ remains differentiable inside the acquisition optimization: on $\mathcal{D}$ it is linear in $\mathbf{x}$, and on Φ it is the total dose returned by the dosing network. In the rare case where every restart of the optimizer returns an infeasible point, we fall back to a candidate drawn uniformly from $\mathcal{C}_t$, by rejection sampling on a Sobol batch of 4096 points; the random acquisition baseline draws its candidates in the same way. The reported best guess maximizes the GP posterior mean over $\{\mathbf{x} : D(\mathbf{x}) \geq D_{\min}\}$ instead: the minimum daily dose applies to it, but the ρ-band does not, since the best guess is a recommended destination rather than a step along the path.

F.2 Implementation details

Acquisition function optimization. We maximize equation 33 with BoTorch's multi-start routine, returning one candidate per iteration: 100 restarts, selected by its initialization heuristic from 256 raw Sobol samples, each run with L-BFGS-B for at most 500 iterations and 150 line search steps per iteration. Log expected improvement is estimated with 100 quasi-Monte Carlo samples, over a baseline pruned to the observed points that can still be optimal; the posterior mean maximized for the best guess is estimated with 1024 samples.

GP surrogate. In the three settings with a scalar observed utility (concentration feedback and the two misspecified settings), the outcomes are standardized before the GP is fitted: they are centred and scaled to unit variance with their empirical mean and standard deviation over the observations available at that iteration, and the posterior is mapped back to the original scale. The utility scales differ across the three settings — an unnormalized squared concentration error in equation 8 against a cumulative dose error in equation 8 — so this keeps one set of GP priors appropriate for all of them. No such transformation is needed in the preference setting, where the GP is fitted to comparisons and the scale of the latent utility is set by the likelihood. Kernel hyperparameters (lengthscales and output scale, under $\Gamma(2, 2)$ priors) are re-estimated at every iteration by maximizing the marginal likelihood, starting from the first acquisition iteration; if the fit does not converge, the previous iteration's values are kept.

Software. The online optimization is implemented with BoTorch (Balandat et al., 2020) and GPyTorch (Gardner et al., 2018). The dosing network is implemented in PyTorch (Paszke et al., 2019), combined with JAX (Bradbury et al., 2018) and diffrax (Kidger, 2022) for the ODE solver.

F.3 Evaluation settings

In the *misspecified feedback* setting, the biased oracle regimen is obtained by a forward pass of the dosing network on the patient's true ϕ, followed by a perturbation of the dose amount at each intake with an additive shift drawn from $\{s\,\Delta_{\text{tab}}\}_{s=-S}^{S}$ with weights $\{(S+1-|s|)/(S+1)^2\}_{s=-S}^{S}$ and $S = 5$. The first dose alone is perturbed with a negative shift, large enough to bring it below the largest of the subsequent doses. The first dose is lowered below the largest of the subsequent doses, which places the oracle outside $\mathcal{M}$: under the selected PK/PD model the model-optimal regimens have a first intake larger than or equal to the subsequent ones;

G PK/PD MODEL FOR LEVODOPA-CARBIDOPA IN PARKINSON'S DISEASE

The use-case is the Levodopa-Carbidopa combination therapy, the standard oral treatment for Parkinson's disease. Levodopa is the active principle: a dopamine precursor that crosses the blood-brain barrier, where it is converted to dopamine. Carbidopa is a peripheral decarboxylase inhibitor that blocks this conversion outside the brain, so that more Levodopa reaches the central nervous system and peripheral side effects are reduced. Because Carbidopa inhibits the peripheral metabolism of Levodopa, the two drugs are pharmacokinetically coupled, and the model accounts for this interaction.

We combine the population pharmacokinetic model of Senek et al. (2018) with the pharmacodynamic model of Westin et al. (2011), whose clinical effect is the treatment response scale (TRS), a score ranging from -3 (*off*, parkinsonism) to $+3$ (dyskinesia), with 0 indicating the optimal response (Nyholm et al., 2005). Together they instantiate the simulator $(f_{\text{PK}}, h, f_{\text{PD}})$ and the population distribution $p(\phi)$ of Section 6.1: f_{PK} is the ODE system governing the drug concentration, f_{PD} is the $E_{\max}$ map from the central Levodopa concentration to the TRS, and ϕ_i collects all patient-specific PK/PD parameters. Here the effect compartment coincides with the central compartment, so that $c_e = c^{\text{LD}}$ and the map h of Section 6.1 is not needed.

Throughout, the subscript i indexes the patient and the superscripts LD and CD refer to Levodopa and Carbidopa. We write $A_\bullet$ for the amount of drug in a compartment; these amounts are the state variables of the ODE system, and the plasma concentration driving the pharmacodynamics is the central Levodopa amount divided by its volume of distribution. To keep the notation light, we suppress the patient subscript i on the PK/PD parameters inside the differential equations and restore it in the parameter models that follow.

G.1 PHARMACOKINETIC (PK) MODEL

Levodopa follows a two-compartment disposition model (a central and a peripheral compartment) and Carbidopa a one-compartment disposition model (a central compartment only). For each drug, the oral dose reaches the central compartment through two parallel absorption routes, a fast and a slow transit chain. A transit chain is a series of compartments through which the dose passes before entering the circulation, and it delays and spreads the absorption. We write $n_1^\bullet$ and $n_2^\bullet$ for the number of compartments in the fast and slow chains, $k_1^\bullet$ and $k_2^\bullet$ for their transit rate constants, d_{LD} and d_{CD} for the administered doses, F for the bioavailability, and $\text{fa}^\bullet$ for the fraction of the dose entering the fast chain.

Levodopa fast-transit chain. The fast chain has n_1^{LD} compartments $A_{1,1}^{\text{LD}}, \dots, A_{1,n_1^{\text{LD}}}^{\text{LD}}$, in which the fast fraction of the dose enters the first compartment,

$$\begin{cases} \dot{A}_{1,1}^{\text{LD}} = -k_1^{\text{LD}} A_{1,1}^{\text{LD}}, & A_{1,1}^{\text{LD}}(0) = F_i \, \text{fa}_i^{\text{LD}} \, d_{\text{LD}}, \\ \dot{A}_{1,j}^{\text{LD}} = k_1^{\text{LD}} A_{1,j-1}^{\text{LD}} - k_1^{\text{LD}} A_{1,j}^{\text{LD}}, & j = 2, \dots, n_1^{\text{LD}}. \end{cases} \tag{34}$$

Levodopa slow-transit chain. The slow chain has n_2^{LD} compartments $A_{2,1}^{\text{LD}}, \dots, A_{2,n_2^{\text{LD}}}^{\text{LD}}$, which receive the complementary fraction of the dose,

$$\begin{cases} \dot{A}_{2,1}^{\text{LD}} = -k_2^{\text{LD}} A_{2,1}^{\text{LD}}, & A_{2,1}^{\text{LD}}(0) = F_i \, (1 - \text{fa}_i^{\text{LD}}) \, d_{\text{LD}}, \\ \dot{A}_{2,j}^{\text{LD}} = k_2^{\text{LD}} A_{2,j-1}^{\text{LD}} - k_2^{\text{LD}} A_{2,j}^{\text{LD}}, & j = 2, \dots, n_2^{\text{LD}}. \end{cases} \tag{35}$$

Levodopa central and peripheral compartments. The central compartment A_C^{LD} (volume V_C^{LD}) collects the output of both chains and exchanges with the peripheral compartment A_P^{LD} (volume V_P^{LD}) at intercompartmental clearance Q^{LD}, while Levodopa is eliminated at clearance CL^{LD},

$$\begin{aligned} \dot{A}_C^{\text{LD}} &= k_1^{\text{LD}} A_{1,n_1^{\text{LD}}}^{\text{LD}} + k_2^{\text{LD}} A_{2,n_2^{\text{LD}}}^{\text{LD}} + \frac{Q^{\text{LD}}}{V_P^{\text{LD}}} A_P^{\text{LD}} - \frac{Q^{\text{LD}} + \text{CL}^{\text{LD}}}{V_C^{\text{LD}}} A_C^{\text{LD}}, \\ \dot{A}_P^{\text{LD}} &= -\frac{Q^{\text{LD}}}{V_P^{\text{LD}}} A_P^{\text{LD}} + \frac{Q^{\text{LD}}}{V_C^{\text{LD}}} A_C^{\text{LD}}. \end{aligned} \tag{36}$$

The Levodopa plasma concentration driving the pharmacodynamics is $c^{\text{LD}} = A_C^{\text{LD}} / V_C^{\text{LD}}$.

Carbidopa transit chains and central compartment. Carbidopa is absorbed through its own fast and slow chains, $A^{\mathrm{CD}}_{1,1},\dots,A^{\mathrm{CD}}_{1,n_1^{\mathrm{CD}}}$ and $A^{\mathrm{CD}}_{2,1},\dots,A^{\mathrm{CD}}_{2,n_2^{\mathrm{CD}}}$, with the same structure as for Levodopa,

$$\begin{cases} \dot{A}^{\mathrm{CD}}_{1,1} = -k_1^{\mathrm{CD}} A^{\mathrm{CD}}_{1,1}, & A^{\mathrm{CD}}_{1,1}(0) = F_i\,\mathrm{fa}_i^{\mathrm{CD}}\,d_{\mathrm{CD}}, \\ \dot{A}^{\mathrm{CD}}_{1,j} = k_1^{\mathrm{CD}} A^{\mathrm{CD}}_{1,j-1} - k_1^{\mathrm{CD}} A^{\mathrm{CD}}_{1,j}, & j=2,\dots,n_1^{\mathrm{CD}}, \\ \dot{A}^{\mathrm{CD}}_{2,1} = -k_2^{\mathrm{CD}} A^{\mathrm{CD}}_{2,1}, & A^{\mathrm{CD}}_{2,1}(0) = F_i\,(1-\mathrm{fa}_i^{\mathrm{CD}})\,d_{\mathrm{CD}}, \\ \dot{A}^{\mathrm{CD}}_{2,j} = k_2^{\mathrm{CD}} A^{\mathrm{CD}}_{2,j-1} - k_2^{\mathrm{CD}} A^{\mathrm{CD}}_{2,j}, & j=2,\dots,n_2^{\mathrm{CD}}. \end{cases} \tag{37}$$

The single (central) Carbidopa compartment A_C^{CD} (volume V_C^{CD}) collects the output of both chains and eliminates Carbidopa at clearance $\mathrm{CL}^{\mathrm{CD}}$,

$$\dot{A}_C^{\mathrm{CD}} = k_1^{\mathrm{CD}} A^{\mathrm{CD}}_{1,n_1^{\mathrm{CD}}} + k_2^{\mathrm{CD}} A^{\mathrm{CD}}_{2,n_2^{\mathrm{CD}}} - \frac{\mathrm{CL}^{\mathrm{CD}}}{V_C^{\mathrm{CD}}} A_C^{\mathrm{CD}}. \tag{38}$$

Disposition parameters and covariates. The disposition parameters (clearances and volumes) depend on patient covariates through an allometric weight term and, for the clearances, additional covariates: the Levodopa clearance depends on the Hoehn-Yahr stage HY_i and, through the metabolic interaction, on the Carbidopa dose d_{CD}, while the Carbidopa clearance depends on the patient age AGE_i,

$$\begin{aligned} \mathrm{CL}_i^{\mathrm{LD}} &= \exp\Big(\log \mathrm{CL}_0^{\mathrm{LD}} + \alpha \log \tfrac{W_i}{70} + \beta_{\mathrm{HY}}(\mathrm{HY}_i - \overline{\mathrm{HY}}) - \log\Big(1 + \tfrac{d_{\mathrm{CD}}}{\beta_{\mathrm{int}}}\Big) + \eta_i^{\mathrm{CL}^{\mathrm{LD}}}\Big), \\ Q_i^{\mathrm{LD}} &= \exp\Big(\log Q_0^{\mathrm{LD}} + \alpha \log \tfrac{W_i}{70} + \eta_i^{Q^{\mathrm{LD}}}\Big), \\ V_{C,i}^{\mathrm{LD}} &= \exp\Big(\log V_{C,0}^{\mathrm{LD}} + \alpha \log \tfrac{W_i}{70} + \eta_i^{V_C^{\mathrm{LD}}}\Big), \\ V_{P,i}^{\mathrm{LD}} &= \exp\Big(\log V_{P,0}^{\mathrm{LD}} + \alpha \log \tfrac{W_i}{70} + \eta_i^{V_P^{\mathrm{LD}}}\Big), \\ \mathrm{CL}_i^{\mathrm{CD}} &= \exp\Big(\log \mathrm{CL}_0^{\mathrm{CD}} + \alpha \log \tfrac{W_i}{70} + \beta_{\mathrm{AGE}}(\mathrm{AGE}_i - \overline{\mathrm{AGE}}) + \eta_i^{\mathrm{CL}^{\mathrm{CD}}}\Big), \\ V_{C,i}^{\mathrm{CD}} &= \exp\Big(\log V_{C,0}^{\mathrm{CD}} + \alpha \log \tfrac{W_i}{70} + \eta_i^{V_C^{\mathrm{CD}}}\Big), \end{aligned} \tag{39}$$

where a subscript 0 denotes the typical (population) value, α is the allometric scaling exponent, β_{HY} and β_{AGE} are covariate coefficients, and 70 kg is the reference body weight. The term $-\log(1 + d_{\mathrm{CD}}/\beta_{\mathrm{int}})$ encodes the reduction in Levodopa clearance caused by Carbidopa, with β_{int} controlling the strength of the interaction.

Absorption parameters. The fast-chain fractions are bounded in $(0,1)$ and modeled on the logit scale, while the transit rate constants are parametrized through the mean transit times $\mathrm{MTT}^{\bullet}$ on the log scale,

$$\begin{aligned} \mathrm{fa}_i^{\mathrm{LD}} &= \mathrm{expit}\Big(\mathrm{logit}(\mathrm{fa}_0^{\mathrm{LD}}) + \eta_i^{\mathrm{fa}^{\mathrm{LD}}}\Big), \\ \mathrm{fa}_i^{\mathrm{CD}} &= \mathrm{expit}\Big(\mathrm{logit}(\mathrm{fa}_0^{\mathrm{CD}}) + \eta_i^{\mathrm{fa}^{\mathrm{CD}}}\Big), \\ k_{r,i}^{\bullet} &= 1/\mathrm{MTT}_{r,i}^{\bullet}, \quad \mathrm{MTT}_{r,i}^{\bullet} = \exp\Big(\log \mathrm{MTT}_{r,0}^{\bullet} + \eta_i^{\mathrm{MTT}_r^{\bullet}}\Big), \quad r \in \{1,2\},\ \bullet \in \{\mathrm{LD},\mathrm{CD}\}. \end{aligned} \tag{40}$$

It was not possible to reconstruct how the bioavailability F_i was parametrised in terms of random effects in the original model, therefore we consider here it fixed at its typical value F_0.

Random effects and inter-patient variability. Each PK/PD parameter carries a patient-specific random effect $\eta_i^{\bullet} \sim \mathcal{N}(0, \omega_{\bullet}^2)$ that captures inter-patient variability. Following common practice, unconstrained parameters (the PD baseline and maximum effect below) use additive normal random effects; positive parameters (clearances, volumes, mean transit times) use log-normal effects, i.e. a normal effect on the log scale; and parameters bounded in $(0,1)$ (bioavailability and fast-chain fractions) use logit-normal effects, i.e. a normal effect on the logit scale.

Pharmacometric reports usually summarize each $\omega_{\bullet}$ as a percentage coefficient of variation, $\mathrm{CV}\%_{\bullet} = 100\ \mathrm{Std}[\cdot]/\,\mathrm{E}[\cdot]$, the ratio of the standard deviation to the mean of the parameter distribution. For

a log-normal parameter this has the closed form $\mathrm{CV}\%_\bullet = 100\sqrt{\exp(\omega_\bullet^2) - 1}$; for a logit-normal parameter no closed form exists and the two moments are evaluated numerically. We take the fixed-effect and random-effect estimates from Senek et al. (2018) for the PK model and Westin et al. (2011) for the PD model.

G.2 Pharmacodynamic (PD) model

The PD model maps the central Levodopa concentration c^{LD} to an effect compartment c through:

$$\dot{c} = k_{\mathrm{EO}} \left(\frac{c_{\mathrm{LD}}}{V_C^{\mathrm{LD}}} - c \right). \tag{41}$$

The effect concentration is then mapped to the TRS through a sigmoidal $E_{\max}$ model with a baseline and a Hill coefficient,

$$\mathrm{TRS}_i(c) = \mathrm{BASE}_i + \frac{E_{\max,i}\, c^\gamma}{c^\gamma + \mathrm{EC}_{50,i}^\gamma}, \qquad c = c^{\mathrm{LD}}, \tag{42}$$

where BASE_i is the drug-free baseline, $E_{\max,i}$ the maximum drug effect, $\mathrm{EC}_{50,i}$ the concentration producing half of $E_{\max,i}$, and γ the Hill coefficient controlling the steepness of the concentration-effect curve. In the notation of Section 6.1, this map is the pharmacodynamic model f_{PD} and TRS_i is the effect $r(t;\phi)$. All parameters carry random effects,

$$\begin{cases} \mathrm{BASE}_i = \mathrm{BASE}_0 + \eta_i^{\mathrm{BASE}}, \\ E_{\max,i} = E_{\max,0} + \eta_i^{E_{\max}}, \\ \mathrm{EC}_{50,i} = \exp\left(\log \mathrm{EC}_{50,0} + \eta_i^{\mathrm{EC}_{50}}\right), \\ \gamma_i = \exp(\log \gamma_0 + \eta_i^\gamma) \end{cases} \tag{43}$$

The fixed-effect and random-effect estimates are taken from Westin et al. (2011).

G.3 PKPD solution under multiple intakes

Let $c_e(s; \phi \mid (\mathbf{d}, \mathbf{t}))$ be the concentration in the effect compartment at time s induced by the regimen $(\mathbf{d}, \mathbf{t})$. In practice, it is computed by integrating the ODE system piecewise between successive intakes, updating the initial conditions at each dosing time, that is, for $t \in (t_i, t_{i+1})$:

$$\mathbf{c}(t_i) = \mathbf{c}(t_i^-) + g(d_i;\phi)\, \mathbf{e}_1, \text{ and } c_e(t) \quad = \left[\mathbf{c}(t_i) + \int_{t_i}^{t} f_{\mathrm{PK}}\big(\mathbf{c}(s);\phi\big)\, ds \right]_e, \tag{44}$$

where $\mathbf{e}_1$ is the first canonical basis vector of $\mathbb{R}^K$, so that the intake d_i enters the dosing compartment while the other compartments carry over their concentrations at t_i^-, and $[\,\cdot\,]_e$ denotes the component corresponding to the effect compartment.

H Intrinsic dimension of the amortized search space

The model-optimal regimen $\mathcal{E}^\star(\phi)$ of Section 6.1 does not depend on all coordinates of ϕ independently. This follows from two observations which we can make directly analysing the PK/PD equations.

First, Carbidopa acts on Levodopa only through its dose amount d_{CD} in the clearance of Levodopa equation 39, and this dose is fixed by the Levodopa regimen due to the pill format. Therefore, the PK parameters of Carbidopa are nuisance parameters for the optimal dosage. Second, the PD parameters BASE, $E_{\max}$, EC_{50} and γ enter the dosing objective only through the scalar target concentration $c_\phi^\star$. The TRS equation 42 is monotone, so the target concentration is unique, which gives only one degree of freedom. Therefore, those 4 parameters are not fully identifiable from the optimal dosage. This gives us the bound $\dim(\mathcal{M}) \leq 7$.

Additional structural non-idenfiability could come from the complex relationship between the PK parameters and the optimal dosing solution, which cannot be ruled out analytically. We therefore

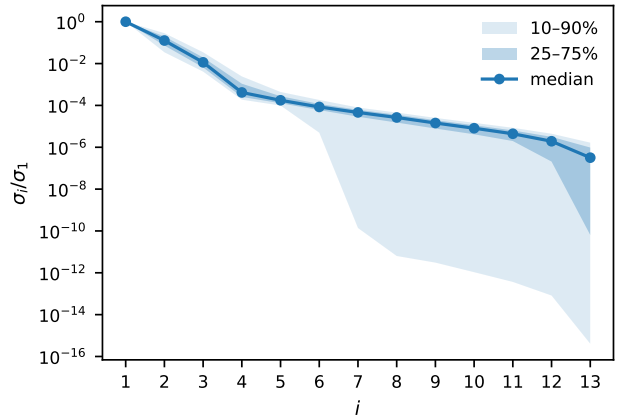

Figure S3: Singular values σ_i of the Jacobian $J_{\mathcal{E}_w}(\phi)$ of the dosing network, divided by the largest one. Median, 25–75% and 10–90% quantiles; log-scale for the y-axis.

study the intrinsic dimension of $\mathcal{M} = \mathcal{E}_w(\Phi)$ looking at the Jacobian of the dosing network, $J_{\mathcal{E}_w}(\phi) = \partial\mathcal{E}_w(\phi)/\partial\phi$. The rank of $J_{\mathcal{E}_w}(\phi)$ is the number of independent directions in which the regimen changes when ϕ is perturbed, that is, the local dimension of $\mathcal{M}$ around $\mathcal{E}_w(\phi)$. In practice, the singular values of $J_{\mathcal{E}_w}(\phi)$ are rarely exactly zero, but several of them can be orders of magnitude smaller than the largest one. Along the corresponding directions the regimen barely changes, so we look at how fast the singular values decay rather than at the exact rank.

Setup. We use the coordinates of the online optimization (Appendix F): ϕ is expressed in the unit box of the $p = 14$ PK/PD parameters searched by BO, and the regimen $\mathcal{E}_w(\phi)$ in the unit-box coordinates of $\mathcal{D}$ given by the right inverse of equation 31. Since this right inverse normalizes $\mathbf{v}$ to sum to one, we drop v_M, so that the regimen has $D = 2M - 1 = 13$ coordinates and $J_{\mathcal{E}_w}(\phi) \in \mathbb{R}^{13\times 14}$. We compute $J_{\mathcal{E}_w}(\phi)$ by automatic differentiation.

Experiment. We draw 10 virtual patients. For each patient, we draw 10^5 parameter vectors ϕ from $p(\phi)$ given the patient's covariates. For each of the 10^6 draws, we compute the singular values $\sigma_1(\phi) \geq \cdots \geq \sigma_{13}(\phi)$ of $J_{\mathcal{E}_w}(\phi)$, i.e. the square roots of the eigenvalues of $J_{\mathcal{E}_w}(\phi)J_{\mathcal{E}_w}(\phi)^\intercal$, and divide them by the largest one, $\sigma_1(\phi)$, to remove the overall scale of the map.

Results. Figure S3 shows the distribution of $\sigma_i(\phi)/\sigma_1(\phi)$ over the draws. The median ratios are 0.13 for $i = 2$, $1.1 \cdot 10^{-2}$ for $i = 3$ and $4.2 \cdot 10^{-4}$ for $i = 4$, and decrease slowly from $1.8 \cdot 10^{-4}$ to $3.2 \cdot 10^{-7}$ for $i = 5, \ldots, 13$. The largest drop between two consecutive median ratios is between $i = 3$ and $i = 4$, by a factor 28. With a threshold of 10^{-3}, inside this range, the number of ratios above the threshold is 3 for 73% of the draws and 4 for 26%. We therefore estimate the intrinsic dimension of $\mathcal{M}$ between 3 and 5.

I BAYESIAN OPTIMIZATION WITH πBOASS

πBOASS uses the same amortized map $\mathcal{E}_w$ as BOASS, but does not restrict the search to its image $\mathcal{M}$. The optimization runs on $\mathcal{D}$, and $\mathcal{E}_w$ is used instead to build a simulator-based prior π over the location of the optimum, combined with the acquisition function as in πBO (Hvarfner et al., 2022). The simulator thus enters the online loop as a prior that decays as observations accumulate, rather than as a hard restriction of the search space. In this sense πBOASS is a soft variant of our method, that we expect to be less sensitive to search-space bias.

Prior. The prior is fitted once per patient and cached across iterations. We draw $N = 10^4$ samples of the optimized PK/PD parameters from their population random-effect distribution and map each draw ϕ_s through the dosing network to a regimen $(\mathbf{d}, \mathbf{t}) = \mathcal{E}_w(\phi_s)$. This sample is the pushforward

of the population distribution $p(\phi)$ through $\mathcal{E}_w$, that is, the distribution of the regimens the network prescribes across the population compatible with the patient's covariates.

Transformation. The pushforward is supported on a bounded set with atoms at its boundary: a fraction of the doses saturate at the single-dose cap $d_{\max}$, and a fraction of the intervals between consecutive intakes sit at the minimum separation Δ. Each regimen $(\mathbf{d}, \mathbf{t})$ is therefore first mapped to unbounded coordinates

$$\mathbf{v} = \Big(\underbrace{\operatorname{logit}\big(\epsilon_d + (1-2\epsilon_d)\, d_i/d_{\max}\big)}_{i=1,\dots,M},\ \underbrace{\log\big(t_{j+1} - t_j - \Delta + s\big)}_{j=1,\dots,M-1} \Big) \in \mathbb{R}^{2M-1}, \tag{45}$$

with $\epsilon_d = \Delta_{\text{tab}}/d_{\max}$ a squashing margin of one tablet increment and $s = 1$ min a fixed shift. The first intake carries no information, since $t_1 = t_0$ for every regimen, which leaves $2M - 1 = 13$ free coordinates.

Mixture fit. We standardize the transformed samples per coordinate, $\mathbf{z} = (\mathbf{v} - \hat{\boldsymbol{\mu}})/\hat{\boldsymbol{\sigma}}$, and fit by expectation-maximization a Gaussian mixture with K full-covariance components,

$$\pi(\mathbf{z}) = \sum_{k=1}^{K} w_k\, \mathcal{N}\big(\mathbf{z}\,;\, \boldsymbol{\mu}_k,\, \boldsymbol{\Sigma}_k + \varrho\, \mathbf{I}\big), \tag{46}$$

keeping the best of 10 restarts. The two parameters ϱ and K are fixed for all patients and selected offline, before the experiments. The floor ϱ, added to the covariance of every component during fitting, bounds how narrow a component can become: we set $\varrho = 10^{-2}$, which corresponds to a standard deviation of one tenth of the population spread of each coordinate. Since the density concentrates in lower-dimensional region, the regularization provided by ϱ is crucial to prevent components from collapsing.

In order to select K, we consider a grid of possible values from 1 to 30 and select K by BIC, over the ten representative patients. We selected $K = 9$, the median of the per-patient selections, so that the complexity of the prior does not depend on the patient's own sample.

Combination with the acquisition function. At acquisition iteration n, the prior is combined with the base acquisition function α as $\alpha(\mathbf{x})\,\big(\pi(\mathbf{x}) + \epsilon_\pi\big)^{\gamma_n}$, evaluated in logarithmic form, with $\gamma_n = \beta/n$. The constant $\epsilon_\pi = 10^{-12}$ added to the density is the one used in the original implementation (Hvarfner et al., 2022, Appendix F.1), and it helps preventing numerical issues when the prior density is extremely low.

Calibration of β. The schedule γ_n sets the weight of the prior relative to the data-driven term: β fixes the tilt at the first acquisition iteration, and the tilt decays as $1/n$ as observations accumulate. We calibrate β on the same ten patients, measuring the average log-density the prior assigns to a plausible regimen above its floor,

$$g \;=\; \frac{1}{S} \sum_{s=1}^{S} \log \pi\big(\mathcal{E}_w(\phi_s)\big) - \log \epsilon_\pi, \tag{47}$$

estimated on the pushforward samples $\phi_s \overset{\text{iid}}{\sim} p(\phi)$. We then consider the median of $\hat{g}$ over the ten patients. A plausible regimen is then favoured over an implausible one by $\gamma_n\, \hat{g} = \beta\, \hat{g}/n$ units of log-acquisition at iteration n. We set that at $n = 1$ the prior effect is $\kappa = 30$, giving $\beta = \kappa/\hat{g} = 0.808$, so that the prior dominates during the first iterations and its tilt becomes comparable to the range of the log-acquisition function after a few of them. The stronger setting $\beta = 10$ is the value used by Hvarfner et al. (2022) in their experiments; with the prior fitted here it amounts to $\kappa \approx 370$, and the prior dominates over a much longer initial phase (πBOASS-strong).

Acquisition optimization. Half of the restarts of the multi-start optimizer used for optimizing the acquisition function are drawn from the fitted prior and the rest from a Sobol sequence, both filtered to the feasible set equation 32. Drawing part of the restarts from the prior helps guide the optimizer towards regions favored by the prior, which can improve the efficiency of the search.

Table S2: Median [10th, 90th percentile] number of acquisition days (out of 47) on which the first prescribed dose is strictly lower than all other doses, per method and well-specified setting for the Levodopa-Carbidopa use-case in Section 6

	BOASS-*	πBOASS-on-$\mathcal{D}$	πBOASS-on-$\mathcal{D}$-strong	BO-on-$\mathcal{D}$
preference feedback	0.0 [0.0, 0.0]	0.0 [0.0, 2.0]	0.0 [0.0, 0.0]	1.0 [0.0, 13.1]
concentration feedback	0.0 [0.0, 0.0]	0.0 [0.0, 4.0]	0.0 [0.0, 3.1]	1.5 [0.0, 7.0]

J SUPPLEMENTARY RESULTS

For the two well-specified settings in Section 6.5, we also count the iterations at which the tested regimen has a first dose smaller than the maintenance doses (*low-first-dose*); Table S2 reports the counts. A low first dose is a sufficient condition for a regimen to lie outside $\mathcal{M}$. When the oracle follows the PK/PD model, this provides a proxy criterion to identify a subset of the pharmacologically incoherent regimens, that is, regimens a clinician would likely not prescribe. By construction, our method never tests such regimens, while the other methods test a small but non-zero number of them. All methods are initialised in $\mathcal{M}$.

K SYNTHETIC EXPERIMENTS

Section 4 splits the simple regret of BOASS into two terms: the online optimization error, that is the gap between the best point found so far and the best point of the amortized search space $\mathcal{M}$, and the search-space bias, that is the gap between the best point of $\mathcal{M}$ and the best point of $\mathcal{D}$. The synthetic benchmarks below control them separately, and vary one ingredient at a time.

We use two benchmarks with complementary roles.

- **Branin embedded in an ambient space** (Section K.1). The ambient objective, its optimum and the exact value of the search-space bias are known in closed form. We study the followings:
  - We vary the ambient dimension D and the search space bias δ, and measure how the gain obtained by searching a two-dimensional space compares with the cost of the bias at a fixed evaluation budget;
  - We train a NN with different sample size to induce different degrees of amortization error, measuring the impact in terms of regret (Section K.1.2);
- **Objective drawn from a Gaussian process** (Section K.2). The objective is a sample from the prior used by the surrogate, so the GP model is well-specified in the ambient space, and the reference value is the best value found on $\mathcal{M}$, so no search-space bias enters. What varies is the geometry of the embedding: the intrinsic dimension p, the ambient dimension D, and how far the embedding is from being injective. This is the setting in which the two kernel choices of Section B.5 differ, so this is the benchmark we use to compare them.

Table S3 lists the methods used across all benchmarks. The Branin benchmark compares BOASS-k_Φ with BO-on-$\mathcal{D}$ and high-dimensional BO methods (TuRBO (Eriksson et al., 2019), SAASBO (Eriksson & Jankowiak, 2021)), while the Gaussian process benchmark compares the three restricted methods.

K.1 BRANIN EMBEDDED IN AN AMBIENT SPACE

Objective on the surface. The construction embeds the two-dimensional Branin function B in an ambient decision space of dimension D. The intrinsic parameter is $\phi = (\phi_1, \phi_2) \in \Phi = [-1, 1]^2$. It is mapped to the conventional Branin domain by :

$$T(\phi) = \begin{pmatrix} 7.5\phi_1 + 2.5 \\ 7.5(\phi_2 + 1) \end{pmatrix}, \tag{48}$$

Table S3: Methods compared in the synthetic experiments.

Method	Search space and GP input	Role
BO-on-$\mathcal{D}$	Search $x \in \mathcal{D}$; GP input is x	Unrestricted reference.
BOASS-$k_{\mathcal{D}}$	Search $\phi \in \Phi$; GP input is $\mathcal{E}(\phi)$	Proposed method with the pullback kernel $k_{\mathcal{D}}\big(\mathcal{E}(\phi), \mathcal{E}(\phi')\big)$.
BOASS-k_{Φ}	Search $\phi \in \Phi$; GP input is ϕ	Same search space, but the kernel is defined on the simulator parameters and ignores the induced geometry.
BOASS-random-k_{Φ}	Draw ϕ uniformly and evaluate $\mathcal{E}(\phi)$	Separates the benefit of restricting the search space from the benefit of the acquisition function.
TuRBO	Search $x \in \mathcal{D}$; GP input is x	Unrestricted high-dimensional BO reference (Eriksson et al., 2019)
SAASBO	Search $x \in \mathcal{D}$; GP input is x	Unrestricted high-dimensional BO reference (Eriksson & Jankowiak, 2021)

so that $T_1(\phi) \in [-5, 10]$ and $T_2(\phi) \in [0, 15]$. The global minimum is $B_{\min} \simeq 0.397887$. The implementation uses the shifted and scaled cost

$$\widetilde{B}(\phi) = \frac{B(T(\phi)) - B_{\min}}{50}. \tag{49}$$

Thus the minimum of $\widetilde{B}$ is zero.

Embedding. The ambient space is $\mathcal{D} = [-1, 1]^D$ and the surface is the image of the ideal embedding $\mathcal{E}^\star(\phi) = (\phi, m(\phi))$, with $m : \Phi \to \mathbb{R}^{D-2}$ smooth. We use the fixed-geometry family

$$m_D^{\text{fix}}(\phi) = a \begin{pmatrix} \sin(\pi\phi_1) \\ \cos(\pi\phi_2) \\ \sin\big(\pi(\phi_1 + \phi_2)\big) \\ \cos\big(\pi(\phi_1 - \phi_2)\big) \\ 0_{D-6} \end{pmatrix}, \qquad a = 0.15, \tag{50}$$

defined for $D \geq 6$. Increasing D only appends zero coordinates, so the pairwise distances on the surface, and therefore the problem seen by a method restricted to it, do not change. Only the dimension seen by BO-on-$\mathcal{D}$ grows.

Search-space bias. The bias is introduced by moving the optimum away from the surface. Let $e_1 = (1, 0, \ldots, 0)^\top \in \mathbb{R}^{D-2}$. This vector selects the first of the $D - 2$ auxiliary coordinates. Consequently, $\delta e_1 = (\delta, 0, \ldots, 0)^\top$ shifts the target surface only along the ambient coordinate x_3. We define the ambient minimization objective as:

$$f_\delta(x) = \widetilde{B}(x_{1:2}) + \lambda \left\| x_{3:D} - m(x_{1:2}) - \delta e_1 \right\|_2^2, \qquad x \in [-1, 1]^D, \tag{51}$$

with $\lambda = 1$ and $m = m_D^{\text{fix}}$. The global optimum has again cost zero, but lies on the shifted surface $\{x : x_{3:D} = m(x_{1:2}) + \delta e_1\}$. The simulator-defined ideal embedding remains $\mathcal{E}^\star(\phi) = (\phi, m(\phi))$, but the ideal surface $\mathcal{M}^\star$ does not contain this optimum when $\delta > 0$, so its global simple regret has the exact floor $\Delta_{\text{model}} = \lambda\delta^2$. The two terms of the regret are thus controlled separately: D acts only on the problem faced by BO-on-$\mathcal{D}$, and δ only on the bias of the restricted search.

Common protocol. Observations are $y_t = f_\delta(x_t) + \eta_t$, with $\eta_t \sim \mathcal{N}(0, 0.01^2)$. Unless stated otherwise, the surrogates use a Matérn-$5/2$ covariance with the noise variance held at its true value $\sigma^2 = 0.01^2$, and BOASS and BO-on-$\mathcal{D}$ propose the next evaluation by minimizing the lower confidence bound $\text{LCB}_t(z) = \mu_{t-1}(z) - 2\sigma_{t-1}(z)$. Since f_δ attains the value 0 on $\mathcal{D}$, we report simple regret as the best noiseless value observed so far.

K.1.1 Comparison with high-dimensional BO methods and bias analysis

This experiment measures how the search-space bias trades off against the ambient dimension, and compares BOASS with BO methods designed for high-dimensional spaces. The embedding $\mathcal{E}^\star$ is injective, so two distinct ϕ never produce the same design and the ambiguity that the pullback kernel is meant to resolve does not arise here. We therefore report BOASS-k_Φ only, and leave the comparison between the two kernels to Section K.2.

TuRBO (Eriksson et al., 2019) and SAASBO (Eriksson & Jankowiak, 2021) adapt Bayesian optimization to high-dimensional spaces through different mechanisms. TuRBO-1 restricts the search to a single axis-aligned trust region, whose side lengths are proportional to the ARD lengthscales of a GP fit only to the local data; candidates are selected by Thompson sampling inside the region. The region size is adapted rather than monotonically shrunk: doubled after τ_{succ} consecutive improvements, halved after τ_{fail} failures, and restarted from a fresh design once it falls below $L_{\min}$. Thus, exploration comes from restarts rather than from the acquisition function. SAASBO instead keeps a single global GP and places a half-Cauchy hierarchical prior on the inverse squared lengthscales, shrinking most coordinates to irrelevance and concentrating the model on a sparse axis-aligned subspace inferred from data, with hyperparameters sampled by NUTS. Both methods search the full ambient space $\mathcal{D}$, so, like BO-on-$\mathcal{D}$, they are not subject to the bias floor $\lambda\delta^2$.

Setting. We use $D \in \{6, 8, 16, 32\}$ and $\delta \in \{0, 0.1, 0.25, 0.5, 1\}$. Each method has a budget of 60 evaluations, and each pair (D, method) is repeated over 25 independent seeds. All methods start from a single random point of $\mathcal{M}$, common to all methods, obtained by drawing ϕ uniformly in Φ and mapping it by $\mathcal{E}^\star$.

Methods. BOASS-k_Φ and BO-on-$\mathcal{D}$ estimate the amplitude and the length scale of their kernel by maximum marginal likelihood. TuRBO-1 uses an ARD Matérn-$5/2$ kernel, whose length scales set the shape of the trust region, and Thompson sampling as acquisition function. We run it with two pairs of failure and success tolerances $(\tau_{\text{fail}}, \tau_{\text{succ}})$. With TuRBO-1-Restart (tolerances $(3, 10)$), the trust region converges after an average of 35 iterations, after which it restarts with a new initial random point on $\mathcal{M}$, while with TuRBO-1 (tolerances $(8, 3)$), the region shrinks slowly and a run has no restart for 99.6% of experiments when the budget is 60. SAASBO is fitted by NUTS and uses q-log expected improvement as acquisition function (Ament et al., 2023). Random search on Φ draws ϕ uniformly and evaluates $\mathcal{E}^\star(\phi)$.

Results. Tables S4 and S5 give the simple regret after 60 evaluations, as mean with standard error and as median with quantile range. Table 1 in the main paper is a subset of Table S4, and reports only TuRBO-1, which was consistently better than TuRBO-1-Restart. When $\delta = 0$, the optimum lies on $\mathcal{M}$ and BOASS-k_Φ has the lowest regret in every dimension. When $\delta > 0$, its regret cannot go below $\lambda\delta^2$, and the methods that search $\mathcal{D}$ pass it in low dimension. The dimension from which BOASS-k_Φ is ahead grows with δ: $D \geq 8$ for $\delta = 0.1$, $D \geq 16$ for $\delta = 0.25$, and $D = 32$ for $\delta = 0.5$. At $\delta = 1$, SAASBO has the lowest median regret in every dimension.

Regret curves. Figures S4 and S5 follow BOASS-k_Φ and BO-on-$\mathcal{D}$ along the whole run, in a slightly different setting: $D \in \{6, 8, 12, 16\}$, a finer grid of biases $\delta \in \{0.01, 0.05, 0.1, 0.5, 1\}$, a budget of 80 evaluations, and 50 paired seeds. All models estimate the amplitude and lengthscale, fixing one lengthscale common to all dimensions. The floor $\lambda\delta^2$ is the horizontal asymptote of every BOASS-k_Φ curve, drawn as a dotted line in Figure S4. Since the restricted problem does not depend on D, the BOASS-k_Φ curves are the same in all panels, and all differences between panels come from BO-on-$\mathcal{D}$. Figure S5 reports, for $T \in \{10, 20, 40, 80\}$, the mean over seeds of $r_T^{\text{BO-on-}\mathcal{D}} - r_T^{\text{BOASS-}k_\Phi}$, positive when BOASS-$k_\Phi$ is ahead. For $\delta \leq 0.1$, BOASS-k_Φ is ahead at every budget and every dimension. At $\delta = 0.5$, BO-on-$\mathcal{D}$ passes it between 5 and 30 evaluations for $D = 6$, and between 20 and 40 for $D = 8$. This is the trade-off of Section 4: the restriction pays as long as $\lambda\delta^2$ stays below the regret BO-on-$\mathcal{D}$ can reach within the budget.

K.1.2 Learned embedding and amortization error

In the previous experiments, the BO search space was defined using the optimal embedding $\mathcal{E}^\star$. However, our BOASS method assumes that the embedding is learned using a neural network (Algorithm

Table S4: Effect of search-space bias and ambient dimension on the simple regret of BO methods in the optimization of the Branin function. Simple regret at iteration 60, in units of 10^{-3}: mean $\pm$ standard error of the mean over 25 seeds. Blocks correspond to the five values of δ. $\Delta_{\mathcal{M}} = \delta^2$ (in units of 10^{-3}) is the smallest regret attainable by a method restricted to Φ; a dash marks the methods that search $\mathcal{D}$ directly, for which no such bias applies. Within each block, **bold** marks, for each D column, the lowest mean together with every method whose $\hat{m} \pm$ s.e.m. interval overlaps that of the lowest.

Method	$\Delta_{\mathcal{M}}$	$D = 6$	$D = 8$	$D = 16$	$D = 32$
$\delta = 0$					
BOASS-k_Φ	0	$\mathbf{6.3 \pm 2.5}$	$\mathbf{6.3 \pm 2.5}$	$\mathbf{6.3 \pm 2.5}$	$\mathbf{6.3 \pm 2.5}$
BO-on-$\mathcal{D}$	–	37.2 ± 7.1	74.7 ± 18.0	149.1 ± 26.0	687.5 ± 120.0
SAASBO	–	$\mathbf{9.0 \pm 2.0}$	21.6 ± 3.7	177.0 ± 36.0	489.7 ± 72.0
TuRBO-1	–	$\mathbf{7.7 \pm 1.6}$	31.2 ± 7.6	275.2 ± 48.0	609.8 ± 99.0
TuRBO-1-Restart	–	57.4 ± 17.0	118.6 ± 26.0	437.5 ± 90.0	648.6 ± 110.0
Random search on Φ	0	12.1 ± 2.3	12.1 ± 2.3	12.1 ± 2.3	12.1 ± 2.3
$\delta = 0.1$					
BOASS-k_Φ	10	16.3 ± 2.5	$\mathbf{16.3 \pm 2.5}$	$\mathbf{16.3 \pm 2.5}$	$\mathbf{16.3 \pm 2.5}$
BO-on-$\mathcal{D}$	–	40.8 ± 7.5	68.2 ± 17.0	150.7 ± 24.0	651.7 ± 100.0
SAASBO	–	$\mathbf{11.4 \pm 2.2}$	25.4 ± 4.9	150.0 ± 46.0	464.5 ± 68.0
TuRBO-1	–	$\mathbf{8.5 \pm 2.4}$	33.7 ± 10.0	333.4 ± 58.0	662.1 ± 130.0
TuRBO-1-Restart	–	94.3 ± 41.0	120.5 ± 26.0	472.8 ± 94.0	730.1 ± 140.0
Random search on Φ	10	22.1 ± 2.3	22.1 ± 2.3	22.1 ± 2.3	22.1 ± 2.3
$\delta = 0.25$					
BOASS-k_Φ	62.5	68.8 ± 2.5	68.8 ± 2.5	$\mathbf{68.8 \pm 2.5}$	$\mathbf{68.8 \pm 2.5}$
BO-on-$\mathcal{D}$	–	41.5 ± 11.0	81.0 ± 20.0	154.7 ± 22.0	761.8 ± 110.0
SAASBO	–	$\mathbf{11.4 \pm 2.3}$	$\mathbf{19.0 \pm 3.0}$	108.3 ± 17.0	534.0 ± 59.0
TuRBO-1	–	$\mathbf{8.3 \pm 1.6}$	28.5 ± 6.2	292.2 ± 54.0	636.7 ± 110.0
TuRBO-1-Restart	–	40.7 ± 9.5	66.1 ± 14.0	504.3 ± 86.0	638.4 ± 110.0
Random search on Φ	62.5	74.6 ± 2.3	74.6 ± 2.3	74.6 ± 2.3	74.6 ± 2.3
$\delta = 0.5$					
BOASS-k_Φ	250	256.3 ± 2.5	256.3 ± 2.5	256.3 ± 2.5	$\mathbf{256.3 \pm 2.5}$
BO-on-$\mathcal{D}$	–	48.1 ± 8.2	69.7 ± 14.0	$\mathbf{184.1 \pm 31.0}$	853.8 ± 92.0
SAASBO	–	$\mathbf{8.8 \pm 2.0}$	$\mathbf{22.2 \pm 4.2}$	$\mathbf{140.1 \pm 24.0}$	637.2 ± 57.0
TuRBO-1	–	$\mathbf{7.3 \pm 1.2}$	40.3 ± 8.0	308.3 ± 51.0	764.7 ± 110.0
TuRBO-1-Restart	–	33.7 ± 9.4	76.9 ± 13.0	502.7 ± 78.0	868.3 ± 130.0
Random search on Φ	250	262.1 ± 2.3	262.1 ± 2.3	262.1 ± 2.3	262.1 ± 2.3
$\delta = 1$					
BOASS-k_Φ	1000	1006.0 ± 2.5	1006.0 ± 2.5	1006.0 ± 2.5	1006.0 ± 2.5
BO-on-$\mathcal{D}$	–	52.5 ± 9.9	88.2 ± 20.0	201.2 ± 35.0	1250.0 ± 65.0
SAASBO	–	$\mathbf{13.7 \pm 2.3}$	$\mathbf{22.1 \pm 3.4}$	$\mathbf{125.0 \pm 20.0}$	$\mathbf{705.2 \pm 64.0}$
TuRBO-1	–	$\mathbf{12.7 \pm 1.9}$	50.6 ± 14.0	602.4 ± 66.0	1190.0 ± 93.0
TuRBO-1-Restart	–	37.4 ± 11.0	173.5 ± 43.0	926.8 ± 130.0	1333.0 ± 120.0
Random search on Φ	1000	1012.0 ± 2.3	1012.0 ± 2.3	1012.0 ± 2.3	1012.0 ± 2.3

A). To study the effect of amortization error due, e.g., to poor training of the network, we train a neural network $\widehat{\mathcal{E}}$ to learn the embedding with varying amounts of training data.

Setting. The objective is f_δ, as defined in Section K.1, with $D = 10$ and $\delta = 0.1$, so a search restricted to $\mathcal{M}^\star$ has the floor $\lambda\delta^2 = 0.01$. We vary the size of the training set, $n_{\text{train}} \in \{25, 50, 200, 1000\}$.

For each replicate, we draw n_{train} Sobol points in Φ and fit a multilayer perceptron to the full ideal embedding $\mathcal{E}^\star(\phi) = (\phi, m(\phi)) \in \mathbb{R}^{10}$. The network has two hidden layers of 128 units with $\tanh$ activations. It is trained with Adam (learning rate 10^{-3}, weight decay 10^{-5}), with early stopping on a 15% validation split. Its outputs are clipped to $\mathcal{D} = [-1, 1]^{10}$.

Table S5: Effect of search-space bias and ambient dimension on the simple regret of BO methods in the optimization of the Branin function. Simple regret at iteration 60, in units of 10^{-3}: median with [lower, upper] quantiles over 25 seeds, to three digits. Blocks correspond to the five values of δ. $\Delta_{\mathcal{M}} = \delta^2$ (in units of 10^{-3}) is the smallest regret attainable by a method restricted to Φ; a dash marks the methods that search $\mathcal{D}$ directly, for which no such bias applies. Within each block, **bold** marks the lowest median in each D column.

Method	$\Delta_{\mathcal{M}}$	$D = 6$	$D = 8$	$D = 16$	$D = 32$
$\delta = 0$					
BOASS-k_Φ	0	**0.11 [0.01, 30.9]**	**0.11 [0.01, 30.9]**	**0.11 [0.01, 30.9]**	**0.11 [0.01, 30.9]**
BO-on-$\mathcal{D}$	–	31.6 [5.00, 101]	35.0 [10.8, 217]	100 [49.6, 294]	583 [116, 1450]
SAASBO	–	6.51 [2.16, 19.1]	12.0 [5.15, 43.4]	93.0 [33.5, 447]	460 [116, 882]
TuRBO-1	–	4.35 [2.03, 20.3]	18.6 [4.76, 53.0]	175 [56.7, 707]	446 [104, 1300]
TuRBO-1-Restart	–	26.6 [9.82, 108]	60.2 [26.4, 328]	280 [71.2, 1170]	489 [97.3, 1380]
Random search on Φ	0	8.90 [2.54, 23.7]	8.90 [2.54, 23.7]	8.90 [2.54, 23.7]	8.90 [2.54, 23.7]
$\delta = 0.1$					
BOASS-k_Φ	10	10.1 [10.0, 40.9]	**10.1 [10.0, 40.9]**	**10.1 [10.0, 40.9]**	**10.1 [10.0, 40.9]**
BO-on-$\mathcal{D}$	–	32.1 [5.98, 95.8]	36.9 [8.05, 176]	107 [70.4, 225]	519 [121, 1290]
SAASBO	–	8.58 [2.86, 30.8]	11.9 [6.06, 55.1]	85.1 [33.1, 332]	378 [126, 1010]
TuRBO-1	–	**4.01 [1.39, 15.5]**	17.2 [5.25, 54.7]	295 [43.4, 835]	457 [115, 1430]
TuRBO-1-Restart	–	17.0 [6.67, 247]	79.2 [17.3, 285]	352 [92.0, 1060]	443 [84.4, 1710]
Random search on Φ	10	18.9 [12.5, 33.7]	18.9 [12.5, 33.7]	18.9 [12.5, 33.7]	18.9 [12.5, 33.7]
$\delta = 0.25$					
BOASS-k_Φ	62.5	62.6 [62.5, 93.4]	62.6 [62.5, 93.4]	**62.6 [62.5, 93.4]**	**62.6 [62.5, 93.4]**
BO-on-$\mathcal{D}$	–	31.4 [3.12, 77.0]	36.8 [8.18, 238]	92.9 [66.3, 283]	583 [169, 1540]
SAASBO	–	5.90 [2.00, 27.0]	**14.0 [6.23, 37.7]**	89.2 [33.2, 190]	522 [179, 897]
TuRBO-1	–	**3.96 [2.53, 21.6]**	16.4 [6.27, 64.9]	183 [56.6, 658]	422 [145, 1450]
TuRBO-1-Restart	–	14.1 [6.37, 102]	31.9 [17.4, 141]	369 [132, 1020]	499 [125, 1430]
Random search on Φ	62.5	71.4 [65.0, 86.2]	71.4 [65.0, 86.2]	71.4 [65.0, 86.2]	71.4 [65.0, 86.2]
$\delta = 0.5$					
BOASS-k_Φ	250	250 [250, 281]	250 [250, 281]	250 [250, 281]	**250 [250, 281]**
BO-on-$\mathcal{D}$	–	33.5 [5.08, 98.9]	47.8 [9.19, 200]	102 [67.9, 449]	710 [345, 1490]
SAASBO	–	5.32 [2.12, 25.0]	**15.2 [5.48, 55.7]**	**86.8 [49.8, 350]**	546 [359, 1080]
TuRBO-1	–	**5.31 [1.86, 15.0]**	23.8 [5.62, 105]	227 [110, 625]	604 [300, 1410]
TuRBO-1-Restart	–	22.3 [4.06, 48.0]	55.0 [16.6, 168]	410 [133, 967]	633 [293, 1650]
Random search on Φ	250	259 [253, 274]	259 [253, 274]	259 [253, 274]	259 [253, 274]
$\delta = 1$					
BOASS-k_Φ	1000	1000 [1000, 1030]	1000 [1000, 1030]	1000 [1000, 1030]	1000 [1000, 1030]
BO-on-$\mathcal{D}$	–	32.6 [31.3, 109]	35.7 [21.8, 248]	114 [65.4, 448]	1210 [868, 1720]
SAASBO	–	**8.19 [2.91, 32.5]**	**14.5 [7.95, 43.6]**	**86.1 [53.5, 237]**	**682 [393, 888]**
TuRBO-1	–	10.4 [3.65, 26.3]	29.7 [14.0, 63.8]	685 [223, 994]	1160 [656, 1940]
TuRBO-1-Restart	–	17.0 [3.51, 64.9]	59.9 [14.2, 396]	1010 [159, 1780]	1120 [688, 2110]
Random search on Φ	1000	1010 [1000, 1020]	1010 [1000, 1020]	1010 [1000, 1020]	1010 [1000, 1020]

Methods and protocol. We compare four methods. BOASS-$k_{\mathcal{D}}$, BOASS-k_Φ, BOASS-$k_{\mathcal{D}}$-oracle, and BO-on-$\mathcal{D}$. BOASS-$k_{\mathcal{D}}$ and BOASS-k_Φ differ only in the GP input, $\widehat{\mathcal{E}}(\phi)$ or ϕ respectively, and evaluate the objective at $\widehat{\mathcal{E}}(\phi)$. BOASS-$k_{\mathcal{D}}$-oracle uses $\mathcal{E}^\star$ instead, and BO-on-$\mathcal{D}$ is the unrestricted BO reference.

The surrogates, the noise and the LCB acquisition are as in Section K.1. The optimization budget is 80 iterations, with the acquisition minimized over 4096 candidates per step. Each configuration is repeated over 20 seeds. To each seed corresponds a trained network, the initial candidate, and the noise.

Results. Figure S6 shows the median simple regret and interquartile range, one panel per n_{train}. The floor $\lambda\delta^2$ is drawn as a dotted line. The oracle and BO-on-$\mathcal{D}$ do not depend on $\widehat{\mathcal{E}}$, so they give similar curves at every n_{train} and serve as fixed references. The approximation error incurs larger regret for BOASS-$k_{\mathcal{D}}$ and BOASS-k_Φ compared to BOASS-$k_{\mathcal{D}}$-oracle , since they rely $\widehat{\mathcal{E}}(\phi)$ instead

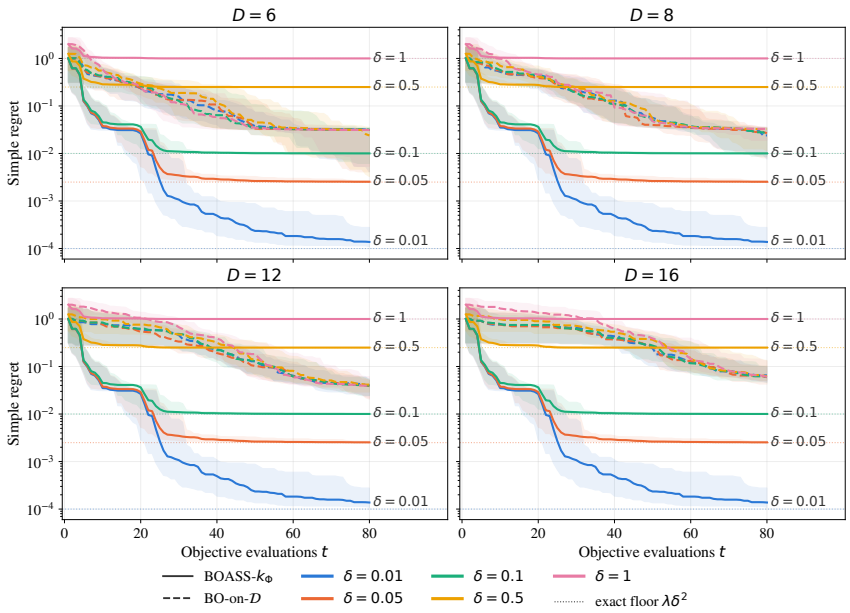

Figure S4: Effect of search-space bias and ambient dimension on the simple regret of BOASS in the optimization of the Branin function. Median simple regret of BOASS-k_Φ (solid lines) and BO-on-$\mathcal{D}$ (dashed lines), interquartile range highlighted. Each plot corresponds to a dimension of $\mathcal{D}$. Within each plot, the results for the different bias parameters δ are plotted.

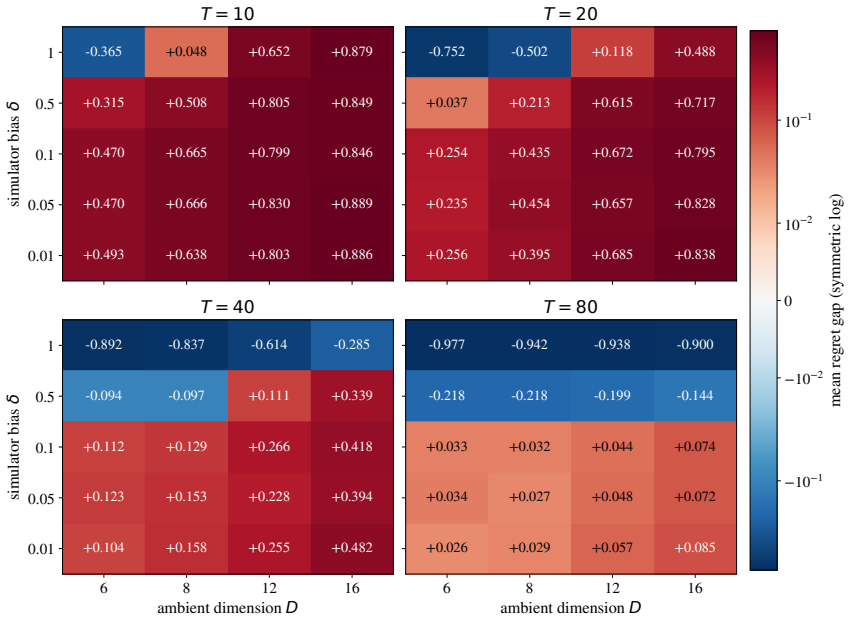

Figure S5: Comparison of simple regret in the optimization of the Branin function between BOASS-k_Φ and BO-on-$\mathcal{D}$, depending on search-space bias and ambient dimension. Each plot corresponds to a dimension of $\mathcal{D}$. Within each plot, each cell is the gap between the simple regret of BO-on-$\mathcal{D}$ and that of BOASS-k_Φ (higher means BOASS-k_Φ has a lower regret).

of $\mathcal{E}^\star(\phi)$. More precisely, for a small training set, both BOASS-$k_\mathcal{D}$ and BOASS-k_Φ it a regret lower bound that is higher than the floor $\lambda\delta^2$, but still have better or similar performance compared to BO-on-$\mathcal{D}$. This can be explained by the gains from a restricted search space that compensate for the

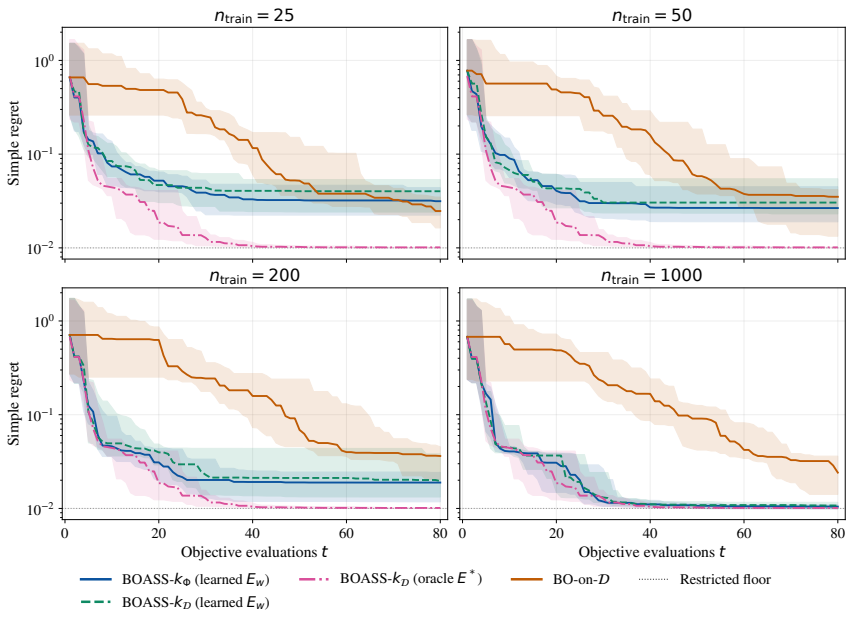

Figure S6: Effect of amortization error on the simple regret of BOASS during the optimization of the Branin function, with a learned embedding $\widehat{\mathcal{E}}$. Lines correspond to the median simple regret, interquartile range highlighted. Each panel corresponds to a training-set size n_{train}. The dotted line is the floor $\lambda\delta^2$ of the search on $\mathcal{M}^\star$.

amortization bias, which is consistent with our regret analysis in Section 4. As the size of the training set increases, the amortization gap vanishes, so that the performance of BOASS-$k_{\mathcal{D}}$ and BOASS-k_Φ reaches the one of BOASS-$k_{\mathcal{D}}$-oracle.

K.2 Objective drawn from a Gaussian process

Setting. The Branin objective is not a sample from the GP prior that is used when fitting the GP regression surrogate model. To remove any effect that may be due to the surrogate GP kernel being misspecified, we therefore use a second benchmark in which the objective is drawn from the same Gaussian process prior the surrogate assumes.

Let $\Phi = [-1, 1]^p$, and $\mathcal{D} = [-1, 1]^D$. Let $f \sim \text{GP}(0, k_D)$ be the objective function defined in $\mathcal{D}$. We define

$$\mathcal{E}_D : \Phi \to \mathcal{D} \tag{52}$$

$$\phi \mapsto W g(\phi) \tag{53}$$

where $g(\phi) = (\phi + 1)(\phi - 1)\sin(a\pi\phi + b)/2 + \phi$ acts coordinatewise, and a, b are scalar hyperparameters.

The function g maps $[-1, 1]$ onto (surjectively) $[-1, 1]$, and parameter a dictates the level of non-injectivity of the function. To ensure full rank of W each row is sampled independently from a Dirichlet distribution, $W_i \sim \text{Dir}(\alpha = (1, \ldots, 1))$. Since W has full column rank almost surely, the non-injectivity of $\mathcal{E}_D$ comes from g alone.

The kernel k_D is a Matérn-$5/2$ kernel whose lengthscale is set to one half of the median pairwise distance of $\mathcal{E}_D(\Phi)$, so that f_D does not become rougher as D grows. Observations are corrupted by Gaussian noise $N(0, 0.01^2)$.

What we vary. We run every combination of $p \in \{2, 4, 8\}$ and $D \in \{p, 10, 20\}$, with the linear embedding $\mathcal{E}_D(\phi) = W\phi$ and with $a \in \{0.5, 0.7, 0.9, 1.1\}$. At each experiment we sample a random value of b, uniformly in $[0, \pi]$. The pair (p, D) sets the dimension of the search space and the dimension of the design space it is mapped into, while a sets how much the embedding folds, which

we summarize by the fraction ν defined below. Note that $a = 0.5$ makes g curved while keeping it injective, which separates the effect of curvature from the effect of folding. No search-space bias is introduced, since the reference value is the best value found on $\mathcal{M}$.

Methods and protocol. We compare the three restricted methods of Table S3: BOASS-$k_{\mathcal{D}}$, BOASS-k_{Φ}, and BOASS-random-k_{Φ}. Both BOASS methods use GP-LCB[2], $\mu_t - \beta\sigma_t$ with $\beta = 2$ fixed, minimised at each step over 2048 fresh candidates. Each run has 5 initial Sobol points, common to the three methods. The surrogates are a Matérn-$5/2$ Kernel with a single lengthscale. The signal variance s^2 and the length scale ℓ are re-estimated by maximum marginal likelihood at every step, while the noise variance is held at its true value $\sigma^2 = 0.01^2$. The prior mean is zero, which is exact here. Each configuration is repeated over 50 replicates, and within a replicate the three methods share the same realization of f.

Sampling the objective. We instantiate the GP sample f sequentially, by sampling, for each new query, $f_{t+1} \sim \mathcal{N}\left(\mu_t(\mathbf{x}_{t+1}), \sigma_t^2(\mathbf{x}_{t+1})\right)$, where μ_t and σ_t are the posterior mean of variance of the GP given $\{(\mathbf{x}_s, f_s)\}_{s \le t}$, the samples pooled across the evaluated methods. Pooling the samples ensure that all methods share the same GP sample. . The observations are corrupted by i.i.d. Gaussian noise, that is $y_{t+1} = f_{t+1} + \epsilon_{t+1}$ with $\epsilon_{t+1} \sim \mathcal{N}(0, \sigma^2)$.

By the chain rule for Gaussian conditionals, this approach leads to an exact draw from the finite-dimensional distribution of the prior.

Non-injectivity of $\mathcal{E}_D$. For each configuration, we report ν, the fraction of $\Phi = [-1, 1]^p$ whose image is shared with some other point,

$$\nu \;=\; \frac{\lambda(\{\phi \in \Phi \;:\; \exists\, \phi' \neq \phi,\; \mathcal{E}_D(\phi') = \mathcal{E}_D(\phi)\})}{\lambda(\Phi)},$$

with λ the Lebesgue measure. Because W has full column rank p almost surely, $\mathcal{E}_D(\phi) = \mathcal{E}_D(\phi')$ holds iff $g(\phi) = g(\phi')$, and g acts coordinatewise; hence ϕ has a twin iff *at least one* of its coordinates does. Writing

$$q_i \;=\; \tfrac{1}{2}\,\lambda\big(\{u \in [-1, 1] \;:\; g_i^{-1}(g_i(u)) \neq \{u\}\}\big)$$

for the folded fraction of the i-th coordinate map, independence of the coordinates under the uniform measure on Φ gives

$$\nu \;=\; 1 - \prod_{i=1}^{p} (1 - q_i)\,.$$

Each q_i is evaluated numerically: g_i is split into monotone branches at its turning points, each branch covering an interval of values, and u is counted whenever $g_i(u)$ lies in the value range of a branch other than its own. We use a uniform grid of $2 \times 10^4 + 1$ points, at which the estimate is converged to five decimal places. At fixed a, ν grows with p, since a point is folded as soon as one of its coordinates is.

Results. We report, for each replicate, the gap between the best value a method has found and the best value found by any of the three methods on that replicate. The best method of a replicate has a gap of zero by construction, so only comparisons within a configuration are meaningful. Tables S6 and S7 give the mean gap and its standard error over the 50 replicates, after 30 and 60 iterations.

When the function $\mathcal{E}_D$ is not linear, BOASS-$k_{\mathcal{D}}$ has the lowest mean gap in every configuration after 30 and 60 iterations. The two kernels are indistinguishable when the embedding is linear and $p = 2$, and they separate as soon as g is curved, already at $a = 0.5$ where the map is still injective. The distance between them grows with p: at $p = 8$ the mean gap of BOASS-k_{Φ} is about twice that of BOASS-$k_{\mathcal{D}}$ after 30 iterations, and more after 60, whether the embedding folds or not. The ambient dimension D has a smaller effect than p, in line with the analysis of Section 4, where the rate depends on the intrinsic dimension of $\mathcal{M}$ rather than on D. Random search on Φ is behind both BOASS variants everywhere, so the gain does not come from the restriction of the search space alone.

[2] GP-UCB for minimization

Table S6: Comparison of BOASS variants in the optimization of a sample from a Gaussian process. Gap to the best method after 30 iterations, in units of 10^{-2}: mean $\pm$ standard error of the mean over $N = 50$ paired replicates. On each replicate the gap of a method is its best objective value found minus the lowest value found by any of the three methods on that same replicate; the best method therefore has a gap of 0 by construction, and only within-row contrasts are interpretable. **Bold**: lowest mean in the row, together with every method whose $\hat{m} \pm \mathrm{s.e.m.}$ interval overlaps that of the lowest.

p	D	ν	BOASS-$k_{\mathcal{D}}$	BOASS-k_{Φ}	BOASS-random-k_{Φ}
no fold ($\mathcal{E}_D(\phi) = W\phi$)					
2	2	0.0%	$\mathbf{1.80 \pm 1.50}$	$\mathbf{2.70 \pm 1.60}$	23.01 ± 4.80
2	10	0.0%	$\mathbf{4.58 \pm 1.70}$	$\mathbf{3.63 \pm 1.50}$	27.35 ± 4.40
2	20	0.0%	$\mathbf{6.86 \pm 2.70}$	$\mathbf{6.36 \pm 2.60}$	27.20 ± 4.80
4	4	0.0%	$\mathbf{17.34 \pm 4.40}$	32.80 ± 4.40	59.70 ± 5.90
4	10	0.0%	$\mathbf{25.08 \pm 4.50}$	$\mathbf{18.26 \pm 3.20}$	77.67 ± 5.70
4	20	0.0%	$\mathbf{17.47 \pm 4.40}$	27.96 ± 6.00	70.30 ± 6.30
8	8	0.0%	$\mathbf{31.44 \pm 5.30}$	62.61 ± 5.90	136.50 ± 7.70
8	10	0.0%	$\mathbf{43.86 \pm 6.30}$	75.59 ± 7.00	145.10 ± 8.90
8	20	0.0%	$\mathbf{44.97 \pm 6.40}$	67.48 ± 6.50	144.10 ± 8.40
$a = 0.5$					
2	2	0.0%	$\mathbf{1.76 \pm 0.83}$	7.55 ± 3.00	24.51 ± 4.10
2	10	0.0%	$\mathbf{9.55 \pm 3.40}$	20.35 ± 4.90	30.84 ± 4.80
2	20	0.0%	$\mathbf{5.04 \pm 2.50}$	11.53 ± 3.70	25.60 ± 4.60
4	4	0.0%	$\mathbf{16.50 \pm 5.00}$	34.98 ± 5.60	58.51 ± 7.50
4	10	0.0%	$\mathbf{16.39 \pm 3.60}$	37.00 ± 6.10	67.79 ± 5.70
4	20	0.0%	$\mathbf{17.32 \pm 4.00}$	$\mathbf{24.05 \pm 4.20}$	67.08 ± 6.10
8	8	0.0%	$\mathbf{29.31 \pm 4.30}$	61.30 ± 5.60	118.80 ± 8.50
8	10	0.0%	$\mathbf{30.68 \pm 4.50}$	77.13 ± 7.20	117.80 ± 9.20
8	20	0.0%	$\mathbf{42.72 \pm 6.00}$	73.93 ± 6.70	127.70 ± 8.80
$a = 0.7$					
2	2	25.7%	$\mathbf{4.29 \pm 2.00}$	14.54 ± 4.60	37.14 ± 6.20
2	10	25.7%	$\mathbf{4.62 \pm 2.60}$	13.55 ± 4.00	36.91 ± 5.40
2	20	25.7%	$\mathbf{0.16 \pm 0.09}$	10.22 ± 4.10	28.84 ± 4.80
4	4	25.7%	$\mathbf{13.42 \pm 3.90}$	31.43 ± 4.20	57.10 ± 6.40
4	10	25.7%	$\mathbf{16.74 \pm 3.90}$	47.88 ± 6.60	65.14 ± 7.10
4	20	25.7%	$\mathbf{14.96 \pm 3.70}$	35.73 ± 6.50	68.55 ± 6.20
8	8	40.4%	$\mathbf{29.32 \pm 5.50}$	81.83 ± 7.00	118.50 ± 9.20
8	10	40.4%	$\mathbf{33.64 \pm 5.80}$	72.24 ± 7.00	115.50 ± 9.00
8	20	40.4%	$\mathbf{38.47 \pm 5.30}$	75.48 ± 6.40	113.30 ± 8.00
$a = 0.9$					
2	2	43.6%	$\mathbf{2.16 \pm 1.20}$	12.57 ± 3.90	31.16 ± 5.70
2	10	43.6%	$\mathbf{4.87 \pm 2.70}$	13.52 ± 3.80	31.94 ± 5.20
2	20	43.6%	$\mathbf{3.38 \pm 2.00}$	11.43 ± 3.60	27.82 ± 5.20
4	4	46.4%	$\mathbf{11.46 \pm 3.50}$	39.52 ± 4.60	51.43 ± 5.50
4	10	46.4%	$\mathbf{13.35 \pm 3.10}$	40.31 ± 5.20	66.67 ± 5.20
4	20	46.4%	$\mathbf{8.87 \pm 2.60}$	28.44 ± 4.30	57.38 ± 5.50
8	8	68.0%	$\mathbf{28.01 \pm 4.70}$	81.38 ± 9.00	118.70 ± 9.10
8	10	68.0%	$\mathbf{27.15 \pm 4.30}$	63.07 ± 5.60	109.30 ± 7.30
8	20	68.0%	$\mathbf{40.06 \pm 6.00}$	81.85 ± 6.60	127.30 ± 8.90
$a = 1.1$					
2	2	67.8%	$\mathbf{4.84 \pm 2.80}$	15.90 ± 3.80	32.10 ± 5.90
2	10	67.8%	$\mathbf{2.59 \pm 2.10}$	22.97 ± 5.30	35.80 ± 5.60
2	20	67.8%	$\mathbf{5.84 \pm 2.80}$	17.75 ± 4.60	30.41 ± 5.20
4	4	87.3%	$\mathbf{8.15 \pm 2.60}$	33.69 ± 4.30	37.91 ± 3.90
4	10	87.3%	$\mathbf{13.08 \pm 3.40}$	45.76 ± 6.40	63.14 ± 6.40
4	20	87.3%	$\mathbf{27.56 \pm 6.30}$	45.01 ± 7.60	75.84 ± 7.30
8	8	98.3%	$\mathbf{33.17 \pm 6.40}$	85.22 ± 8.30	111.60 ± 8.80
8	10	98.3%	$\mathbf{33.48 \pm 5.40}$	87.27 ± 6.60	130.10 ± 8.00
8	20	98.3%	$\mathbf{38.85 \pm 6.10}$	88.63 ± 7.40	125.00 ± 8.90

Table S7: Comparison of BOASS variants in the optimization of a sample from a Gaussian process. Gap to the best method after 60 iterations, same specifications as in Table S6.

p	D	ν	BOASS-$k_{\mathcal{D}}$	BOASS-k_{Φ}	BOASS-random-k_{Φ}
no fold ($\mathcal{E}_D(\phi) = W\phi$)					
2	2	0.0%	$\mathbf{1.77 \pm 1.50}$	$\mathbf{0.54 \pm 0.38}$	7.65 ± 2.20
2	10	0.0%	$\mathbf{4.45 \pm 1.70}$	$\mathbf{2.83 \pm 1.40}$	9.43 ± 2.10
2	20	0.0%	$\mathbf{6.41 \pm 2.60}$	$\mathbf{6.15 \pm 2.60}$	$\mathbf{10.62 \pm 3.00}$
4	4	0.0%	$\mathbf{12.35 \pm 4.00}$	$\mathbf{19.84 \pm 4.40}$	34.01 ± 4.70
4	10	0.0%	17.49 ± 3.70	$\mathbf{9.62 \pm 2.50}$	47.09 ± 5.20
4	20	0.0%	$\mathbf{10.93 \pm 3.70}$	$\mathbf{7.35 \pm 2.50}$	44.08 ± 5.40
8	8	0.0%	$\mathbf{9.36 \pm 3.40}$	29.33 ± 5.10	95.65 ± 7.20
8	10	0.0%	$\mathbf{18.04 \pm 5.30}$	29.43 ± 5.10	100.70 ± 7.70
8	20	0.0%	$\mathbf{15.68 \pm 3.70}$	29.33 ± 4.40	97.40 ± 7.80
$a = 0.5$					
2	2	0.0%	$\mathbf{1.74 \pm 0.83}$	6.81 ± 2.90	12.84 ± 2.90
2	10	0.0%	$\mathbf{9.41 \pm 3.40}$	$\mathbf{15.99 \pm 4.60}$	17.29 ± 3.70
2	20	0.0%	$\mathbf{4.94 \pm 2.50}$	$\mathbf{9.59 \pm 3.50}$	10.77 ± 3.00
4	4	0.0%	$\mathbf{7.12 \pm 2.90}$	20.11 ± 4.80	28.68 ± 4.60
4	10	0.0%	$\mathbf{8.67 \pm 2.80}$	16.74 ± 5.00	40.63 ± 4.50
4	20	0.0%	$\mathbf{11.70 \pm 3.80}$	$\mathbf{12.82 \pm 4.00}$	41.03 ± 5.00
8	8	0.0%	$\mathbf{10.27 \pm 2.70}$	31.26 ± 4.40	77.00 ± 7.20
8	10	0.0%	$\mathbf{9.34 \pm 3.00}$	45.72 ± 6.10	77.97 ± 7.30
8	20	0.0%	$\mathbf{16.27 \pm 4.40}$	42.80 ± 6.70	90.87 ± 8.30
$a = 0.7$					
2	2	25.7%	$\mathbf{4.12 \pm 2.10}$	$\mathbf{5.84 \pm 1.90}$	14.93 ± 2.90
2	10	25.7%	$\mathbf{3.69 \pm 2.40}$	$\mathbf{8.92 \pm 3.10}$	23.23 ± 4.80
2	20	25.7%	$\mathbf{0.03 \pm 0.01}$	7.65 ± 3.90	14.26 ± 3.90
4	4	25.7%	$\mathbf{7.91 \pm 2.80}$	15.87 ± 3.30	30.40 ± 4.00
4	10	25.7%	$\mathbf{8.31 \pm 3.00}$	24.29 ± 4.60	43.37 ± 5.90
4	20	25.7%	$\mathbf{8.92 \pm 3.00}$	21.18 ± 4.90	36.97 ± 5.00
8	8	40.4%	$\mathbf{7.97 \pm 2.50}$	51.86 ± 6.90	79.82 ± 8.80
8	10	40.4%	$\mathbf{8.13 \pm 2.40}$	47.16 ± 6.00	73.57 ± 7.70
8	20	40.4%	$\mathbf{15.43 \pm 3.80}$	41.30 ± 5.80	80.56 ± 7.50
$a = 0.9$					
2	2	43.6%	$\mathbf{2.10 \pm 1.20}$	$\mathbf{4.76 \pm 1.50}$	20.01 ± 5.10
2	10	43.6%	$\mathbf{4.83 \pm 2.70}$	$\mathbf{10.53 \pm 3.70}$	23.16 ± 5.00
2	20	43.6%	$\mathbf{3.32 \pm 2.00}$	$\mathbf{4.47 \pm 2.50}$	14.32 ± 3.90
4	4	46.4%	$\mathbf{6.70 \pm 2.70}$	23.35 ± 4.30	32.05 ± 4.40
4	10	46.4%	$\mathbf{9.75 \pm 2.80}$	25.45 ± 4.80	36.42 ± 3.90
4	20	46.4%	$\mathbf{4.95 \pm 2.10}$	12.33 ± 3.10	33.82 ± 4.20
8	8	68.0%	$\mathbf{9.32 \pm 2.70}$	53.59 ± 8.20	86.77 ± 8.70
8	10	68.0%	$\mathbf{9.34 \pm 3.00}$	42.80 ± 4.80	70.03 ± 6.70
8	20	68.0%	$\mathbf{10.39 \pm 3.70}$	45.01 ± 5.10	86.90 ± 7.80
$a = 1.1$					
2	2	67.8%	$\mathbf{4.61 \pm 2.80}$	$\mathbf{10.56 \pm 3.30}$	15.99 ± 3.30
2	10	67.8%	$\mathbf{2.48 \pm 2.10}$	13.24 ± 4.30	25.25 ± 4.90
2	20	67.8%	$\mathbf{5.80 \pm 2.80}$	$\mathbf{9.78 \pm 3.10}$	21.16 ± 4.40
4	4	87.3%	$\mathbf{3.12 \pm 1.70}$	23.25 ± 3.90	26.17 ± 3.30
4	10	87.3%	$\mathbf{9.22 \pm 2.90}$	26.04 ± 5.20	44.96 ± 5.90
4	20	87.3%	$\mathbf{15.90 \pm 4.60}$	$\mathbf{21.94 \pm 5.60}$	53.12 ± 6.50
8	8	98.3%	$\mathbf{8.18 \pm 2.50}$	58.61 ± 7.50	72.53 ± 8.00
8	10	98.3%	$\mathbf{13.49 \pm 3.60}$	55.22 ± 6.60	91.12 ± 7.30
8	20	98.3%	$\mathbf{8.38 \pm 2.80}$	59.35 ± 6.70	92.40 ± 8.30