

# Few-Shot Closed-Loop Performance Optimization with Bayesian Meta-Learning

Chakrabarty, Ankush

TR2022-160 December 09, 2022

## Abstract

Bayesian optimization (BO) has been widely adopted for optimizing closed-loop performance in data-limited settings, especially for systems with unmodeled dynamics or performance functions. The BO algorithm efficiently trades-off exploration and exploitation by leveraging uncertainty estimates using surrogate models learned using data from the target closed-loop system to be optimized. The convergence rate of BO can be greatly improved if the underlying surrogate model can accurately predict the target system performance despite having very limited data. In this paper, we propose the use of Bayesian meta-learning to generate an initial surrogate model based on data collected from closed-loop performance optimization tasks performed on a variety of systems that are different to the target system. To this end, we employ deep kernel networks (DKNs) which are simple to train and which comprise encoded Gaussian process models that integrate seamlessly with classical BO. The effectiveness of our proposed DKN-BO approach for speeding up closed-loop performance optimization is demonstrated using a well-studied uncertain nonlinear system with unknown dynamics and an unmodeled performance function.

*IEEE Conference on Decision and Control (CDC) 2022*



# Few-Shot Closed-Loop Performance Optimization with Bayesian Meta-Learning

Ankush Chakrabarty<sup>†</sup>

**Abstract**—Bayesian optimization (BO) has been widely adopted for optimizing closed-loop performance in data-limited settings, especially for systems with unmodeled dynamics or performance functions. The BO algorithm efficiently trades-off exploration and exploitation by leveraging uncertainty estimates using surrogate models learned using data from the target closed-loop system to be optimized. The convergence rate of BO can be greatly improved if the underlying surrogate model can accurately predict the target system performance despite having very limited data. In this paper, we propose the use of Bayesian meta-learning to generate an initial surrogate model based on data collected from closed-loop performance optimization tasks performed on a variety of systems that are different to the target system. To this end, we employ deep kernel networks (DKNs) which are simple to train and which comprise encoded Gaussian process models that integrate seamlessly with classical BO. The effectiveness of our proposed DKN-BO approach for speeding up closed-loop performance optimization is demonstrated using a well-studied uncertain nonlinear system with unknown dynamics and an unmodeled performance function.

**Index Terms**—Bayesian optimization, kernel methods, neural networks, latent variables, few-shot optimization.

## I. INTRODUCTION

Performance-driven optimization of closed-loop control systems often involves optimizing system-specific performance functions. Such performance functions often lack a complete mathematical representation or exhibit such high model complexity that standard model-based techniques cannot be deployed without oversimplifying assumptions or introducing conservativeness. As argued in [1], Bayesian optimization (BO) possesses some clear advantages that have led to its adoption for closed-loop performance optimization [2]–[5]. Namely, BO is a gradient-free, global optimization method that works well in data-limited settings while concurrently learning a probabilistic approximation of the underlying performance function from data. This has led BO to be implemented in a range of engineering applications, including energy systems [6]–[8], vehicles [9], robotics [10], and aerospace [11].

In the performance-driven BO literature, a standard assumption is that the target system to be optimized is available for closed-loop experimentation so that, at design time, one can evaluate the performance of the system by changing some closed-loop system parameters such as controller gains or setpoints. Furthermore, it is often the case that there is very limited data available from the target system at the start of

the BO procedure. Therefore, a learned surrogate model of the performance function constructed with this limited data will initially be inaccurate, and will need to be improved in subsequent BO iterations by exploring the search space [12]. In many industrial applications, however, there is often data collected from previous performance optimization tasks on similar, but not identical, closed-loop control systems. We posit that this previously collected data, combined with the limited target system data, can warm-start BO by estimating a good initial surrogate model, potentially improving the convergence of BO for the target task.

To this end, we propose the use of few-shot meta-learning (or just meta-learning for brevity), wherein a machine learning model is trained on a variety of related ‘source’ optimization tasks so that it can make accurate predictions for a new target task, in spite of limited data points available from that target task. This typically involves solving a bi-level optimization problem, where the outer level is dedicated to extracting task-independent features across a set of source tasks, and the inner level is devoted to quickly adapting to a specific optimization task. A widely used meta-learning framework that employs this bi-level optimization for training deep neural networks is model-agnostic meta-learning (MAML) [13]. However, as shown recently in [14], MAML is not always easy to train and often exhibits numerical instabilities. In addition, MAML does not generate uncertainty estimates around the predictions since it is a deterministic algorithm. Consequently, we adopt a Bayesian meta-learning approach in this work, which builds on the methodology proposed in [15]–[17]. To the best of our knowledge, this paper is the first to investigate the effectiveness of meta-learning for closed-loop system optimization.

Our main **contribution** in this work is to propose a meta-learned Bayesian optimization methodology for rapidly optimizing closed-loop performance of systems with dynamics and performance functions whose mathematical representations are unknown, but can be evaluated by experiment/simulation. Specifically, we employ Bayesian meta-learning to systematically learn from data collected from a variety of systems that are different to a target system to be optimized, and leverage this disparate dataset to generate a good surrogate model for the target system performance function, despite having very few data samples from the target system. Our Bayesian meta-learning framework is implemented using deep kernel networks (DKNs) which has the advantage of requiring single-level optimization for training as the inner training loop is replaced by a Gaussian process

(GP) base kernel. Additionally, our approach is simple to implement using standard deep learning toolkits such as PyTorch. The choice of GPs as the base learner lends itself to classical BO, where the default choice of probabilistic surrogate model is a GP. We demonstrate via simulation experiments that this performance-driven meta-learned BO can generate near-optimal solutions with fewer online experiments than classical BO for a class of parameter-uncertain systems with unmodeled dynamics, without requiring explicit parameter estimation. While meta-learning has recently been explored in the context of adaptive control in [?], [18], [19], our work differs from these works in a few key aspects: (i) we propose meta-learning for optimizing closed-loop performance rather than controller gain adaptation for tracking/regulation, and (ii) we adopt Bayesian deep learning for meta-learning and instead of adaptive controllers; and, (iii) we do not assume any modeling knowledge on the closed-loop dynamics, controller structure, or performance function.

The rest of the paper is organized as follows. The closed-loop performance optimization problem is formalized in Section II and the meta-learning setup is explained therein. In Section III, we provide a brief overview of BO and deep kernel networks, and describe how they are used for meta-learned BO in Section IV. We demonstrate the effectiveness of the proposed approach using a numerical example in Section V, and conclude in Section VI.

## II. MOTIVATION

We consider a class of stable closed-loop systems of the form

$$x_+ = f(x, r, \theta) \quad (1)$$

where  $x, x_+ \in \mathbb{R}^{n_x}$  denote the current and updated state of the system, respectively,  $r \in \mathcal{R} \subset \mathbb{R}^{n_r}$  is a control parameter to be tuned (e.g., a setpoint or controller gain), and  $\theta \in \Theta \subset \mathbb{R}^{n_\theta}$  denotes unknown system parameters. The admissible sets of control parameters and system parameters:  $\mathcal{R}$  and  $\Theta$ , respectively, are known. Additionally, we assume that for each  $\theta \in \Theta$  and each  $r \in \mathcal{R}$ , the closed-loop system (1) is globally asymptotically stable to a parameter-dependent equilibrium state  $x^\infty(r, \theta)$ , and that the map  $x^\infty(\cdot, \cdot)$  is continuous on  $\mathcal{R} \times \Theta$ . To determine how the closed-loop system performs, we define a continuous function  $J : \mathcal{R} \times \Theta \rightarrow \mathbb{R}$ . The performance output  $J(r, \theta)$  is measured. Note that for our purposes, it suffices that  $f$  and  $J$  are smooth and can be evaluated either by simulation or experiments—however, we do not assume access to a mathematical representation of these functions, nor can we compute their gradients analytically.

For a target system (1) with unknown system parameter  $\theta$ , our *target optimization task* is to compute the control parameter

$$r^* := \arg \max_{r \in \mathcal{R}} J(r, \theta) \quad (2)$$

that optimizes the performance of the closed-loop system (1) directly from data. We will do this without estimating  $\theta$ ,

which would prove difficult since  $f$  is unknown, and  $x$  is not measured.

In most industrial applications that involve optimization of closed-loop performance, it is often the case that the optimization task has been performed before, most likely for similar, but not necessarily identical, systems. While solving these *source optimization tasks*, one has likely generated optimization-relevant data that can prove valuable to the target optimization task, even if the target system is not exactly the same as one of the systems encountered before. In BO-based tuning, such source data (data obtained while solving source optimization tasks) is typically ignored, and the target system is optimized from scratch [1]. This can lead to slow convergence of BO since the algorithm’s initial estimate of the target performance function, constructed with very few target data points, is often inaccurate or highly uncertain. In this work, we propose meta-learning from a variety of source data and integrating the information extracted from the source data to target optimization task. In the sequel, we will show that meta-learning from disparate sources of data can enable optimization of closed-loop performance in a few-shot manner (i.e., with very limited target system data) and thereby lead to significant acceleration of the target optimization procedure.

**Example 1.** As a motivating example, consider a common task in smart building control: to minimize energy consumption of a newly constructed target building by tuning controllable variables in heating/cooling systems [6]. At design time, it is likely that very little data has been collected from the target building, and using classical BO will require first exploring the search space until a good estimate of the energy function is learned. Conversely, we posit that the designer could have access to larger quantities of source data from other buildings that are similar in architectural style, location, occupancy, and HVAC equipment. Even if the energy functions learned from those source tasks are not identical to the energy function of the target building, a learner can extract information (e.g., curvature) about classes of energy functions, which can be used to rapidly optimize energy in the target building. ■

Concretely, we assume that we have access to a dataset  $\mathcal{D}^S \triangleq \{\mathcal{D}_k^S\}_{k=1}^{N_S}$ , collected from  $N_S \in \mathbb{N}$  source optimization tasks performed in the past, where

$$\mathcal{D}_k^S = \{r_{t,k}, J(r_{t,k}, \theta_k)\}_{t=1}^{T_k}. \quad (3)$$

That is, the  $k$ -th source optimization task generated a data sequence of  $(r, J)$  pairs over  $T_k \in \mathbb{N}$  optimization iterations. Each  $k$ -th source data pair consists of the control parameter  $r_{t,k}$ , and the corresponding performance output  $J(r_{t,k}, \theta_k)$  evaluated on a source system with dynamics modeled by (1) with system parameter  $\theta_k \in \Theta$ , where each  $\theta_k$  is unique. We reiterate that we do not assume access to the source systems or the system parameters  $\theta_k$ . For the target task, our objective is to compute (2) by learning a variety of performance functions from the source dataset  $\mathcal{D}^S$  and generating a good estimate of the performance function for the target

optimization task, despite having access to a very limited target dataset

$$D^T = \{r_t, J(r_t, \theta)\}_{t=1}^T,$$

where  $T \ll \min_k T_k$ .

To this end, we propose the use of Bayesian meta-learning to learn a distribution of performance functions from the source data, and predict the target performance function by conditioning on the target data. With successful meta-learning, we expect that the estimated target task performance function will be significantly more accurate than an initial estimate that uses only the limited target dataset, and in turn, expect this to improve the convergence of BO. Our meta-learning framework comprises a deep kernel network (DKN). Deep kernel networks comprise of a latent encoder and a final Gaussian process layer. The latent encoder takes an input  $r$  and transforms it to a latent variable of user-defined size, thereby lending itself to target systems with a large number of control parameters. Since the final layer is a Gaussian process regressor, it integrates seamlessly into classical BO procedures and well-studied acquisition functions can be used with the DKN without modification.

### III. PRELIMINARIES

In this section, we explain how classical Bayesian optimization is typically used to solve (2) for the closed-loop system (1) for a fixed  $\theta$ , assuming the absence of source task data. We also describe deep kernel networks that will be used in the next section for meta-learning.

#### A. Bayesian Optimization (BO)

Classical BO comprises two main components: (1) a probabilistic surrogate model that maps  $r \rightarrow J$  for  $r \in \mathcal{R}$ , and (2) an acquisition functions that balances exploration and exploitation to propose new candidate control parameters  $r$  that are likely to optimize  $J$ , based on the predictions of the surrogate model. Since BO starts with no initial data available for the system under consideration, classical BO typically starts with some random sampling on  $\mathcal{R}$ . For each such control parameter  $r_i$ ,  $i \in \mathbb{N}$ , the closed-loop system (1) is allowed to achieve a steady-state after some suitable wait time  $t_f$ , after which a measurement  $J$  is taken. This steady-state performance output acts as a proxy for the objective function value for the corresponding  $r_i$ , that is,  $J_i = J(t_0 + it_f)$ . This enables us to obtain an initial set of  $(r_t, J_t)$  pairs from which a probabilistic surrogate model can be trained to estimate  $J$  on  $\mathcal{R}$ .

In particular, classical BO methods utilize Gaussian process (GP) regression [20] to construct the surrogate model. The Gaussian process framework works on the principle that the performance  $J$  is a random variable and the joint distribution of all  $J_i$  for the  $n$  training data instances is a multivariate Gaussian distribution. This implies that the GP can be characterized entirely by a mean function  $\mu : \mathcal{R} \rightarrow \mathbb{R}$  and a covariance function  $K : \mathcal{R} \times \mathcal{R} \rightarrow \mathbb{R}$ ; that is,  $J \sim \mathcal{N}(\mu(r), K(r, r))$ . A common choice is to set  $\mu \equiv 0$ , and to represent the covariance function using kernels  $K$  such as squared exponential kernels or Matérn kernels, which are

parameterized by hyperparameters  $\gamma$  such as length-scales and variances. At inference, for predicting the performance of a new instance  $\hat{r}$ , the assumption that the predictions are jointly Gaussian with respect to the training instances  $J = [J_0 \dots J_i \dots J_n]^T$  yields an estimate  $\hat{J}$  with posterior mean and variance given by

$$\begin{aligned} \mathbb{E}[\hat{J}] &= \hat{K}^T K^{-1} J, \\ \text{Var}[\hat{J}] &= \bar{K} - \hat{K}^T K^{-1} \hat{K} \end{aligned} \quad (4)$$

where  $K = K(J, J|\gamma) + \sigma_w^2 I$  with  $\sigma_w$  being an estimate of the variance of the noise corrupting the performance output measurement,  $\hat{K} = K(J, \hat{J}|\gamma)$ , and  $\bar{K} = K(\hat{J}, \hat{J}|\gamma)$ . Clearly, the accuracy of the predictions depends on the choice of kernel and the kernel hyperparameters  $\gamma$ . In classical BO,  $\gamma$  is computed by maximizing a log-marginal likelihood function

$$\mathcal{L}(\gamma) = -\frac{1}{2} \log |K_+| - \frac{1}{2} J^T K_+^{-1} J + C_0, \quad (5)$$

where  $C_0$  is a constant. Although this problem is non-convex in  $\gamma$ , one can employ stochastic gradient ascent to search for optimal hyperparameters.

The selection of the next best candidate for  $r$  in classical BO is performed via an acquisition function  $A : \mathcal{R} \rightarrow \mathbb{R}$ . The acquisition function uses the predictive distribution given by the GP to compute the utility of performing an evaluation of the objective at each  $r_i$ , given the data obtained thus far. The next  $r_i$  at which the objective has to be evaluated is obtained by solving the maximization problem  $r_{i+1} := \arg \max_{\mathcal{R}} A(r)$ . Commonly used acquisition function include expected improvement (EI) and upper confidence bound (UCB) [12]. With the new control parameter candidate  $r_{i+1}$ , we wait  $t_f$  seconds to obtain  $J_{i+1}$  and iteratively perform the surrogate model retraining and the acquisition function evaluation steps with the training set augmented each BO iteration with the new pair  $(r_{i+1}, J_{i+1})$ .

#### B. Deep Kernel Networks (DKNs)

While commonly used kernel functions have shown good performance in some applications, there are advantages to directly learning the kernel function from data. This is the idea proposed in deep kernel learning [21], where a base kernel  $K(\cdot, \cdot|\gamma)$  with hyperparameters  $\gamma$  is used to transform the control parameters to a latent space via a non-linear map  $\mathcal{F}_\omega(r)$  induced by a deep neural network, parameterized by weights  $\omega$ .

The deep kernel

$$\tilde{K}(r, r'|\gamma, \omega) := K(\mathcal{F}_\omega(r), \mathcal{F}_\omega(r')|\gamma) \quad (6)$$

is trained by maximizing a log marginal likelihood, similar to (5), to concurrently compute  $\gamma$  and  $\omega$ . A noteworthy benefit of deep kernels is that they can be defined on a latent space whose dimension may be significantly lower than the dimensionality of the input, allowing the learning to be applied on problems where the dimensionality of the control parameters  $n_r$  is large. A simple example of the DKN architecture is shown in Fig. 1.

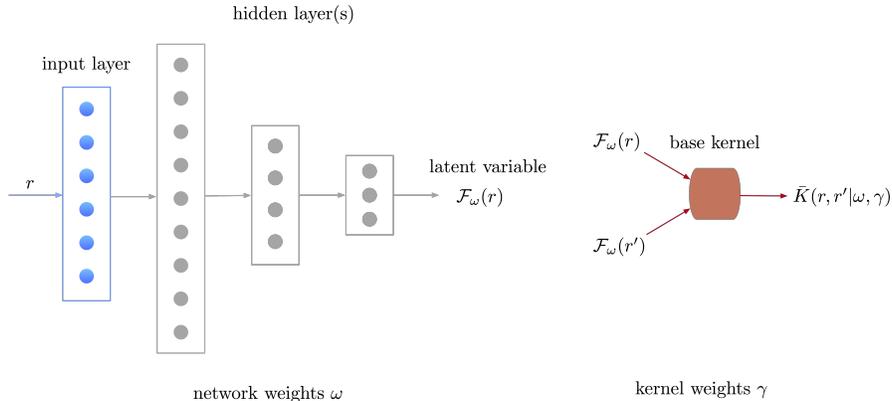


Fig. 1. Architecture of the deep kernel network used for Bayesian meta-learning. The latent variable representation is learned by the neural network from multiple source tasks, and fixed during inference. The base kernel is used to estimate the mean and variance of a performance function.

#### IV. META-LEARNED BAYESIAN OPTIMIZATION WITH DEEP KERNEL NETWORKS (DKN-BO)

Recall that  $\mathcal{D}^S$  is a collection of data obtained from source tasks, where the data of the  $k$ -th source task is described in (3). Most recent Bayesian meta-learning paradigms [16], [17] work under the assumption that all the source tasks (and the target task) share a set of task-independent hidden parameters  $\omega$  and  $\gamma$  (the reasons for using two symbols here will be explained shortly), and a set of task-specific parameters  $\{\rho_k\}_{k=1}^{N_S}$ . Given an input  $r_*$  from an unseen task, the Bayesian treatment for inference requires estimating the task-independent parameters  $\omega$  and  $\gamma$  from the source task dataset, and then estimating two distributions: the posterior of the task-specific parameters  $\pi(\rho_k | r_*, \mathcal{D}^S, \omega, \gamma)$  and the posterior for the output  $\pi(J_* | r_*, \omega, \gamma)$ . However, as argued in [15], handling these two distributions requires sampling or amortized distributions, resulting in significantly complex network architectures, and bi-level learning has been shown to require higher-order derivative information, in spite of which the training might suffer from instabilities. By using deep kernel networks, one can marginalize out the task-specific parameters with kernel functions defined in the latent space, which leads to a simple architecture (see Fig. 1), allows for quantification of uncertainty during inference, all the while avoiding the numerical conditioning issues that are reported to occur [14] during training via bi-level optimization (e.g., in MAML).

##### A. Training

The marginal likelihood of the Bayesian meta-learning module conditioned on  $\omega$  and  $\gamma$  is given by

$$\pi(J|r, \mathcal{D}^S, \omega) = \prod_{k=1}^{N_S} \pi(J|r, \mathcal{D}_k^S, \omega, \gamma), \quad (7)$$

where

$$\pi(J|r, \mathcal{D}_k^S, \omega, \gamma) \approx \int \prod_{k=1}^{N_S} \pi(J|r, \mathcal{D}_k^S, \omega, \gamma, \rho_t) d\rho_t$$

represents the marginalization over sets of task-specific parameters. We approximate the Bayesian integral above using a Gaussian process prior with kernel  $\gamma$ ; this is a commonly used approach for tackling similar integrals that are implicitly solved via Bayesian quadrature methods, see for example, [15], [22].

Consequently, one can cast the meta-learning problem as learning the parameters  $\omega$  and  $\gamma$  that characterize a deep kernel network (6), which is tantamount to optimizing (maximizing) the log-marginal likelihood

$$\mathcal{L}_{\text{DKN}}(\omega, \gamma) = \sum_{k=1}^{N_S} -\frac{1}{2} \log |\tilde{K}| - \frac{1}{2} \mathbf{J}_k^\top \tilde{K}^{-1} \mathbf{J}_k + C_1, \quad (8)$$

where  $\tilde{K} := \tilde{K}(r_k, r'_k | \gamma, \omega)$ ,  $\mathbf{J}_k$  is a vector of labels of the  $k$ -th source task, and  $C_1$  is a constant scalar.

In each training iteration, a source task is selected at random from the set of source tasks, and a random subset of batch data from this source task is selected. The log-marginal likelihood is computed using (7), and well-known variants of stochastic gradient descent methods are used to minimize the negative of  $\mathcal{L}_{\text{DKN}}$ . In other words, the learnable parameters are updated by backpropagation

$$\begin{aligned} \omega &\leftarrow \omega + \beta_\omega \frac{\partial}{\partial \omega} \mathcal{L}_{\text{DKN}}(\omega, \gamma), \\ \gamma &\leftarrow \gamma + \beta_\gamma \frac{\partial}{\partial \gamma} \mathcal{L}_{\text{DKN}}(\omega, \gamma), \end{aligned}$$

where  $\beta_\omega$  and  $\beta_\gamma$  are the step sizes of the neural network component and the base kernel component of the deep kernel network, respectively. By optimizing the entire DKN in mini-batches, we can also limit the computational expenditure of the determinant and inverse operations during loss function evaluation, without losing guarantees of convergence of the training loss (up to a small neighborhood of a critical point of the loss function), as demonstrated in [23, Theorem 3.1]. Let  $\omega^*$  and  $\gamma^*$  denote the parameters of the DKN learned at the termination of the training loop.

## B. Inference

As in the few-shot optimization setting, we assume that very few data points  $\mathcal{D}^T$  are initially available from a target task  $T$ . To predict the control performance  $J$  on  $\mathcal{R}$  for the target task, and subsequently compute a next candidate  $r$  with Bayesian optimization, we rely primarily on the optimized task-independent parameters  $\omega^*$  and  $\gamma^*$ , as the trained DKN is expected to have learned a suitable surrogate model of  $r \mapsto J$  from the set of source tasks. By conditioning on the target data, we encourage the predictions of the DKN

$$\pi(J_*|r_*, \mathcal{D}^T, \omega^*, \gamma^*)$$

to be biased towards to the target task. Under the assumption that the target task is similar to the source tasks, we expect the predictions of the DKN conditioned on limited target task data to be more accurate and exhibit lower uncertainty than a regressor trained only on the target data without meta-learning.

Analogous to the classical BO algorithm described in Section III-A, we use the DKN surrogate to predict the control performance at samples in  $\mathcal{R}$ , and use these predictions to optimize an acquisition function. Optimizing the acquisition function yields the next best control input candidate, given the target task data collected so far. In subsequent DKN-BO iterations, even if  $\omega^*$  and  $\gamma^*$  are kept constant, since the kernel covariance matrix  $K$  is appended with new target task data (obtained by evaluating the performance  $J$  of the target task with candidate control inputs  $r$ ), the DKN predictions gradually tighten around the true performance function of the target task. A full pseudocode of DKN-BO is provided in Algorithm 1.

The authors in [24] propose a method for introducing scale invariance during training of DKNs so that the target task labels do not need to be normalized to the unit line (which would be impossible, since the target task’s maximum and minimum is unknown). In particular, the proposed method involves computing  $J_{\min}$  and  $J_{\max}$ , which are the minimum and maximum performance output observed over all source tasks, respectively. Each time a batch is sampled during training, the labels are rescaled to  $(J - J_{\min}) / (J_{\max} - J_{\min})$ , which, over a large number of training iterations allows the learner to learn trends from the source tasks that are scale independent, enabling inference for the unseen target task without normalization. Empirically, we have noticed improvements in DKN-BO performance when retraining the base kernel parameters  $\gamma$  as in classical BO while keeping the neural network weights, i.e.  $\omega$ , fixed. However, this approach is closer in spirit to transfer learning than meta-learning, and thus, we do not retrain the base kernel in this work to avoid online training expenditure.

---

### Algorithm 1 Meta-Learned Bayesian Optimization with Deep Kernel Networks (DKN-BO)

---

**Require:**  $\mathcal{R} \leftarrow$  set of control inputs  
**Require:**  $\mathcal{D}^S \leftarrow$  source task dataset  
**Require:**  $\mathcal{D}^T \leftarrow$  target task dataset  
**Require:**  $\beta_\omega, \beta_\gamma \leftarrow$  step sizes  $\triangleright$  default:  $10^{-4}, 10^{-3}$   
**Require:** latent vector dimension  $\triangleright$  default: 8  
**Require:** mini-batch size for training  $\triangleright$  default: 32  
**Require:** number of training iterations  $\triangleright$  default: 10000  
**Require:**  $\omega, \gamma \leftarrow$  random initial DKN parameters  
**Require:**  $\mathcal{F}_\omega \leftarrow$  latent space encoder  
**Require:**  $K(\cdot, \cdot | \gamma) \leftarrow$  base kernel

---

#### Training the DKN

---

```

1: for each training iteration do
2:    $k \leftarrow$  choose random source task index
3:   for each mini-batch do
4:     sample random subsets of data from  $\mathcal{D}_k^S$ 
5:     compute DKN loss using (8)
6:     update  $\omega$  and  $\gamma$  using SGD/Adam
7:   end for
8: end for
9:  $\omega^*, \gamma^* \leftarrow$  trained DKN weights

```

---

#### Few-Shot BO

---

```

10: initialize surrogate model with trained DKN
11: for each DKN-BO iteration do
12:    $\mu, \sigma \leftarrow$  predict performance function using DKN
   conditioned on  $\mathcal{D}^T$ 
13:   evaluate acquisition function with predictions on  $\mathcal{R}$ 
14:    $r_+ \leftarrow$  next best BO candidate from acq. func.
15:    $J_+ \leftarrow$  evaluate closed-loop performance with  $r_+$ 
16:   Append target task dataset with  $(r_+, J_+)$ 
17: end for
18: return  $r^* \leftarrow$  control parameter for best performance
   observed

```

---

## V. SIMULATION RESULTS

We consider the following nonlinear system, previously studied in [25], [26]:

$$\dot{x}_1 = x_2 - \theta_1 x_1, \quad (9a)$$

$$\dot{x}_2 = -\theta_2 x_1^2 + \kappa(x, r), \quad (9b)$$

$$J = 1 - \theta_1 x_1 - \theta_2 x_1^2, \quad (9c)$$

where

$$\kappa(x, r) = -6x_1 + (\theta_1 - 5)(x_2 - \theta_1 x_1) + \theta_2 x_1^2 + r,$$

is a control policy that renders the closed-loop system stable. For the target system, the true parameters are  $\theta = [2, 5]$  and the optimal control parameter is  $r^* = -1.2$  at which the system attains its maximum  $J^* = 1.2$ ; this is not known to the designer. We choose  $\mathcal{R} = [-10, 10]$  and  $\Theta = [1, 6]^2$ .

In order to generate the source dataset, we extract  $N_S = 20$  unique system parameters from  $\Theta$  using a 2D uniform random distribution. We evaluate the performance output  $J$

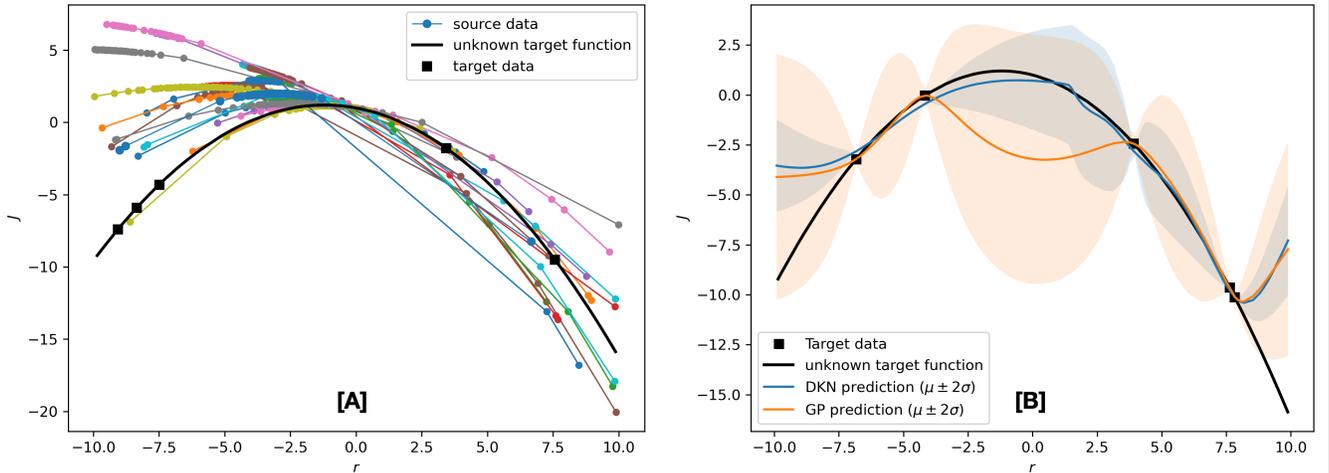


Fig. 2. Data points (without noise) collected from source tasks shown with circles, limited data from target task shown with squares, and unknown target performance function shown with a continuous black line. Note that no target data point is at the optimal value.

for a given  $r$  by forward simulating each system realized by sampling from  $\Theta$  for  $t_f = 10$  s, which provides enough time for the transient response to fade. For each sampled system parameter, we solve the source optimization task of maximizing the performance function given in (9c) within  $T_k \equiv 50$  iterations, which depends on  $\theta$  and therefore has a variety of optimizers and optimal values. The source data collected is illustrated in Fig. 2[A] using colored circles joined by thin continuous lines, and we observe that indeed the performance functions are not identical to the target task performance function which is shown with the continuous thick black line. Our proposed approach is agnostic to the optimization strategy used for optimizing each source task; in fact, a mix of random search and BO are used for source task optimization and source data collection. The target task data is collected by evaluating the performance of the target system at  $T = 5$  randomly selected  $r$  values.

Our DKN architecture involves a latent encoder with 100 neurons per hidden layer, with 4 hidden layers following the input layer. We select a latent dimension of 10, and the activation functions are ReLUs. The latent output is passed to a GP with a scaled Matern-3/2 kernel as the covariance and a constant mean kernel. The entire pipeline is implemented using PyTorch [27] and GPyTorch [28]. The DKN is trained for 10000 iterations with the learning rate scheduled at:  $\beta_\omega = 10^{-3}$  and  $\beta_\gamma = 10^{-2}$  for the first 2000 iterations, and  $\beta_\omega = 10^{-4}$  and  $\beta_\gamma = 10^{-3}$  thereafter. We train using the Adam algorithm and fix the batch-size at 8. The weights corresponding to the the highest log marginal likelihood was saved as  $\omega^*$  and  $\gamma^*$ .

Fig. 2[B] illustrates the effectiveness of meta-learning for predicting the target performance function. With the same limited target data, the trained DKN generates a more accurate mean estimate of the true performance function, as compared to a GP trained for 1000 iterations with a squared-exponential kernel: see the blue (DKN) and orange (GP) continuous lines. Furthermore, it is encouraging that the

uncertainty bands for the DKN are less pronounced than the GP in the areas that are most relevant for optimization.

Fig. 3 demonstrates an instance of DKN-BO for optimizing control performance. We start the target system at the origin and simulate the system forward in time, collecting performance outputs every 10 s for  $r$  value candidates generated by the DKN-BO procedure. The state variation  $x_1$  and  $x_2$  shows that we do not assume that we can reset the experiment each time a new  $r$  candidate is selected, and the closed-loop performance is optimized in an online manner. The corresponding variation of the control parameter  $r$ . The bottom right subplot, which shows the performance output, illustrates clearly that DKN-BO finds the optimal value within 100 s, which is comparable to the performance reported in [26], even though we do not have access to any model information and do not estimate  $\theta$ .

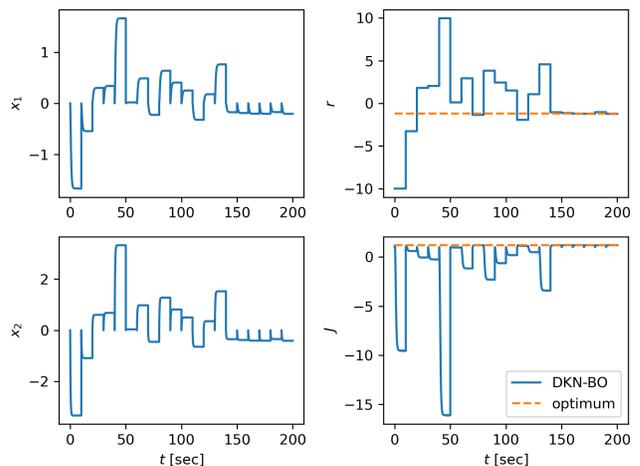


Fig. 3. States  $x$ , controller parameter  $r$ , and performance output  $J$  of the unknown target system during performance optimization by DKN-BO.

We also repeated our experiments 100 times with varying

target datasets, and report the statistics of the simple regret decay in Fig. 4. We observe that DKN-BO performs well in a few-shot setting, and converges to a median regret that is two orders of magnitude below classical GP-BO within 10 iterations. Furthermore, the DKN-BO algorithm is robust to the target dataset as is evident from its tighter confidence intervals compared with GP-BO.

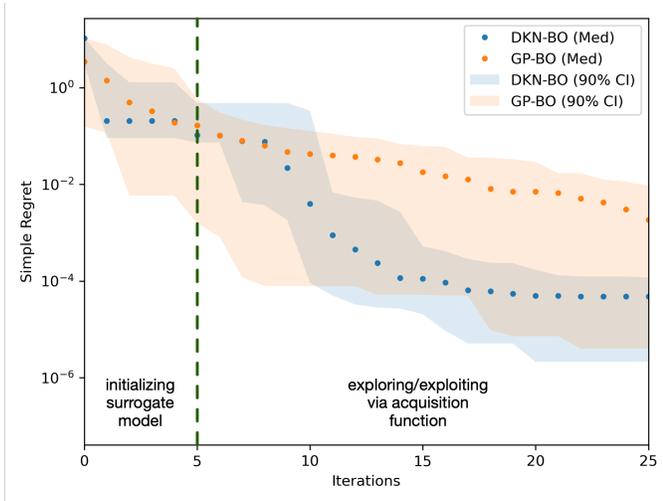


Fig. 4. Regret decay of target task’s control performance cost for meta-learned DKN-BO (without base kernel retraining) and classical BO (median and 90% confidence interval). The benefit of meta-learning is clearly seen through the fast convergence of DKN-BO after 10 initial random samples are provided.

## VI. CONCLUSIONS

This paper provides a highly generalizable, computationally simple, and systematic framework for leveraging prior data for optimizing unseen closed-loop performance optimization tasks, even if the prior data is not from the task to be optimized. This is done in a few-shot manner using meta-learned Bayesian optimization and deep kernel networks as surrogate models. The latent encoding of the DKN enables a task-independent representation of the controlled inputs to the closed-loop system, and the base kernel allows conditioning for task-specific predictions. In future work, we will compare this approach with transfer learning, where the architecture is retrained using target task data acquired during the DKN-BO procedure.

## REFERENCES

- [1] M. Neumann-Brosig, A. Marco, D. Schwarzmann, and S. Trimpe, “Data-efficient autotuning with bayesian optimization: An industrial control study,” *IEEE Transactions on Control Systems Technology*, vol. 28, no. 3, pp. 730–740, 2019.
- [2] R. R. Duivendoorn, F. Berkenkamp, N. Carion, A. Krause, and A. P. Schoellig, “Constrained Bayesian optimization with particle swarms for safe adaptive controller tuning,” *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 11 800–11 807, 2017.
- [3] C. König, M. Khosravi, M. Maier, R. S. Smith, A. Rupenyan, and J. Lygeros, “Safety-aware cascade controller tuning using constrained Bayesian optimization,” *arXiv preprint arXiv:2010.15211*, 2020.
- [4] S. Bansal, R. Calandra, T. Xiao, S. Levine, and C. J. Tomlin, “Goal-driven dynamics learning via bayesian optimization,” in *2017 IEEE 56th Annu. Conf. on Decis. and Control (CDC)*. IEEE, 2017, pp. 5168–5173.

- [5] J. A. Paulson and A. Mesbah, “Data-driven scenario optimization for automated controller tuning with probabilistic performance guarantees,” *IEEE Contr. Syst. Lett.*, vol. 5, no. 4, pp. 1477–1482, 2020.
- [6] A. Chakrabarty, C. Danielson, S. Bortoff, and C. Laughman, “Accelerating self-optimization control of refrigerant cycles with bayesian optimization and adaptive moment estimation,” *Applied Thermal Engineering*, 2021.
- [7] W. Xu, C. N. Jones, B. Svetozarevic, C. R. Laughman, and A. Chakrabarty, “VABO: Violation-aware bayesian optimization for closed-loop control performance optimization with unmodeled constraints,” in *Proc. American Control Conference*, 2022, pp. 1–6.
- [8] M. Khosravi, A. Eichler, N. Schmid, R. S. Smith, and P. Heer, “Controller tuning by Bayesian optimization an application to a heat pump,” in *2019 18th Eur. Control Conf. (ECC)*. IEEE, 2019, pp. 1467–1472.
- [9] A. Pal, L. Zhu, Y. Wang, and G. G. Zhu, “Multi-objective stochastic Bayesian optimization for iterative engine calibration,” in *Proc. of the Amer. Control Conf.* IEEE, 2020, pp. 4893–4898.
- [10] F. Berkenkamp, A. Krause, and A. P. Schoellig, “Bayesian optimization with safety constraints: safe and automatic parameter tuning in robotics,” *Machine Learning*, pp. 1–35, 2021.
- [11] R. Lam, M. Poloczek, P. Frazier, and K. E. Willcox, “Advances in bayesian optimization with applications in aerospace engineering,” in *2018 AIAA Non-Deterministic Approaches Conference*, 2018, p. 1656.
- [12] J. Snoek, H. Larochelle, and R. P. Adams, “Practical Bayesian optimization of machine learning algorithms,” in *Proc. NeurIPS*, 2012, p. 2951–2959.
- [13] C. Finn, P. Abbeel, and S. Levine, “Model-agnostic meta-learning for fast adaptation of deep networks,” in *Proceedings of the 34th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, D. Precup and Y. W. Teh, Eds., vol. 70. PMLR, 06–11 Aug 2017, pp. 1126–1135.
- [14] A. Antoniou, H. Edwards, and A. Storkey, “How to train your MAML,” *arXiv preprint arXiv:1810.09502*, 2018.
- [15] M. Patacchiola, J. Turner, E. J. Crowley, M. F. O’Boyle, and A. J. Storkey, “Bayesian meta-learning for the few-shot setting via deep kernels,” in *Proc. NeurIPS*, 2020, pp. 16 108–16 118.
- [16] J. Yoon, T. Kim, O. Dia, S. Kim, Y. Bengio, and S. Ahn, “Bayesian model-agnostic meta-learning,” in *Proc. NeurIPS*, 2018, pp. 7343–7353.
- [17] C. Finn, K. Xu, and S. Levine, “Probabilistic model-agnostic meta-learning,” in *Proc. NeurIPS*, 2018, pp. 9537–9548.
- [18] M. O’Connell, G. Shi, X. Shi, and S.-J. Chung, “Meta-learning-based robust adaptive flight control under uncertain wind conditions,” *arXiv preprint arXiv:2103.01932*, 2021.
- [19] S. M. Richards, N. Azizan, J.-J. Slotine, and M. Pavone, “Adaptive-control-oriented meta-learning for nonlinear systems,” *arXiv preprint arXiv:2103.04490*, 2021.
- [20] C. K. Williams and C. E. Rasmussen, *Gaussian Processes For Machine Learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.
- [21] A. G. Wilson, Z. Hu, R. Salakhutdinov, and E. P. Xing, “Deep kernel learning,” in *Artificial Intelligence and Statistics (AISTATS)*. PMLR, 2016, pp. 370–378.
- [22] J. Courts, J. Hendriks, A. Wills, T. B. Schön, and B. Ninness, “Variational state and parameter estimation,” *IFAC-PapersOnLine*, vol. 54, no. 7, pp. 732–737, 2021.
- [23] H. Chen, L. Zheng, R. Al Kontar, and G. Raskutti, “Stochastic gradient descent in correlated settings: A study on gaussian processes,” *NeurIPS*, vol. 33, pp. 2722–2733, 2020.
- [24] M. Wistuba and J. Grabocka, “Few-shot Bayesian optimization with deep kernel surrogates,” in *Proc. ICLR*, 2021.
- [25] M. Guay and T. Zhang, “Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties,” *Automatica*, vol. 39, no. 7, pp. 1283–1293, 2003.
- [26] D. Nesic, A. Mohammadi, and C. Manzie, “A framework for extremum seeking control of systems with parameter uncertainties,” *IEEE Trans. Automatic Control*, vol. 58, no. 2, pp. 435–448, 2012.
- [27] A. Paszke, S. Gross *et al.*, “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” pp. 8024–8035, 2019.
- [28] J. R. Gardner, G. Pleiss, D. Bindel, K. Q. Weinberger, and A. G. Wilson, “GPpyTorch: Blackbox Matrix-Matrix Gaussian Process Inference with GPU Acceleration,” in *Advances in Neural Information Processing Systems*, 2018.