Active Learning in Gaussian Process State Space Model

Hon Sum Alec Yu[1,2], Dingling Yao[1],
Christoph Zimmer[1], Marc Toussaint[2], and Duy Nguyen-Tuong[1]

[1] Bosch Center for Artificial Intelligence, 71272 Renningen, Germany
honsumalec.yu@de.bosch.com
[2] Learning & Intelligent System Lab, Technische Universität Berlin, Berlin, Germany

Abstract. We investigate active learning in Gaussian Process state-space models (GPSSM). Our problem is to actively steer the system through latent states by determining its inputs such that the underlying dynamics can be optimally learned by a GPSSM. In order that the most informative inputs are selected, we employ mutual information as our active learning criterion. In particular, we present two approaches for the approximation of mutual information for the GPSSM given latent states. The proposed approaches are evaluated in several physical systems where we actively learn the underlying non-linear dynamics represented by the state-space model.

Keywords: Active Learning · Gaussian Process State Space Model · Mutual Information.

1 Introduction

State-space models (SSMs) are a compact representation of dynamical systems as a set of input, output and state variables where the transition could be deterministic or stochastic. SSMs are widely used in practice, where applications range from aerospace-related work (e.g. [16]), medicine (e.g. [1]) to economics and finance (e.g. [52]). The two important aspects of SSMs are *learning* and *control* and they have been a fruitful research field in classical engineering (e.g. [17,29]). Learning in SSM is about finding optimal hyperparameters so that the dynamical system is accurately captured whereas control refers to optimising a specific objective function based on inputs and states of a given SSM. The latter is often referred as *optimal control*. During the last decade, both aspects of SSMs also attracted the attention from the machine learning community (e.g. [18,41]). To capture the uncertainty of system dynamics, Bayesian models are often employed in the context of SSM. For example, a Gaussian Process (GP) prior [42] is mostly used to describe the transition between states, which leads to family of Gaussian Process State-Space Models (GPSSMs) with an impressive amount of progress has been made (e.g. [12, 13, 19, 20, 27, 28, 31, 38, 49], and more). Learning and inference of GPSSM are an active research topic in machine learning community and this paper focuses on *learning* of GPSSM.

While different approaches of GPSSM mainly address the modelling issues, the question about data-efficient learning of GPSSM remains largely unanswered. Such questions go naturally to the field of Active Learning (AL), which is the process of strategically selecting and generating new data for supervised learning. The purpose of AL is to learn the proposed model with the least amount of data possible. Though the majority of recent papers in this field focused on classification problems (e.g. [3,46]) and rather few were for regression (e.g. [7,44]), Schreiter et al. [43] applied AL in GP regression and Zimmer et al. [53] extended the scope to learning time-series models. Both papers employed the entropy criterion for exploration, while using a greedy selection based on the maximum predictive variance of the GP. In context of state-space models for predictive control, Buisson et al. [6] proposed an AL approach for learning the transition dynamics via optimising the trajectory. For AL in GPSSM, Capone et al. [8] was the first to introduce the concept here. However, in both cases, Buisson et al. [6] and Capone et al. [8] simplified the settings by assuming the states are observable and measurable. By doing so, the resulting state-space models can be actively learned in the same way as standard GP regression models in a supervised manner.

In contrast to previous work, we present an AL strategy on GPSSM without relying on the assumption that the states are observed. That is, we actively steer the system through the latent state-space by determining its inputs such that the underlying dynamics can be optimally learned by a GPSSM. For the information criterion, we employ an approximated measure of the mutual information. We propose and discuss two different approaches for the approximation of mutual information in the context of unmeasurable, latent state-space.

This paper is on *learning* a stochastic nonlinear dynamical system actively from noisy data via probabilistic SSM. Its main contributions include:

- We derive tractable approximate mutual information estimates for GPSSM with latent states based on the state-of-the-art learning scheme and the technique of approximate Gaussian integral.
- We propose an AL strategy for GPSSM with latent states based on inputs.
- We conduct experiments for different examples to demonstrate the usefulness of our strategy in physical systems, which is a typical application for SSM.

The reminder of this paper is organised as follow: We give a brief overview of backgrounds in section 2 and introduce GPSSM in section 3. Section 4 shows our active learning strategies and section 5 evaluates our concepts on several test scenarios. A conclusion is given in section 6. For brevity, we leave detailed proof of propositions, experimental details, partial code and few further remarks in supplementary materials, which will be uploaded in due course.

2 Background

AL is itself a broad topic and here we refer readers to, for example, Settles [45] and Dasgupta [10] for an in-depth survey of the basic algorithmic and theoretical

ideas. There are many different paradigms in AL. For example, *Bayesian Active Learning* incorporates the Bayesian framework and is often referred to AL with GP models (e.g. [26,50]), because GPs naturally carry the uncertainty measures. Specifically, Houlsby et al. [26] applied AL in GP classification problems, where they defined an acquisition function which estimates the quantity of mutual information between model predictions and model parameters. Other variants include *Batch Active Learning* and *Deep Bayesian Active Learning*. While Batch AL refers to making multiple queries in parallel (e.g. [26, 30]), Deep Bayesian AL explores techniques for actively learning deep Bayesian models, e.g. Bayesian convolutional neural network [22]. For AL with regression models, Seo et al. [44] proposed a test point rejection strategy based on the posterior variance. Krause et al. [32] proposed an optimal exploration strategy for GP regression models, while providing bounds on the advantage of using AL. Srinivas et al. [47] further showed convergence rate for AL in GPs under specific kernels. Schreiter et al. [43] extended AL in GP regression by introducing constraints and Zimmer et al. [53] extended constrained AL in GP regression to time-series. They proposed using the determinant of the predictive covariance matrix defined in the GP as optimality criterion along with theoretical guarantees.

In dynamics modelling, SSM has been a long-lasting topic with enormous literature across different disciplines (e.g. [17, 29, 52]). In machine learning, the combination between SSM and GP began arguably from Wang et al. [51], where they learned the latent state of GPSSM via maximum a posteriori. Turner et al. [49] extended it to learning the transition function. Frigola et al. [20] presented a Bayesian treatment based on particle Markov-Chain-Monte-Carlo and Frigola et al. [19] further introduced a variational inference scheme to overcome the computational complexity of Monte-Carlo based methods. This becomes the state-of-the-art and due to the fact that this approach depends heavily on the approximated function, most recent work introduced their own variational inference scheme with their own advantages and disadvantages (e.g. [12, 27, 28, 38]). For example, Doerr et al. [12] and Ialongo et al. [28] introduced their specific variational inference scheme to account for the dependence between transition function and latent states, which were treated as independent in earlier work.

To the best of our knowledge, AL and GPSSM have not been discussed together until Buisson et al. [6] first brought the topic of AL in learning the transition function using GP. They stated that the dynamical problem, i.e. learning the transition function, is fundamentally different from a static AL in GP problem, because we need to steer the system to certain states through the unknown transition function with a sequence of actions, which is the only component we can control. For each round of exploration, the next input is actively picked so that the model error can be maximally reduced. They proposed their AL strategy using the maximum entropy. Another recent work on AL in GPSSM is by Capone et al. [8]. They proposed taking the most informative data point while employing the mutual information between the most informative point and reference points. The states, in the existing work on AL in GPSSM, are hitherto assumed to be observable such that the model can be trained in a fashion similar

to GP regression. Namely, in Capone et al. [8], the fact that states are observable simplify the estimation of the mutual information, resulting in two optimisation problems to solve in parallel.

Our paper differs from the previous work by presenting an AL strategy on GPSSM with states that are *latent*, resulting in a different estimation of the mutual information as exploration criterion while extending the scope of AL for GPSSM. We leave further quantitative discussion in section 4.

3 Gaussian Process State-Space Model with Inputs

We consider a discrete-time sequence of T observations $\boldsymbol{y}_{1:T} \equiv \{\boldsymbol{y}_t\}_{t=1}^T$, where each observed point $\boldsymbol{y}_t \in \mathcal{Y} \subseteq \mathbb{R}^{d_y}$ is generated by a corresponding latent variable $\boldsymbol{x}_t \in \mathcal{X} \subseteq \mathbb{R}^{d_x}$ and the previous step control $\boldsymbol{c}_{t-1} \in \mathcal{C} \subseteq \mathbb{R}^{d_c}$. The collection of latent states and controls are denoted by $\boldsymbol{x}_{0:T} \equiv \{\boldsymbol{x}_t\}_{t=0}^T$ and $\boldsymbol{c}_{0:T} \equiv \{\boldsymbol{c}_t\}_{t=0}^T$, respectively. These latent variables and controls are assumed to satisfy the Markov property, meaning that any $\boldsymbol{x}_{t+1}$ can be generated by only conditioning on $\boldsymbol{x}_t$, $\boldsymbol{c}_t$ and the transition function $f : \mathbb{R}^{d_x} \times \mathbb{R}^{d_c} \to \mathbb{R}^{d_x}$. To align with previous work, we use Gaussian distribution for both transition and observation density function. We also simplify the mean of the observation function by using linear mapping and tackle multivariate latent states by placing a GP prior on each dimension separately, as in, for instance, Ialongo et al. [28].

GPSSM is defined to be a probabilistic SSM with a GP prior over the transition function f, specified by

$$f \sim \mathcal{GP}(m(\cdot), k(\cdot,\cdot)), \tag{1}$$

$$\boldsymbol{x}_0 \sim p(\boldsymbol{x}_0), \tag{2}$$

$$\boldsymbol{x}_t | f(\boldsymbol{x}_{t-1}, \boldsymbol{c}_{t-1}) \sim \mathcal{N}(\boldsymbol{x}_t | f(\boldsymbol{x}_{t-1}, \boldsymbol{c}_{t-1}), \boldsymbol{Q}), \tag{3}$$

$$\boldsymbol{y}_t | \boldsymbol{x}_t \sim \mathcal{N}(\boldsymbol{y}_t | \boldsymbol{C}\boldsymbol{x}_t + \boldsymbol{d}, \boldsymbol{R}), \tag{4}$$

where in equation 1, f is a GP governed by a given mean function $m(\cdot)$ and positive definite covariance function $k(\cdot,\cdot)$. The initial state $p(\boldsymbol{x}_0) = \mathcal{N}(\boldsymbol{x}_0 | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ in equation 2 is assumed known. $\boldsymbol{C}$ and $\boldsymbol{d}$ are parameters in the linear mapping, whereas $\boldsymbol{Q}$ and $\boldsymbol{R}$ are covariance matrices which capture process and observation noise, respectively. The advantages of linear mapping are that linear mean wont limit the range of systems that can be modelled and this reduces the non-identifiabilities between transitions and emissions. As a convention from previous work, we denote $\boldsymbol{f}_t \equiv f(\boldsymbol{x}_{t-1}, \boldsymbol{c}_{t-1})$ and equation 3 becomes

$$\boldsymbol{x}_t | \boldsymbol{f}_t \sim \mathcal{N}(\boldsymbol{x}_t | \boldsymbol{f}_t, \boldsymbol{Q}). \tag{5}$$

For brevity, we write $\tilde{\boldsymbol{x}}_* = (\boldsymbol{x}_*, \boldsymbol{c}_*)$, where $*$ could be a specific index or a collection of indices. Thus, the matrix of covariance functions are denoted by $K_{i:j} := (k(\tilde{\boldsymbol{x}}_s, \tilde{\boldsymbol{x}}_t))_{s,t=i}^j$. For convenience, we write $k(\tilde{\boldsymbol{x}}_{i:j}, \tilde{\boldsymbol{x}}_k) \equiv (k(\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{x}}_k), \cdots, k(\tilde{\boldsymbol{x}}_j, \tilde{\boldsymbol{x}}_k))$ as a vector collection of kernel evaluations. It holds that $k(\tilde{\boldsymbol{x}}_{i:j}, \tilde{\boldsymbol{x}}_k)^T = k(\tilde{\boldsymbol{x}}_k, \tilde{\boldsymbol{x}}_{i:j})$.

The control inputs $\boldsymbol{c}_t$ can be viewed as an augmented latent state. Since $\boldsymbol{c}_t$ is not a random variable, we always condition on this quantity. Most previous

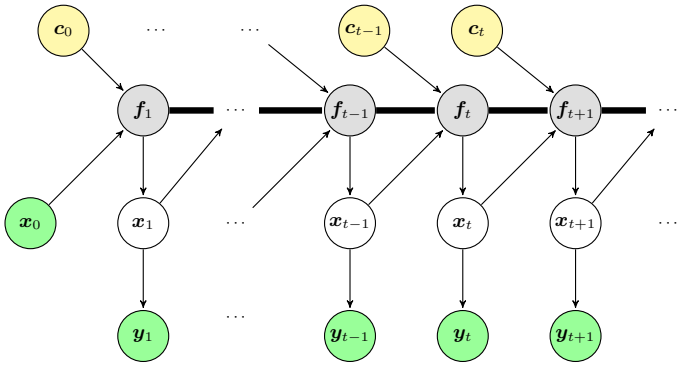

Fig. 1. Graphical model of GPSSM with control inputs as defined in equation 1 - 4. The observations $\boldsymbol{y}_t$ and initial state $\boldsymbol{x}_0$ are assumed known (filled in green). Control inputs $\boldsymbol{c}_t$ are independent variables (filled in yellow) whereas latent variables are denoted by $\boldsymbol{x}_t$. We use a thick, straight line to connect $\boldsymbol{f}_t \equiv f(\boldsymbol{x}_{t-1}, \boldsymbol{c}_{t-1})$ (filled in gray) to show that all variables are fully connected.

papers related to this model were either studied without controls or, if a control is involved, by assuming that this quantity was randomly distributed [36]. We, however, construct an AL strategy based on $\boldsymbol{c}_t$ to learn the GPSSM efficiently. Hence, we purposely keep this term in our definitions. Figure 1 shows a graphical model of such a setting. We would like to readdress here that $\boldsymbol{c}_t$ is a known quantity where our task is to optimise the value to select. This should not be confused with control in other contexts such as that in RL.

To complete the structure of GPSSM, let us define $\boldsymbol{f}_{1:T} \equiv \{\boldsymbol{f}_t\}_{t=1}^T$ and we can write the joint density as follows (e.g. chapter 3 in Frigola [18]),

$$p(\boldsymbol{y}_{1:T}, \boldsymbol{x}_{0:T}, \boldsymbol{f}_{1:T}) = p(\boldsymbol{x}_0) \prod_{t=1}^{T} p(\boldsymbol{y}_t|\boldsymbol{x}_t) p(\boldsymbol{x}_t|\boldsymbol{f}_t) p(\boldsymbol{f}_t|\boldsymbol{x}_{0:t-1}, \boldsymbol{f}_{1:t-1}). \tag{6}$$

In particular, from equation 1, the last factor in equation 6 is derived in the same manner as the posterior in GP regression, given by

$$p(\boldsymbol{f}_t|\boldsymbol{x}_{0:t-1}, \boldsymbol{f}_{1:t-1}) = \mathcal{N}(\boldsymbol{f}_t|\mathcal{M}_{t-1}, \mathcal{K}_{t-1}), \tag{7}$$

where

$$\mathcal{M}_{t-1} = m(\tilde{\boldsymbol{x}}_{t-1}) + k(\tilde{\boldsymbol{x}}_{t-1}, \tilde{\boldsymbol{x}}_{0:t-2}) K_{0:t-2}^{-1} (\boldsymbol{f}_{1:t-1} - m(\tilde{\boldsymbol{x}}_{0:t-2}))^T,$$
$$\mathcal{K}_{t-1} = k(\tilde{\boldsymbol{x}}_{t-1}, \tilde{\boldsymbol{x}}_{t-1}) - k(\tilde{\boldsymbol{x}}_{t-1}, \tilde{\boldsymbol{x}}_{0:t-2}) K_{0:t-2}^{-1} k(\tilde{\boldsymbol{x}}_{0:t-2}, \tilde{\boldsymbol{x}}_{t-1}),$$

and the term $\boldsymbol{f}_{1:t-1} - m(\tilde{\boldsymbol{x}}_{0:t-2})$ is written as

$$\boldsymbol{f}_{1:t-1} - m(\tilde{\boldsymbol{x}}_{0:t-2}) \equiv (\boldsymbol{f}_1 - m(\tilde{\boldsymbol{x}}_0), \cdots, \boldsymbol{f}_{t-1} - m(\tilde{\boldsymbol{x}}_{t-2})).$$

3.1 Learning and Prediction in GPSSM

In learning the posterior, our target is to compute

$$p(\boldsymbol{x}_{0:t}, \boldsymbol{f}_{1:t}|\boldsymbol{y}_{1:t}) = p(\boldsymbol{x}_{0:t}, \boldsymbol{f}_{1:t}, \boldsymbol{y}_{1:t})/p(\boldsymbol{y}_{1:t}).$$

The major challenge is the computation of $p(\boldsymbol{y}_{1:t})$. Throughout this paper, we are going to apply *variational inference* to train GPSSM, which is a technique based on making assumptions about the posterior over latent variables that leads to a tractable lower bound, often called *evidence lower bound* (ELBO). There are two reasons for choosing this approach, which is currently state-of-the-art in training GPSSM. First, variational inference is computationally efficient. Second, this approach will lead to an approximation of the posterior, which can be carried to the AL strategy, as we will present further in section 4.

The first step of using variational inference in GPSSM is to introduce M ($M \ll T$) inducing points $\boldsymbol{u} = \boldsymbol{u}_{1:M} = \{\boldsymbol{u}_i\}_{i=1}^{M}$ with their corresponding inputs. This is referred as sparse GP technique (see e.g. [21,48]) and the joint density is

$$p(\boldsymbol{y}_{1:T}, \boldsymbol{x}_{0:T}, \boldsymbol{f}_{1:T}, \boldsymbol{u}) = p(\boldsymbol{x}_0)p(\boldsymbol{u}) \prod_{t=1}^{T} p(\boldsymbol{y}_t|\boldsymbol{x}_t)p(\boldsymbol{x}_t|\boldsymbol{f}_t)p(\boldsymbol{f}_t|\boldsymbol{x}_{0:t-1}, \boldsymbol{f}_{1:t-1}, \boldsymbol{u}). \tag{8}$$

Then, the ELBO $\mathcal{L}_t$ to the log marginal likelihood of $p(\boldsymbol{y}_{1:t})$ in equation 6, which is based on KL divergence (see e.g. chapter 8.5 in Cover [9]), is given by

$$\log(p(\boldsymbol{y}_{1:t})) = \mathcal{L}_t + \mathrm{KL}\left[q(\boldsymbol{x}_{0:t}, \boldsymbol{f}_{1:t}) \| p(\boldsymbol{x}_{0:t}, \boldsymbol{f}_{1:t}|\boldsymbol{y}_{1:t})\right]$$

for $t = 1, \cdots, T$, where $q(\cdot)$ is the distribution function to approximate $p(\cdot)$. With further derivations from variational inference methodology, $\mathcal{L}_t$ is given by

$$\mathcal{L}_t = \int q(\boldsymbol{x}_{0:t}, \boldsymbol{f}_{1:t}, \boldsymbol{u}) \log\left(\frac{p(\boldsymbol{y}_{1:t}, \boldsymbol{x}_{0:t}, \boldsymbol{f}_{1:t}, \boldsymbol{u})}{q(\boldsymbol{x}_{0:t}, \boldsymbol{f}_{1:t}, \boldsymbol{u})}\right) d\boldsymbol{x}_{0:t} d\boldsymbol{f}_{1:t} d\boldsymbol{u}. \tag{9}$$

Analogously to Ialongo et al. [28], we set

$$q(\boldsymbol{x}_{0:t}, \boldsymbol{f}_{1:t}, \boldsymbol{u}) = q(\boldsymbol{u})q(\boldsymbol{x}_0) \prod_{i=1}^{t} q(\boldsymbol{x}_i|\boldsymbol{f}_i)p(\boldsymbol{f}_i|\boldsymbol{f}_{1:i-1}, \boldsymbol{x}_{0:i-1}, \boldsymbol{u}) \tag{10}$$

which leads to equation 17 of [28]

$$\begin{aligned}\mathcal{L}_T = &\int \sum_{t=1}^{T} q(\boldsymbol{x}_{0:T}) \log\left(p(\boldsymbol{y}_t|\boldsymbol{x}_t)\right) d\boldsymbol{x}_{0:T} - \sum_{t=1}^{T} \int q(\boldsymbol{f}_t)\, \mathrm{KL}\left[q(\boldsymbol{x}_t|\boldsymbol{f}_t) \| p(\boldsymbol{x}_t|\boldsymbol{f}_t)\right] d\boldsymbol{f}_t \\ &- \mathrm{KL}\left[q(\boldsymbol{x}_0) \| p(\boldsymbol{x}_0)\right] - \mathrm{KL}\left[q(\boldsymbol{u}) \| p(\boldsymbol{u})\right].\end{aligned} \tag{11}$$

Under sparse GP approximation, we can specify a free Gaussian density on the function values $\boldsymbol{u}$ giving $q(\boldsymbol{u}) = \mathcal{N}(\boldsymbol{u}|\boldsymbol{\mu}_{\boldsymbol{u}}, \boldsymbol{\Sigma}_{\boldsymbol{u}})$. The only term that requires further specification is $q(\boldsymbol{x}_i|\boldsymbol{f}_i)$ and [28] presented $q(\boldsymbol{x}_i|\boldsymbol{f}_i) = \mathcal{N}(\boldsymbol{x}_i|\boldsymbol{A}_{i-1}\tilde{\boldsymbol{f}}_{i-1} +$

$\boldsymbol{b}_{i-1}, \boldsymbol{S}_{i-1})$, where $\tilde{\boldsymbol{f}}_i$ is a free variational parameter. $\boldsymbol{A}_i$, $\boldsymbol{b}_i$ and $\boldsymbol{S}_i$ are other free variational parameters depending on the choice of $\tilde{\boldsymbol{f}}_i$. This generalises a number of previous work. For example, by setting both $\tilde{\boldsymbol{f}}_i = \boldsymbol{x}_i$ in $q(\cdot|\cdot)$ and $q(\boldsymbol{x}) = q(\boldsymbol{x}_0)$, we recover the expression from [19] and the optimal $q(\boldsymbol{x})$ can be solved by calculus of variations. If we set $\tilde{\boldsymbol{f}}_i = \boldsymbol{x}_i$ only, then we have the Gaussian factorised approximation of [38]. In [28], they proposed setting $\tilde{\boldsymbol{f}}_i = k(\boldsymbol{x}_i, \boldsymbol{u})K_{\boldsymbol{uu}}^{-1}\boldsymbol{u}$ and the training turned out to be marginally better if a sufficiently large amount of iterations are allowed.

We choose to set only $\tilde{\boldsymbol{f}}_i = \boldsymbol{x}_i$ throughout this paper, which is the approach by [38]. This is because their approach has shown to have a stable improvement in model accuracy during the training phase. Also, its linear factorising nature, given a fixed and often small number of allowed iterations, is also beneficial for training time. Such setting leads to $q(\boldsymbol{x}_i|\boldsymbol{f}_i) = \mathcal{N}(\boldsymbol{x}_i|\boldsymbol{A}_{i-1}\boldsymbol{x}_{i-1} + \boldsymbol{b}_{i-1}, \boldsymbol{S}_{i-1})$, with the free parameters $\boldsymbol{S}_{i-1} = (\boldsymbol{Q}^{-1} + \boldsymbol{C}^T\boldsymbol{R}^{-1}\boldsymbol{C})^{-1}$, $\boldsymbol{A}_{i-1} = \boldsymbol{S}_{i-1}\boldsymbol{Q}^{-1}$ and $\boldsymbol{b}_{i-1} = \boldsymbol{S}_{i-1}\boldsymbol{C}^T\boldsymbol{R}^{-1}(\boldsymbol{y}_t - \boldsymbol{d})$.

The prediction with GPSSM can be done rather cheaply. Suppose we would like to predict the new observation $\boldsymbol{y}^\star$ based on a new control input $\boldsymbol{c}^\star$, by defining all training data as $\mathcal{D} \equiv \{\boldsymbol{y}_{1:T}, \boldsymbol{x}_{0:T}\}$, we can show easily [19] that the predictive distribution $p(\boldsymbol{y}^\star|\boldsymbol{c}^\star, \mathcal{D})$ is given by

$$p(\boldsymbol{y}^\star|\boldsymbol{c}^\star, \mathcal{D}) = \mathcal{N}(\boldsymbol{y}^\star|\boldsymbol{C}f(\boldsymbol{x}_T, \boldsymbol{c}^\star) + \boldsymbol{d}, \boldsymbol{R} + \boldsymbol{CQC}^T). \tag{12}$$

In practice, we sample $f(\boldsymbol{x}_T, \boldsymbol{c}^\star)$, where the latent states $\boldsymbol{x}_{1:T}$ are already estimated via training. Then, we predict the new observation based on drawing samples from the distribution. This is the approach used by [28].

4 Active Learning Strategies

The previous section presents the modelling of GPSSM and in this section, we are interested in acquiring a strategy to actively learn the GPSSM by steering the system through the latent state-space. We employ AL as a query strategy to pick the most informative new control input $\boldsymbol{c}_t^*$ while observing $\boldsymbol{y}_t$, and to learn the model as we explore. Contrary to the two previous work [6,8], we treat the states as latent and unknown and $\boldsymbol{c}_t$ is the only controllable variable. For the exploration criterion, we use approximated mutual information. The usage of mutual information for AL has been well motivated by, for example, Krause et al. [33]. They pointed out that the mutual information might lead to a more accurate model than the differential entropy. This quantity was also shown to be the same as minimising the expected uncertainty of the model [15].

First, let us recall from chapter 8.5 of Cover [9] that the mutual information between any two sets of random variables $\boldsymbol{Y}$ and $\boldsymbol{F}$ with joint density $p(\boldsymbol{Y}, \boldsymbol{F})$ is defined as

$$I(\boldsymbol{Y}; \boldsymbol{F}) = \int p(\boldsymbol{Y}, \boldsymbol{F}) \log \left(\frac{p(\boldsymbol{Y}, \boldsymbol{F})}{p(\boldsymbol{Y})p(\boldsymbol{F})} \right) d\boldsymbol{Y} d\boldsymbol{F}$$

and a well-known relationship between mutual information and differential entropy $h(\cdot)$ is given by $I(\boldsymbol{Y};\boldsymbol{F}) = h(\boldsymbol{Y}) - h(\boldsymbol{Y}|\boldsymbol{F})$.

For steering the system through latent state-space while gathering information for efficiently learning the model, a sensible quantity for exploration is the mutual information between the latest observations $\boldsymbol{y}_{t+1}$ and latest predicted transition functions $\boldsymbol{f}_{t+1}$. This quantity is an extension to the active strategy from Buisson et al. [6] by using mutual information as a criterion. That is,

$$\boldsymbol{c}_t^* = \underset{\boldsymbol{c}_t \in \mathcal{C}}{\operatorname{argmax}}\, I(\hat{\boldsymbol{y}}_{t+1}; \boldsymbol{f}_{t+1}), \tag{13}$$

where $\boldsymbol{c}_t$ is placed within the term $\boldsymbol{f}_{t+1} \equiv f(\boldsymbol{x}_t, \boldsymbol{c}_t)$. We use $\hat{\boldsymbol{y}}_{t+1}$ because this term is provided from predicting GPSSM with a given new control $\boldsymbol{c}_t$. Altogether, $\boldsymbol{c}_t$ is the only independent variable.

This expression is intractable and there have been enormous efforts in estimating it (e.g. [2, 25, 34, 39, 40]). However, most previous work assumed that we have little or no information on both entropies in the mutual information expression and this comes down to the fact that we do not know the corresponding probability density function (pdf). Their approach includes deriving bounds but extra functions such as *critic* or *variational parameters* are often required (e.g. section 2 in [40]). However, should we have some information about this quantity, it is much better to make use of this known information instead. Therefore, we derive our estimate for the mutual information by directly approximating the pdf. The main advantages of our method versus others are: (1) As discussed in the previous paragraph, other methods often require extra functions called *critic* or *variational parameters*, or some free parameters to set. Finding the best settings of these extra functions and parameters could be complicated at times. Mcallester et al. [37] also showed that there are statistical limitations in various bounds of mutual information and a better approach is to measure this quantity as a difference in entropies. (2) Our mathematical derivation is purely via inequalities and numerical approximations, resulting an approximated closed form. Hence, the final expression derived can be easily implemented.

The quantity $I(\boldsymbol{y}_t; \boldsymbol{f}_t)$ is often studied in SSM (e.g. [15]), motivated by the assumption of Markov property. However, if the Markov assumption is not completely valid, as there are (weak) dependencies between the states, the mutual information between all observations $\boldsymbol{y}_{1:t+1}$ and all predictions from the transition function $\boldsymbol{f}_{1:t+1}$ within a time-horizon might present a better quantity. To the best of our knowledge, we are the first to propose using mutual information between all variables in time, and the AL strategy presented in equation 13 can be viewed as a special case of this general formulation. Thus, another more general AL strategy is

$$\boldsymbol{c}_t^* = \underset{\boldsymbol{c}_t \in \mathcal{C}}{\operatorname{argmax}}\, I(\boldsymbol{y}_{1:t}, \hat{\boldsymbol{y}}_{t+1}; \boldsymbol{f}_{1:t+1}). \tag{14}$$

Again, the term $\hat{\boldsymbol{y}}_{t+1}$ is provided from predicting GPSSM. This also leads to the third advantage of our method: (3) To the best of our knowledge, we are the first

input : Initial T observations $\boldsymbol{y}_{1:T}$ and control inputs $\boldsymbol{c}_{0:T}$. Initial state $\boldsymbol{x}_0$.
output: An optimised GPSSM after taking N exploration steps, the final data set including the explored observations $\boldsymbol{y}_{1:T+N}$ and the corresponding controls $\boldsymbol{c}_{0:T+N}$.

Train the initial GPSSM **for** $t = T, T+1, \cdots, T+N$ **do**
1. Solve $\boldsymbol{c}_t^*$ via equation (14) (or (13));
2. Evaluate $\boldsymbol{y}_{t+1}$ from the system;
3. Update the training data set;
4. Retrain GPSSM to update hyperparameters;
5. $t = t + 1$;

end

Algorithm 1: *Active Learning in GPSSM* on control through approximation of mutual information

to estimate the total mutual information by leveraging the inference scheme of the model. This turns out to be computationally efficient as well.

For clarity, we refer $I(\hat{\boldsymbol{y}}_{t+1}; \boldsymbol{f}_{t+1})$ as *latest* mutual information (latMI) and $I(\boldsymbol{y}_{1:t}, \hat{\boldsymbol{y}}_{t+1}; \boldsymbol{f}_{1:t+1})$ as *total* mutual information (totMI). In the remainder of this section, we present how the two quantities can be approximated for GPSSM with latent states. We first present an estimation scheme for the *latest* mutual information to align with previous work. Then, we proceed to the estimation of our proposed *total* mutual information. Given the estimates of mutual information, we can formulate the AL strategy for GPSSM, as summarised in algorithm 1.

4.1 Computation of latest Mutual Information $I(\boldsymbol{y}_t; \boldsymbol{f}_t)$

To estimate the *latest* mutual information $I(\boldsymbol{y}_t; \boldsymbol{f}_t)$, we employ the Gaussian approximate integral derived by Girard [23]. Here, for two random variables $\boldsymbol{x}$ and $\boldsymbol{y}$ with $p(\boldsymbol{y}|\boldsymbol{x}) = \mathcal{N}(\boldsymbol{y}|\mu(\boldsymbol{x}), \sigma^2(\boldsymbol{x}))$ and $p(\boldsymbol{x}|\boldsymbol{u}, \boldsymbol{\Sigma}_{\boldsymbol{x}}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{u}, \boldsymbol{\Sigma}_{\boldsymbol{x}})$ for some mean $\boldsymbol{u}$ and variance $\Sigma_{\boldsymbol{x}}$, the Gaussian approximation yields

$$\int p(\boldsymbol{y}|\boldsymbol{x}) p(\boldsymbol{x}|\boldsymbol{u}, \boldsymbol{\Sigma}_{\boldsymbol{x}}) d\boldsymbol{x} \approx \mathcal{N}(M(\boldsymbol{u}, \boldsymbol{\Sigma}_{\boldsymbol{x}}), V(\boldsymbol{u}, \boldsymbol{\Sigma}_{\boldsymbol{x}})),$$

where $M(\boldsymbol{u}, \boldsymbol{\Sigma}_{\boldsymbol{x}})$ and $V(\boldsymbol{u}, \boldsymbol{\Sigma}_{\boldsymbol{x}})$ are integral functions to be evaluated (see supplementary materials for details). Based on this approximation, $I(\boldsymbol{y}_t; \boldsymbol{f}_t)$ can be approximately computed as in the following proposition.

Proposition 1. *Given the definition of GPSSM from equation 1 - 4, as well as the notation of $\mathcal{M}_t$ and $\mathcal{K}_t$ defined in equation 7. We define the approximation of the following integrals as*

$$\begin{aligned} \mathcal{N}(\boldsymbol{f}_1|M_1, V_1) &:= \mathcal{N}(\boldsymbol{f}_1|M(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0), V(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)) \\ &\approx \int \mathcal{N}(\boldsymbol{f}_1|\mathcal{M}_0, \mathcal{K}_0) \mathcal{N}(\boldsymbol{x}_0|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) d\boldsymbol{x}_0 \end{aligned}$$

and, recursively, for all $t = T, T+1, \cdots, T+N$,

$$\mathcal{N}(\boldsymbol{f}_t|M_t, V_t) \approx \int \mathcal{N}(\boldsymbol{f}_t|\mathcal{M}_{t-1}, \mathcal{K}_{t-1}) \mathcal{N}(\boldsymbol{x}_{t-1}|M_{t-1}, V_{t-1} + \boldsymbol{Q}) d\boldsymbol{x}_{t-1}.$$

Then, the latest mutual information is approximately

$$I(\boldsymbol{y}_t; \boldsymbol{f}_t) \approx \frac{1}{2} \log \left(\frac{\det(\boldsymbol{R} + \boldsymbol{C}(V_t + \boldsymbol{Q})\boldsymbol{C}^T)}{\det(\boldsymbol{R} + \boldsymbol{C}\boldsymbol{Q}\boldsymbol{C}^T)} \right). \tag{15}$$

Proof (sketch). Based on the definition of mutual information, the numerator in equation 15 relies on the approximation via *moment matching*, whereas the denominator in equation 15 can be computed directly.

4.2 Computation of total Mutual Information $I(\boldsymbol{y}_{1:t}; \boldsymbol{f}_{1:t})$

The ELBO $\mathcal{L}_t$ given by equation 11 can be utilised to estimate the *total* mutual information. Since this quantity is estimated via samples, this quantity is sample driven and the following proposition presents our computational approach.

Proposition 2. *Given the definition of GPSSM from equation 1 - 4, as well as the expression of the ELBO $\mathcal{L}_t$ from equation 11, for $t = T, T+1, \cdots, T+N$, if S samples are drawn, the s-th sample estimate ($s = 1, \cdots, S$) of the mutual information between observations $\boldsymbol{y}_{1:t}$ and the prediction $\boldsymbol{f}_{1:t}$, denoted by i_s, is bounded by*

$$i_s \leq \sum_{i=1}^{t} \log(\mathcal{N}_s(\boldsymbol{y}_i | \boldsymbol{C}\boldsymbol{f}_i + \boldsymbol{d}, \boldsymbol{R} + \boldsymbol{C}\boldsymbol{Q}\boldsymbol{C}^T)) - \mathcal{L}_{t,s}, \tag{16}$$

where $\mathcal{N}_s(\cdot)$ and $\mathcal{L}_{t,s}$ are the s-th sample estimate of the normal distribution and ELBO $\mathcal{L}_t$, respectively. The total mutual information is then approximated by $I(\boldsymbol{y}_{1:t}; \boldsymbol{f}_{1:t}) \approx \frac{1}{S} \sum_{s=1}^{S} i_s$.

Proof (sketch). The proof employs mainly the definition of mutual information. The intractable part is bounded via ELBO, and the other term can be computed directly. Since it is a sample based approach, the bound from a sample becomes an approximation after taking all samples.

Remark 3. In computing equation 16, the sampling nature comes from the specific value of $\boldsymbol{x}_0$ drawn, and this value will be different between samples. Therefore, the evaluated i_s will also be different.

The efficiency of this approach heavily relies on the inference of GPSSM and it is possible to use another form of ELBO by defining $q(\cdot)$ differently or even other inference techniques such as a MCMC based approaches. While different approaches will incur their corresponding computational effort, it is better to keep the same approach in training GPSSM and computing total mutual information in order to avoid extra computational cost, as we reuse the posterior with trained hyperparameters in estimating the total mutual information.

5 Experiments

In this section, our experiments demonstrate how our proposed AL strategies help in learning GPSSM. From our discussion during previous sections, the two

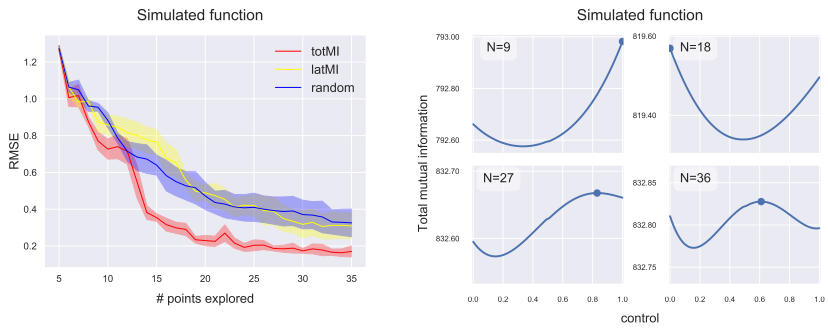

Fig. 2. The *left* diagram shows the results of the simulated function: Mean and standard deviation of RMSE for 30 exploration steps (starting with 5 initial points) are collected from independent trials. latMI is presented in yellow, totMI in red, and random exploration in blue. The right figures shows where the maximum *total* mutual information is attained when 9, 18, 27, 36 points are explored in one particular trial.

latest work from Buisson et al. [6] and Capone et al. [8] assumed known states, so we cannot compare against them in the setting with unobservable states we focus on. Therefore, the reasonable benchmark we can compare to is the usual setting where models are learned with randomly distributed control inputs (e.g. [36]). We first perform an experiment on a simulated function, then on few non-linear physical problems.[3] The overall goal is to obtain a high accuracy for the GPSSM with as few selected points as possible.

5.1 Simulated Function

We first consider a modified kink function based on [28]. Our goal is to learn the dynamics of the system within a range of $\mathcal{X} = [-3.0, 1.1]$ for the latent states and $\mathcal{C} = [0, 1]$ for the control. We begin with 5 given points and explore 30 steps.

The left diagram in Figure 2 depicts the RMSE of training sessions applying random selection, *latest* mutual information (latMI) and *total* mutual information (totMI) strategy, respectively. Each session consists of independent runs. Soon after collecting the first control input, both AL show results not worse than random but latMI is not significant, while totMI is more profound.

To illustrate why we need our AL strategies, especially totMI, we also present a snapshot of the position of the maximum totMI as we explore in the right of Figure 2 in a trial. The selected diagram indicates that there is no trivial most informative control where we can attain maximum totMI. Therefore, optimisation in equation 14 (or 13) is necessary. These diagrams could change as we run different trials, and such evaluation will be more complicated as we increase the dimension of controls or the dynamics become more complex.

[3] The ground base of GPSSM code is based on Ialongo et al. [28], re-engineered in `GPflow` 2.1 [35].

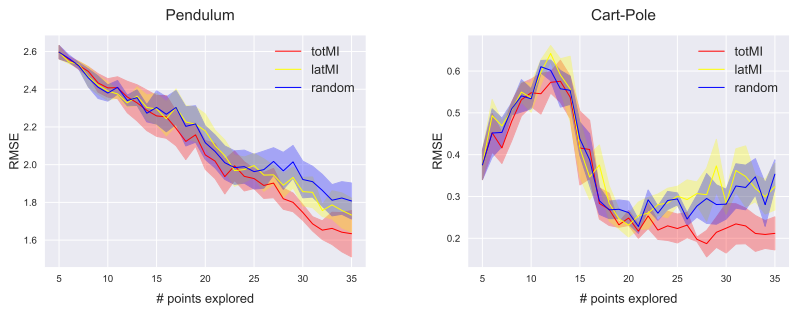

Fig. 3. The left diagram shows the results of the pendulum: Mean and standard deviation of RMSE for 30 exploration steps (starting with 5 initial points) are collected from independent trials. latMI is presented in yellow, totMI in red, and random exploration in blue. The right diagram shows the results of the cart-pole: Again, mean and standard deviation of RMSE for 30 exploration steps (starting with 5 initial points) are collected from independent trials.

5.2 Pendulum and Cart-Pole

In these experiments, our aim is to learn the physics of the pendulum and cart-pole by actively controlling the input torque and the force of the cart, respectively. We evaluate the state-space model's accuracy via the angular position obtained from the model against the ground-truth, which is calculated via the equation of motions.

Figure 3 depicts the RMSE of training sessions applying random selection, *latest* mutual information and *total* mutual information strategy for pendulum (left) and cart-pole (right), respectively. Both experiments consist of independent runs. For the pendulum, both totMI and latMI strategy show advantages against the random exploration. However, totMI shows even better results as the exploration proceeds. For cart-pole, all strategies have, in the initial exploration phase, a raise in RMSE, which is subsequently reduced, as the model becomes more stable. latMI is indeed unable to attain noticeable improvement but totMI strategy still shows favourable performance via a consistently lower RMSE.

Since the physics of both systems are common enough that we can actually measure the states, we also compare our work directly with the AL strategy by Capone et al. [8], even though the fairness of such a comparison is debatable.

5.3 Twin-Rotor Aerodynamical System

As a more realistic setting, we test our AL strategy on a Twin-Rotor Aerodynamical System (TRAS). This is a typical design for control experiments, and its behaviour resembles that of a helicopter. There are two rotors – one horizontal and another vertical – joined by a beam. There is also a counter-weight, which determines a stable equilibrium position. When the system is switched off, the main rotor is lowered. The motion of the system is controlled by the two

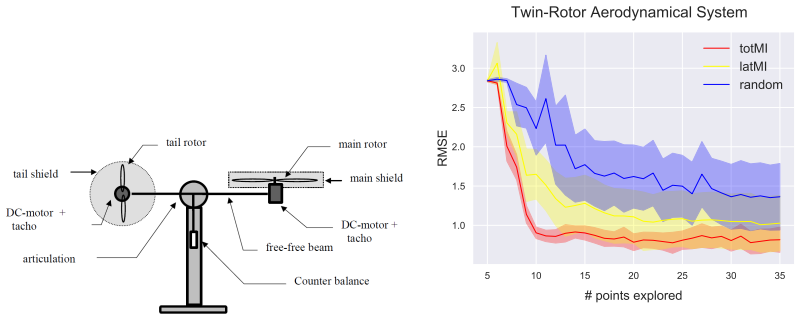

Fig. 4. The left diagram shows an example of a Twin-Rotor Aerodynamical System, taken from [4]. The right diagram shows the results of the Twin-Rotor Aerodynamical System: Mean and standard deviation of RMSE for 30 exploration steps (starting with 5 initial points) are collected from independent trials. latMI is presneted in yellow; totMI in red and random exploration in blue.

motor supply voltages to each motor. We demonstrate a simple example of such a system in the left diagram of Figure 4.

Our aim is to learn the physics of the TRAS by actively control the two voltages to the system. We evaluate the model's accuracy via the angular position measured from the model against which is calculated via the equation of motions on both rotors. The physics of the TRAS is based on the description by [4]. The right diagram of Figure 4 depicts the RMSE of training sessions applying random selection, latMI and totMI strategy, respectively. Each session consists of independent runs. Both latMI and totMI strategies show advantage against random exploration in presence of more sophisticated dynamics. However, GPSSM with totMI reaches an acceptable accuracy with much lower number of data points.

6 Discussion

We study AL in GPSSM by proposing a tractable mutual information estimates in order to select the most informative control inputs to maximise our model accuracy as we explore. Both *latest* and *total* mutual information are proposed. The empirical results show that *total* mutual information strategy outperforms the other approach and random exploration. There is a promising potential of our approach towards more complex industrial settings. Since we leverage ELBO to derive our estimates, we also expect our propositions to be applicable to future work on inference of GPSSM and other probabilistic models with latent space.

Another direction of extension is to look into the use of maximal mutual information in reinforcement learning (RL) community, as such strategy has been used in planning and control. For instance, Ding et al. [11] applied such strategy in Partially Observable Markov Decision Process (POMDP) framework in order to efficiently learn the RL model. However, there are key differences between the two themes. First, the word *control* is interpreted very differently.

In GPSSM, this refers to an augmented latent state and each input is simply a deterministic real number whereas in POMDP, this is a mapping between an action taken in certain state to maximise the expected return or reward. Second, AL in GPSSM has a very specific problem formulation whereas efficient learning in POMDP is a general framework with a large room of freedom to specify different components, leading to various research topics (e.g. [5, 14, 24]).

References

1. Alaa, A.M., van der Schaar, M.: Attentive state-space modeling of disease progression. In: Advances in Neural Information Processing Systems. pp. 11338–11348 (2019)
2. Alemi, A., Poole, B., Fischer, I., Dillon, J., Saurous, R.A., Murphy, K.: Fixing a broken elbo. In: International Conference on Machine Learning. pp. 159–168. PMLR (2018)
3. Amin, K., Cortes, C., DeSalvo, G., Rostamizadeh, A.: Understanding the effects of batching in online active learning. In: International Conference on Artificial Intelligence and Statistics. pp. 3482–3492 (2020)
4. Anonymous: Two rotor aero-dynamical system. user's manual. Tech. rep., Inteco, www.inteco.com.pl (2006)
5. Berkenkamp, F., Nurchetta, M., Schoellig, A.P., Krause, A.: Safe model-based reinforcement learning with stability guarantees. Advances in neural information processing systems (2017)
6. Buisson-Fenet, M., Solowjow, F., Trimpe, S.: Actively learning gaussian process dynamics. In: Learning for Dynamics and Control. pp. 5–15. PMLR (2020)
7. Cai, W., Zhang, M., Zhang, Y.: Batch mode active learning for regression with expected model change. In: IEEE Transactions on Neural Networks and Learning Systems. vol. 28, pp. 1668–1681. IEEE (2016)
8. Capone, A., Noske, G., Umlauft, J., Beckers, T., Lederer, A., Hirche, S.: Localized active learning of gaussian process state space models. In: Learning for Dynamics and Control. pp. 490–499. PMLR (2020)
9. Cover, T.M., Thomas, J.A.: Elements of information theory. John Wiley & Sons, 2nd edn. (2006)
10. Dasgupta, S.: Two faces of active learning. Theoretical computer science **412**(19), 1767–1781 (2011)
11. Ding, Y., Clavera, I., Abbeel, P.: Mutual information maximization for robust plannable representations. arXiv preprint arXiv:2005.08114 (2020)
12. Doerr, A., Daniel, C., Schiegg, M., Nguyen-Tuong, D., Schaal, S., Toussaint, M., Sebastian, T.: Probabilistic recurrent state-space models. In: International Conference on Machine Learning. pp. 1280–1289 (2018)
13. Eleftheriadis, S., Nicholson, T.F., Deisenroth, M.P., Hensman, J.: Identification of gaussian process state space models. In: Advances in neural information processing systems. pp. 5309–5319 (2017)
14. Engel, Y., Mannor, S., Meir, R.: Reinforcement learning with gaussian processes. In: Proceedings of the 22nd international conference on Machine learning. pp. 201–208 (2005)
15. Ertin, E., Fisher, J.W., Potter, L.C.: Maximum mutual information principle for dynamic sensor query problems. In: Information processing in sensor networks. pp. 405–416. Springer (2003)

16. Faruqi, F.A.: State space model for autopilot design of aerospace vehicles. Tech. rep., Weapons Systems Division, Defence Science and Technology Organisation (DSTO), Edinburgh (Australia) (2007)
17. Friedland, B.: Control system design: an introduction to state-space methods. Courier Corporation (2012)
18. Frigola, R.: Bayesian time series learning with Gaussian processes. Ph.D. thesis, University of Cambridge (2015)
19. Frigola, R., Chen, Y., Rasmussen, C.E.: Variational gaussian process state-space models. In: Advances in neural information processing systems. pp. 3680–3688 (2014)
20. Frigola, R., Lindstenk, F., Schön, T.B., Rasmussen, C.E.: Bayesian inference and learning in gaussian process state-space models with particle mcmc. In: Advances in Neural Information Processing Systems. pp. 3156–3164 (2013)
21. de G. Matthews, A.G., Hensman, J., Turner, R., Ghahramani, Z.: On sparse variational methods and the kullback-leibler divergence between stochastic processes. In: Artificial Intelligence and Statistics. pp. 231–239 (2016)
22. Gal, Y., Islam, R., Ghahramani, Z.: Deep bayesian active learning with image data. In: International Conference on Machine Learning. pp. 1183–1192 (2017)
23. Girard, A.: Approximate methods for propagation of uncertainty with Gaussian process models. Ph.D. thesis, Citeseer (2004)
24. Grande, R., Walsh, T., How, J.: Sample efficient reinforcement learning with gaussian processes. In: International Conference on Machine Learning. pp. 1332–1340. PMLR (2014)
25. Hoffman, M.D., Johnson, M.J.: Elbo surgery: yet another way to carve up the variational evidence lower bound. In: Workshop in Advances in Approximate Bayesian Inference, NIPS. vol. 1, p. 2 (2016)
26. Houlsby, N., Huszár, F., Ghahramani, Z., Lengyel, M.: Bayesian active learning for classification and preference learning. arXiv preprint arXiv:1112.5745 (2011)
27. Ialongo, A.D., van der Wilk, M., Rasmussen, C.E.: Closed-form inference and prediction in gaussian process state-space models. Advances in Neural Information Processing Systems, Workshop (2017)
28. Ialongo, A.D., Wilk, M.V.D., Hensman, J., Rasmussen, C.E.: Overcoming mean-field approximations in recurrent gaussian process models. In: International Conference on Machine Learning. pp. 2931–2940 (2019)
29. Kirk, D.E.: Optimal control theory: an introduction. Courier Corporation (2004)
30. Kirsch, A., van Amersfoort, J., Gal, Y.: Batchbald: Efficient and diverse batch acquisition for deep bayesian active learning. In: Advances in Neural Information Processing Systems. pp. 7024–7035 (2019)
31. Ko, J., Fox, D.: Gp-bayesfilters: Bayesian filtering using gaussian process prediction and observation models. Autonomous Robots **27**(1), 75–90 (2009)
32. Krause, A., Guestrin, C.: Nonmyopic active learning of gaussian processes: an exploration-exploitation approach. In: Proceedings of the 24th international conference on Machine learning. pp. 449–456 (2007)
33. Krause, A., Singh, A., Guestrin, C.: Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. Journal of Machine Learning Research **9**(Feb), 235–284 (2008)
34. Lombardi, D., Pant, S.: Nonparametric k-nearest-neighbor entropy estimator. Physical Review E **93**(1), 013310 (2016)
35. Matthews, A.G.D.G., Wilk, M.V.D., Nickson, T., Fujii, K., Boukouvalas, A., León-Villagrá, P., Ghahramani, Z., Hensman, J.: Gpflow: A gaussian process library

using tensorflow. The Journal of Machine Learning Research **18**(1), 1299–1304 (2017)
36. Mattos, C.L.C., Damianou, A., A, G.B., Lawrence, N.D.: Latent autoregressive gaussian processes models for robust system identification. IFAC-PapersOnLine **49**(7), 1121–1126 (2016)
37. McAllester, D., Stratos, K.: Formal limitations on the measurement of mutual information. In: International Conference on Artificial Intelligence and Statistics. pp. 875–884 (2020)
38. McHutchon, A.J.: Nonlinear modelling and control using Gaussian processes. Ph.D. thesis, University of Cambridge (2015)
39. Pérez-Cruz, F.: Estimation of information theoretic measures for continuous random variables. In: Advances in neural information processing systems. pp. 1257–1264 (2009)
40. Poole, B., Ozair, S., Oord, A.V.D., Alemi, A.A., Tucker, G.: On variational bounds of mutual information. In: International Conference on Machine Learning. pp. 5171–5180 (2019)
41. Rangapura, S.S., Seeger, M.W., Gasthaus, J., Stella, L., Wang, Y., Januschowski, T.: Deep state space models for time series forecasting. In: Advances in neural information processing systems. pp. 7785–7794 (2018)
42. Rasmussen, C.E., Williams, C.K.: Gaussian processes for machine learning. The MIT Press, Cambridge, MA (2006)
43. Schreiter, J., Nguyen-Tuong, D., Eberts, M., Bischoff, B., Markert, H., Toussaint, M.: Safe exploration for active learning with gaussian processes. Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML/PKDD) **9286**, 133–147 (8 2015)
44. Seo, S., Wallat, M., Graepel, T., Obermayer, K.: Gaussian process regression: Active data selection and test point rejection. In: Mustererkennung 2000, pp. 27–34. Springer (2000)
45. Settles, B.: Active learning literature survey. Computer Sciences Technical Report 1648, University of Wisconsin–Madison (2009)
46. Shui, C., Zhou, F., Gagné, C., Wang, B.: Deep active learning: Unified and principled method for query and training. In: International Conference on Artificial Intelligence and Statistics. pp. 1308–1318. PMLR (2020)
47. Srinivas, N., Krause, A., Kakade, S.M., Seeger, M.W.: Information-theoretic regret bounds for gaussian process optimization in the bandit setting. IEEE Transactions on Information Theory **58**(5), 3250–3265 (2012)
48. Titsias, M.: Variational learning of inducing variables in sparse gaussian processes. In: Artificial Intelligence and Statistics. pp. 567–574 (2009)
49. Turner, R., Deisenroth, M., Rasmussen, C.: State-space inference and learning with gaussian processes. In: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. pp. 868–875 (2010)
50. Twomey, N., Diethe, T., Flach, P.: Bayesian active learning with evidence-based instance selection. In: Workshop on Learning over Multiple Contexts, European Conference on Machine Learning (ECML15) (2015)
51. Wang, J., Hertzmann, A., Fleet, D.J.: Gaussian process dynamical models. In: Advances in neural information processing systems. pp. 1441–1448 (2006)
52. Zeng, Y., Wu, S.: State-space models: Applications in economics and finance, vol. 1. Springer (2013)
53. Zimmer, C., Meister, M., Nguyen-Tuong, D.: Safe active learning for time-series modeling with gaussian processes. In: Advances in Neural Information Processing Systems. pp. 2730–2739 (2018)