Active Sequential Learning with Tactile Feedback

Hannes P. Saal
University of Edinburgh
Edinburgh, UK
hannes.saal@ed.ac.uk

Jo-Anne Ting
University of British Columbia
Vancouver, BC, Canada
jting@acm.org

Sethu Vijayakumar
University of Edinburgh
Edinburgh, UK
sethu.vijayakumar@ed.ac.uk

Abstract

We consider the problem of tactile discrimination, with the goal of estimating an underlying state parameter in a sequential setting. If the data is continuous and high-dimensional, collecting enough representative data samples becomes difficult. We present a framework that uses active learning to help with the sequential gathering of data samples, using information-theoretic criteria to find optimal actions at each time step. We consider two approaches to recursively update the state parameter belief: an analytical Gaussian approximation and a Monte Carlo sampling method. We show how both active frameworks improve convergence, demonstrating results on a real robotic hand-arm system that estimates the viscosity of liquids from tactile feedback data.

1 INTRODUCTION

Learning from tactile sensory data is challenging due to the sparse, high-dimensional nature of touch. Consider a representative example of searching for keys in a large purse without looking inside. The task is to locate the keys as fast as possible (i.e., perform efficient exploration of the purse), with each exploratory finger movement giving tactile feedback about the object being felt (or lack thereof). To accomplish this task, one needs to i) be able to discriminate between keys and other objects/materials based on only tactile information; ii) know how to efficiently explore the purse (probing corners where objects tend to be stored); and iii) have tactile sensitivity to a wide range of stimuli. These three elements correspond to estimating a state parameter (that encodes relevant features of keys), performing a sequence of efficient actions to reduce the search time, and having an accurate and robust observation/sensory model (the mapping of actions to observations).

If we consider the problem of haptic tactile discrimination, e.g., performing the task above with a high degree-of-freedom (DOF) humanoid robot, the space of possible actions is continuous and high-dimensional, and collecting enough representative data samples such that the state parameter is identifiable becomes a difficult task. One could randomly and uniformly sample the space of possible actions, but this would be extremely time-consuming. If we could collect data samples that are more "informative", we would be able to estimate the state parameter in a reasonable amount of time.

In this paper, we use active learning to help with the sequential gathering of data samples in order to efficiently and accurately estimate the state parameter. We consider continuous actions, state parameters and observations. We assume that the observation model is learnt in an offline (training) phase and the estimation of the state parameter is done during run/test time. The state parameter can be time-varying (in which case, a filter would be used for recursive estimation), but we focus on the scenario where the state parameter value is fixed. Our proposed framework is straightforward to extend to filters.

The belief of the state parameter can be updated sequentially using Bayes' rule, but the resulting posterior distribution is analytically intractable when the observation model is nonlinear. We consider two approaches for approximating the posterior: i) an analytical Gaussian approximation (previously applied by (Girard, Rasmussen, Candela, & Murray-Smith, 2003; Deisenroth, Huber, & Hanebeck, 2009) to fil-

Appearing in Proceedings of the 13^{th} International Conference on Artificial Intelligence and Statistics (AISTATS) 2010, Chia Laguna Resort, Sardinia, Italy. Volume 9 of JMLR: W&CP 9. Copyright 2010 by the authors.

ters); and ii) Monte Carlo (MC) sampling (Doucet, Freitas, & Gordon, 2001). The first approach is fast but fails to capture multi-modal distributions, while the second is computationally involved, though more accurate.

We incorporate active learning to both belief updating methods, using information-theoretic criteria to find optimal actions in a sequential setting. Our proposed active framework relies on mutual information (MI) to improve the speed of convergence to the true state parameter value over its passive equivalent. When active sequential updating with Gaussian approximations is used, the MI measure can be calculated analytically. However, when active sequential updating MC sampling is used, calculation of MI becomes intractable in high dimensions and difficult to optimize. To address this, we consider an alternative information-theoretic criterion, quadratic information information (QMI), e.g., Torkkola (2003), that allows simplifying some of the computations. Both active sequential updating methods use a Gaussian process prior to estimate the nonlinear sensory/observation model.

We also evaluate the active sequential framework on a robotic anthropomorphic arm with a three-finger hand equipped with tactile sensor arrays, demonstrating the first implementation of active sequential learning on a real robotic system with continuous actions, state parameters and observations.

2 ACTIVE LEARNING

Different strategies may be used in order to decide which sample is most informative. For example, one could adopt an uncertainty sampling strategy (Lewis & Gale, 1994), querying samples that maximize an uncertainty measure (e.g., entropy or MI). Another approach is the query-by-committee algorithm (Seung, Opper, & Sompolinsky, 1992) that queries in "controversial" parts of the space, and various measures can be used to measure the amount of disagreement between committee members, e.g., entropy (Dagan & Engelson, 1995), average Kulback-Leibler (KL) divergence (McCallum & Nigam, 1998), etc. Yet another class of query strategies focuses on variance minimization and selects the sample that minimizes the generalization error by minimizing the variance of the model (Cohn, Ghahramani, & Jordan, 1996). Other approaches attempt to minimize the model's generalization error directly, e.g., (Roy & McCallum, 2001; Zhu, Lafferty, & Ghahramani, 2003), sometimes using uncertainty sampling as a fail-safe backup strategy (Guo & Greiner, 2007).

Information-theoretic criteria have been used in various fields for optimal experimental design (analogous to determining the optimal samples in active learning), e.g., (Lindley, 1956; Mackay, 1992; Denzler & Brown, 2002; Lewi, Butera, & Paninski, 2009). Paninski (2005) showed that an information-maximization strategy is asymptotically more efficient than non-adaptive, independent and identically distributed strategies such as random uniform sampling.

2.1 MUTUAL INFORMATION

MI is a measure from probability and information theory that captures the mutual dependence of two variables $\mathbf{y}$ and $\boldsymbol{\theta}$. If $\boldsymbol{\theta}$ and $\mathbf{y}$ are independent, their MI will be zero. MI can also be viewed as the KL divergence between the joint density (of $\boldsymbol{\theta}$ and $\mathbf{y}$) and the product of the individual marginal densities.

Let us assume the following notation: $\boldsymbol{\theta} \in \mathbb{R}^{d_{\boldsymbol{\theta}}}$ is the state parameter that we want to estimate (e.g., position, inertia, viscosity of an object); $\mathbf{x} \in \mathbb{R}^{d_{\mathbf{x}}}$ is the action taken; and $\mathbf{y} \in \mathbb{R}^{d_{\mathbf{y}}}$ is observed sensory data given a particular $\{\boldsymbol{\theta}, \mathbf{x}\}$. Figure 3 shows an example scenario with a robotic anthropomorphic arm that we use in our evaluations. We assume that observations $\mathbf{y}$ are a (nonlinear) function of the actions $\mathbf{x}$ taken and state parameter $\boldsymbol{\theta}$, as the observation model below describes:

$$\mathbf{y} = f(\mathbf{x}, \boldsymbol{\theta}) + \epsilon_{\mathbf{y}} \tag{1}$$

where $\{\mathbf{x}, \boldsymbol{\theta}, \mathbf{y}\}$ are all continuous, and $\epsilon_{\mathbf{y}}$ is observation noise. Both $\mathbf{x}$ and $\mathbf{y}$ are potentially sparse and high-dimensional.

For active learning, we are interested in determining the optimal action $\mathbf{x}^*$ to take during test time such that the MI between $\mathbf{y}$ and $\boldsymbol{\theta}$, $\mathbf{I}(\boldsymbol{\theta}; \mathbf{y}|\mathbf{x})$, is maximized, i.e., $\mathbf{x}^* = \arg\max_{\mathbf{x} \in \mathbf{X}} \mathbf{I}(\boldsymbol{\theta}; \mathbf{y}|\mathbf{x})$, where:

$$\mathbf{I}(\boldsymbol{\theta}; \mathbf{y}|\mathbf{x}) = \iint p(\boldsymbol{\theta}, \mathbf{y}|\mathbf{x}) \log \frac{p(\boldsymbol{\theta}, \mathbf{y}|\mathbf{x})}{p(\boldsymbol{\theta}) p(\mathbf{y}|\mathbf{x})} d\mathbf{y} d\boldsymbol{\theta} \tag{2}$$

where $p(\boldsymbol{\theta}, \mathbf{y}|\mathbf{x})$ is the joint probability distribution of $\boldsymbol{\theta}$ and $\mathbf{y}|\mathbf{x}$; and $p(\boldsymbol{\theta})$ and $p(\mathbf{y}|\mathbf{x})$ are the marginal probability distributions of $\boldsymbol{\theta}$ and $\mathbf{y}$, respectively. Eq. (2) is analytically tractable only if $p(\boldsymbol{\theta})$ and $p(\mathbf{y}|\mathbf{x})$ are both Gaussian distributions and if $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})$ depends linearly on $\boldsymbol{\theta}$. As a result, when any nonlinear function f with high-dimensional observations $\mathbf{y}$ is adopted, MI is intractable, and its optimization numerically unstable.

Previous methods have attempted to solve the integral in Eq. (2) in some of following ways: MC sampling (as used by Denzler and Brown (2002) for object recognition in active camera control) or

numerical approximation techniques (e.g., Gaussian quadrature rules, as done by Amari, Cichocki, and Yang (1994) for Independent Component Analysis estimation and by Fookes and Maeder (2004) for medical image registration). These methods tend to be computationally prohibitive even when the dimensionality of $\mathbf{y}$ becomes moderately high, limiting the practical use of active learning with MI. In the following sections, we take a first step towards tackling this problem.

3 ACTIVE SEQUENTIAL LEARNING

3.1 GAUSSIAN PROCESS OBSERVATION MODEL

Let us assume an observed N-sample training data set $D = \{\boldsymbol{\tau}^i, \boldsymbol{\chi}^i, \mathbf{t}^i\}_{i=1}^N$, where $\boldsymbol{\tau}$ are the state parameters, $\boldsymbol{\chi}$ are the actions taken, and $\mathbf{t}$ are the corresponding sensory observations. We place a Gaussian Process (GP) (Williams & Rasmussen, 1995) prior over the observation model in Eq. (1):

$$y_m(\boldsymbol{\theta}, \mathbf{x}) \sim \mathrm{GP}(0, \mathbf{K}_m) \tag{3}$$

where $m = 1, ..., d_{\mathbf{y}}$[1], $d_{\mathbf{y}}$ is the dimensionality of $\mathbf{y}$, and the covariance function $\mathbf{K}_m \in \mathbb{R}^{N\times N}$ is a Gaussian kernel with matrix elements $k_m(\mathbf{z}_p, \mathbf{z}_q)$:

$$k_m(\mathbf{z}_p, \mathbf{z}_q) = \alpha_m^2 \exp\left\{\frac{1}{2}(\mathbf{z}_p - \mathbf{z}_q)^T \mathbf{H}_m^{-1}(\mathbf{z}_p - \mathbf{z}_q)\right\} + \sigma_m^2 \delta_{\mathbf{z}_p}(\mathbf{z}_q) \tag{4}$$

where $\mathbf{z} \in \mathbb{R}^{d_{\boldsymbol{\theta}}+d_{\mathbf{x}}}$ is the vector $[\boldsymbol{\theta}\ \mathbf{x}]^T$; $\mathbf{H}_m = \begin{pmatrix} \mathbf{H}_m^{\boldsymbol{\theta}} & 0 \\ 0 & \mathbf{H}_m^{\mathbf{x}} \end{pmatrix}$; $\mathbf{H}_m^{\boldsymbol{\theta}}$ and $\mathbf{H}_m^{\mathbf{x}}$ are diagonal matrices; and $\delta_{\mathbf{z}_p}$ is a Dirac delta function centered at $\mathbf{z}_p$. Note that $\mathbf{H}_m$ is block diagonal since $\mathbf{x}$ and $\boldsymbol{\theta}$ are assumed to be independent, simplifying the model considerably. The set of GP hyperparameters to be optimized is $\gamma = \left\{\alpha_m^2, \sigma_m^2, \mathbf{H}_m^{\boldsymbol{\theta}}, \mathbf{H}_m^{\mathbf{x}}\right\}_{m=1}^{d_{\mathbf{y}}}$.

3.2 BELIEF UPDATING

Assuming that the observation model in Eq. (1) is learnt accurately in an offline training phase, we are faced with the task of estimating the particular value of $\boldsymbol{\theta}^*$ at run/test time. As data samples are sequentially observed/collected, the estimate of $\boldsymbol{\theta}^*$ is recursively updated. Once an action $\mathbf{x}_t$ has been taken and the corresponding $\mathbf{y}_t$ observed at time step t, the posterior distribution over $\boldsymbol{\theta}$ can then be updated using Bayes' rule:

$$p_t(\boldsymbol{\theta}|\mathbf{y}_t, \mathbf{x}_t) = \frac{p(\mathbf{y}_t|\boldsymbol{\theta}, \mathbf{x}) p_{t-1}(\boldsymbol{\theta})}{p(\mathbf{y}_t|\mathbf{x}_t)} \tag{5}$$

where $p_{t-1}(\boldsymbol{\theta})$ is the distribution over $\boldsymbol{\theta}$ from the previous time step), $p(\mathbf{y}_t, \mathbf{x}_t|\boldsymbol{\theta}) \equiv p(\mathbf{y}_t|\boldsymbol{\theta}, \mathbf{x}_t)$ is the likelihood of the current observation conditioned on the action that was taken, and $p(\mathbf{y}_t|\mathbf{x}_t)$ is the marginal of $\mathbf{y}$.

This update is analytically intractable when f is non-linear. Fast approximations can be used, but they might fail if f is highly nonlinear On the other hand, it can be be numerically approximated with sequential MC sampling techniques, though at a computational cost. In the following we present active extensions for both fast approximate Gaussian updating and MC sampling updating.

3.2.1 Analytical Gaussian Updating

Past work has proposed the use of a GP to model $\mathbf{y}$ in the context of a filter, e.g., (Girard et al., 2003; Deisenroth et al., 2009), arriving at a Gaussian approximation of the marginal $p(\mathbf{y}|\mathbf{x})$. The expressions for the marginal mean of $\mathbf{y}$, the marginal variance of $\mathbf{y}$, and cross-covariance between $\boldsymbol{\theta}$ and $\mathbf{y}$ (denoted by $\mathbf{m}(\mathbf{x}) \in \mathbb{R}^{d_{\mathbf{y}}}$, $\mathbf{S}(\mathbf{x}) \in \mathbb{R}^{d_{\mathbf{y}}\times d_{\mathbf{y}}}$, and $\mathbf{C}(\mathbf{x}) \in \mathbb{R}^{d_{\mathbf{y}}\times d_{\boldsymbol{\theta}}}$, respectively) can all be evaluated analytically. We use the following recursive update equations for the mean and variance of $\boldsymbol{\theta}$ at time step t, which appear in similar for in the extended Kalman filter or the unscented filter:

$$\boldsymbol{\mu}_t = \boldsymbol{\mu}_{t-1} + \mathbf{C}_t^T \mathbf{S}_t^{-1}(\mathbf{y}_t^{obs} - \mathbf{m}_{t-1})$$
$$\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}_{t-1} - \mathbf{C}_t^T \mathbf{S}_t^{-1} \mathbf{C}_t$$

where $\boldsymbol{\mu}_t$ is the mean of the Gaussian over $\boldsymbol{\theta}$ at time step t and $\boldsymbol{\Sigma}$ is the covariance. $\mathbf{m}$, $\mathbf{S}$, and $\mathbf{C}$ can be evaluated analytically as follows (Girard et al., 2003; Deisenroth et al., 2009):

$$m_m = \mathbf{q}_m(\mathbf{x})^T \mathbf{K}_m^{-1} \mathbf{t}_m$$
$$S_{mn} = \mathbf{t}_m^T (\mathbf{K}_m^{-1})^T \mathbf{Q}_{mn}(\mathbf{x}) \mathbf{K}_n^{-1} \mathbf{t}_n - m_m m_n + \delta(m-n)\left(\alpha_m^2 - tr(\mathbf{K}_m^{-1}\mathbf{Q}_{mm})\right)$$
$$C_{mn} = \mathbf{Z}_m^T(\mathbf{x}) \mathbf{K}_m^{-1} \mathbf{t}_m - \mu_n m_m$$

Here, m_m is the m-th coefficient of $\mathbf{m}$; S_{mn} is the (m,n)-th entry of the matrix $\mathbf{S}$; C_{mn} is the (m,n)-th entry of the matrix $\mathbf{C}$; and $\mathbf{t}_m \in \mathbb{R}^N$ is the m-th dimension of observed tactile data for all training samples. The derived expressions for $\mathbf{q}_m$, $\mathbf{Q}_{mm}$, and $\mathbf{Z}_m$ are given in the supplementary material.

We introduce an active component to the selection of action $\mathbf{x}_t$ at time step t, by noting that the MI can be

[1] By assuming independent $\mathbf{y}$ dimensions, $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})$ factorizes as $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x}) = \prod_{m=1}^{d_{\mathbf{y}}} p(\mathbf{y}_m|\boldsymbol{\theta}, \mathbf{x})$.

calculated analytically when $p(\boldsymbol{\theta})$ is Gaussian. We find the optimal action $\mathbf{x}_t^* = \text{argmax}_{\mathbf{x}}\ \mathbf{I}(\boldsymbol{\theta};\mathbf{y}|\mathbf{x}) = \text{argmax}_{\mathbf{x}}\ |\mathbf{C}(\mathbf{x})\mathbf{S}(\mathbf{x})\mathbf{C}(\mathbf{x})^T|$ by performing gradient ascent on $\mathbf{I}(\boldsymbol{\theta};\mathbf{y}|\mathbf{x})$, with the gradient defined as $\nabla_{\mathbf{x}}|\mathbf{C}(\mathbf{x})\mathbf{S}(\mathbf{x})\mathbf{C}(\mathbf{x})^T|$. Expressions for the gradient of $\mathbf{C}$ and $\mathbf{S}$ are given in the supplementary material.

3.2.2 MC Sampling-based Updating

The second approach is to approximate $p_t(\boldsymbol{\theta})$ with a set of particles, which comes with higher computational complexity, but allows to accurate updates even when f is highly nonlinear.

We can also numerically approximate Eq. (5) by sequential MC sampling methods—using P samples, with each sample p having a value $\boldsymbol{\theta}_p$ and a corresponding weight w_p ($0 \leq w_p \leq 1$ and $\sum_{p=1}^{P} w_p = 1$). We evaluate Eq. (2) in order to find the optimal action $\mathbf{x}_t^*$ to take at time step t. Unfortunately, when the nonlinear observation model of Eq. (1) is adopted, the MI is intractable for high dimensional $\mathbf{y}$, and its optimization is numerically unstable. As an alternative to MI, we consider QMI, $\mathbf{I}_Q(\boldsymbol{\theta};\mathbf{y}|\mathbf{x})$:

$$\begin{aligned}\mathbf{I}_Q(\boldsymbol{\theta};\mathbf{y}|\mathbf{x}) &= \iint \left(p(\boldsymbol{\theta},\mathbf{y}|\mathbf{x}) - p(\boldsymbol{\theta})p(\mathbf{y}|\mathbf{x})\right)^2 d\mathbf{y}d\boldsymbol{\theta} \\ &= \iint \Big[\left(p(\boldsymbol{\theta})p(\mathbf{y}|\boldsymbol{\theta},\mathbf{x})\right)^2 + p(\boldsymbol{\theta})^2 p(\mathbf{y}|\mathbf{x})^2 \\ &\qquad -2p(\mathbf{y}|\boldsymbol{\theta},\mathbf{x})p(\boldsymbol{\theta})^2 p(\mathbf{y}|\mathbf{x})\Big]\, d\mathbf{y}d\boldsymbol{\theta} \\ &:= V_1 + V_2 - 2V_3 \end{aligned} \tag{6}$$

where the terms $p(\boldsymbol{\theta},\mathbf{y}|\mathbf{x})$, $p(\boldsymbol{\theta})$ and $p(\mathbf{y}|\mathbf{x})$ are as defined for Eq. (2). QMI is not a strict measure of information (since it violates the principle of additivity), but it is always positive. Torkkola (2003) shows that if we are interested in finding the distribution that maximizing the KL divergence (equivalent to maximizing mutual information), then maximizing a quadratic divergence measure is equivalent to *maximizing a lower bound to the KL divergence*.

The advantage of QMI is that the integral over $\mathbf{y}$ in Eq. (6) can be solved *analytically*[2] if Gaussian densities are assumed for $p(\mathbf{y}|\boldsymbol{\theta},\mathbf{x})$. We can evaluate Eq. (6) to get:

$$V_1 = \sum_{p=1}^{P} w_p^2 (4\pi)^{-\frac{d_{\mathbf{y}}}{2}} |\boldsymbol{\Phi}(\boldsymbol{\theta}_p,\mathbf{x})|^{-\frac{1}{2}}$$

$$V_2 = \left[\sum_{p=1}^{P} w_p^2\right] \sum_{a=1}^{P}\sum_{b=1}^{P} w_a w_b G(\boldsymbol{\nu}_{ab},\boldsymbol{\Phi}_{ab})$$

$$V_3 = \sum_{a=1}^{P}\sum_{b=1}^{P} w_a^2 w_b G(\boldsymbol{\nu}_{ab},\boldsymbol{\Phi}_{ab})$$

where:

$$\begin{aligned}\boldsymbol{\nu}_{ab} &= \boldsymbol{\nu}(\boldsymbol{\theta}_a,\mathbf{x}) - \boldsymbol{\nu}(\boldsymbol{\theta}_b,\mathbf{x}) \\ \boldsymbol{\Phi}_{ab} &= \boldsymbol{\Phi}(\boldsymbol{\theta}_a,\mathbf{x}) + \boldsymbol{\Phi}(\boldsymbol{\theta}_b,\mathbf{x}) \\ G(\mathbf{r},\mathbf{R}) &= (2\pi)^{-0.5d_{\mathbf{y}}}|\mathbf{R}|^{-0.5}\exp(-0.5\mathbf{r}^T\mathbf{R}^{-1}\mathbf{r})\end{aligned}$$

and $\boldsymbol{\nu}(\boldsymbol{\theta}_a,\mathbf{x})$ and $\boldsymbol{\Phi}(\boldsymbol{\theta}_a,\mathbf{x})$ are the GP predictive mean and variance at particle $\boldsymbol{\theta}_a$ and action $\mathbf{x}$. At each time step t, we find the optimal action $\mathbf{x}_t^* = \text{argmax}_{\mathbf{x}}\mathbf{I}_Q(\boldsymbol{\theta},\mathbf{y}|\mathbf{x})$ by performing gradient ascent on the QMI, with the gradient defined as:

$$\nabla_{\mathbf{x}}V_1 = \sum_{p=1}^{P} w_p^2 (4\pi)^{-\frac{d_{\mathbf{y}}}{2}} |\boldsymbol{\Phi}_p|^{-\frac{1}{2}} \left[-\frac{1}{2}\left(\boldsymbol{\Phi}_p^{-1}\frac{\partial \boldsymbol{\Phi}_p}{\partial \mathbf{x}}\right)\right]$$

$$\nabla_{\mathbf{x}}V_2 = \left[\sum_{p=1}^{P} w_p^2\right] \sum_{a=1}^{P}\sum_{b=1}^{P} w_a w_b \frac{\partial G(\boldsymbol{\nu}_{ab},\boldsymbol{\Phi}_{ab})}{\partial \mathbf{x}}$$

$$\nabla_{\mathbf{x}}V_3 = \sum_{p=1}^{P}\sum_{a=1}^{P} w_p^2 w_a \frac{\partial G(\boldsymbol{\nu}_{pa},\boldsymbol{\Phi}_{pa})}{\partial \mathbf{x}}$$

where the derivatives of $G(\boldsymbol{\nu}_{ab},\boldsymbol{\Phi}_{ab})$ and $\boldsymbol{\Phi}_p$ are straightforward to derive.

3.3 COMPLEXITY

During run time, the analytical Gaussian update is $O(d_{\mathbf{y}}^3) + O((d_{\boldsymbol{\theta}} + d_{\mathbf{x}})d_{\mathbf{y}}^2 N^2)$ at each time step. The additional computational burden due to the information maximization in the active case is linear in the number of steps the gradient ascent takes.

For MC sampling-based belief updating, the computational complexity of the active update at each time step during test time will be $O(d_{\mathbf{y}}P^2N)$. As the complexity is linear in P in the passive case, this represents significantly higher computational requirements. However, by using QMI we avoid having to evaluate integrals over $\mathbf{y}$ using sampling, as done in other approaches (Denzler & Brown, 2002).

It is worth noting, that during optimization each successive step along the gradient improves the informativeness of the selected action. Therefore the optimization does not have to be run until convergence for time-critical applications, while still providing benefits compared to the passive case.

4 EXPERIMENTAL RESULTS

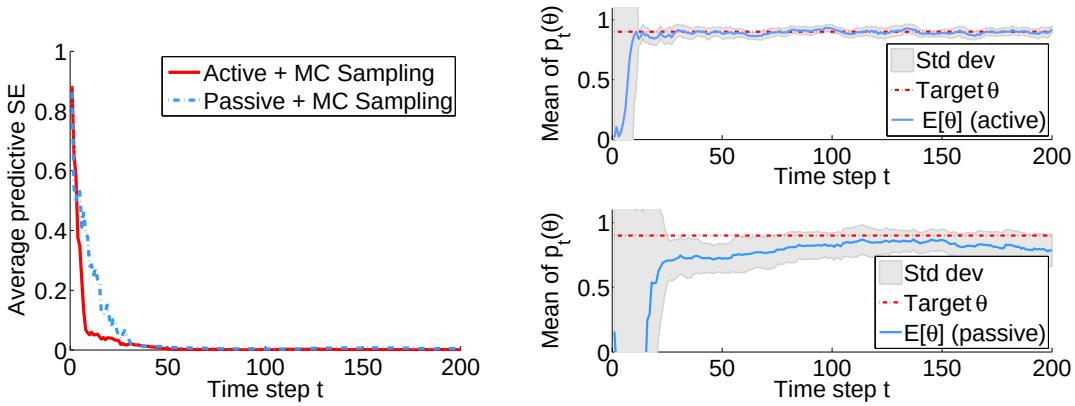
4.1 SYNTHETIC DATA

We applied the active Gaussian and MC sampling-based belief updates in Sections 3.2.1 and 3.2.2 to

[2]Given a particular action $\mathbf{x}$ and a Gaussian $p(\mathbf{y}|\boldsymbol{\theta},\mathbf{x})$, the marginal $p(\mathbf{y}|\mathbf{x})$ is a mixture of Gaussians. Each product of Gaussians gives an unnormalized Gaussian, which can then be integrated over $\mathbf{y}$.

Table 1: Converged average predictive squared error values after 200 time steps for active and non-active belief updating approaches, averaged over 10 trials, for different observation models (functions i–iii).

	Function i		Function ii		Function iii	
Observation noise	low	high	low	high	low	high
Gaussian belief updates with random $\mathbf{x}$	0.0001	0.0017	5.0304	15.94	-	-
Active Gaussian belief updates	0.0001	0.0005	0.0018	10.27	-	-
MC sampling-based belief updates with random $\mathbf{x}$	0.0004	0.0018	0.0013	0.0062	0.0035	0.0833
Active MC sampling-based belief updates	0.0002	0.0019	0.0006	0.0011	0.0032	0.0594



(a) Average predictive SE

(b) Estimated $p_t(\boldsymbol{\theta})$

Figure 1: a) Average predictive squared error (SE) using posterior mean of $\boldsymbol{\theta}$ over all time steps, averaged over 10 trials, for function ii (with high observation noise); b) Estimated posterior $p_t(\boldsymbol{\theta})$ over all time steps t for a sample trial with (active and passive) MC sampling-based belief updates for function ii (with high observation noise).

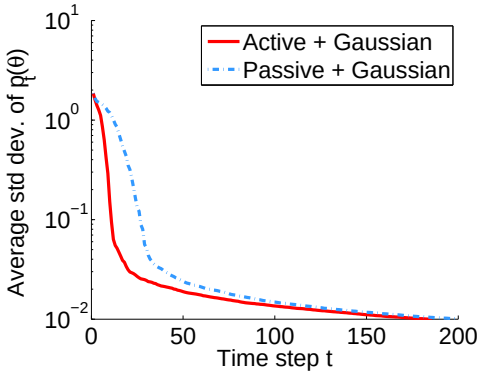
synthetic data sets, for the purpose of estimating a particular value of $\boldsymbol{\theta}^*$ at runtime. We compared them with their passive equivalents.

Table 1 lists the four different frameworks evaluated: a) Gaussian belief updates with random actions, b) active Gaussian belief updates, c) MC sampling-based belief updates with random actions, and d) active MC sampling-based belief updates. We considered the following three observation models/functions: i) $y = \exp(-(\theta - x)^2)$; ii) $y = \cos(\theta)\sin(\theta x)$; iii) $y = \exp(-(\theta_1 - x_1)^2) + \exp(-0.5(\theta_2 - x_2)^2)$. Functions i and ii both assume scalar θ and x, while in function iii, both $\boldsymbol{\theta}$ and $\mathbf{x}$ are 2-dimensional. We generated data sets, each having 1000 samples, from functions i-iii with either a small amount or a larger amount of observation noise[3]. The "high" and "low" headers in Table 1 indicate scenarios with high and low observation noise, respectively, for functions i-iii. We optimized the location of $M = 500$ pseudo-inputs during training to learn the observation model over $\mathbf{y}$. The MC sampling-based belief updates used a population of $P = 500$ samples, with the active version using a downsampled population of 250 samples during optimization of $\mathbf{x}_t^*$ to increase the speed of simulations.

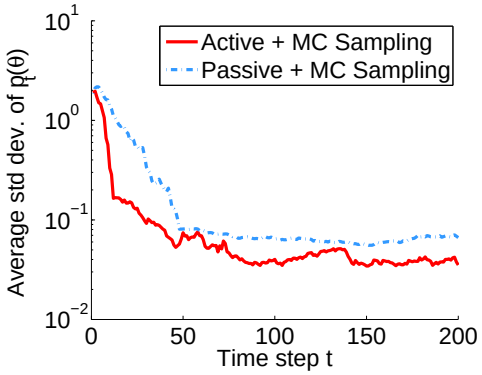
Table 1 shows that the converged average predictive squared error (MSE) (using the posterior mean of $\boldsymbol{\theta}$) after 200 time steps for the four frameworks, evaluated on functions i-iii. Results are averaged over 10 trials for each framework-observation model combination. Function i appears to be a relatively easy data set (with all methods attaining low predictive errors), while functions ii and iii appear to be a little more interesting. Function ii is highly nonlinear—with multiple peaks—in $\boldsymbol{\theta}$. In this scenario, analytical Gaussian belief updating fails to capture the multimodal distribution of $p_t(\boldsymbol{\theta})$, with the occasional failure (i.e., convergence to an incorrect $\boldsymbol{\theta}^*$) inflating the MSE values[4]. For this reason, we omit the pre-

[3]We parameterize noise with the coefficient of determination, $r^2 = (\sigma_y^2 - \sigma_{res}^2)/\sigma_y^2$, where σ_{res}^2 is the variance of the residual error. We added noise scaled to the variance of the noiseless $\bar{\mathbf{y}}$ such that $\sigma_{noise}^2 = c\sigma_{\bar{\mathbf{y}}}^2$, where $c = 1/r^2 - 1$. We set $r^2 = 0.995$ and $r^2 = 0.95$ for the low and higher noise cases, respectively.

[4]Even though reported results are averaged over 10



(a) Gaussian belief updates



(b) MC sampling-based belief updates

Figure 2: Average standard deviations of posterior $p_t(\boldsymbol{\theta})$ over all time steps t, averaged over 10 trials. Figures 2(a) and 2(b) are for functions i and ii, respectively (both with high observation noise).

dictive error values for Gaussian belief updating on function iii since they are high for similar reasons.

Figure 1(a) shows the predictive squared error in $\boldsymbol{\theta}$ for a particular trial with MC sampling-based belief updates and function ii as an observation model. Figure 1(b) shows the corresponding estimated posterior $p_t(\boldsymbol{\theta})$ over all time steps t. Plots for analytical Gaussian belief updating with function ii were omitted due to high predictive errors. Similar trends in the speed of convergence can be observed for active and passive MC sampling belief updates on functions i and ii. These were omitted due to lack of space.

Figure 2 shows the standard deviation of the posterior $p_t(\boldsymbol{\theta})$, averaged over all 10 trials, for both Gaussian and MC sampling-based belief updates. Figure 2(a) shows the average posterior standard deviations for analytical Gaussian belief updates on function i (for both active and passive cases). MC sampling-based belief updates produce similar curves that were omitted from the plot for clarity. Figure 2(b) shows the average posterior standard deviations for function ii, with results from the Gaussian belief updates omitted since they failed. These figures demonstrate that adding an active component to sequential data selection/collection appears to speed up convergence to $\boldsymbol{\theta}^*$.

trials, the range of average predictive error values hold for Gaussian belief updates when averaged over 100 trials.

4.2 ESTIMATION OF VISCOSITY FROM TACTILE FEEDBACK

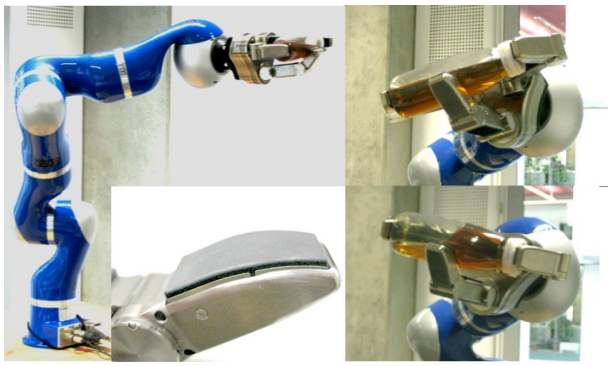


Figure 3: Schunk SDH 7-DOF hand on a DLR 7-DOF arm. $\mathbf{y}$ is observed tactile data from sensors on fingers, $\mathbf{x}$ describes the arm's actions (e.g., joint angles, shaking frequency), and $\boldsymbol{\theta}$ is the viscosity of the bottle's liquid.

We explored the problem of viscosity estimation from tactile data using a robotic anthropomorphic system as shown in Figure 3. This consists of a 7 degree-of-freedom robot arm with a 7 degree-of-freedom three-fingered hand attached. The fingers are equipped with 6 tactile arrays containing 486 texels in total

The robot's task was to determine the viscosity of liquids $\boldsymbol{\theta}$ in bottles by shaking the containers at different frequencies $\mathbf{x}$. Observations $\mathbf{y}$ come from tactile arrays mounted on the robot's fingers. The shaking frequency is either selected at random (passive case) or maximized to be most informative with respect to the current belief of $\boldsymbol{\theta}$ (active case).

For gathering the training data, we took bottles containing three different liquids and recorded the tactile responses while shaking the bottles at a range of frequencies (from 0.3 to 1.1 Hz) for 5 seconds each. The three liquids had viscosities of 1 cst (water), 130 cst (motor oil), and 1300 cst (glycerine). These values were transformed to log_{10} space, yielding values of 0, 2.07 and 3.07, respectively. The bottles used for the three liquids were identical, and the content was matched for weight. A standard initial grip was

used for all bottles.

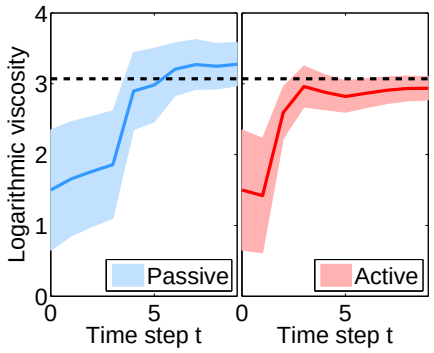
To deal with the high-dimensional $\mathbf{y}$, we preprocessed the tactile data, applying PCA and retaining the principal component (containing time-varying data). We calculated the Fourier transform of the data, and fitted individual GPs to each Fourier component (23 in total). The resulting model maps the joint space of viscosity and shaking frequency $[\boldsymbol{\theta}\ \mathbf{x}]^T \in \mathbb{R}^2$ to the preprocessed tactile space (23D). We collected 270 training points in total.

At the start of run/test time, we placed an initial broad Gaussian prior over the viscosity $\boldsymbol{\theta}$ and ran the experiment for 10 time steps. In the first step, a shaking frequency of 0.5 Hz was always chosen, as the sensor values had to be normalized to account for slight variations in grip force and grip location. In subsequent steps, shaking frequencies were either selected randomly (passive case), or by gradient ascent on the mutual information (active case). For Gaussian-based updates, optimization for a time step took less than 0.5 seconds, with a negligible impact on the total run time (see Figure 4(a) for results from a sample trial). In contrast, for MC sampling-based updates, optimization for a time step took between 15 and 24 seconds for a set of 500 particles. While computations could be speeded up by parallelization, the MC sampling-based solution remains infeasible for problems requiring near real-time decisions.

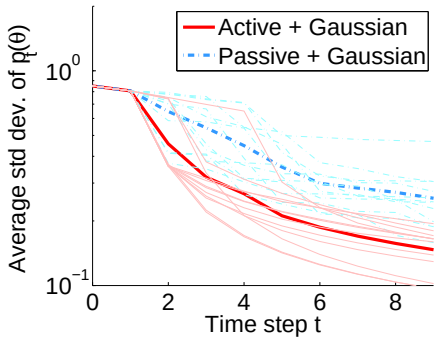
We considered 4 liquids, 3 of which were used in training, with the fourth having a viscosity of 30 cst (1.47 in log space). The fourth liquid was not part of the training set. For each liquid, we performed 3 trials of Gaussian-based updates and 2 trials of sampling-based updates—for both the active and passive cases.

For Gaussian-based updates, the predictive mean squared error—averaged over all four liquids and trials—for the passive case after 10 time steps was 0.37, compared to 0.07 for the active case. Estimation worked equally well for the liquids used in training as well for the newly introduced liquid. Errors for the MC sampling-based approach were generally higher, with 0.42 and 0.34 for the passive and active cases, respectively. The posterior standard deviation after 10 steps for Gaussian updates was 0.26 for the passive case, compared to 0.15 for the active case. Figure 4(b) confirms this result, showing a faster convergence in the active case. MC sampling-based updates show a similar trend, with posterior standard deviations taking a value of 0.31 for the passive and 0.24 for the active case.

The slightly worse performance of sampling-based updates seemed to be mainly due to outlier particles affecting the overall statistics, as well as the low number of particles used. We found the posterior densities to be mostly unimodal, explaining the good performance for the fast Gaussian approach on this task.



(a) Single trials for passive (left) and active (right) Gaussian belief update of viscosity over all time steps, where the liquid tested on was glycerine.



(b) Standard deviation of posterior viscosity over time for passive (blue) and active (red) Gaussian update.

Figure 4: a) Single trial results for $p_t(\boldsymbol{\theta})$ for all t; b) Average standard deviations of posterior viscocity over all time steps t.

5 DISCUSSION

We described an active sequential framework for continuous actions, state parameters and observations, exploring two variants that used either Gaussian or MC sampling updates. We use mutual information to improve the speed of convergence and rely on quadratic mutual information to simplify the optimization of actions in MC sampling updates. We evaluated the proposed active framework on a real robotic system, with the task of determining viscosity of different liquids from tactile feedback under different shaking frequencies. We demonstrated not

only successful estimation of liquid viscosities, but also a faster convergence towards the viscosity estimate (over the passive setting). Thus, with limited added computational burden, a significant speed-up in convergence was achieved. The extension of the active framework to filters (i.e., where the state parameter is time-varying) is straightforward and involves learning a model of the system dynamics.

Deciding between Gaussian and MC sampling updates involves a speed-accuracy trade-off. While the former is fast, it fails to capture multi-modal distributions (as seen by failures on synthetic data sets in Section 4.1). On the other hand, MC sampling updates are computationally restrictive but more powerful. Even though sequential MC sampling may be computationally involved, there do exist scenarios where data collection may be costly (e.g., collecting data with the Mars Rover), such that it is acceptable to wait for an informative data sample. Future work will explore extensions to observations with correlated dimensions and to higher dimensional data.

References

Amari, S., Cichocki, A., & Yang, H. H. (1994). A new learning algorithm for blind source separation. In *Advances in Neural Information Processing Systems* (pp. 757–763). MIT Press.

Cohn, D. A., Ghahramani, Z., & Jordan, M. I. (1996). Active learning with statistical models. *Journal of Artificial Intelligence Research*, *4*, 129–145.

Dagan, I., & Engelson, S. (1995). Committee-based sampling for training probabilistic classifiers. In *International Conference on Machine Learning* (pp. 150–157). Morgan Kaufmann.

Deisenroth, M. P., Huber, M. F., & Hanebeck, U. D. (2009). Analytic moment-based gaussian process filtering. In *International Conference on Machine Learning*.

Denzler, J., & Brown, C. M. (2002). Information theoretic sensor data selection for active object recognition and state estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(2), 145–157.

Doucet, A., Freitas, N. de, & Gordon, N. (2001). *Sequential Monte Carlo methods in practice*. Springer.

Fookes, C., & Maeder, A. J. (2004). Quadrature-based image registration method using mutual information. In *IEEE International Symposium on Biomedical Imaging* (p. 728-731). IEEE.

Girard, A., Rasmussen, C. E., Candela, J. Q., & Murray-Smith, R. (2003). Gaussian process priors with uncertain inputs – application to multiple-step ahead time series forecasting. In *Advances in Neural Information Processing Systems* (p. 545-552). MIT Press.

Guo, Y., & Greiner, R. (2007). Optimistic active learning using mutual information. In *International Joint Conference on Artificial Intelligence* (pp. 823–829). AAAI Press.

Lewi, J., Butera, R. J., & Paninski, L. (2009). Sequential optimal design of neurophysiology experiments. *Neural Computation*, *21*(3), 619–687.

Lewis, D. D., & Gale, W. A. (1994). A sequential algorithm for training text classifiers. In *ACM SIGIR Conference on Research and Development* (pp. 3–12). Springer-Verlag.

Lindley, D. (1956). On a measure of information provided by an experiment. *Annals of Mathematical Statistics*, *29*, 986–1005.

Mackay, D. J. C. (1992). Information-based objective functions for active data selection. *Neural Computation*, *4*, 589–603.

McCallum, A., & Nigam, K. (1998). Employing EM in pool-based active learning for text classification. In *International Conference on Machine Learning* (pp. 359–367). Morgan Kaufmann.

Paninski, L. (2005). Asymptotic theory of information-theoretic experimental design. *Neural Computation*, *17*(7), 1480–1507.

Roy, N., & McCallum, N. (2001). Toward optimal active learning through sampling estimation of error reduction. In *International Conference on Machine Learning* (pp. 441–448). Morgan Kaufmann.

Seung, H. S., Opper, M., & Sompolinsky, H. (1992). Query by committee. In *ACM Workshop on Computational Learning Theory* (pp. 287–294).

Torkkola, K. (2003). Feature extraction by non-parametric mutual information maximization. *Journal of Machine Learning Research*, *3*.

Williams, C. K. I., & Rasmussen, C. E. (1995). Gaussian processes for regression. In *Advances in Neural Information Processing Systems*. MIT Press.

Zhu, X., Lafferty, J., & Ghahramani, Z. (2003). Combining active learning and semi-supervised learning using gaussian fields and harmonic functions. In *ICML Workshop on the Continuum from Labeled to Unlabeled Data* (pp. 58–65).