Variational Information Planning for Sequential Decision Making

Abstract

We consider the setting of sequential decision making where, at each stage, potential actions are evaluated based on expected reduction in posterior uncertainty, given by mutual information (MI). As MI typically lacks a closed form, we propose an approach which maintains variational approximations of, both, the posterior and MI utility. Our planning objective extends an established variational bound on MI to the setting of sequential planning. The result, variational information planning (VIP), is an efficient method for sequential decision making. We further establish convexity of the variational planning objective and, under conditional exponential family approximations, we show that the optimal MI bound arises from a relaxation of the well-known exponential family moment matching property. We demonstrate VIP for sensor selection, experiment design, and active learning, where it meets or exceeds methods requiring more computation, or those specialized to the task.

1 Introduction

Bayesian machine learning research has paid much attention to developing posterior inference algorithms for models where posterior calculation is computationally prohibitive. Little focus, by contrast, has been given to methods for decision making based on the results of inference. Such problems are of paramount importance in reinforcement learning (RL), however the context and details of RL planning differ substantially from our own setting.

In this paper we explore methods for sequential decision making, where action choices drive data collection. We further consider the setting of *information planning*, where decisions are generated by maximizing the mutual information (MI) utility (Williams, 2007).

Preliminary work. Under review by AISTATS 2019. Do not distribute.

The setting most closely resembles Bayesian experiment design (Lindley, 1956), where experiments are chosen to minimize uncertainty over a quantity of interest. Indeed, MI has long been used as a design utility since it is a measure of expected reduction in posterior uncertainty (Blackwell, 1950; Bernardo, 1979).

Unlike experiment design, which typically assumes the cost of a measurement dominates that of inference, our focus is on high throughput sequential decision systems. Where the former relies on Monte Carlo inference and MI estimates during planning, we present a comprehensive approach to inference and planning based on efficient variational approximations.

Our approach, which we call variational information planning (VIP), maintains a series of variational approximations to the posterior and MI utility. For the planning stage, VIP extends a well-known lower bound of MI (Barber and Agakov, 2004) to the sequential setting. During planning, the MI lower bound is optimized over an auxiliary distribution, which approximates the posterior under a hypothesized measurement. We demonstrate that this optimization is convex for a certain class of auxiliary models. We establish optimality conditions for the natural parameters of this family, and show that they are a relaxation of the well known moment matching conditions.

A core challenge we face is that variational approximations of posterior uncertainty can be arbitrarily poor (Giordano et al., 2015; Turner and Sahani, 2011), despite their good predictive accuracy. Since MI is a measure of uncertainty, a naive variational approximation will tend to yield poor decisions. To address these issues, our auxiliary distributions belong to a set that, conditioned on a hypothesized observation, are in the exponential family. This set enables nonlinear dependence on the conditioning variable, and is thus strictly larger than the set of joint exponential family distributions.

In our experiments we demonstrate that VIP is sufficiently flexible to apply in a variety of problem instances such as nonlinear target tracking in a sensor network, experiment design, and active learning. Moreover, VIP meets or exceeds the accuracy of methods based on exact inference, MCMC requiring more computation, or specialized variational approximations.

2 Sequential Information Planning

Consider a model of latent variables x and observations $\mathcal{Y}_{t-1} = \{y_1, \ldots, y_{t-1}\}$. At each time $t-1$ a discrete *action* $a_{t-1} \in \{1, \ldots, A\}$ parameterizes the likelihood, denoted $p_{a_{t-1}}(y_{t-1} \mid x)$. Let $\mathcal{D}_{t-1} = \{\mathcal{Y}_{t-1}, \mathcal{A}_{t-1}\}$ be the set of observations and chosen actions $\mathcal{A}_{t-1} = \{a_1, \ldots, a_{t-1}\}$ at time $t-1$. The posterior is then,

$$p(x \mid \mathcal{D}_{t-1}) \propto p(x) \prod_{i=1}^{t-1} p_{a_i}(y_i \mid x) \tag{1}$$

The goal of sequential information planning is to choose the sequence of actions $\mathcal{A}$ that minimize entropy of the posterior (1). Specifically, at time t, an action a_t is selected to maximize the posterior mutual information,

$$\begin{aligned} a_t^* &= \arg\max_a \; I(X; Y_t \mid \mathcal{D}_{t-1}) \\ &= \arg\max_a \; H(X \mid \mathcal{D}_{t-1}) - H_a(X \mid Y_t, \mathcal{D}_{t-1}) \end{aligned} \tag{2}$$

Once an action is selected, new observations are drawn from the appropriate likelihood model $y_t \sim p_{a_t}(\cdot \mid x)$ and the posterior is updated.

Calculating the posterior MI in Eqn. (2) is complicated for two reasons. First, the entropy terms involve expectations under the posterior distribution (1). Second, calculating the conditional entropy $H(X \mid Y, \mathcal{D})$ requires evaluation of the posterior predictive distribution $p(y \mid \mathcal{D})$ as in,

$$H(X \mid Y, \mathcal{D}) = \mathbb{E}\left[-\log \frac{p(x, y \mid \mathcal{D})}{p(y \mid \mathcal{D})}\right],$$

where we have dropped explicit indexing on time. One approach is to estimate this over samples $\{y_t^i\} \sim p_a(y \mid \mathcal{D}_{t-1})$. The resulting empirical plug-in estimator of MI is,

$$\hat{I}_a = -\frac{1}{N} \sum_{i=1}^{N} \log \frac{p_a(y_t^i \mid x^i)}{\frac{1}{M}\sum_{j=1}^{M} p_a(y_t^i \mid x^{ij})}. \tag{3}$$

Independent samples $\{x^{ij}\}_{j=1}^{M} \sim p(x \mid \mathcal{D}_{t-1})$ are required for each action, and observation sample, to ensure estimates are independent, thus increasing sample complexity. While the estimator (3) is consistent, it is biased. Moreover, bias is known to decay slowly (Zheng et al., 2018; Rainforth et al., 2018).

3 Variational Information Planning

Given the challenges associated with sample-based estimates of MI, we shift attention to a more efficient variational approach. We begin by introducing the MI bound at the core of VIP and show how this bound can be extended to sequential decision making. We then demonstrate the calculations necessary to apply VIP in a model where observations are conditionally independent. We conclude by showing how VIP is applied in a model of annotations for static data, a common scenario in the related active learning problem.

3.1 Variational Information Bound

Setting aside the details of sequential planning for the moment, we have that for any valid distribution $\omega(x \mid y)$, the following is a lower bound on MI,

$$I(X; Y) \geq H(X) + \mathbb{E}_p[\log \omega(X \mid Y)]. \tag{4}$$

This well-known bound is the result of Gibbs' inequality, and has been independently explored in various contexts (Barber and Agakov, 2004; Mohamed and Rezende, 2015; Gao et al., 2016; Chen et al., 2018). The dual problem, then, is to maximize this bound w.r.t. $\omega(x \mid y)$, which we refer to as the *auxiliary distribution*.

In the sequential setting, at each planning stage t we require the posterior mutual information, $I(X, Y_t \mid \mathcal{D}_{t-1})$. Direct application of the bound (4) results in,

$$H(X \mid \mathcal{D}_{t-1}) + \mathbb{E}_p[\log \omega(X \mid Y) \mid \mathcal{D}_{t-1}],$$

which involves expectations over the posterior distribution $p(x, y_t \mid \mathcal{D}_{t-1})$. Thus, application of this bound to sequential decision making requires further approximation, which we now discuss.

At a high-level, VIP extends the bound (4) to the sequential setting through a series of approximating distributions. First, given past observations and actions $\mathcal{D}_{t-1} = \{\mathcal{A}_{t-1}, \mathcal{Y}_{t-1}\}$, we form a variational approximation to the posterior $q(x) \approx p(x \mid \mathcal{D}_{t-1})$. Then, we construct a local approximation over the joint distribution $\hat{p}_a(x, y_t) \approx p(x, y_t \mid \mathcal{D}_{t-1})$ for each hypothesized action $a = 1, \ldots, A$. Finally, we perform planning by optimizing a lower bound with respect to an auxiliary distribution $\omega(x \mid y_t)$ and select the maximizing action a_t for the next time step. See Fig. 1 for an illustration.

3.2 Conditionally Independent Observations

We begin with the simple model in Eqn. 1, where observations y are independent conditioned on latent x. Let $\mathcal{D}_{t-1} = \{\mathcal{A}_{t-1}, \mathcal{Y}_{t-1}\}$ be the set of past actions and observations, assume that we have a variational approximation of the posterior $q(x) \approx p(x \mid \mathcal{D}_{t-1})$. We then form a local approximation of the distribution

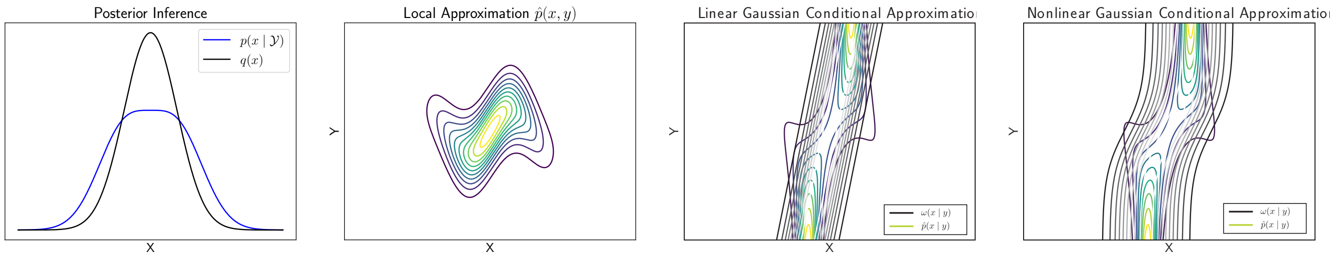

Figure 1: **Distributional approximations.** *Left:* Given observations $\mathcal{Y}$ the posterior is approximated with a tractable family $q(x) \approx p(x \mid \mathcal{Y})$. *Center-Left:* To consider a new observation y, a local approximation is formed $\hat{p}(x, y) = q(x)p(y \mid x)$ using the forward model $p(y \mid x)$. *Center-Right:* VIP optimizes a lower bound on MI w.r.t. a distribution $\omega(x \mid y)$ approximating the conditional $\hat{p}(x \mid y)$. We use a linear Gaussian approximation in this case. *Right:* Directly parameterizing the conditional $\omega(x \mid y)$ allows nonlinear functions of the conditioning variable y, allowing for better approximations and tighter bounds.

over a future measurement at time t,

$$\hat{p}_a(x, y_t) \equiv q_{t-1}(x)p_a(y_t \mid x). \tag{5}$$

Here, $p_a(y_t \mid x)$ is the true likelihood under the hypothesized action a. The distribution $\hat{p}(\cdot)$ is analogous to the *augmented distribution* at each stage of EP inference, which we will exploit in the next section. We can then bound the MI under $\hat{p}(\cdot)$ as,

$$H_{\hat{p}}(X) + \max_{a,\,\omega} \mathbb{E}_{\hat{p}_a}\left[\log \omega(X \mid Y_{t+1})\right]. \tag{6}$$

We have moved the marginal entropy outside the optimization since it is constant in this model, and so can be ignored for planning. The bound can be evaluated in parallel for a discrete set of actions $1, \ldots, A$.

Fig. 1 illustrates the role of each approximation in a single planning stage, and how the approximations relate to the target distributions. To be clear, Eqn. (6) bounds mutual information w.r.t. the local approximate distribution $\hat{p}(\cdot)$, not MI under the true posterior. The conditions under which $\hat{p}(\cdot)$ is a reliable surrogate for posterior MI are the same as those conditions that must hold for variational inference to be effective.

3.3 Annotation Models

The conditionally independent model of the previous section is simple and does not apply in all settings. Consider a different, more complicated distribution, often arising in active learning (Settles, 2012). Specifically, consider a joint distribution over latent x, data $\{z_d\}_{d=1}^D$ which are fixed, and annotations $\{y_d\}_{d=1}^D$,

$$p(x, y, z) = p(x) \prod_{d=1}^{D} p(z_d \mid x)p(y_d \mid z_d).$$

For example, y_d may be a discrete class assignment for an image z_d.

The objective at each learning stage selects the most informative annotation, $\max_d I(X; Y_d \mid \mathcal{D}_{t-1})$, where $\mathcal{D}_{t-1}$ represents the set of data and previous annotations after $t-1$ learning rounds. MI is computed with respect to the distribution,

$$p(x, y_d \mid \mathcal{D}_{t-1}) \propto p(x \mid \mathcal{D}_{t-1} \setminus \{z_d\})p(z_d \mid x)p(y_d \mid z_d).$$

Here, $p(x \mid \mathcal{D}_{t-1} \setminus \{z_d\})$ represents the posterior distribution after removing z_d from the data. This step is algebraic, but intuitively avoids double counting z_d.

To form a local approximation we appeal to our connection with the EP augmented distribution. We assume an EP-like posterior approximation which is a product of factor approximations (messages): $q(x) \propto \prod_{d=1}^{D} \psi_d(x)$. The *cavity distribution* $q^{\setminus d}(x) \propto q(x)/\psi_d(x)$ expresses the posterior approximation having removed z_d. Our local approximation is then,

$$\hat{p}(x, y_d) \propto q^{\setminus d}(x)p(z_d \mid x)p(y_d \mid z_d). \tag{7}$$

The MI lower bound is then identical to (6). More complicated models with nuisance variables that must be integrated out for planning can be handled in the same manner. We consider such a setting for labeled LDA active learning example in Sec. 6.3.

4 Optimization for Conditional Exponential Families

During planning we optimize the bound (6), with respect to the auxiliary distribution $\omega(x \mid y)$. This optimization can be complicated for arbitrary distributions. In this section we consider the optimization for the class auxiliary distribution which are in the exponential family, when conditioned on a hypothesized measurement. This flexible family allows for nonlinear dependence on the observation y as illustrated in Fig. 1 (right). We show that the resulting optimization is convex in the exponential family natural parameters. Also, optimality conditions yield a relaxation of the moment matching property for exponential families.

4.1 Optimizing the Auxiliary Distribution

Consider the set of conditional distributions in the exponential family $\omega \in \mathcal{W}$ with PDF,

$$\omega_\theta(x \mid y) = h(x) \exp\left(\theta(y)^T \phi(x,y) - A(\theta(y))\right), \quad (8)$$

with natural parameters $\theta(y)$ a function of the conditioning variable, sufficient statistics $\phi(x,y)$, base measure $h(x)$ and log-partition function $A(\theta(y))$. Optimizing the bound in Eqn. (6) is equivalent to minimizing the cross entropy,

$$\theta^*(y) = \arg\min_\theta \; J(\theta) \equiv \mathbb{E}_{\hat{p}}[-\log \omega_\theta(x \mid y)]. \quad (9)$$

Convexity of $J(\theta)$ can be established by explicit calculation of the Hessian. Alternatively, by adding $-H(\hat{p})$ we have the following problem, which is equivalent to $J(\theta)$ up to constant terms,

$$\theta^*(y) = \arg\min_\theta \; \mathbb{E}_{\hat{p}_y}\left[\mathrm{KL}(\hat{p}_{x|y} \,\|\, \omega_\theta)\right] \quad (10)$$

For brevity we have introduced the shorthand $\hat{p}_{x|y} \equiv \hat{p}(x \mid y)$. For any realization $Y = y$ the KL term is convex in $\theta(y)$, a well known property of the exponential families (Wainwright and Jordan, 2003). Eqn. (10) is then a convex combination of convex functions, thus convexity holds.

The optimal parameter function $\theta^*(y)$ is given by the stationary point condition,

$$\mathbb{E}_{\hat{p}_y}\left[\mathbb{E}_{\omega_{\theta^*}}[\phi(x,y) \mid Y = y]\right] = \mathbb{E}_{\hat{p}}[\phi(x,y)]. \quad (11)$$

This is a weaker condition than the standard moment matching property of exponential families, which typically minimizes KL. Under (11) moments of $\omega(x \mid y)$ must match *in expectation* w.r.t. the marginal distribution $p(y)$, but need not be equal for any particular realization $Y = y$.

4.2 Parameter Function Optimization

Stationary conditions (11) are in terms of a function $\theta(y)$ which is assumed to be parametric. Let η be parameters of the function, denoted $\theta_\eta(y)$. Stationary conditions in terms of parameters η are then,

$$\mathbb{E}_{p_y}\left[(D_\eta\theta)^T \mathbb{E}_{\omega_\eta}[\phi(x,y)]\right] = \mathbb{E}_{p_y}\left[(D_\eta\theta)^T \mathbb{E}_{p_{x|y}}[\theta(x,y)]\right]$$

where $D_\eta\theta$ is the Jacobian matrix of partial derivatives. If $\theta(y)$ is convex in the parameters η then the optimization Eqn. (10) remains convex.

5 Evaluating the MI Bound

The previous section characterized natural parameters maximizing the MI bound (6) w.r.t. the auxiliary distribution. Planning, however, requires the value of this bound at its optimum. For some models this evaluation is straightforward, but others require estimation. We begin with a discussion of computing the bound for complex models. We conclude with a class of models for which evaluation is simple, and corresponds to the standard moment matching property.

5.1 Bound Estimation

To simplify the discussion, we focus on the conditionally independent model with PDF (1). Recall the local approximation $\hat{p}(x,y) = q(x)p(y \mid x)$, where we drop explicit time indexing for brevity. The relevant term in the bound (6) is the conditional cross entropy,

$$\mathbb{E}_{\hat{p}}[-\log \omega(x \mid y)] \approx -\frac{1}{N}\sum_{i=1}^N \mathbb{E}_{\hat{p}_{y|x^i}}[\log \omega(x^i \mid y)]$$

where samples $\{x^i\}_{i=1}^N \sim q(x)$. Since $q(x)$ is a tractable distribution, this step can be done efficiently.

The expectation $\mathbb{E}_{y|x^i}[\cdot]$ is with respect to the forward model (likelihood), and can often be computed in closed-form. For some models, however, this term must be approximated, and requires simulation of the forward model. This step is also efficient, assuming a Bayesian network, but leads to a higher variance estimate. Both estimators are consistent by the LLN.

5.2 Moment Matching Solution

Under some conditions the value of the MI bound (6) takes a simple form at its optimum. To see this, we first establish that standard moment matching of the auxiliary distribution is optimal for some models. We then show that the value of the bound at this moment matching solution is equivalent to calculating entropy under the auxiliary distribution, which is tractable.

One class of models for which the bound is easily calculated are those where the marginal $\hat{p}(y)$ is in the exponential family, for example if y is a discrete label. Then, consider the following joint exponential family,

$$\omega_\eta(x,y) = h(x,y) \exp\left\{\eta^T \phi(x,y) - A(\eta)\right\}.$$

Furthermore, consider the parameters η^* satisfying the moment matching property,

$$\mathbb{E}_{\hat{p}}[\phi(x,y)] = \mathbb{E}_{\omega_\eta}[\phi(x,y)]. \quad (12)$$

Moment matching, combined with the assumption that $\hat{p}(y)$ is in the exponential family, implies that the marginal can be exactly calculated $\omega_\eta(y) = \hat{p}(y)$. Using this equivalence, and rewriting (12), we have:

$$\mathbb{E}_{\hat{p}}[\phi(x, y)] = \mathbb{E}_{\hat{p}_y}[\mathbb{E}_{\omega_{x|y}}[\phi(x, y) \mid Y = y]], \tag{13}$$

where $\omega_{\eta^*}(x \mid y) = \omega_{\eta^*}(x, y) \div \int \omega_{\eta^*}(x, y)\mathrm{d}x$. Eqn. (13) is the optimality condition (11) of the MI lower bound. This establishes that standard moment matching is optimal for the class of models where $\hat{p}(y)$ is in the exponential family.

We now establish that the moment matching solution η^* leads to a simple form of the bound (6). By direct calculation, the cross entropy $H_{\hat{p}}(\omega_{\eta^*}(x, y))$ equals,

$$\mathbb{E}_{\hat{p}}[-\log h(x, y)] - \eta^T \mathbb{E}_{\omega_{\eta^*}}[\phi(x, y)] + A(\eta). \tag{14}$$

For distributions with constant base measure $h(x, y)$ we have that, $H_{\hat{p}}(\omega_{\eta^*}(x, y)) = H(\omega_{\eta^*}(x, y))$. By similar logic for the marginal entropy, and by applying the entropy chain rule, we have that:

$$H_{\hat{p}}(\omega_{\eta^*}(x \mid y)) = H(\omega_{\eta^*}(x, y)) - H(\omega_{\eta^*}(y)). \tag{15}$$

The l.h.s. is the relevant conditional entropy term from the MI bound (6). The r.h.s. is the entropy of the joint and marginal distributions $\omega(\cdot)$ at the optimal parameters, which is closed form. We have thus shown that the aforementioned MI approximation is equivalent to optimizing the variational lower bound.

6 Experimental Results

We demonstrate VIP in a variety of contexts including sensor selection, Bayesian experiment design, and active learning. The comparison in each setting is to MCMC inference, or exact numerical inference when possible, along with empirical mean estimation of the MI for planning. In this way, our motivation is to demonstrate comparable accuracy using more efficient variational methods. For the more complex case of LLDA we in fact observe sustained improvements over baseline.

6.1 Sensor Selection

We begin with estimating target position in a network of sensors, each with fixed position. Due to communication constraints we can draw measurements from only a single sensor at each time. At each planning stage we must draw measurements from the most informative sensor.

Static Estimation. A stationary target has position drawn from a Gaussian prior $x \sim \mathcal{N}(m, \sigma^2)$. Observations are drawn from one of K sensors, each with fixed position l_k. Sensor noise is modeled as a two-component Gaussian mixture model,

$$y \mid x; k \sim w * \mathcal{N}(0, v_0) + (1 - w) * \mathcal{N}(x, v_k(x)).$$

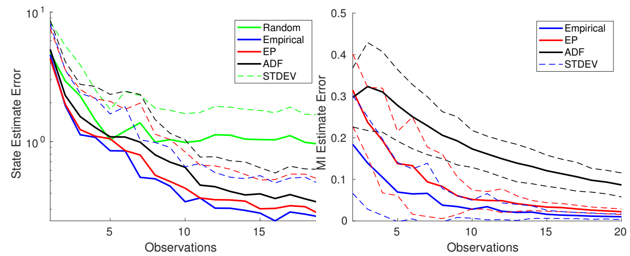

Figure 2: **Static target estimation.** Estimation of a 1D target position in a network of $K = 10$ equally spaced sensors. Mean (solid) plus STDEV (dashed) over 20 realizations. *Left:* Variational planning based on EP inference yields comparable error in state estimate compared to exact inference with empirical MI estimates. ADF inference yields higher error. *Right:* Empirical estimates based on the numerical posterior most accurately estimate MI. VIP bound gap is consistently lower for more accurate EP posterior estimates as compared to ADF.

The mixture consists of a noise distribution with fixed variance v_0 and an observation model with noise variance increasing with relative distance: $v_k(x) = |l_k - x| + v_1$.

Dynamical System. We extend the setting above to a dynamical system with nonlinear dynamics $p(x_t \mid x_{t-1})$ frequently used in the sequential Monte Carlo literature (Kitagawa, 1996; Gordon et al., 1993; Cappé et al., 2007),

$$\mathcal{N}\left(0.5x_{t-1} + 25x_{t-1}/(1 + x_{t-1}^2) + 8\cos(1.2t), \sigma_u^2\right).$$

To keep consistent with prior work we model observations as, $y_t \mid x_t; k \sim \mathcal{N}(x_t^2/20, v_{tk}(x_t))$. The variance function $v_{tk}(x_t)$ is identical to our static example. To demonstrate flexibility additionally replace variational inference with a particle filter. Fig. 3 demonstrates an example scenario.

6.1.1 Results

In both models, static and dynamic, we optimize the MI lower bound over a linear Gaussian approximation $\omega(x \mid y_t)$, which can be solved in closed-form. We compare the impact of inference on predictive accuracy by comparing exact numerical calculation, variational inference, and MCMC.

State predictions competitive with empirical. In both cases VIP produces state estimates with similar or better accuracy to empirical planning, depending on the chosen posterior approximation. In the static estimation model we find similar accuracy between exact numerical inference with empirical planning and EP inference with VIP planning (Fig. 2; left). In the tracking model we also consider exact inference with VIP planning, for which median error is lowest. When comparing particle filter inference VIP and em-

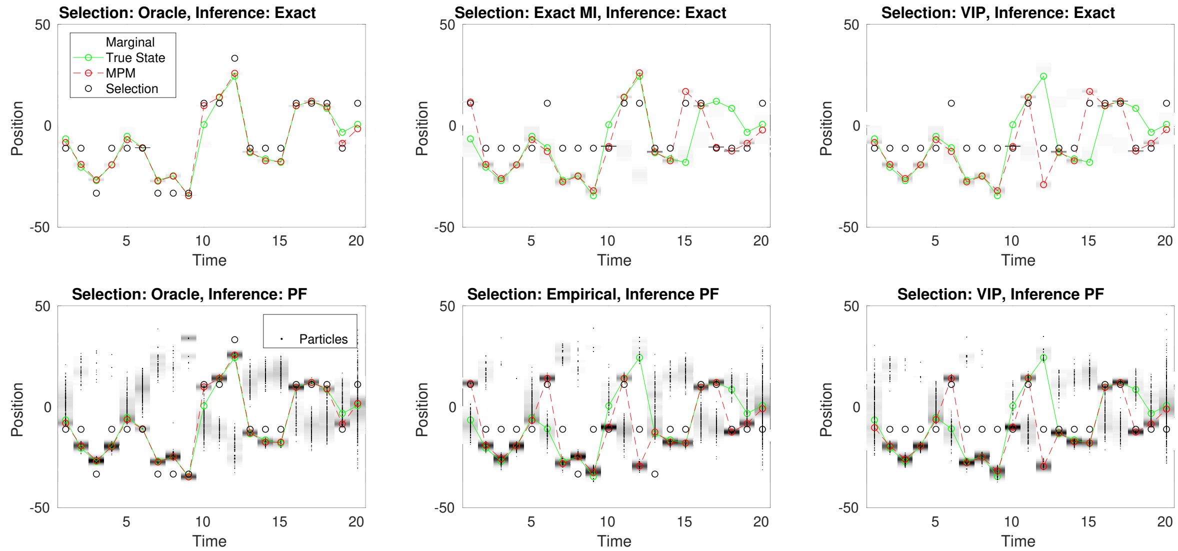

Figure 3: **Nonlinear tracking example** in a field of $K = 10$ equally spaced, stationary, sensors. The best comparison is to numerical approximation to MI and the posterior distribution (*top-center*). For reference, we have also included an *oracle* which selects the sensor closest to the true target location (*left column*). In typical cases such as this one, we see that VIP state error is comparable to empirical estimation under the same posterior approximation. However, VIP shows lower accuracy when planning is computed against an approximate posterior, in this case particle filtering (PF).

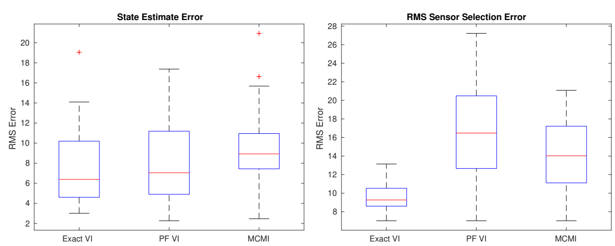

Figure 4: **Nonlinear tracking for 20 random trials.** *Left:* Exact inference with variational planning yields the lowest RMS state error. Particle filter inference with 500 particles and variational planning (PFVI) yields lower median error compared to MC estimates of information (MCMI), though wider error quantiles. *Right:* Again, Exact VI shows the lowest RMS error of the selected sensor position w.r.t. optimal, whereas PFVI and MCMI both have higher error, again with PFVI having larger quantiles.

pirical planning accuracy are comparable, with the former showing slightly lower median error (Fig. 4; left).

Planning is sensitive to posterior accuracy. We also find that estimates of the state estimate, and of the MI calculation, are sensitive to accuracy of the posterior approximation. In the static case we compare against assumed density filtering (ADF), which is more numerically stable than EP but tends to produce less accurate posterior approximations. Planning based on the ADF posterior produces less accurate state estimates (Fig. 2; left) and higher error MI estimates (Fig. 2; right). Similarly, use of particle filter inference in the tracking model increases error for both the state (Fig. 4; left) and MI calculation (Fig. 4; right).

6.2 Gene Regulatory Network

Next, we demonstrate VIP for Bayesian experimental design in the discovery of causal gene interactions. Specifically, we develop our approach for the sparse linear model of Steinke et al. (2007); Seeger (2008).

Let $x \in \mathbb{R}^n$ represent the deviation of n gene expression levels from steady state. The matrix $A \in \mathbb{R}^{n \times n}$ represents causal interactions, with sparse entries drawn independently from a Laplace distribution. The model is,

$$p(A, x) \propto \prod_{i=1}^{n} \mathcal{N}(u_i \mid a_i^T x, \sigma^2) \prod_{j=1}^{n} \text{Laplace}(a_{ij} \mid \lambda) \quad (16)$$

where a_i is the i^{th} row of matrix A. Here u represents an external control vector (perturbation). Interventions include up regulating $u_i > 0$, down regulation $u_i < 0$ and no intervention $u_i = 0$ for the i^{th} gene. Observe that the joint (16) is defined up to normalization since the likelihood term is not a normalized distribution of x.

As exact inference is infeasible, we instead compare to the method developed in Steinke et al. (2007); Seeger (2008), which maintains a mean field Gaussian posterior approximation using EP, $q(A) = \prod_i p_i^{(0)}(a_i) \prod_j \widetilde{t}_{ij}(a_{ij})$. Here, $p_i^{(0)}(a_i)$ approximates the base measure $N(u_i \mid a_i^T x, \sigma^2)$ and $\widetilde{t}_{ij}(a_{ij})$ the Laplace factors.

Given past observations $\mathcal{D}$ and a new control and observation pair $\{x_*, u_*\}$, maximizing MI is (approximately) equivalent to maximizing $\mathbb{E}_{x_*}[\text{KL}(q' \,\|\, q)]$,

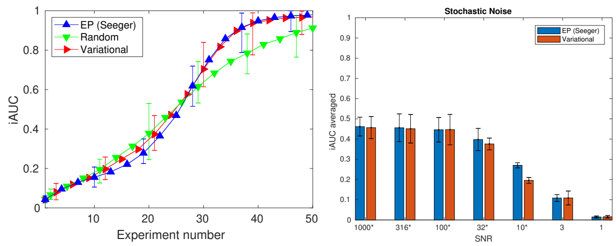

Figure 5: **Sparse linear model.** Average +/- STDEV computed for $n = 50$ nodes over 20 random networks. *Left:* AUC for edges with true weights $|a_{ij}| > 0.1$ plotted for each intervention experiment. VIP shows similar results to Steinke and Seeger with slight improvement for less than 25 experiments. *Right:* Average AUC at varying noise levels are broadly similar to Steinke and Seeger. The dip at moderate SNR levels may be due to our choice of a simple approximating family (linear Gaussian).

where approximation comes from the update posterior $q'(A) \approx p(A \mid \mathcal{D} \cup \{x_*, u_*\})$. The previous authors approximate expectation over samples $\{x_*^k\}$ from the predictive distribution $p(x_* \mid \mathcal{D}, u_*)$. Since this approach requires updating the EP posterior $q'(A)$ for each sample x_*^k, which is prohibitive, they instead propose a non-iterative approximation which only updates the Gaussian base measure $p^{(0)}(A)$ at each sample.

Similar results to specialized method. We optimize the MI bound over a linear Gaussian approximation for 20 random trials and report area under the curve (AUC) for edge prediction (Fig. 5; left). Steinke and Seeger approximate MI by updating only the base measure $p^{(0)}$, which involves a moment matching projection $\mathbb{E}_{q'}[\phi(A)] = \mathbb{E}_{\hat{p}}[\phi(A)]$, where $\hat{p}(A)$ is the augmented distribution, which is similar to the moment matching solution discussed in Sec. 5.2. As a result, VIP reduces to a solution similar to that of Steinke and Seeger, and results are comparable for more than 30 interventions. Results remain similar across varying noise levels (Fig. 5; right), albeit with a slight drop in accuracy at moderate SNR. We hypothesize that this intermediate region depends more strongly on good MI approximations, and that VIP would benefit from a more flexible approximating distribution than the linear Gaussian one chosen.

Improved AUC for fewer interventions. Despite similarity in the methods, we do observe small improvements at early interventions. In particular, Steinke and Seeger observe that their proposed experimental design approach performs poorly for few interventions and propose a hybrid method which randomly selects the first 20 interventions, then performs information guided selection thereafter. Random selection still performs well in this regime.

6.3 Active Learning for Labeled LDA

Labeled LDA (LLDA) is one of several proposed semi-supervised extensions to LDA (Blei et al., 2003). The standard unsupervised LDA model is given by,

$$\begin{aligned} \theta_d &\sim \text{Dirichlet}(\alpha), && \text{For } d = 1, \ldots, D \\ \psi_k &\sim \text{Dirichlet}(\beta_k), && \text{For } k = 1, \ldots, K \\ z_{dn} \mid \theta_d &\sim \text{Cat}(\theta_d), && \text{For } n = 1, \ldots, N_d \\ w_{dn} \mid z_{dn}, \psi &\sim \text{Cat}(\psi_{z_{dn}}) \end{aligned}$$

LLDA augments the standard LDA model with semi-supervised annotations for each word (Flaherty et al., 2005). We model annotations as noisy observations of the true topic assignment: $y_{dn} \mid z_{dn} \sim \text{Cat}(\pi_{z_{dn}})$. In this way, annotations induce a preferred ordering of topic labels in the posterior, producing interpretable topics. There are variations on this approach (Ramage et al., 2009) which model annotations at the document level, but we do not consider them here.

We perform active learning, which falls under the annotation model of Sec. 3.3. At each learning stage t, the planner selects an annotation $y_{d^*n^*}$ maximizing MI over the topics,

$$(d^*, n^*) = \underset{(d,n)}{\arg\max}\; I(\Psi, Y_{dn} \mid \mathcal{W}, \mathcal{Y}_{t-1}), \tag{17}$$

given the observed corpus of words $\mathcal{W}$ and previous annotations $\mathcal{Y}_{t-1}$. We demonstrate active learning on the "bars" data (Griffiths and Steyvers, 2004) in which topics can be arranged and visualized as a set of vertical and horizontal *bars* (see Fig. 6). To amplify the effect of poorly informed selections we additionally model a *rare* topic by setting a non-symmetric Dirichlet prior on topic proportions: $\alpha = (0.05, 1, 1, \ldots, 1)^T$.

Empirical MI estimates based on Gibbs samples involve averages across independent MCMC chains, and are thus sensitive to topic label switching (Stephens, 2000). We found that alignment to any individual sample (e.g. MAP) produced poor results, both in terms of the posterior mean estimate of the topics as well as entropy estimates. We instead relabel Gibbs samples to minimize absolute error w.r.t. the ground truth topics. Such an approach would not be feasible in typical cases as it requires knowledge of the true topics, but represents the best achievable estimate over the given samples. Our variational planning approach does not suffer similar issues as samples are drawn from a consistent topic posterior distribution.

Lower topic error compared to Monte Carlo. Fig. 7 reports total variation (TV) error across topics $\sum_k \|\psi_k - \hat{\psi}_k\|_1$, based on the posterior mean estimate $\hat{\psi}$. The TV error is invariant to topic ordering as we solve a bipartite matching to compute the lowest TV

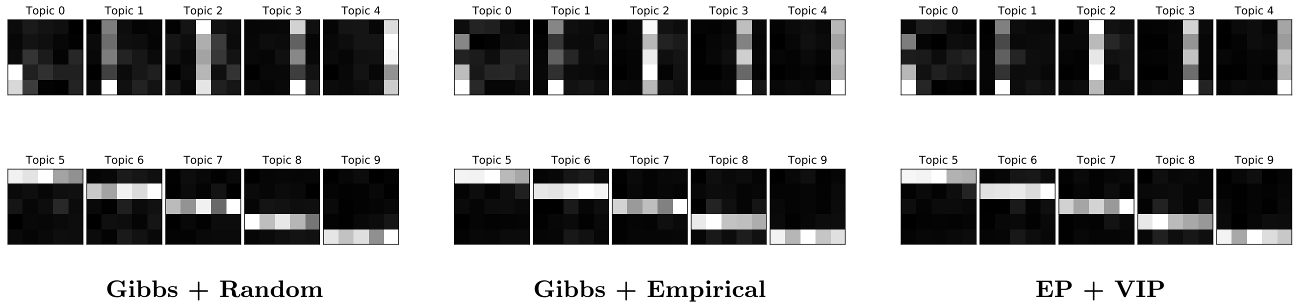

Figure 6: **Learned LLDA topics** from a corpus of $D = 50$ documents, each with $N_d = 25$ words drawn from the bars topics with a $W = 25$ word vocabulary. We model Topic 0 as a *rare* topic (see text). Gibbs estimates are averaged over 1000 samples drawn from parallel chains. Topic estimates under EP inference with selection using VIP (*right*) are broadly similar to Gibbs when using empirical MI estimates for selection (*center*), though at reduced computational cost. Gibbs estimates have higher noise in low probability regions. Annotation based on random selection (*left*) performs poorly regardless of the inference method – Gibbs shown.

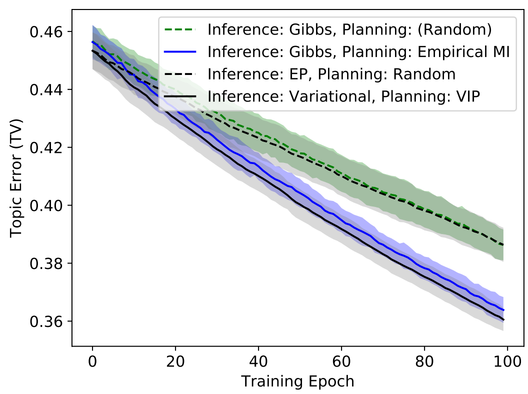

Figure 7: **Labeled LDA.** Total variation error (absolute error) for all topics on the "bars" dataset across 10 random trials. Variational planning with fully parameterized softmax shows consistent improvement over Gibbs on average (solid). Gibbs estimates show tighter standard deviation (shaded) for planning based on MI estimates. Both inference methods perform similarly for random selection.

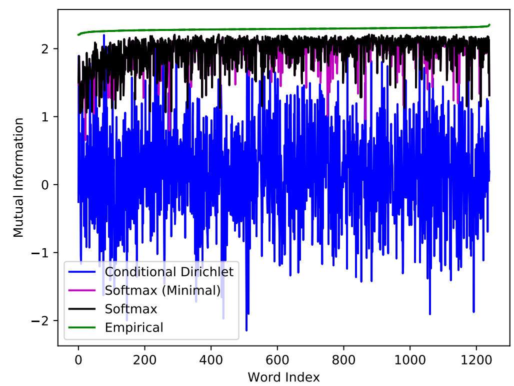

Figure 8: **LLDA Bound Comparison.** Variational MI estimates for annotations using three approximating families, ranked in order of increasing value by empirical estimate of MI. An ideal approximation would show monotonically increasing values with little gap. The conditional Dirichlet model is a poor approximation. Two variations on the softmax distribution provide tighter bounds, with the fully parameterized softmax being slightly preferred.

error across labels. Increasing MCMC samples from the reported 1000 samples does not lead to significant improvements. We instead suspect that poor MCMC results are due to a combination of MI estimator bias, inference local optima, and topic label switching.

More flexible approximations lead to better bounds. We consider three forms of the auxiliary distribution $\omega(\cdot)$ in the MI lower bound: a conditional Dirichlet, $\omega(\psi \mid y = j) = \prod_k \text{Dirichlet}(\psi_k \mid \gamma_{kj})$, a *minimal* softmax where class conditionals depend on a single topic $\omega(y = k \mid \psi) \propto \exp(w_k^T \text{vec}(\psi_k) + w_{0k})$, and where class conditionals weight all topics $\omega(y = k \mid \psi) \propto \exp(w_k^T \text{vec}(\psi) + w_{0k})$. Fig. 8 compares variational MI estimates to that of an empirical estimate for each of these families. We find that the larger parameter set of this final family leads to more accurate bounds overall.

7 Conclusion

We have introduced VIP, an approach to sequential decision making which leverages the efficiency of variational approximations to produce high quality decisions at timescales that would be infeasible with existing sample-based methods. Using the same basic methodology we have shown that VIP can be easily adapted to a variety of contexts, and performs comparably to sample-based approaches requiring more computation and to methods specialized for a particular problem, as in the regulatory network example. We look forward to future applications of VIP in domains where decisions are time-sensitive.

References

D. Barber and F. Agakov. The IM algorithm: a variational approach to information maximization. *NIPS*, 16:201, 2004.

J. M. Bernardo. Expected Information as Expected Utility. *Ann. Stat.*, 7(3):686–690, May 1979.

D. Blackwell. Comparison of experiments. In J. Neyman, editor, *2nd BSMSP*, pages 93–102, Berkeley, CA, August 1950. UC Berkeley.

D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

O. Cappé, S. J. Godsill, and E. Moulines. An overview of existing methods and recent advances in sequential monte carlo. *Proceedings of the IEEE*, 95(5): 899–924, 2007.

J. Chen, L. Song, M. J. Wainwright, and M. I. Jordan. Learning to explain: An information-theoretic perspective on model interpretation. *arXiv preprint arXiv:1802.07814*, 2018.

P. Flaherty, G. Giaever, J. Kumm, M. I. Jordan, and A. P. Arkin. A latent variable model for chemogenomic profiling. *Bioinformatics*, 21(15):3286–3293, 2005.

S. Gao, G. Ver Steeg, and A. Galstyan. Variational information maximization for feature selection. In *Advances in neural information processing systems*, pages 487–495, 2016.

R. J. Giordano, T. Broderick, and M. I. Jordan. Linear response methods for accurate covariance estimates from mean field variational bayes. In *Advances in Neural Information Processing Systems*, pages 1441–1449, 2015.

N. J. Gordon, D. J. Salmond, and A. F. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)*, volume 140, pages 107–113. IET, 1993.

T. L. Griffiths and M. Steyvers. Finding scientific topics. *Proceedings of the National academy of Sciences*, 101(suppl 1):5228–5235, 2004.

G. Kitagawa. Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of computational and graphical statistics*, 5(1): 1–25, 1996.

D. V. Lindley. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 27(4):986–1005, December 1956. ISSN 0003-4851.

S. Mohamed and D. J. Rezende. Variational information maximisation for intrinsically motivated reinforcement learning. In *Advances in neural information processing systems*, pages 2125–2133, 2015.

T. Rainforth, R. Cornish, H. Yang, and A. Warrington. On nesting monte carlo estimators. In *International Conference on Machine Learning*, pages 4264–4273, 2018.

D. Ramage, D. Hall, R. Nallapati, and C. D. Manning. Labeled lda: A supervised topic model for credit attribution in multi-labeled corpora. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1*, pages 248–256. Association for Computational Linguistics, 2009.

M. W. Seeger. Bayesian inference and optimal design for the sparse linear model. *Journal of Machine Learning Research*, 9(Apr):759–813, 2008.

B. Settles. Active learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 6(1):1–114, 2012.

F. Steinke, M. Seeger, and K. Tsuda. Experimental design for efficient identification of gene regulatory networks using sparse bayesian models. *BMC systems biology*, 1(1):51, 2007.

M. Stephens. Dealing with label switching in mixture models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(4):795–809, 2000.

R. E. Turner and M. Sahani. Two problems with variational expectation maximisation for time-series models. *Bayesian Time series models*, 1(3.1):3–1, 2011.

M. J. Wainwright and M. I. Jordan. Graphical models, exponential families, and variational inference. Technical report, UC Berkeley, Dept. of Statistics, 2003.

J. L. Williams. *Information Theoretic Sensor Management*. PhD thesis, MIT, Cambridge, MA, USA, 2007.

S. Zheng, J. Pacheco, and J. Fisher. A robust approach to sequential information theoretic planning. In *International Conference on Machine Learning*, pages 5936–5944, 2018.