Variational Information Planning for Sequential Decision Making

Jason Pacheco
MIT CSAIL

John W. Fisher, III
MIT CSAIL

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 the development of posterior inference algorithms, yet comparatively little attention to methods for decision making based on the results of inference. In this paper we explore sequential decision making based on information theoretic quantities. Specifically, we introduce efficient methods for *information planning*, where decisions are generated by maximizing the mutual information (MI) utility (Williams, 2007).

Our setting resembles Bayesian experiment de-

sign (Lindley, 1956), where experiments are chosen to minimize uncertainty over a quantity of interest. MI has long been used as a design utility in this setting (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 Markov chain Monte Carlo (MCMC), 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 lower bound of MI (Barber and Agakov, 2004) to the sequential setting. The bound is optimized over an auxiliary distribution approximating the expected posterior. We demonstrate that VIP yields a convex optimization for exponential family auxiliary models, leading to efficient planning. 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.

Despite good predictive accuracy, variational approximations of posterior uncertainty can be poor (Giordano et al., 2015; Turner and Sahani, 2011). Thus, a naive variational approach will tend to yield poor planning decisions. We address these issues by defining a class of auxiliary distributions that, when conditioned on future observations, define exponential families. This set allows arbitrary nonlinear dependence on the observation variable, and is thus strictly larger than the set of jointly exponential family models.

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 $\mathcal{Y}_t = \{y_1, \ldots, y_t\}$. At each time t a discrete *action* $a_t \in \{1, \ldots, A\}$ parameterizes the likelihood, denoted $p_{a_t}(y_t \mid x)$. Let $\mathcal{D}_t = \{\mathcal{Y}_t, \mathcal{A}_t\}$ be the set of past observations and chosen actions $\mathcal{A}_t = \{a_1, \ldots, a_t\}$ at time t. The posterior is then,

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

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, action a_t is selected to maximize the posterior MI,

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

New observations are then drawn from the appropriate likelihood $y_t \sim p_{a_t}(\cdot \mid x)$ and the posterior is updated. However, calculating posterior MI in Eqn. (2) is complicated for two reasons. First, entropies involve expectations under the posterior (1). Second, 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

Motivated by the challenges of sample-based MI estimation, we introduce an efficient variational approach. Beginning with a lower bound on MI, we extend this to sequential decision making and formulate the calculations for a model where observations are conditionally independent. Finally, we show how VIP can be applied to a more complex model common in the related setting of active learning.

3.1 Variational Information Bound

For any valid conditional distribution $\omega(x \mid y)$, Gibbs' inequality admits the following lower bound on MI:

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

This bound has been independently explored in various contexts (Barber and Agakov, 2004; Mohamed and Rezende, 2015; Gao et al., 2016; Chen et al., 2018). In the remainder of this paper we refer to $\omega(x \mid y)$ as the *auxiliary distribution*. The dual planning problem maximizes the bound (4) w.r.t. this auxiliary distribution. Each stage of planning requires the posterior mutual information, $I(X, Y_t \mid \mathcal{D}_{t-1})$ bounded by,

$$H(X \mid \mathcal{D}_{t-1}) + \mathbb{E}_p[\log \omega(X \mid Y) \mid \mathcal{D}_{t-1}]. \tag{5}$$

Calculating Eqn. 5 involves expectations over the posterior distribution $p(x, y_t \mid \mathcal{D}_{t-1})$, thus efficient sequential planning requires further approximations. The procedure is most easily understood for a simple model of conditionally independent observations, which we now discuss before moving to more complicated settings.

3.2 Conditionally Independent Observations

Consider the model in Eqn. (1) where observations $y_1, \ldots, y_t$ are independent, conditioned on x. Given the variational approximation $q(x) \approx p(x \mid \mathcal{D}_{t-1})$ we form a local approximation of the distribution over the future measurement at time t,

$$\hat{p}_a(x, y_t) \equiv q_{t-1}(x) p_a(y_t \mid x) \approx p_a(x, y_t \mid \mathcal{D}_{t-1}). \tag{6}$$

Here, $p_a(y_t \mid x)$ is the true likelihood under the hypothesized action a. We then bound the MI under $\hat{p}(\cdot)$ as,

$$H_{\hat{p}}(X) + \max_{a, \omega} \mathbb{E}_{\hat{p}_a}[\log \omega(X \mid Y_t)]. \tag{7}$$

Under this model the marginal entropy $H(X)$ is constant during planning and can be ignored. The bound (7) can be evaluated in parallel for all 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. Eqn. (7) bounds mutual information under the local approximation $\hat{p}(\cdot)$. The conditions ensuring Eqn. (7) is a reliable surrogate to Eqn. (5) are the same as those for variational inference to be effective.

3.3 Semi-Supervised Annotation Model

We now consider a more complicated model consisting of semi-supervised *annotations* $\{y_n\}_{n=1}^{N}$, a fixed set
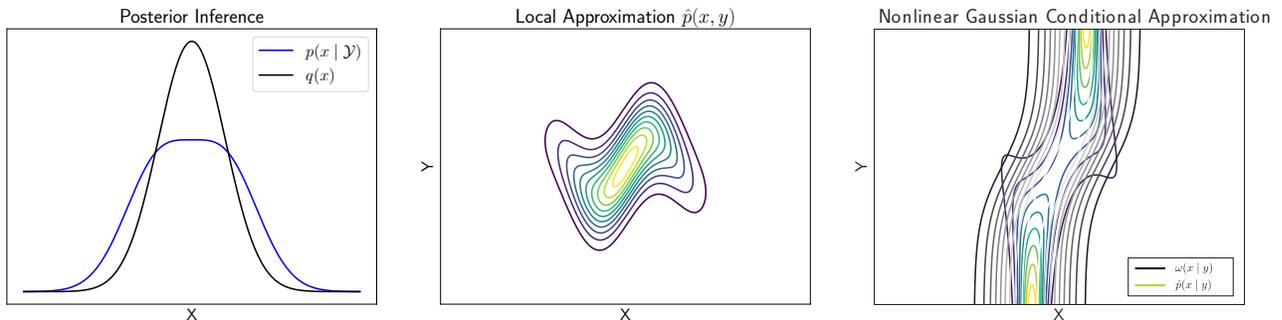
Figure 1: **Variational information planning steps.** *Left:* Given observations $\mathcal{Y}$ the posterior is approximated with a tractable family $q(x) \approx p(x \mid \mathcal{Y})$. *Center:* To consider a new observation y, a local approximation is formed $\hat{p}(x, y) = q(x)p(y \mid x)$ using the forward model. *Right:* The auxiliary distribution $\omega(x \mid y)$ minimizes $\mathrm{KL}(\hat{p}(x \mid y) \parallel \omega(x \mid y))$ to bound $I(X; Y)$. When the auxiliary distribution is in the exponential family we achieve efficient updates, yet allow nonlinear dependence on the conditioning variable y, thus yielding tighter MI bounds.

of data $\{z_n\}_{n=1}^N$, and latent quantities x. The joint distribution is given by,

$$p(x, y, z) = p(x) \prod_{n=1}^N p(z_n, y_n \mid x).$$

Variations of this model are common in active learning contexts (Settles, 2012). For example, y_n may be a class label for data element z_n. Each learning stage selects the most informative annotation y_{n^*} maximizing posterior MI:

$$n^* = \arg\max_n \; I(X; Y_n \mid \mathcal{D}_{t-1}) \qquad (8)$$

Here $\mathcal{D}_{t-1}$ is the set of all data $z_1, \ldots, z_N$ and the currently observed annotations. To form a local approximation $\hat{p}(\cdot)$ we assume a posterior approximation that is a product of nonnegative normalizeable factors,

$$q(x) \propto \prod_{n=1}^N \psi_n(x) \qquad (9)$$

In expectation propagation (EP) parlance, factors $\psi(x)$ can be interpreted as messages in a factor graph. Similarly, EP defines the concept of a *cavity distribution* $q^{\setminus n}(x) \propto q(x)/\psi_n(x)$, which expresses the posterior approximation having removed z_n. Our local approximation is then analogous to the EP *augmented distribution*,

$$\hat{p}(x, y_n) \propto q^{\setminus n}(x)p(z_n, y_n \mid x). \qquad (10)$$

The MI lower bound is then identical to (7). More complicated models with nuisance variables that must be integrated out for planning can be handled in a similar manner, with additional marginalization. We consider such a setting for the labeled LDA active learning example in Sec. 6.3.

4 Optimization for Conditional Exponential Families

Optimization of the bound (7) with respect to the auxiliary distribution $\omega(x \mid y)$ can be complicated in general. In this section we consider optimization for the class of auxiliary distributions 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 and that 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 having density,

$$\omega_\theta(x \mid y) = h(x) \exp\left(\theta(y)^T \phi(x, y) - A(\theta(y))\right), \quad (11)$$

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. (7) is equivalent to minimizing the cross entropy,

$$\theta^*(y) = \arg\min_\theta \; J(\theta) \equiv \mathbb{E}_{\hat{p}}[-\log \omega_\theta(x \mid y)]. \qquad (12)$$

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

$$\theta^*(y) = \arg\min_\theta \; \mathbb{E}_{\hat{p}_y}\left[\mathrm{KL}(\hat{p}_{x|y} \parallel \omega_\theta)\right] \qquad (13)$$

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. (13) is then a convex combination of convex functions, thus convexity holds.

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

$$\mathbb{E}_{\hat{p}_y}\left[\mathbb{E}_{\omega_{\theta^*}}[\phi(x,y) \mid Y = y]\right] = \mathbb{E}_{\hat{p}}[\phi(x,y)]. \qquad (14)$$

This is a weaker condition than the standard moment matching property of exponential families, which typically minimizes KL. Under (14) 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 (14) 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,

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

where $D_\eta\theta$ is the Jacobian matrix of partial derivatives. If $\theta(y)$ is convex in the parameters η then the optimization Eqn. (13) remains convex.

In principle, the map $\theta_\eta(y)$ can be any parametric function, for example a neural network with parameters η. Indeed, in related work Chen et al. (2018) optimize the MI bound (4) w.r.t. a neural network map for feature selection tasks. However, such an approach violates the convexity properties above and leads to computation that is prohibitive for sequential decision making tasks.

5 Evaluating the MI Bound

The previous section characterized natural parameters maximizing the MI bound (7) 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 can be done in closed form, and corresponds to the standard moment matching property.

5.1 Empirical 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 (7) 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. 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 MI bound (7) is easily optimized and evaluated by standard moment matching of the auxiliary distribution. One such class occurs when the marginal $\hat{p}(y)$ is in the exponential family. Note that $\hat{p}(x,y)$ need not be jointly exponential, for example the condition holds if y is discrete .

Now, 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,

$$\mathbb{E}_{\hat{p}}[\phi(x,y)] = \mathbb{E}_{\omega_\eta}[\phi(x,y)]. \qquad (15)$$

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 (15), we have:

$$\mathbb{E}_{\hat{p}}[\phi(x,y)] = \mathbb{E}_{\hat{p}_y}[\mathbb{E}_{\omega_{x|y}}[\phi(x,y) \mid Y = y]], \qquad (16)$$

where $\omega_{\eta^*}(x \mid y) = \omega_{\eta^*}(x,y) \div \int \omega_{\eta^*}(x,y)\mathrm{d}x$. Eqn. (16) is the optimality condition (14) of the MI lower bound. This solution also leads to a simple form of the MI bound (7). By direct calculation, the cross entropy $H_{\hat{p}}(\omega_{\eta^*}(x,y))$ equals,

$$\mathbb{E}_{\hat{p}}[-\log h(x,y)] - \eta^T\mathbb{E}_{\omega_{\eta^*}}[\phi(x,y)] + A(\eta). \qquad (17)$$

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:

$$H_{\hat{p}}(\omega_{\eta^*}(X \mid Y)) = H_{\omega_{\eta^*}}(X,Y) - H_{\omega_{\eta^*}}(Y). \qquad (18)$$

The l.h.s. is the relevant conditional entropy term from the MI bound (7). The r.h.s. is the entropy of the

joint and marginal distributions $\omega_{\eta^*}(\cdot)$ at the optimal parameters, which is closed form.

To summarize, we have shown one sufficient condition, namely that $\hat{p}(y)$ is an exponential family, which leads to efficient evaluation of the MI bound. We further conjecture that broader conditions exist which lead to the same moment-matching optimization and evaluation of the variational MI bound.

6 Experimental Results

We demonstrate VIP in a variety of contexts including sensor selection, Bayesian experiment design, and active learning. The primary comparison in most settings 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.

We consider estimation in, both, static and dynamic settings. In both cases we optimize the MI lower bound over a linear Gaussian auxiliary distribution $\omega(x \mid y_t) = \mathcal{N}(ay_t, \sigma^2)$, 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.

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,

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

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 to a nonlinear dynamical system frequently used in the sequential Monte Carlo literature (Kitagawa, 1996; Gordon et al., 1993; Cappé et al., 2007),

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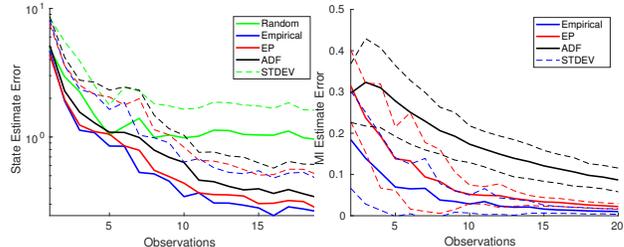

Figure 2: **Static target estimation.** Estimation of a 1D target position in a network of $K = 10$ equally spaced sensors. Mean (solid) plus STDEV (dashed) over 20 realizations. *Left:* Variational planning based on EP inference yields comparable error in state estimate compared to exact inference with 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.

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.

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 empirical 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

Steinke et al. (2007), and later Seeger (2008), introduced a heuristic approach to sequential experi-
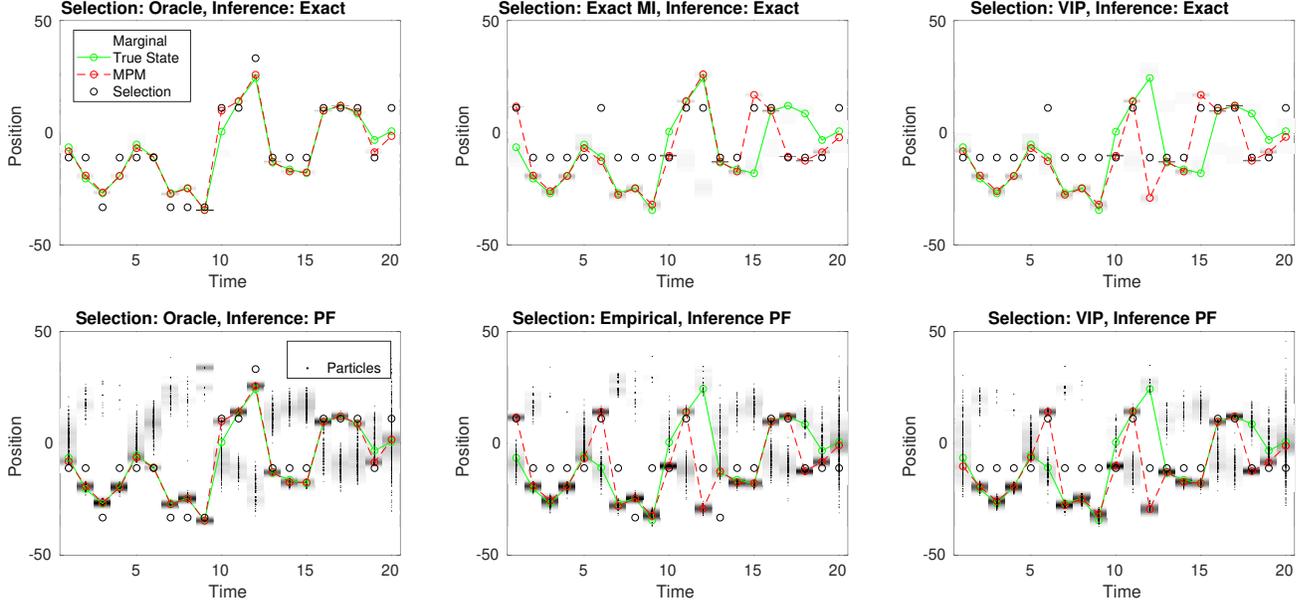
Figure 3: **Nonlinear tracking example** in a field of $K = 10$ equally spaced, stationary, sensors. The best comparison is to numerical approximation to MI and the posterior distribution (*top-center*). For reference, we have also included an *oracle* which selects the sensor closest to the true target 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).
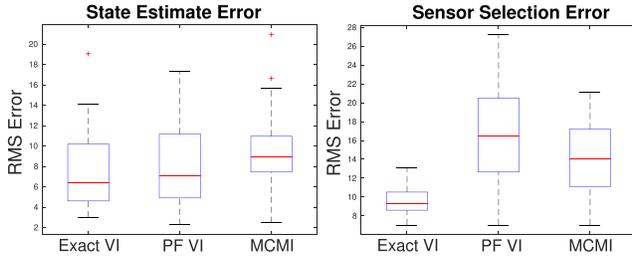

Figure 4: **Nonlinear tracking for 20 random trials.** *Left:* Exact inference with variational planning yields the lowest RMS state error. Particle filter inference with 500 particles and variational planning (PFVI) yields lower median error compared to MC estimates of information (MCMI), though wider 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.
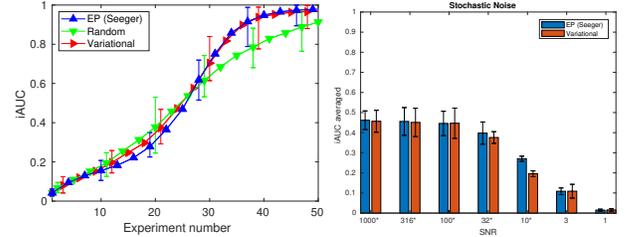

Figure 5: **Sparse linear model.** Average +/- STDEV computed for $n = 50$ nodes over 20 random networks. *Left:* AUC for edges with true weights $|a_{ij}| > 0.1$ plotted for each intervention experiment. VIP shows similar results to Steinke and Seeger with slight improvement for less than 25 experiments. *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).

ment design which shares components of our current method. Their approach relies on EP inference and related approximations of the MI utility. Using an example from these papers, we compare VIP for the discovery of causal gene interactions.

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 i.i.d. from a Laplace distribution,

$$p(A, x) \propto \prod_{i=1}^{n} \mathcal{N}(u_i \mid a_i^T x, \sigma^2) \prod_{j=1}^{n} \text{Laplace}(a_{ij} \mid \lambda) \quad (19)$$

where a_i is the ith 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 ith gene. Observe that the joint (19) is unnormalized since the likelihood term is not a normalized distribution of x.

We compare to the variational method of Steinke et al. (2007) and later Seeger (2008) which maintain 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 $\tilde{t}_{ij}(a_{ij})$ the Laplace factors. Given past observations $\mathcal{D}$ and a new control and observation pair $\{x_*, u_*\}$, maximizing MI is (approximately) equivalent to maximizing $\mathbb{E}_{x_*}[\mathrm{KL}(q' \| q)]$, 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 a semi-supervised extension of LDA (Blei et al., 2003). For K topics, D documents $d = 1, \ldots, D$, and N_d words per document, the standard unsupervised LDA model is given by,

$$\theta_d \sim \mathrm{Dirichlet}(\alpha) \quad \text{Topic proportions}$$
$$\psi_k \sim \mathrm{Dirichlet}(\beta_k) \quad \text{Topics}$$
$$z_{dn} \mid \theta_d \sim \mathrm{Cat}(\theta_d) \quad \text{Topic Labels}$$
$$w_{dn} \mid z_{dn}, \psi \sim \mathrm{Cat}(\psi_{z_{dn}}) \quad \text{Words}$$

As shown in Fig. 6, LLDA introduces semi-supervised annotations for each word (Flaherty et al., 2005). We
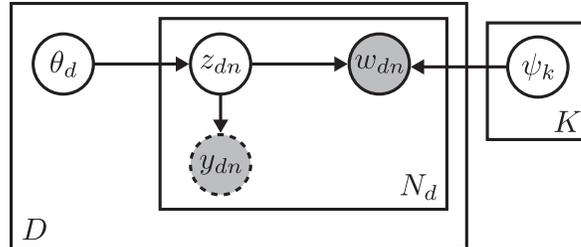

Figure 6: **Labeled LDA** graphical model augments standard LDA with semi-supervised annotations y_{dn} shown as shaded dashed nodes. Each stage of active learning selects unlabeled item (d, n) maximizing mutual information $I(\Psi, Y_{dn})$ between topics Ψ and annotations Y, as in Eqn. (20).

model annotations as noisy observations of the true topic assignment,

$$y_{dn} \mid z_{dn} \sim \mathrm{Cat}(\pi_{z_{dn}}).$$

This choice ensures interpretable topics as the posterior concentrates on a preferred ordering of labels. Other supervised LDA variations exist (Ramage et al., 2009) but we do not consider them here.

Our evaluation uses the "bars" dataset (Griffiths and Steyvers, 2004) where topics can be visualized as a set of vertical and horizontal *bars* (see Fig. 7). To amplify the effect of informative annotations we model a *rare* topic with a non-symmetric Dirichlet prior over topic proportions: $\alpha = (0.05, 1, 1, \ldots, 1)^T$.

Active Learning At each learning stage t the planner selects an annotation $y_{d^* n^*}$ maximizing posterior MI:

$$(d^*, n^*) = \underset{(d,n)}{\arg\max} \ I(\Psi, Y_{dn} \mid \mathcal{W}, \mathcal{Y}_{t-1}), \quad (20)$$

with observed annotations $\mathcal{Y}_{t-1}$ and $\mathcal{W}$ the corpus of words. Eqn. (20) is an instance of the annotation model (Sec. 3.3) as integration over nuisance variables induces dependence among annotations. We choose a softmax auxiliary distribution:

$$\omega(y = k \mid \psi) \propto \exp(w_k^T \mathrm{vec}(\psi_k) + w_{0k}),$$

which has a number of beneficial properties: first, it ensures convexity of the planning objective (20) in the parameters w (Sec. 4). Second, the marginal distribution $p(y)$ is in the exponential family (discrete) thus ensuring simple evaluation of the MI bound (Sec. 5).

Lower topic error compared to MCMC. We compare VIP using EP inference (Broderick et al., 2013) to empirical MI planning based on Gibbs samples. Observe that MI estimates require averaging over topic samples, and thus must account for topic label switching. We found that any reasonable alignment (e.g.
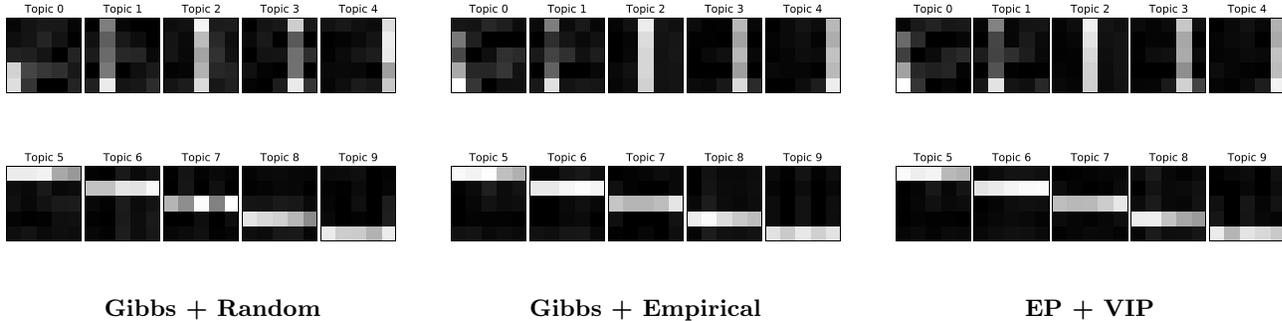
Figure 7: **Learned LLDA topics** from a corpus of $D = 50$ documents, each with $N_d = 25$ words drawn from the bars topics with a $W = 25$ word vocabulary. We model Topic 0 as a *rare* topic (see text). Gibbs estimates are averaged over 1000 samples drawn from parallel chains. Topic estimates under 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.
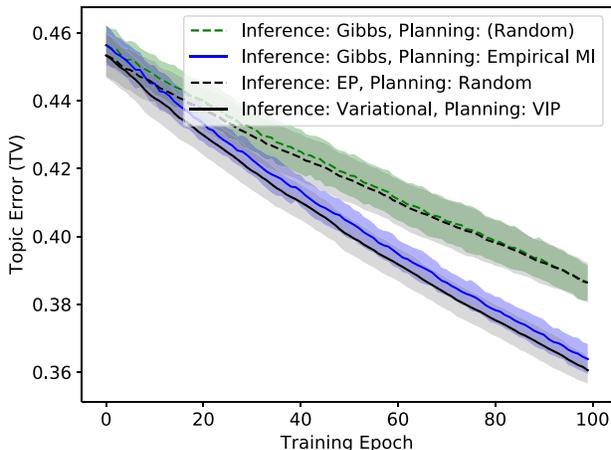
7 Conclusion

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.

Figure 8: **Total variation (TV) 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.

MAP) resulted in poor estimates (Stephens, 2000). We instead relabel Gibbs samples to minimize absolute error w.r.t. the *true topics*, which would not be feasible in practice. Fig. 8 reports total variation (TV) error across topics $\sum_k \|\psi_k - \hat{\psi}_k\|_1$, based on the posterior mean estimate $\hat{\psi}$. We found that increasing from the reported 1000 Gibbs samples did not lead to significant improvements. We instead suspect that poor MCMC results are due to a combination of MI estimator bias and local optima.

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.

T. Broderick, N. Boyd, A. Wibisono, A. C. Wilson, and M. I. Jordan. Streaming variational bayes.

In *Advances in Neural Information Processing Systems*, pages 1727–1735, 2013.

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. Wainwright, and M. Jordan. Learning to explain: An information-theoretic perspective on model interpretation. In *International Conference on Machine Learning*, pages 882–891, 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.