
Active Learning of Continuous-time Bayesian Networks through Interventions

Dominik Linzner^{1,2} Heinz Koepl^{1,3}

Abstract

We consider the problem of learning structures and parameters of Continuous-time Bayesian Networks (CTBNs) from time-course data under minimal experimental resources. In practice, the cost of generating experimental data poses a bottleneck, especially in the natural and social sciences. A popular approach to overcome this is Bayesian optimal experimental design (BOED). However, BOED becomes infeasible in high-dimensional settings, as it involves integration over all possible experimental outcomes. We propose a novel criterion for experimental design based on a variational approximation of the expected information gain. We show that for CTBNs, a semi-analytical expression for this criterion can be calculated for structure and parameter learning. By doing so, we can replace sampling over experimental outcomes by solving the CTBNs master-equation, for which scalable approximations exist. This alleviates the computational burden of integrating over possible experimental outcomes in high-dimensions. We employ this framework in order to recommend interventional sequences. In this context, we extend the CTBN model to conditional CTBNs in order to incorporate interventions. We demonstrate the performance of our criterion on synthetic and real-world data.

Learning directed dependencies in multivariate data is a fundamental problem in science and has application across many disciplines such as in the natural and social sciences, finance, and engineering (Acerbi et al., 2014; Schadt et al., 2005). However, large amounts of data are needed in order

*Equal contribution ¹Department of Engineering and Information Technology, TU Darmstadt, Germany ²The Why Company GmbH, Berlin, Germany ³Department of Biology, TU Darmstadt, Germany. Correspondence to: Dominik Linzner <dlinzner90@gmail.com>, Heinz Koepl <heinz.koepl@bcs.tu-darmstadt.de>.

to learn these dependencies. This is a problem when data is acquired under limited resources, which is the case in dedicated experiments, e.g. in molecular biology or psychology (Steinke et al., 2007; Zechner et al., 2012; Liepe et al., 2013; Myung & Pitt, 2015; Dehghannasiri et al., 2015; Prangemeier et al., 2018). Active learning schemes pave a principled way to design sequential experiments such that the required resources are minimized.

The framework of Bayesian optimal experimental design (BOED) (Chaloner & Verdinelli, 1987; Ryan et al., 2016) allows for the design of active learning schemes, which are provably (Lindley, 1956; Sebastiani & Wynn, 2000) one-step optimal. However, BOED becomes infeasible in high-dimensional settings, as it involves integration over possible experimental outcomes.

While active learning schemes have been previously applied in order to learn dependency structures (Tong & Koller, 2001; Eaton & Murphy, 2007; He & Geng, 2008; Lindgren et al., 2018) or parameters (Rubenstein et al., 2017) of probabilistic graphical models from snapshot or static data, active learning schemes for longitudinal and especially temporal data is as of yet under-explored. Dynamic Bayesian networks offer an appealing framework to formulate structure learning for temporal data within the graphical model framework (Koller & Friedman, 2010). The fact that the time granularity of the data can often be very different from the actual granularity of the underlying process motivates the extension to continuous-time Bayesian networks (CTBN) (Nodelman et al., 1995), where no time granularity of the unknown process has to be assumed.

In this manuscript, we present an active learning scheme for learning CTBNs from interventions. We make the following contributions: (i) We derive a criterion for active learning suited for the case when the space of possible experimental outcomes is high-dimensional in Section 2. (ii) In Section 3, we extend CTBNs to incorporate interventions, thereby introducing conditional CTBNs (cCTBNs). (iii) We discuss pooling of interventional data in cCTBNs in Section 4. (iv) We derive semi-analytical expressions of our design criterion from Section 2 for parameter and structure learning in Section 5. (v) We demonstrate the performance of our approach on synthetic and real-world data in section 6.

1. Background

Interventions. Consider a directed acyclic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with nodes $\mathcal{V} \equiv \{1, \dots, N\}$ and edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. The parent-set of node n is then defined as $\text{par}^{\mathcal{G}}(n) \equiv \{m \mid (m, n) \in \mathcal{E}\}$. Conversely, we define the *child-set* $\text{ch}^{\mathcal{G}}(n) \equiv \{m \mid (n, m) \in \mathcal{E}\}$. Consider a joint distribution over a set of random variables over a countable domain $p(X_1, \dots, X_N)$, characterized by a Bayesian network with graph \mathcal{G} , i.e.,

$$p(X_1, \dots, X_N) = \prod_{i=1}^N p(X_i \mid X_{\text{par}^{\mathcal{G}}(i)}).$$

Interventions as popularized in (Pearl, 2000) are denoted via a do-operation and correspond to a change to the model. In our example, an intervention on variable X_k , $\text{do}(X_k = x_k)$ would change the distribution to

$$p(X_1, \dots, X_N \mid \text{do}(X_k = x_k)) = \mathbb{1}(X_k = x_k) \tilde{p}(X_{\neg k}), \quad (1)$$

with $\mathbb{1}(\cdot)$ the indicator function and $\tilde{p}(X_{\neg k}) \equiv \prod_{i=1, i \neq k}^N p(X_i \mid X_{\text{par}^{\mathcal{G}}(i)})$, where notation $\neg k$ means all variables except X_k . It corresponds to a model where the conditional distribution of every variable remains unchanged, but the value of variable $X_k = x_k$ is fixed (Pearl, 2000; Spirtes, 2010). This do-operation contrasts traditional conditioning

$$p(X_{\neg k} \mid X_k = x_k) = \frac{p(X_k = x_k \mid X_{\text{par}^{\mathcal{G}}(k)}) \tilde{p}(X_{\neg k})}{p(X_k = x_k)},$$

where conditioning on $X_k = x_k$ also affects the parents of node k (instead of only its children). Interventions can also be modelled as external condition variables (Pearl, 2000) whose effects are equivalent to do-operations. A problem that is encountered when learning from interventions is data-pooling (Eberhardt, 2008): Under what conditions can observations gathered under different interventions be used to learn about the original model. We later show that CTBNs allow for data-pooling naturally.

Conditional Continuous-time Markov Chain. We introduce a conditional Continuous-time Markov process (cCTMC) (Nodelman et al., 1995; Norris, 1997) by a tuple $(\mathcal{S}, \mathcal{I}, W, s_0)$. It defines a Markov process $\{S(t)\}_{t \geq 0}$ through a transition intensity matrix $W : \mathcal{S} \times \mathcal{S} \times \mathcal{I} \rightarrow \mathbb{R}$ over a countable state space \mathcal{S} given a countable space of external conditions (interventions) \mathcal{I} and initial states s_0 , which may also be dependent on the external condition $s_0 : \mathcal{I} \rightarrow \mathcal{S}$. For the sake of conciseness, we will often adopt shorthand notations of the type $p_{t'-t}(s' \mid s, i) \equiv p(S(t') = s' \mid S(t) = s, i)$, with $s, s' \in \mathcal{S}, i \in \mathcal{I}$. Given a condition, its time evolution can be understood as a usual continuous-time Markov

chain (CTMC) with the (infinitesimal) transition probability $p_h(s' \mid s, i) = \mathbb{1}(s = s') + h W(s, s', i) + o(h)$, for some time-step h with $\lim_{h \rightarrow 0} o(h)/h = 0$. We note that any intensity matrix W fulfills $W(s, s, i) = -\sum_{s' \neq s} W(s, s', i)$ for any condition i . In the continuous-time limit $h \rightarrow 0$, the cCTMCs marginal probabilities can be shown to follow the Chapman–Kolmogorov-, or master-equation

$$\frac{d}{dt} p_t(s \mid s_0, i) = \sum_{s' \neq s} [W(s', s, i) p_t(s' \mid s_0, i) - W(s, s', i) p_t(s \mid s_0, i)]. \quad (2)$$

Bayesian Optimal Experimental Design. The objective of BOED is to find the design (intervention i) that maximizes the expected information gain $\text{EIG}(i)$ about different models Θ (Lindley, 1956; Chaloner & Verdinelli, 1987). In our case Θ will be rate matrices W , or their induced graph-structures, of continuous-time Markov processes (see also Section 3). The $\text{EIG}(i)$ corresponds to the expected Kullback–Leibler (KL) divergence between prior $p(\Theta)$ and posterior $p(\Theta \mid D)$ after intervention i under uncertain experimental outcome (data) D . The objective of this task can then be formulated as

$$i^* = \arg \max_{i \in \mathcal{I}} \text{EIG}(i), \quad (3)$$

$$\text{EIG}(i) = \text{E} [\text{KL}(p(\Theta \mid D, i) \parallel p(\Theta))],$$

with the expectation taken with respect to $p(D \mid i)$. Unfortunately, in practice, (4) is notoriously hard to evaluate (Foster et al., 2019) rendering it impractical for our purposes.

2. Variational Box-Hill Criterion for Active Learning

Evaluating the EIG is intractable, due to the repeated model posterior evaluations for all possible outcomes D . For this reason, various approximation techniques have been employed (Lewi et al., 2009; Rainforth et al., 2018; Foster et al., 2019). Recently, promising advances via variational approximations of the EIG have been made (Foster et al., 2019). In the following, we want to take a similar route, while making use of our model assumptions. One way to do this is to upper-bound the EIG, which can be rewritten as $\text{EIG}(i) = \text{E} \left[\ln \frac{p(D \mid \Theta, i)}{p(D \mid i)} \right]$, with the expectation with respect to $p(D, \Theta \mid i)$. By lower-bounding the marginal likelihood $\ln p(D \mid i) \geq \text{E} [\ln p(D \mid \Theta, i) - \text{KL}[q_{\kappa}(\Theta) \parallel p(\Theta)]]$, the expectation subject to some, yet unspecified, variational distribution $q_{\kappa}(\Theta)$ over models, one arrives at an upper-bound to the EIG, which takes the form of a weighted KL-divergence between different possible models Θ

$$\text{EIG}(i) \leq \text{E} \left[\text{E} [\text{KL}(p(D \mid \Theta, i) \parallel p(D \mid \Theta', i))] + \text{KL}(q_{\kappa}(\Theta) \parallel p(\Theta)) \equiv \text{VBHC}(i, \kappa). \quad (4)$$

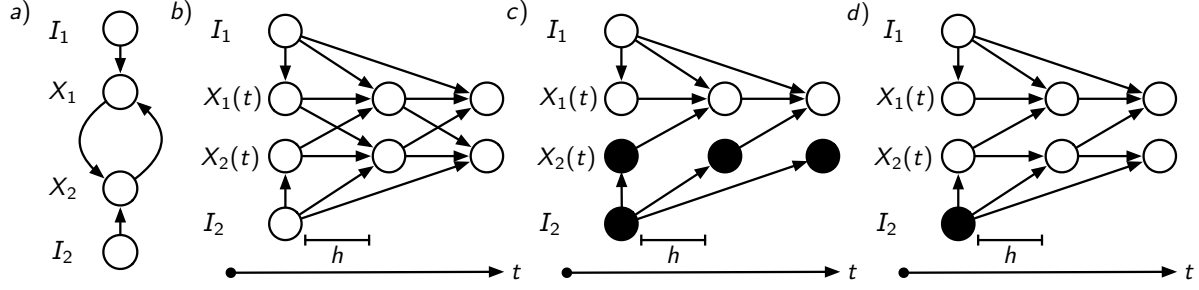


Figure 1. a) Conditional CTBN with two nodes X_1 and X_2 and local interventions I_1 and I_2 . Black means for interventions $I_n \neq 0$ and for $X_n(t) = x_n^0$ for all t . b), c) and d) are the h -discretized (unrolled) cCTBN for different interventions. b) Unrolled cCTBN without intervention. c) Unrolled cCTBN with perfect intervention on X_2 and d) with imperfect intervention.

Here the outer expectation is w.r.t $p(\Theta)$, the inner one w.r.t $q_\kappa(\Theta')$. We will refer to this quantity as the *Variational Box-Hill Criterion* (VBHC). By setting $q_\kappa(\Theta) = p(\Theta)$, which we emphasize is *not* the minimizer of the VBHC (and thus not the best approximation to the EIG), one recovers the classical Box-Hill criterion (BHC) (Box & Hill, 1967), $\text{BHC}(i) \equiv \mathbb{E}[\mathbb{E}[\text{KL}(p(D | \Theta, i) || p(D | \Theta', i))]]$, with expectations w.r.t $p(\Theta)$ and $p(\Theta')$. The BHC has been used in different contexts (Reilly, 1970; Daniel et al., 1996; Myung & Pitt, 2015; Ng & Chick, 2004) as a criterion for design for model discrimination experiments, as it can be computed analytically for some models (e.g. for Gaussian distributions). To our surprise, it hasn't received much attention otherwise. In contrast to the BHC, the VBHC allows for minimization, and a thereby tightening the upper-bound w.r.t κ before selecting an intervention

$$i^* = \arg \max_{i \in \mathcal{I}} \min_{\kappa} \text{VBHC}(i, \kappa). \quad (5)$$

To the best of our knowledge, we are the first to apply the BHC to the discrimination of different (continuous-time) Markov chains. Further, we note that our derivation via variational inference is so far missing in the literature, and we will demonstrate later that it can improve on the classical criterion. The VBHC can be related to a popular variational estimator for the mutual information (MI) (Poole et al., 2019), where, however, the marginal likelihood is directly replaced by a variational distribution. From a computational perspective, the (V)BHC corresponds to the following simplification: While computing the EIG (or the MI), the posterior computation over all models has to be performed for each experimental outcome, the (V)BHC only requires a single posterior computation per model. This is especially helpful in settings where (repeated) posterior computation becomes prohibitively expensive.

3. Model: Conditional Continuous-time Bayesian Networks

Definition. Analogous to continuous-time Bayesian Networks (Nodelman et al., 1995), we define Conditional Continuous-time Bayesian Networks (cCTBN) as an N -component cCTMC over factorizing state-spaces $\mathcal{S} = \mathcal{X}_1 \times \dots \times \mathcal{X}_N$, evolving jointly as a CTMC given any condition i . In this work we consider local interventions, thus we assume $\mathcal{I} = \mathcal{I}_1 \times \dots \times \mathcal{I}_N$. We state explicitly that single component states and interventions are entries of the states and interventions of the global cCTMC i.e. $s = (x_1, \dots, x_N)$ for $s \in \mathcal{S}$ with $x_n \in \mathcal{X}_n$ and $i = (i_1, \dots, i_N)$ for $i \in \mathcal{I}$ with $i_n \in \mathcal{I}_n$ for all $n \in \{1, \dots, N\}$. For cCTBNs the parent configuration can be summarized via a directed, but possibly cyclic graph structure $G = (\mathcal{V}, \mathcal{E})$ and in general depends on the current intervention. We note that the graph of a cCTBN can be unrolled on an infinitesimal time-grid (with spacing h) to reveal the interaction graph \mathcal{G} in time, which is again acyclic (Cohn et al., 2010; Linzner & Koeppl, 2018). The effect of unrolling is illustrated in figure 1 a) and b). In this manuscript, our goal is learning the possibly cyclic graph G .

The n 'th nodes' process $\{X_n(t)\}_{t \geq 0}$ depends on its current state $x_n \in \mathcal{X}_n$, its condition $i_n \in \mathcal{I}_n$ and of all its parents $U_n(t) = u_n$ taking values in $\mathcal{U}_n^G \equiv \times_{m \in \text{par}^G(n)} \mathcal{X}_m$, with \times denoting the Cartesian product. For a cCTBN, the global transition matrix $p_h(s' | s, i)$ then factorizes over nodes $p_h(s' | s, i) = \prod_{n=1}^N p_h(x'_n | x_n, u_n, i_n)$, into local conditional transition probabilities. We define local transition rates $\Lambda_n^i : \mathcal{X}_n \times \mathcal{X}_n \times \mathcal{U}_n \rightarrow \mathbb{R}$ for each condition $i \in \mathcal{I}_n$. In the following, we write compactly $\Lambda_n^i(x, x', u) \equiv \Lambda_n(x_n, x'_n, u_n, i_n)$. Subsequently, we can express the local conditional transition probabilities as $p_h(x'_n | x_n, u_n, i_n) = \mathbb{1}(x = x') + h\Lambda_n^i(x', x, u) + o(h)$. Lastly, we mention that an equivalent cCTMC can be constructed by amalgamation (El-Hay et al., 2011). We can define the global transition rate-matrix W through

these sets $W = \{G, \Lambda\}$. This can be used to solve the Chapman-Kolmogorov equation of a cCTBN by transforming it into an equivalent cCTMC.

Properties. cCTBNs present an extension to CTBNs as, given any condition i , a cCTBN is a CTBN. Paths of a CTBN $S^{[0,T]} = \{X_1^{[0,T]}, \dots, X_N^{[0,T]}\}$ assume values in the space of piece-wise constant (cadlag) functions. The path-likelihood of a cCTBN given a condition i is the likelihood of a CTBN (Nodelman et al., 2003)

$$p(S^{[0,T]} | \Lambda, G, i, s_0) = \prod_{n=1}^N p(X_n^{[0,T]} | X_{\text{par}^G(n)}^{[0,T]}, \Lambda_n, i_n) \mathbb{1}(X_n(0) = x_n^0), \quad (6)$$

where we introduced the vector-valued path of the cCTMC $S^{[0,T]} = \{S(t) | 0 \leq t \leq T\}$ and its components $X_n^{[0,T]} = \{X_n(t) | 0 \leq t \leq T\}$ for all $n \in \{1, \dots, N\}$ and the (vector-valued) initial state of the system $s_0 = (x_1^0, \dots, x_N^0)$. The conditional path-likelihood $p(X_n^{[0,T]} | X_{\text{par}^G(n)}^{[0,T]}, \Lambda_n, i_n)$ of an individual node n can in turn be expressed in terms of the path-statistics (Nodelman et al., 2003), the number of transitions of node n from state x to x' given the parents state u denoted by $M_n(x, x', u)$ and $T_n(x, u)$, which denotes the amount of time node n spent in state x

$$p(X_n^{[0,T]} | X_{\text{par}^G(n)}^{[0,T]}, \Lambda_n, i_n) = \prod_{x, x' \neq x, u} \Lambda_n^i(x, x', u)^{M_n(x, x', u)} e^{-T_n(x, u) \Lambda_n^i(x, x', u)}. \quad (7)$$

In (Nodelman et al., 2003) it was shown that a marginal likelihood for the structure of a CTBN can be calculated in closed form under the assumption of independent gamma priors over the rates

$$p(X_n^{[0,T]} | X_{\text{par}^G(n)}^{[0,T]}, i_n) \propto \prod_{x, x' \neq x, u} \Gamma(\bar{\alpha}_n^i(x, x', u)) \bar{\beta}_n^i(x, u)^{-\bar{\alpha}_n^i(x, x', u)}, \quad (8)$$

where $\bar{\alpha}_n^i(x, x', u) = M_n(x, x', u) + \alpha_n^i(x, x', u)$ and $\bar{\beta}_n^i(x, u) = T_n(x, u) + \beta_n^i(x, u)$ and $\alpha_n^i(x, x', u)$, $\beta_n^i(x, u)$ being the hyper-parameters of the gamma priors.

Interventions. Without loss of generality, we denote 0 as no intervention. This identifies the model with $i = 0$ as the un-intervened on or *original* model.

Perfect interventions. A perfect intervention, illustrated in figure 1 c), corresponds to Pearls do-operation. While it is known (Pearl, 2000), that interventions can be modelled as additional variables, correspondence in the case of cCTBNs can found directly. By setting $\Lambda_n^i(x, x', u) = 0$ for the intervened node n with $i_n \neq 0$, (7) evaluates to 1 iff $M_n(x, x', u) = 0$ for all $x, x' \in \mathcal{X}_n$ and $u \in \mathcal{U}_n^G$, and 0 otherwise. Together with the indicator for the initial condition

in (6) this evaluates to an indicator $\mathbb{1}(X_n^{[0,T]} = x_n^0)$ in the path-likelihood (6). One then recovers the same relationship as in the static model (1)

$$p(S^{[0,T]} | \Lambda, G, i_{-k} = 0, i_k \neq 0, s_0) = \mathbb{1}(X_k^{[0,T]} = x_k^0) \tilde{p}(X_{-k}^{[0,T]}),$$

with $\tilde{p}(X_{-k}^{[0,T]}) \equiv \prod_{n=1, n \neq k}^N p(X_n^{[0,T]} | X_{\text{par}^G(n)}^{[0,T]}, \Lambda_n, i_n) \mathbb{1}(X_n(0) = x_n^0)$ and i_{-k} being the vector-valued intervention without the k 'th component.

Imperfect interventions, as studied for example in (Eaton & Murphy, 2007), do not fix the state of the intervened on n 'th node, but only corresponds to a change to its path-likelihood. In cCTBNs, this is reflected by $\Lambda_n^i(x, x', u)$ being an arbitrary function of the condition $i_n \neq 0$. Here the dependency of node n on its parents can but doesn't have to be broken. Imperfect interventions are illustrated in figure 1 d).

4. Experimental Sequence Likelihood of a cCTBN

We want to calculate the likelihood of a sequence of observations $\mathcal{H} \equiv \{S^{[0,T],k} | 0 \leq k \leq K\}$ collected in K different experiments under (possibly) K different interventions $\Pi \equiv \{i^k | 0 \leq k \leq K\}$.

Data-pooling for cCTBNs. Given an experimental sequence \mathcal{H} under local interventions i , the likelihood of this sequence can be expressed in terms of node-wise likelihoods, independent on the order of the sequence, i.e.

$$p(\mathcal{H} | \Lambda, G, \Pi) = \prod_{n=1}^N \prod_{i \in \mathcal{I}_n} \prod_{x, x' \neq x, u} \Lambda_n^i(x, x', u)^{\bar{M}_n(x, x', u, i)} \exp[-\bar{T}_n(x, u) \Lambda_n^i(x, x', u)], \quad (9)$$

with sufficient statistics $\bar{T}_n(x, u, j) = \sum_{k=1}^K \mathbb{1}(i_n^k = j) T_n^k(x, u)$ and $\bar{M}_n(x, x', u, j) = \sum_{k=1}^K \mathbb{1}(i_n^k = j) M_n^k(x, x', u)$. This can be directly seen by considering the joint likelihood

$$p(\mathcal{H} | \Lambda, G, \Pi) = \prod_{k=1}^K \prod_{n=1}^N p(X_n^{[0,T],k} | X_{\text{par}^G(n)}^{[0,T],k}, \Lambda_n, i_n^k).$$

Inserting the likelihood (7) of $X_n^{[0,T],k} | X_{\text{par}^G(n)}^{[0,T],k}, \Lambda_n, i_n^k$, the above claim follows after definition of the statistics.

The above result allows us to use samples generated under interventions for the estimation of the original CTBN with $i = 0$, if each node was observed under condition $i = 0$ at least once. It reveals the penalty of performing an experiment under a perfect intervention: The statistics of

intervened-on node remain unchanged – the node is unobserved. We give a simple example that interventional data can be more informative than complete observations in the presence of time-scale separation.

Example: Interventional Data can be more informative than complete observations. Consider a minimal example of a two node CTBN $X \rightarrow Y$ with time-scale separation. For simplicity, we assume $\Lambda_X(x, x') = \varepsilon \rightarrow 0$, but the argument translates to finite rates. Further, we assume $\Lambda_X(x', x) > 0$ and $X(0) = x$. Then $X(t) = x'$ can not be observed for any finite $t \geq 0$ and the posterior over any rate $\Lambda_Y(y, y', x') > 0$ remains the prior and can thus not be learned. To see this we compute the posterior

$$\frac{p(\Lambda_Y(y, y', x') | Y^{[0,T]}, X^{[0,T]} = x) = p(Y^{[0,T]} | X^{[0,T]} = x, \Lambda_Y(y, y', x')) p(\Lambda_Y(y, y', x'))}{p(Y^{[0,T]} | X^{[0,T]} = x)},$$

however, by checking the likelihood (7), we find $p(Y^{[0,T]} | X^{[0,T]} = x, \Lambda_Y(y, y', x')) = p(Y^{[0,T]} | X^{[0,T]} = x)$, because the number of transitions $M_Y(y, y', x') = 0$, thus $p(\Lambda_Y(y, y', x') | Y^{[0,T]}, X^{[0,T]} = x) = p(\Lambda_Y(y, y', x'))$ and no information has been gained. On the other hand, an intervention allows us to set $X^{[0,T]} = x'$, thus, for the observation time T being sufficiently large, we will have $M_Y(y, y', x') > 0$ almost surely and $p(Y^{[0,T]} | \text{do}(X^{[0,T]} = x'), \Lambda_Y(y, y', x')) \neq p(Y^{[0,T]} | \text{do}(X^{[0,T]} = x'))$. Thus we have a finite information gain.

5. Active Learning of CTBNs

In the following, we will use (V)BHC for choosing sequences of interventions for repeated experiments. To facilitate this, we now derive semi-analytical expressions for the (V)BHC. We restrict ourselves to perfect interventions only. To avoid clutter, we define the set of nodes not subject to an intervention $\aleph \equiv \{n \in \{1, \dots, N\} | i_n = 0\}$. In the following, experimental outcomes are possible paths of a cCTBN $D = S^{[0,T]}$ and $\Lambda^0 \equiv \Lambda^{i=0}$ are the rates of the original model. As for the computation of the (V)BHC we marginalize over possible future outcomes, we introduce the notation $\hat{S}^{[0,T]}$ for a possible outcome path $\hat{S}^{[0,T]} \sim p(S^{[0,T]} | \Lambda, G, i, s_0)$ and its statistics, the number of transitions $\hat{M}_n(x, x', u)$ and dwell times $\hat{T}_n(x, u)$ of the n 'th node for all $x, x' \in \mathcal{X}_n$ and $u \in \mathcal{U}_n$.

Active Parameter Learning. The VBHC can be calculated as an expected KL-divergence between different models. In the following, we derive the criterion for optimal rate-estimation, given the structure G . The KL-divergence between two cCTBNs, with different rate-matrices Λ and

Λ' under the same intervention i and same graph G read

$$\begin{aligned} \text{KL} \left(p(S^{[0,T]} | \Lambda, G, i) || p(S^{[0,T]} | \Lambda', G, i) \right) = \\ \sum_{n \in \aleph} \sum_{x, x', u} \{ \mathbb{E}[\hat{M}_n(x, x', u) | \Lambda, G, i] \ln \frac{\Lambda_n^0(x, x', u)}{\Lambda_n'^0(x, x', u)} \\ - \mathbb{E}[\hat{T}_n(x, u) | \Lambda, G, i] \{ \Lambda_n^0(x, x', u) - \Lambda_n'^0(x, x', u) \} \}. \end{aligned}$$

The expected path-statistics of possible outcomes \hat{M}_n and \hat{T}_n , can be calculated analytically and are given by the solution $p_t(s | s_0, i)$ of the Chapman-Kolmogorov equation (2)

$$\mathbb{E}[\hat{T}_n(x, u) | \Lambda, G, i] = \quad (10)$$

$$\sum_s \mathbb{1}(s_n = x) \mathbb{1}(s_{\text{par}^G(n)} = u) \int_0^T dt p_t(s | s_0, i),$$

$$\mathbb{E}[\hat{M}_n(x, x', u) | \Lambda, G, i] = \quad (11)$$

$$\mathbb{E}[\hat{T}_n(x, u) | \Lambda, G, i] \Lambda_n^0(x, x', u).$$

While the exact solution of the master-equation is a limiting factor for larger graphs, we note that scalable methods that approximate it exist (Cohn et al., 2010; El-Hay et al., 2010; Linzner & Koepl, 2018). For a derivation of the expected moments, see Appendix B.2, for a derivation of the KL-divergence B.1.

Under a Gamma prior for Λ , the posterior $p(\Lambda | \mathcal{H}, G)$, calculated using (9), will be a Gamma distribution due to conjugacy $p(\Lambda | \mathcal{H}, G) = \prod_{n, x, x', u} \text{Gam}(\Lambda_n(x, x', u) | \bar{\alpha}_n(x, x', u), \bar{\beta}_n(x, u))$, with $\bar{\alpha}_n(x, x', u) = \bar{M}_n(x, x', u, i = 0) + \alpha_n(x, x', u)$ the posterior transition counts and $\bar{\beta}_n(x, u) = \bar{T}_n(x, u, i = 0) + \beta_n(x, u)$ the posterior waiting times. Consequently, we choose the same parametric form for $q_\kappa(\Lambda) = \prod_{n, x, x', u} \text{Gam}(\Lambda_n(x, x', u) | \alpha_n^\kappa(x, x', u), \beta_n^\kappa(x, u))$, with $\alpha_n^\kappa(x, x', u)$ and $\beta_n^\kappa(x, u)$ being variational parameters. We finally arrive at the semi-analytical expression for the VBHC in terms of the variational parameters and the expected path-statistics

$$\text{VBHC}(i, \kappa) = \text{KL}(q_\kappa(\Lambda) || p(\Lambda | \mathcal{H}, G)) \quad (12)$$

$$\begin{aligned} - \int d\Lambda p(\Lambda | \mathcal{H}, G) \sum_{n \in \aleph} \sum_{x, x', u} \\ \times \{ \mathbb{E}[\hat{M}_n(x, x', u) | \Lambda, G, i] \\ \times (\ln \Lambda - \psi^{(0)}(\alpha_n^\kappa(x, x', u)) + \ln \beta_n^\kappa(x, u)) \\ - \mathbb{E}[\hat{T}_n(x, u) | \Lambda, G, i] \left(\Lambda - \frac{\alpha_n^\kappa(x, x', u)}{\beta_n^\kappa(x, u)} \right) \}, \end{aligned}$$

where $\psi^{(k)}$ is the di-gamma function of k 'th order. The KL-divergence $\text{KL}(q_\kappa(\Lambda) || p(\Lambda))$ between two Gamma distributions has a closed form expression, – see Appendix

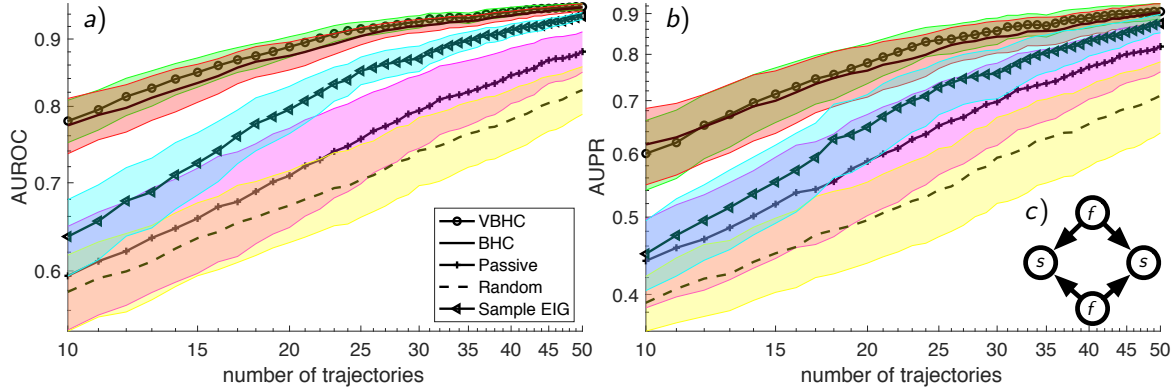


Figure 2. Results of active structure learning from synthetic data. We plotted mean (line) and variance (areas) after repeating structure learning for 500-times for each design. We plotted a) AUROC and b) AUPR for a different number of sequential experiments. In c) we denoted the ground truth graph, where f denotes fast and s slow nodes.

B.3. We can compute the gradients $\partial_{\kappa} \text{VBHC}(i, \kappa)$ analytically, which facilitates optimization of the bound for large systems.

Active Structure Learning. The KL-divergence between two marginal CTBNs reads

$$\begin{aligned} \text{KL} \left(p(S^{[0,T]} | G, i) || p(S^{[0,T]} | G', i) \right) = \\ \sum_{n \in \mathbb{N}} \sum_{x, x'} \sum_{u \in \mathcal{U}_n^G} \sum_{u' \in \mathcal{U}_n^{G'}} \mathbb{E} \left[\mathbb{E} \left[\ln \frac{\Gamma(\tilde{\alpha}_n(x, x', u))}{\Gamma(\tilde{\alpha}_n(x, x', u'))} \middle| \Lambda, G, i \right] \right] \\ + \mathbb{E} \left[\mathbb{E} \left[\tilde{\alpha}_n(x, x', u') \ln \tilde{\beta}_n(x, u') \middle| \Lambda, G, i \right] \right] \\ - \mathbb{E} \left[\mathbb{E} \left[\tilde{\alpha}_n(x, x', u) \ln \tilde{\beta}_n(x, u) \middle| \Lambda, G, i \right] \right], \end{aligned}$$

with the outer expectation w.r.t $p(\Lambda | \mathcal{H}, G)$ and the inner one w.r.t the path-likelihood $p(S^{[0,T]} | \Lambda, G, i, s_0)$. Further we defined short-hands $\tilde{\alpha}_n(x, x', u) \equiv \hat{M}_n(x, x', u) + \alpha_n(x, x', u)$ and $\tilde{\beta}_n(x, u) \equiv \hat{T}_n(x, u) + \beta_n(x, u)$. The posterior over structures can be computed in closed form, see (8), by marginalization $p(G | \mathcal{H}) = \int d\Lambda p(\Lambda | \mathcal{H}, G) p(G)$, with any prior $p(G)$ (we assume a uniform categorical). It is natural to assume a categorical distribution $q_{\kappa}(G) = \text{Cat}(\kappa_G)$. Then the VBHC for structure learning reads

$$\begin{aligned} \text{VBHC}(i, \kappa) = \text{KL} (q_{\kappa}(G) || p(G | \mathcal{H})) \\ - \mathbb{E}[\mathbb{E}[\text{KL}(p(S^{[0,T]} | G, i) || p(S^{[0,T]} | G', i))], \end{aligned} \quad (13)$$

with the inner expectation w.r.t $q_{\kappa}(G)$, the outer w.r.t $p(G' | \mathcal{H})$. As the moments $\mathbb{E}[\ln \hat{M}_n | \Lambda, G, i]$ and $\mathbb{E}[\ln \hat{T}_n | \Lambda, G, i]$ can not be evaluated in closed form, we approximate the KL-divergence by a first-order expansion around the moments (10) and (11), see Appendix B.4 for details.

Discussion. We emphasize, that the expressions (12) and (13), enable us to perform active parameter and structure learning in high-dimensions, as the integration over outcomes is performed analytically. As mentioned in section 2, the corresponding BHC is computed by setting $q_{\kappa}(\Lambda) = p(\Lambda | \mathcal{H}, G)$ in (12) and $q_{\kappa}(G) = p(G | \mathcal{H})$ in (13). For the (V)BHC for parameter- and structure learning, only the integral over $p(\Lambda | \mathcal{H}, G)$ can not be evaluated analytically, as for each realization of Λ , the full master-equation needs to be solved. We thus approximate this integral by N_S posterior samples. We summarized the computational steps mentioned above in algorithmic format in Appendix A.

To highlight the computational benefit of the (V)BHC, we compare inference complexity in the number of nodes N and the cardinality of local state spaces $|\mathcal{X}|$. For the EIG, one needs to integrate the posterior over parameters for all possible paths the network can take. The complexity here is two-fold: (i) The number of possible paths scales as $|\mathcal{X}|^N$ and (ii) the complexity of a posterior-update, which is dominated by computing conditional summary statistics for transitions is $N|\mathcal{X}|^N$. Thus, in total the complexity of calculating the EIG is given by the product $N|\mathcal{X}|^{2N}$. For the (V)BHC calculation of the conditional summary statistics is only performed once. Thus, we have a sum instead of a product $N|\mathcal{X}|^N + |\mathcal{X}|^N$.

6. Experiments

We evaluate the performance of interventions selected by different designs for sequential experiments. In all considered scenarios, interventions are tuples of node and state pairs, $i = ((m, x_m), (j, x_j), \dots)$ with $m, j \in \{1, \dots, N\}$ and $x_m \in \mathcal{X}_m, x_j \in \mathcal{X}_j$, which corresponds to a do-operation

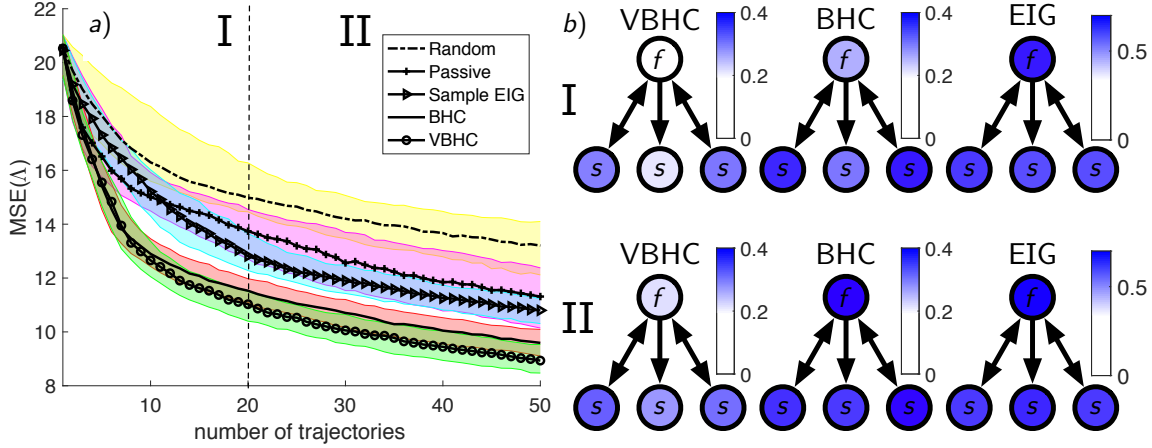


Figure 3. a) Results of active parameter learning in from synthetic data. We plotted mean (line) and 25-75% confidence intervals of the mean-squared error of the estimated rates for various experimental designs after repeating parameter learning for 500 times. In b) we denote the graph, where f denotes fast and s slow nodes. Color-scales denote the marginal probability of intervening on this node in the k 'th experiment for regions I and II, see a).

of setting the m 'th node into state x_m and the j 'th node into state x_j . In the following, we search the optimal intervention in the set of all possible interventions of this type. In order to select the optimal intervention, we compare the EIG (4) and the (V)BHC (5) given (13) for structure and (12) for parameter learning, and also compare to random or no interventions (passive). Neither the EIG, nor the (V)BHC can be computed completely in analytical form. We employ a (nested) monte-carlo scheme for the computation of the EIG, see Appendix A. For the computation of the (V)BHC, we approximate the integration over rates via posterior sampling. We discuss using sample estimates of design criteria in Appendix C.1. For the problem sizes considered, the (exact) solution of the master-equation has a negligible contribution to the run-times, compared to the repeated posterior computations. Thus, we throughout compare EIG and (V)BHC given the same number of posterior samples N_S . Minimization of the VBHC is feasible using standard Matlab optimizers as gradients $\partial_{\kappa} \text{VBHC}(i, \kappa)$ are computed analytically.

6.1. Synthetic Data

For synthetic experiments, we learn structures and parameters from randomly generated CTBNs with $L = 4$ nodes with binary state-spaces. As the amount of data needed to identify structures and parameters varies strongly with the underlying ground truth, we display results for fixed ground truths. We generate problems, which contain features that can be identified and exploited by an active learning scheme. As shown in this manuscript, one way to do this is by inducing time-scale separation. For structure and parameter learning, we thus create problems containing fast and slow

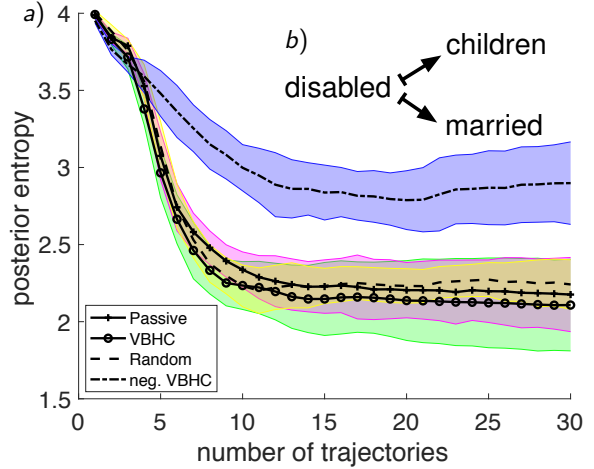


Figure 4. a) Mean and variance (area) of the evolution of the posterior entropy in BHPS data-set for 100 repetitions. b) Sketch of the underlying network, where "disabled" is a root node.

nodes. In both cases, we exhaustively search in the space of all possible interventions, including targeting multiple nodes or performing no interventions. For each trajectory, an initial state s_0 is drawn at random, which is accessible to the design (although a distribution of initial states could also be learned). We set $N_S = 10$ and the length of each trajectory fixed to be $\tau = 3$ a.u..

Active Structure Learning. We generate ground truth rates by sampling rate-parameters from independent Gamma-distributions $\Lambda^* \sim \text{Gam}(\Lambda \mid \alpha_{f/s}, \beta_{f/s})$ with two differ-

ent types hyper-parameters f and s , referring to fast and slow nodes, respectively. We used $\alpha_f = 5$, $\alpha_s = 1/5$ and $\beta_f = \beta_s = 1$. We then perform structure learning via exhaustive scoring as in (Nodelman et al., 2003). As a metric for structure learning, we employ the area under the Receiver-Operator-Characteristic curve (AUROC) and Precision-Recall curve (AUPR), which are frequently used to quantify the performance of (structure) classifiers. For an unbiased classifier, both metrics should approach one in the limit of infinite data. In figure 2 a) and b), respectively, we show the results for structure learning. As AUROC and AUPR can only be calculated w.r.t a ground truth, they can not be made an objective for a design.

Active Parameter Learning. For parameter learning, we fix the ground truth structure and the ground truth rates across experiments. Following CTBN literature (Cohn et al., 2010; El-Hay et al., 2011), we chose the ground truth rates Λ^* as scaled softmax functions $\Lambda_n^*(x, x', u) = r_{f/s} \text{softmax}(\gamma \sum_{y \in u} \mathbb{1}(x = y))$ for the n 'th node, with $\gamma = 3$ and $r_s = 1/5$ and $r_f = 5$ for slow and fast nodes, respectively. We quantify the performance of parameter learning by the posterior averaged mean-squared-error (MSE), defined by $\text{MSE}(\Lambda) = E[(\Lambda - \Lambda^*)^2]$, with the expectation subject to the current rate-posterior $p(\Lambda \mid \mathcal{H}, G)$. The results of this experiment are shown in figure 3 a). Similar to the case of structure learning, the ground truth rate-matrix is unavailable during the experiment, and the MSE can thus not be used as a design objective. We also investigate the qualitative behavior of the designs in figure 3 b), where we provide the marginal probability of intervening on a node in the k 'th step of the experiment. For the (V)BHC it can be seen, that recommended interventions follow the intuition of exploiting time-scale separation, as interventions targeting slow nodes pointing on the fast node have higher probability.

6.2. Real-World Data

British Household Data-set. We apply our method to the British Household Panel Survey (BHPS) (on [Micro-social Change](#), 2003). This data-set has been collected yearly from 1991 to 2002, thus consisting of 11 time-points. Each of the 1535 participants was questioned about several facts (variables) of their life. We choose 3 variables "marital status", "has children under 12" and "is registered disabled" of those facts. We had no means of intervening on this data-set directly, but can use it as a proof of concept in order to verify our method on real world data. For this, we identified "is registered disabled" as a root node, meaning there is no edge from "marital status" or "has children under 12" to this node. We confirmed this by performing network inference on the full data-set, see inlet 4 b). We then selected for a subset where the variable "is registered disabled" remains in one state during the complete trajectory. We interpreted these

cases as interventions, which is valid as conditioning on a root node is the same as intervening on a root node. We then simulated an experiment, where one either draws a data-point from the full-set or from the subset of interventions on "is registered disabled". We performed this experiment for passive, random, VBHC designs and negative VBHC design (always pick the worst possible intervention) for structure learning with $N_S = 40$ posterior samples as recommenders for interventions. For details on the processing of this (incomplete) data-set, we refer to the Appendix C.2. As for the real-world data-set no ground truth structure is available, we track the evolution of the posterior entropy over structures for 100 independent runs, see figure 4 a). In Appendix C.3, we show that for all designs the inferred network converge against the one inferred using the full data-set (using AUROC and AUPR as metrics). We note that the effect of active learning can be expected to be small in this synthetic scenario, as we were only able to intervene on a single node.

7. Conclusion

We presented a novel variational criterion for active learning, that alleviates the curse of dimensionality when integrating over all possible experimental outcomes. We presented cCTBNs, a framework to study the effect of interventions on CTBNs. We have shown that our novel criterion can be calculated semi-analytically for cCTBNs and can be used to recommend interventions that speed-up structure and parameter learning on synthetic and real-world data. In this manuscript, we performed exact inference and exhaustive search, limiting us to small graphs. However, many principled approximation techniques for inference (Cohn et al., 2010; El-Hay et al., 2010; 2011; Rao & Teh, 2012), and structure learning (Nodelman et al., 2005; Linzner et al., 2019) are available that allow learning of large scale CTBNs and are compatible with our framework.

Acknowledgements

We thank the anonymous reviewers for helpful comments on the previous version of this manuscript. D. L. and H. K. acknowledge funding by the European Union's Horizon 2020 research and innovation programme (iPC-Pediatric Cure, No. 826121). H. K. acknowledges support by the Hessian research priority programme LOEWE within the project CompuGene.

References

Acerbi, E., Zelante, T., Narang, V., and Stella, F. Gene network inference using continuous time bayesian networks: a comparative study and application to th17 cell differentiation. *BMC Bioinformatics*, 15, 2014.

- Box, G. E. and Hill, W. J. Discrimination among mechanistic models. *Technometrics*, 9(1):57–71, 1967. ISSN 15372723.
- Chaloner, K. and Verdinelli, I. Bayesian experimental design: A review. *Statistical Science*, 2(1):45–54, 1987. ISSN 08834237.
- Cohn, I., El-Hay, T., Friedman, N., and Kupferman, R. Mean field variational approximation for continuous-time bayesian networks. *Journal Of Machine Learning Research*, 11:2745–2783, 2010.
- Daniel, R. D., Steinberg, D. M., and Box, G. Follow-up designs to resolve confounding in multifactor experiments. *Technometrics*, 38(4):303–313, 1996. ISSN 15372723.
- Dehghannasiri, R., Yoon, B. J., and Dougherty, E. R. Optimal experimental design for gene regulatory networks in the presence of uncertainty. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 12(4):938–950, 2015. ISSN 15455963.
- Eaton, D. and Murphy, K. Exact bayesian structure learning from uncertain interventions. In *Journal of Machine Learning Research*, volume 2, pp. 107–114, 2007.
- Eberhardt, F. A sufficient condition for pooling data. *Synthese*, 163(3):433–442, 2008. ISSN 00397857, 15730964.
- El-Hay, T., Cohn, I., Friedman, N., and Kupferman, R. Continuous-time belief propagation. *Proceedings of the 27th International Conference on Machine Learning*, pp. 343–350, 2010.
- El-Hay, T., Kupferman, R., and Friedman, N. Gibbs sampling in factorized continuous-time markov processes. *Proceedings of the 22th Conference on Uncertainty in Artificial Intelligence*, 2011.
- Foster, A., Jankowiak, M., Bingham, E., Horsfall, P., Teh, Y. W., Rainforth, T., and Goodman, N. Variational bayesian optimal experimental design. *Advances in Neural Information Processing Systems*, (NeurIPS):1–18, 2019.
- He, Y. B. and Geng, Z. Active learning of causal networks with intervention experiments and optimal designs. *Journal of Machine Learning Research*, 9:2523–2547, 2008. ISSN 15324435.
- Koller, D. and Friedman, N. *Probabilistic graphical models principles and techniques*. MIT Press, 2010. ISBN 0262013193.
- Lewi, J., Butera, R., and Paninski, L. Sequential optimal design of neurophysiology experiments. *Neural Computation*, (21):619–687, 2009.
- Liepe, J., Filippi, S., Komorowski, M., and Stumpf, M. P. Maximizing the information content of experiments in systems biology. *PLoS Computational Biology*, 9(1), 2013. ISSN 1553734X.
- Lindgren, E. M., Dimakis, A. G., Kocaoglu, M., and Vishwanath, S. Experimental design for cost-aware learning of causal graphs. *Advances in Neural Information Processing Systems*, (NeurIPS):5279–5289, 2018. ISSN 10495258.
- Lindley, D. V. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 27(4):986–1005, 1956. ISSN 0003-4851.
- Linzner, D. and Koepl, H. Cluster variational approximations for structure learning of continuous-time bayesian networks from incomplete data. *Advances in Neural Information Processing Systems*, (NeurIPS):7880–7890, 2018.
- Linzner, D., Schmidt, M., and Koepl, H. Scalable structure learning of continuous-time bayesian networks from incomplete data. *Advances in Neural Information Processing Systems*, (NeurIPS):1–11, 2019.
- Myung, J. I. and Pitt, M. A. Optimal experimental design for model discrimination. *Psychological Review*, 135(2): 612–615, 2015. ISSN 15231747.
- Ng, S. H. and Chick, S. E. Design of follow-up experiments for improving model discrimination and parameter estimation. *Naval Research Logistics*, 51(8):1129–1148, 2004. ISSN 0894069X.
- Nodelman, U., Shelton, C. R., and Koller, D. Continuous time bayesian networks. *Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence*, pp. 378–387, 1995.
- Nodelman, U., Shelton, C. R., and Koller, D. Learning continuous time bayesian networks. *Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence*, pp. 451–458, 2003.
- Nodelman, U., Shelton, C. R., and Koller, D. Expectation maximization and complex duration distributions for continuous time bayesian networks. *Proc. Twenty-first Conference on Uncertainty in Artificial Intelligence*, pp. pages 421–430, 2005.
- Norris, J. R. *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.
- on Micro-social Change, E. R. C. British household panel survey. <http://iserwww.essex.ac.uk/bhps>. Colchester: The Data Archive, 2003.

- Opper, M. and Sanguinetti, G. Variational inference for markov jump processes. *Advances in Neural Information Processing Systems 20*, pp. 1105–1112, 2008.
- Opper, M. and Winther, O. Tractable approximations for probabilistic models: The adaptive thouless-anderson-palmer mean field approach. *Physical Review Letters*, 86(17):3695–3699, 2001.
- Pearl, J. *Causality Second Edition*. 2000. ISBN 0521773628.
- Poole, B., Ozair, S., Van Den Oord, A., Alemi, A. A., and Tucker, G. On variational bounds of mutual information. In *36th International Conference on Machine Learning, ICML 2019*, volume 2019-June, pp. 9036–9049, 2019. ISBN 9781510886988.
- Prangemeier, T., Wildner, C., Hanst, M., and Koepl, H. Maximizing information gain for the characterization of biomolecular circuits. In *Proceedings of the 5th ACM International Conference on Nanoscale Computing and Communication, NANOCOM '18*, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450357111.
- Rainforth, T., Cornish, R., Yang, H., Warrington, A., and Wood, F. On nesting monte carlo estimators. *35th International Conference on Machine Learning, ICML 2018*, 10:6789–6817, 2018.
- Rao, V. and Teh, Y. W. Fast mcmc sampling for markov jump processes and extensions. *Journal of Machine Learning Research*, 14:3295–3320, 2012. ISSN 15324435.
- Reilly, P. M. Statistical methods in model discrimination. *The Canadian Journal of Chemical Engineering*, 48(2): 168–173, 1970. ISSN 00084034.
- Rubenstein, P. K., Tolstikhin, I., Hennig, P., and Schoelkopf, B. Probabilistic active learning of functions in structural causal models. <http://arxiv.org/abs/1706.10234>, 2017.
- Ryan, E. G., Drovandi, C. C., McGree, J. M., and Pettitt, A. N. A review of modern computational algorithms for bayesian optimal design. *International Statistical Review*, 84(1):128–154, 2016. ISSN 17515823.
- Schadt, E. E., Lamb, J., Yang, X., Zhu, J., Edwards, S., Thakurta, D. G., Sieberts, S. K., Monks, S., Reitman, M., Zhang, C., Lum, P. Y., Leonardson, A., Thieringer, R., Metzger, J. M., Yang, L., Castle, J., Zhu, H., Kash, S. F., Drake, T. A., Sachs, A., and Lusi, A. J. An integrative genomics approach to infer causal associations between gene expression and disease. *Nature Genetics*, 37(7): 710–717, jul 2005.
- Sebastiani, P. and Wynn, H. P. Maximum entropy sampling and optimal bayesian experimental design. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 62(1):145–157, 2000. ISSN 13697412.
- Soch, J. and Allefeld, C. Kullback-leibler divergence for the normal-gamma distribution. pp. 0–10, 2016.
- Spirtes, P. Introduction to causal inference. *Journal of Machine Learning Research*, 11:1643–1662, 2010. ISSN 00491241.
- Steinke, F., Seeger, M., and Tsuda, K. Experimental design for efficient identification of gene regulatory networks using sparse bayesian models. *BMC Systems Biology*, 1: 1–15, 2007. ISSN 17520509.
- Studer, L., Zechner, C., Reumann, M., Pauleve, L., Martinez, M. R., and Koepl, H. Marginalized continuous time bayesian networks for network reconstruction from incomplete observations. *Proceedings of the 30th Conference on Artificial Intelligence (AAAI 2016)*, pp. 2051–2057, 2016.
- Tong, S. and Koller, D. Active learning for structure in bayesian networks. *IJCAI International Joint Conference on Artificial Intelligence*, pp. 863–869, 2001. ISSN 10450823.
- Zechner, C., Nandy, P., Unger, M., and Koepl, H. Optimal variational perturbations for the inference of stochastic reaction dynamics. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pp. 5336–5341, 2012.

Supplement: Active Learning of Continuous-time Bayesian Networks through Interventions

A. Algorithms

Algorithm 1 Computation of VBHC for Parameter Learning

Input: Proposed intervention i , current initial state s_0 , desired number of posterior-samples N_S , current belief over parameters $p(\Lambda \mid \mathcal{H}, G)$, initial variational parameters $\alpha^{\kappa,0}, \beta^{\kappa,0}$. As initial values for the optimization, we set the initial values for the optimization to the posterior counts $\alpha^{\kappa,0} = \bar{\alpha}, \beta^{\kappa,0} = \bar{\beta}$, see main-text.

for $n_S = 1 : N_S$ **do**

Draw $\hat{\Lambda}_{n_S} \sim p(\Lambda \mid \mathcal{H}, G)$.

Perform intervention by setting $\hat{\Lambda}_{n_S, n} = 0$ and initial state x_n^0 for all $n \notin \mathbb{N}$.

Calculate W from $\hat{\Lambda}_{n_S}$ and G by amalgamation.

Solve the master-equation main-text (3) subject to W and s_0 and recover $E[\hat{M}(s, s')]$ and $E[\hat{T}(s)]$ using appendix (19) and (18) respectively.

for $n = 1 : N$ **do**

Compute expected statistics $E[\hat{M}_n \mid \hat{\Lambda}_{n_S}, G, i]$ and $E[\hat{T}_n \mid \hat{\Lambda}_{n_S}, G, i]$ from $E[\hat{M}(s, s')]$ and $E[\hat{T}(s)]$, see appendix (16) and (17).

end for

end for

Calculate VBHC(i, κ) via main-text (12) and gradients appendix (20) and (21) with weighted posterior samples replacing $\int dp(\Lambda \mid \mathcal{H}, G)$ with $\sum_{n_S=1}^{N_S} p(\hat{\Lambda}_{n_S} \mid \mathcal{H}, G)$.

Minimize w.r.t κ .

Output: $\min_{\kappa} \text{VBHC}(i, \kappa)$.

Algorithm 2 Computation of VBHC for Structure Learning

1: **Input:** Proposed intervention i , current initial state s_0 , desired number of posterior-samples N_S , current belief over parameters $p(\Lambda \mid \mathcal{H}, G)$ and structures $p(G \mid \mathcal{H})$, initial variational parameters κ .

2: **for** $n = 1 : N$ **do**

3: **for** $n_S = 1 : N_S$ **do**

4: **for** $\text{par}(n)$ in power-set of \mathcal{V} **do**

5: Draw $\hat{G}_{n_S} \sim p(G \mid \mathcal{H}, \text{par}(n))$.

6: Draw $\hat{\Lambda}_{n_S} \sim p(\Lambda \mid \mathcal{H}, \hat{G}_{n_S})$.

7: Perform intervention by setting $\hat{\Lambda}_{n_S, n} = 0$ and initial state x_n^0 for all $n \notin \mathbb{N}$.

8: Calculate W from $\hat{\Lambda}_{n_S}$ and \hat{G}_{n_S} by amalgamation.

9: Solve the master-equation main-text (3) subject to W and recover $E[\hat{M}(s, s')]$ and $E[\hat{T}(s)]$.

10: Compute expected statistics $E[\hat{M}_n \mid \hat{\Lambda}_{n_S}, \hat{G}_{n_S}, i]$ and $E[\hat{T}_n \mid \hat{\Lambda}_{n_S}, \hat{G}_{n_S}, i]$ from $E[\hat{M}(s, s')]$ and $E[\hat{T}(s)]$, see appendix (16) and (17).

11: **end for**

12: **end for**

13: **end for**

14: Calculate VBHC(i, κ) using appendix (26) and gradients appendix (27) with weighted posterior samples replacing $\sum_{G \mid \text{par}(n)} \sum_{G' \mid \text{par}'(n)} p(G \mid \mathcal{H}) q(G')$ with $\sum_{n_S, n'_S} p(\hat{G}_{n_S} \mid \mathcal{H}) q(\hat{G}_{n'_S}) p(\hat{\Lambda}_{n_S} \mid \hat{G}_{n_S})$.

15: Minimize w.r.t κ .

16: **Output:** $\min_{\kappa} \text{VBHC}(i, \kappa)$.

Algorithm 3 Computation of the EIG for Parameter Learning

- 1: **Input:** Proposed intervention i , current initial state s_0 , desired number of posterior-samples N_S , number of path samples N_P , current belief over parameters $p(\Lambda \mid \mathcal{H}, G)$.
 - 2: Set EIG = 0.
 - 3: **for** $n_S = 1 : N_S$ **do**
 - 4: Draw parameter $\hat{\Lambda}_{n_S} \sim p(\Lambda \mid \mathcal{H}, G)$
 - 5: Perform intervention by setting $\hat{\Lambda}_{n_S, n} = 0$ and initial state x_n^0 for all $n \notin \aleph$.
 - 6: **for** $n_P = 1 : N_P$ **do**
 - 7: Draw path $\hat{S}^{[0, T]} \sim p(S^{[0, T]} \mid \hat{\Lambda}_{n_S}, G, i, s_0)$.
 - 8: Set EIG = EIG + $\frac{1}{N_P N_S} \left(\ln p(\hat{\Lambda}_{n_S} \mid \hat{S}^{[0, T]}, \mathcal{H}) - \ln p(\hat{\Lambda}_{n_S} \mid \mathcal{H}) \right)$.
 - 9: **end for**
 - 10: **end for**
 - 11: **Output:** Estimate EIG.
-

Algorithm 4 Computation of the EIG for Structure Learning

- 1: **Input:** Proposed intervention i , current initial state s_0 , number of path samples N_S , current belief over parameters $p(\Lambda \mid \mathcal{H}, G)$ and structures $p(G \mid \mathcal{H})$.
 - 2: Set EIG = 0.
 - 3: **for** $n = 1 : N$ **do**
 - 4: **for** $\text{par}(n)$ in power-set of \mathcal{V} **do**
 - 5: **for** $n_s = 1 : N_S$ **do**
 - 6: Draw $\hat{G}_{n_s} \sim p(G \mid \mathcal{H}, \text{par}(n))$.
 - 7: Draw parameter $\hat{\Lambda}_{n_s} \sim p(\Lambda \mid \mathcal{H}, \hat{G}_{n_s})$.
 - 8: Perform intervention by setting $\hat{\Lambda}_{n_s, n} = 0$ and initial state x_n^0 for all $n \notin \aleph$.
 - 9: Draw path $\hat{S}^{[0, T]} \sim p(S^{[0, T]} \mid \hat{\Lambda}_{n_s}, \hat{G}_{n_s}, i, s_0)$.
 - 10: Set EIG = EIG + $\frac{1}{N_S} \left(\ln p(\text{par}(n) \mid \mathcal{H}, \hat{S}^{[0, T]}) - \ln p(\text{par}(n) \mid \mathcal{H}) \right)$,
see appendix (25).
 - 11: **end for**
 - 12: **end for**
 - 13: **end for**
 - 14: **Output:** Estimate EIG.
-

B. Derivations

All derivations are done for a fixed set of conditions i and respective initial states s_0 . We will omit those in the following derivations for readability.

B.1. Kullback–Leibler divergence between two CTBNs

Evaluation of our design criteria (V)BHC, requires the calculation of the KL-divergence between two cCTBNs.

The likelihood of observing a CTBN path $D = S^{[0,T]}$ is (expressed in terms of its sufficient statistics)

$$p(S^{[0,T]} | \Lambda, G) = \prod_n \prod_{x, x', u} \Lambda_n(x, x', u)^{M_n(x, x', u)} \exp \{-\Lambda_n(x, x', u) T_n(x, u)\}. \quad (14)$$

The KL between two measures is defined via the integration over all paths

$$\text{KL} \left(p(S^{[0,T]} | \Lambda, G) || p(S^{[0,T]} | \Lambda', G) \right) = \int dp(S^{[0,T]} | \Lambda, G) \ln \frac{p(S^{[0,T]} | \Lambda, G)}{p(S^{[0,T]} | \Lambda', G)}.$$

Inserting (14) yields

$$\begin{aligned} \text{KL} \left(p(S^{[0,T]} | \Lambda, G) || p(S^{[0,T]} | \Lambda', G) \right) &= \sum_{n, x, x' \neq x, u} \quad (15) \\ \{ \Lambda'_n(x, x', u) - \Lambda_n(x, x', u) \} \mathbb{E} \left[\hat{T}_n(x, u) | \Lambda, G \right] &- \ln \frac{\Lambda_n(x, x', u)}{\Lambda'_n(x, x', u)} \mathbb{E} \left[\hat{M}_n(x, x', u) | \Lambda, G \right], \end{aligned}$$

with the expectations being taken with respect to the process $p(S^{[0,T]} | \Lambda, G)$. The expected moments can not be calculated from the parametric form of $p(S^{[0,T]} | \Lambda, G)$ directly. Instead, we will construct an ODE for the moments of the CTMC recovered after amalgamation, and recover its expectation as solutions. The moments of the CTBN can then be calculated as projections of the CTMC moments, the dwelling times per state $T(s)$ and the number of transitions $M(s, s')$

$$\hat{M}_n(x, x', u) = \sum_{s, s'} M(s, s') \mathbb{1}(s'_n = x') \mathbb{1}(s_n = x) \mathbb{1}(s_{\text{par}(n)} = u), \quad (16)$$

$$\hat{T}_n(x, u) = \sum_s T(s) \mathbb{1}(s_n = x) \mathbb{1}(s_{\text{par}(n)} = u). \quad (17)$$

B.2. Moment ODEs of a CTMC

Expected Dwelling-times. The expected dwelling-times $\mathbb{E} [T(s)]$ in a state $s \in \mathcal{S}$ of a CTMC are calculated as solution of an ODE. For this, we need to consider the evolution of the stochastic process $T(s, t)$, the dwelling times in state $s \in \mathcal{S}$ up to time t . For this process we can denote transition probabilities, by considering the dynamics of the CTMC

$$\begin{aligned} p(T(s, t+h) = \tau + h | T(s, t) = \tau) &= p(S(t+h) = s, S(t) = s), \\ p(T(s, t+h) = \tau | T(s, t) = \tau) &= 1 - p(S(t+h) = s, S(t) = s). \end{aligned}$$

Thus $T(s, t)$ evolves according to

$$\begin{aligned} p(T(s, t+h) = \tau) &= p(S(t+h) = s, S(t) = s) p(T(s, t) = \tau - h) \\ &+ [1 - p(S(t+h) = s, S(t) = s)] p(T(s, t) = \tau). \end{aligned}$$

For small h , we can expand $p(T(s, t) = \tau - h) = p(T(s, t) = \tau) - h \partial_\tau p(T(s, t) = \tau) + o(h)$, and we arrive at

$$\frac{p(T(s, t+h) = \tau) - p(T(s, t) = \tau)}{h} = -p(S(t) = s) \partial_\tau p(T(s, t) = \tau) + o(h).$$

We can now take the expectation $\mathbb{E} [T(s, t)] = \int_0^\infty d\tau' \tau' p(T(s, t) = \tau')$, and the continuum limit $h \rightarrow 0$ in order to arrive at

$$\partial_t \mathbb{E} [T(s, t)] = -p(S(t) = s) \int_0^\infty d\tau' \tau' \partial_{\tau'} p(T(s, t) = \tau'),$$

which, after integration by parts, reduces to simply

$$\partial_t \mathbb{E} [T(s, t)] = p(S(t) = s).$$

Thus, the expected dwelling-time is given by the solution

$$\mathbb{E} [T(s)] = \int_0^T dt p(S(t) = s). \quad (18)$$

Expected Number of Transitions. Similarly to above, we can compute the expected number of transitions of a CTMC $\mathbb{E} [M(s, s')]$. The computation is analogous to above. We consider the stochastic process $M(s, s', t)$ of transitions from s to s' till time t . Transition probabilities are

$$\begin{aligned} p(M(s, s', t+h) = k \mid M(s, s', t) = k-1) &= p(S(t+h) = s', S(t) = s), \\ p(M(s, s', t+h) = k \mid M(s, s', t) = k) &= 1 - p(S(t+h) = s', S(t) = s). \end{aligned}$$

After inserting the identity $p(S(t+h) = s', S(t) = s) = \mathbb{1}(s = s') + hW(s, s') + o(h)$, we arrive at

$$\begin{aligned} &\frac{p(M(s, s', t+h) = k) - p(M(s, s', t) = k)}{h} \\ &= W(s, s')p(S(t) = s) [p(M(s, s', t) = k-1) - p(M(s, s', t) = k)] + o(h). \end{aligned}$$

The expected number of transitions can be calculated via $\mathbb{E} [M(s, s', t)] = \sum_{k=0}^{\infty} p(M(s, s', t) = k)$. Noticing that $p(M(s, s', t) = k-1) = 0$ for $k < 1$, we can perform an index-shift $k \rightarrow k+1$ and arrive at

$$\begin{aligned} &\frac{\mathbb{E} [M(s, s', t+h)] - \mathbb{E} [M(s, s', t)]}{h} \\ &= W(s, s')p(S(t) = s) [\mathbb{E} [M(s, s', t)] - \mathbb{E} [M(s, s', t)] + 1] + o(h), \end{aligned}$$

and thus in the continuum limit $h \rightarrow 0$ we recover the ODE,

$$\partial_t \mathbb{E} [M(s, s', t)] = W(s, s')p(S(t) = s),$$

with the solution

$$\mathbb{E} [M(s, s')] = W(s, s')\mathbb{E} [T(s)]. \quad (19)$$

B.3. (V)BHC for Parameter Learning

Equipped with the moments derived in the last Section, we can now derive the (V)BHC. The VBHC takes the form of an expected KL-divergence

$$\begin{aligned} \text{VBHC} &= \int d\Lambda \int d\Lambda' p(\Lambda \mid \mathcal{H}, G) q_{\kappa}(\Lambda') \text{KL} \left(p(S^{[0,T]} \mid \Lambda, G) \parallel p(S^{[0,T]} \mid \Lambda', G) \right) \\ &\quad + \text{KL} (q_{\kappa}(\Lambda) \parallel p(\Lambda \mid \mathcal{H}, G)), \end{aligned}$$

with the KL given in appendix (15). As explained in the main-text, we have $p(\Lambda \mid \mathcal{H}, G) = \prod_{n,x,x',u} \text{Gam} (\Lambda_n(x, x', u) \mid \bar{\alpha}_n(x, x', u), \bar{\beta}_n(x, u))$ and choose $q_{\kappa}(\Lambda') = \prod_{n,x,x',u} \text{Gam} (\Lambda_n(x, x', u) \mid \alpha_n^{\kappa}(x, x', u), \beta_n^{\kappa}(x, u))$. As the expected moments in (15) only depend on Λ , we can calculate the integral over Λ' analytically. For this, we notice that the moments

$$\begin{aligned} \mathbb{E} [\Lambda_n(x, x', u)] &= \frac{\alpha_n^{\kappa}(x, x', u)}{\beta_n^{\kappa}(x, u)}, \\ \mathbb{E} [\ln \Lambda_n(x, x', u)] &= \psi^{(0)}(\alpha_n^{\kappa}(x, x', u)) - \ln \beta_n^{\kappa}(x, u), \end{aligned}$$

where the expectation is w.r.t $q_\kappa(\Lambda)$, have a closed form expression. By insertion into (15), we recover the expression from the main-text. Finally, we notice that

$$\text{KL} \left(q_\kappa(\Lambda') \parallel p(\Lambda \mid \mathcal{H}, G) \right) = \sum_{n, x, x', u} \text{KL} \left(\text{Gam}(\alpha_n^\kappa(x, x', u), \beta_n^\kappa(x, u)) \parallel \text{Gam}(\bar{\alpha}_n(x, x', u), \bar{\beta}_n(x, u)) \right),$$

with the KL-divergence between two gamma-distributions (Soch & Allefeld, 2016)

$$\begin{aligned} & \text{KL} \left(\text{Gam}(\alpha_n^\kappa(x, x', u), \beta_n^\kappa(x, u)) \parallel \text{Gam}(\bar{\alpha}_n(x, x', u), \bar{\beta}_n(x, u)) \right) = \\ & \bar{\alpha}_n(x, x', u) \ln \left(\frac{\beta_n^\kappa(x, u)}{\bar{\beta}_n(x, u)} \right) - \ln \left(\frac{\Gamma(\alpha_n^\kappa(x, x', u))}{\Gamma(\bar{\alpha}_n(x, x', u))} \right) \\ & + (\alpha_n^\kappa(x, x', u) - \bar{\alpha}_n(x, x', u)) \psi(\alpha_n^\kappa(x, x', u)) - (\beta_n^\kappa(x, u) - \bar{\beta}_n(x, u)) \frac{\alpha_n^\kappa(x, x', u)}{\beta_n^\kappa(x, u)}. \end{aligned}$$

Gradients. The gradients of the VBHC can be calculated in (semi-)analytical form

$$\begin{aligned} \partial_{\alpha_n^\kappa(x, x', u)} \text{VBHC} &= \int d\Lambda p(\Lambda \mid \mathcal{H}, G) \mathbb{E} \left[\hat{T}_n \mid \Lambda, G \right] \left\{ \Lambda \psi^{(1)}(\alpha_n^\kappa(x, x', u)) - \frac{1}{\beta_n^\kappa(x, u)} \right\} \\ &+ \alpha_n^\kappa(x, x', u) \psi^{(1)}(\alpha_n^\kappa(x, x', u)) - \frac{(\beta_n^\kappa(x, u) - \bar{\beta}_n(x, u))}{\beta_n^\kappa(x, u)} \end{aligned} \quad (20)$$

$$\begin{aligned} \partial_{\beta_n^\kappa(x, u)} \text{VBHC} &= \int d\Lambda p(\Lambda \mid \mathcal{H}, G) \mathbb{E} \left[\hat{T}_n \mid \Lambda, G \right] \left\{ \frac{\alpha_n^\kappa(x, x', u)}{\beta_n^\kappa(x, u)^2} - \frac{\Lambda}{\beta_n^\kappa(x, u)} \right\} \\ &+ \frac{\bar{\alpha}_n(x, x', u)}{\beta_n^\kappa(x, u)} - \frac{\alpha_n^\kappa(x, x', u)}{\beta_n^\kappa(x, u)} + (\beta_n^\kappa(x, u) - \bar{\beta}_n(x, u)) \frac{\alpha_n^\kappa(x, x', u)}{\beta_n^\kappa(x, u)^2}. \end{aligned} \quad (21)$$

If necessary, also higher-order derivatives can be computed in principle.

In all results above, the corresponding BHC expressions are recovered by setting $\beta_n^\kappa(x, u) = \bar{\beta}_n(x, u)$ and $\alpha_n^\kappa(x, x', u) = \bar{\alpha}_n(x, x', u)$.

B.4. (V)BHC for Structure Learning

KL-divergence between Marginal CTBNs. The marginal likelihood of a path $\hat{S}^{[0, T]} \sim p(\hat{S}^{[0, T]} \mid \Lambda, G)$, with statistics $\hat{T}_n(x, u)$ and $\hat{M}_n(x, x', u)$, given a structure and history \mathcal{H} can be calculated via marginalization of (14)

$$\begin{aligned} p(\hat{S}^{[0, T]} \mid G, \mathcal{H}) &= \int d\Lambda p(\Lambda \mid \mathcal{H}, G) \prod_n \prod_{x, x', u} \Lambda_n(x, x', u)^{\hat{M}_n(x, x', u)} \exp \left\{ -\Lambda_n(x, x', u) + \hat{T}_n(x, u) \right\} \\ &\propto \prod_n \prod_{x, x' \neq x, u} \Gamma(\hat{\alpha}_n(x, x', u)) \hat{\beta}_n(x, u)^{-\hat{\alpha}_n(x, x', u)}, \end{aligned} \quad (22)$$

where $\hat{\alpha}_n(x, x', u) = \hat{M}_n(x, x', u) + \bar{\alpha}_n(x, x', u)$ and $\hat{\beta}_n(x, u) = \hat{T}_n(x, u) + \bar{\beta}_n(x, u)$ and $\bar{\alpha}_n(x, x', u) = \alpha_n(x, x', u) + M_n(x, x', u, i=0)$ and $\bar{\beta}_n(x, x', u) = \beta_n(x, u) + T_n(x, u, i=0)$, see main-text. The KL between two measures, in this case CTBNs with different graphs, is defined via the integration over all paths

$$\text{KL} \left(p(S^{[0, T]} \mid G, \mathcal{H}) \parallel p(S^{[0, T]} \mid G', \mathcal{H}) \right) = \int dp(S^{[0, T]} \mid G, \mathcal{H}) \ln \frac{p(S^{[0, T]} \mid G, \mathcal{H})}{p(S^{[0, T]} \mid G', \mathcal{H})}.$$

In order to avoid solving the computationally taxing solution of the marginal master-equation (Studer et al., 2016; Linzner & Koepl, 2018) (which is an integro-differential equation), we can express this in terms of the original path-measure

$$\begin{aligned} & \text{KL} \left(p(S^{[0, T]} \mid G, \mathcal{H}) \parallel p(S^{[0, T]} \mid G', \mathcal{H}) \right) = \\ & \int dp(\Lambda \mid G, \mathcal{H}) \int dp(S^{[0, T]} \mid \Lambda, G) \ln \frac{p(S^{[0, T]} \mid G, \mathcal{H})}{p(S^{[0, T]} \mid G', \mathcal{H})}. \end{aligned}$$

Inserting (22) yields

$$\begin{aligned} \text{KL} \left(p(S^{[0,T]} | G, \mathcal{H}) || p(S^{[0,T]} | G', \mathcal{H}) \right) &= \int dp(\Lambda | G, \mathcal{H}) \int dp(S^{[0,T]} | \Lambda, G) \sum_{n \in \mathbb{N}} \\ &\sum_{u \in \mathcal{U}_n^G} \sum_{u' \in \mathcal{U}_n^{G'}} \sum_{x, x' \neq x} \left[\ln \frac{\Gamma(\hat{\alpha}_n(x, x', u))}{\Gamma(\hat{\alpha}_n(x, x', u'))} + \hat{\alpha}_n(x, x', u) \ln \hat{\beta}_n(x, u') - \hat{\alpha}_n(x, x', u) \ln \hat{\beta}_n(x, u) \right]. \end{aligned} \quad (23)$$

As mentioned in the main-text, exact computation of the integral w.r.t $p(S^{[0,T]} | \Lambda, G)$ is not feasible, due to non-linearity. For this reason we expand this KL around the expected transitions and dwelling times and arrive at

$$\begin{aligned} \text{KL} \left(p(S^{[0,T]} | G, \mathcal{H}) || p(S^{[0,T]} | G', \mathcal{H}) \right) &\approx \int dp(\Lambda | G, \mathcal{H}) \sum_{n \in \mathbb{N}} \sum_{u \in \mathcal{U}_n^G} \sum_{u' \in \mathcal{U}_n^{G'}} \sum_{x, x' \neq x} \\ &\left[\ln \frac{\Gamma(\mathbb{E}[\hat{\alpha}_n(x, x', u)])}{\Gamma(\mathbb{E}[\hat{\alpha}_n(x, x', u')])} + \mathbb{E}[\hat{\alpha}_n(x, x', u)] \ln \mathbb{E}[\hat{\beta}_n(x, u')] - \mathbb{E}[\hat{\alpha}_n(x, x', u)] \ln \mathbb{E}[\hat{\beta}_n(x, u)] \right], \\ &\equiv \mathcal{F}[\kappa, p(G | \mathcal{H})] \end{aligned} \quad (24)$$

with $\mathbb{E}[\hat{\alpha}_n(x, x', u)] \equiv \bar{\alpha}_n(x, x', u) + \mathbb{E}[\hat{M}_n(x, x', u) | \Lambda, G]$ and $\mathbb{E}[\hat{\beta}_n(x, u)] \equiv \bar{\beta}_n(x, u) + \mathbb{E}[\hat{T}_n(x, u) | \Lambda, G]$.

Below, we derive higher-order moments of the transitions and dwelling-times. This allows to compute higher-order approximations of this KL-divergence, under higher computational costs. However, in this work a first order approximation was sufficient to demonstrate effectiveness of our method.

VBHC for Structure Learning. We can then approximate the VBHC by

$$\text{VBHC} \approx \mathcal{F}[\kappa, p(G | \mathcal{H})] + \text{KL} (q_\kappa(G) || p(G | \mathcal{H})),$$

with the KL-divergence, between two categoricals

$$\text{KL} (p(G | \mathcal{H}) || q_\kappa(G)) = \sum_G q_\kappa(G) (\ln q_\kappa(G) - \ln p(G | \mathcal{H})).$$

While the form of \mathcal{F} is compact, it is helpful for computational reasons to re-order this summation into a node-wise form. This is helpful, as it will allow is to compute sample approximations of the VBHC, where only a summation over local parent-sets instead of global graphs needs to be performed

$$\begin{aligned} \mathcal{F}[\kappa, p(G | \mathcal{H})] &= \sum_{n \in \mathbb{N}} \sum_{G, G'} p(G | \mathcal{H}) q_\kappa(G') \int dp(\Lambda | G, \mathcal{H}) \sum_{u \in \mathcal{U}_n^G} \sum_{u' \in \mathcal{U}_n^{G'}} \sum_{x, x' \neq x} \\ &\left[\ln \frac{\Gamma(\mathbb{E}[\hat{\alpha}_n(x, x', u)])}{\Gamma(\mathbb{E}[\hat{\alpha}_n(x, x', u')])} + \mathbb{E}[\hat{\alpha}_n(x, x', u)] \ln \mathbb{E}[\hat{\beta}_n(x, u')] - \mathbb{E}[\hat{\alpha}_n(x, x', u)] \ln \mathbb{E}[\hat{\beta}_n(x, u)] \right]. \end{aligned}$$

The product form of (22) translates to a product posterior, if not broken by the prior, over parent-sets

$$\begin{aligned} p(G | \mathcal{H}) &= \prod_n p(\text{par}^G(n) | \mathcal{H}) \\ &\propto \prod_n p(\text{par}^G(n)) \prod_{x, x' \neq x} \prod_{u \in \mathcal{U}_n^G} \Gamma(\bar{\alpha}_n(x, x', u)) \bar{\beta}_n(x, u)^{-\bar{\alpha}_n(x, x', u)}. \end{aligned} \quad (25)$$

This allows us to rewrite

$$\begin{aligned} \sum_{n \in \mathbb{N}} \sum_{G, G'} p(G) q_\kappa(G') &= \\ \sum_{n \in \mathbb{N}} \sum_{\text{par}(n), \text{par}'(n) \subset \mathcal{V}} p(\text{par}(n) | \mathcal{H}) q_\kappa(\text{par}'(n)) &\sum_{G | \text{par}(n), G' | \text{par}'(n)} p(G | \mathcal{H}) q_\kappa(G'). \end{aligned}$$

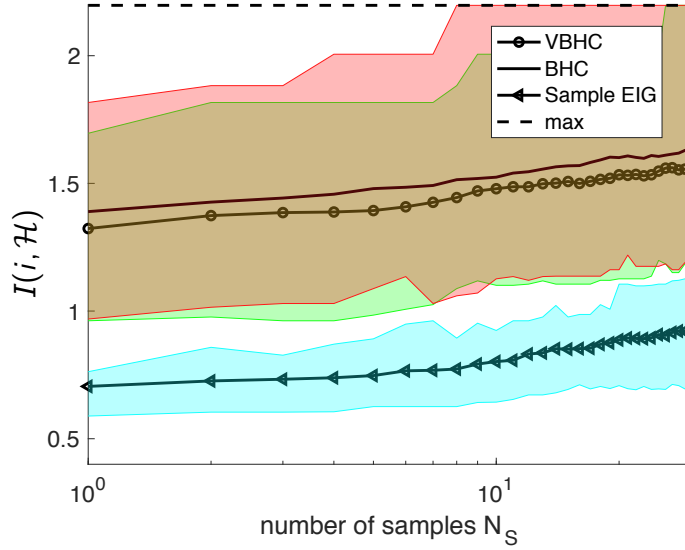


Figure 5. Mutual information between design sample estimates and recommended interventions for different number of samples N_S . Areas denote 25-75% confidence intervals.

We then get the form of the VBHC for structure learning, as used in algorithm 2

$$\begin{aligned} \mathcal{F}[\kappa, p(G | \mathcal{H})] &= \sum_{n \in \mathcal{N}} \sum_{\text{par}(n), \text{par}'(n)} p(\text{par}(n) | \mathcal{H}) q_\kappa(\text{par}'(n)) \\ &\sum_{u \in \mathcal{U}_n^{\text{par}(n)}} \sum_{u' \in \mathcal{U}_n^{\text{par}'(n)'}} \sum_{G | \text{par}(n)} \sum_{G' | \text{par}'(n)} p(G | \mathcal{H}) q_\kappa(G') \int dp(\Lambda | G, \mathcal{H}) \sum_{x, x' \neq x} \\ &\left[\ln \frac{\Gamma(\mathbb{E}[\hat{\alpha}_n(x, x', u)])}{\Gamma(\mathbb{E}[\hat{\alpha}_n(x, x', u')])} + \mathbb{E}[\hat{\alpha}_n(x, x', u')] \ln \mathbb{E}[\hat{\beta}_n(x, u')] - \mathbb{E}[\hat{\alpha}_n(x, x', u)] \ln \mathbb{E}[\hat{\beta}_n(x, u)] \right]. \end{aligned} \quad (26)$$

Similarly, we make the ansatz for $q_\kappa(G) = \prod_n q_\kappa(\text{par}^G(n))$, then the KL-divergence decomposes

$$\text{KL}(p(G | \mathcal{H}) || q_\kappa(G)) = \sum_n \sum_{\text{par}(n)} q_\kappa(\text{par}(n)) (\ln q_\kappa(\text{par}(n)) - \ln p(\text{par}(n) | \mathcal{H})).$$

Gradients. The gradient for the parameter $q_\kappa(\text{par}'(n))$ can be calculated to be

$$\begin{aligned} \partial_{q_\kappa(\text{par}'(n))} \text{VBHC} &= 1 + \ln q_\kappa(\text{par}'(n)) - \ln p(\text{par}'(n) | \mathcal{H}) + \sum_{\text{par}(n)} p(\text{par}(n) | \mathcal{H}) \\ &\sum_{u \in \mathcal{U}_n^{\text{par}(n)}} \sum_{u' \in \mathcal{U}_n^{\text{par}'(n)'}} \sum_{G | \text{par}(n)} \sum_{G' | \text{par}'(n)} p(G | \mathcal{H}) q_\kappa(G') \int dp(\Lambda | G, \mathcal{H}) \sum_{x, x' \neq x} \\ &\left[\ln \frac{\Gamma(\mathbb{E}[\hat{\alpha}_n(x, x', u)])}{\Gamma(\mathbb{E}[\hat{\alpha}_n(x, x', u')])} + \mathbb{E}[\hat{\alpha}_n(x, x', u')] \ln \mathbb{E}[\hat{\beta}_n(x, u')] - \mathbb{E}[\hat{\alpha}_n(x, x', u)] \ln \mathbb{E}[\hat{\beta}_n(x, u)] \right]. \end{aligned} \quad (27)$$

C. Experiments

C.1. Additional Experiments

Sample Estimates of Design Criteria. We want to investigate the viability of using sample estimates of different criteria for active learning of CTBNs. One basic requirement on such an estimate is that its recommendations actually depend

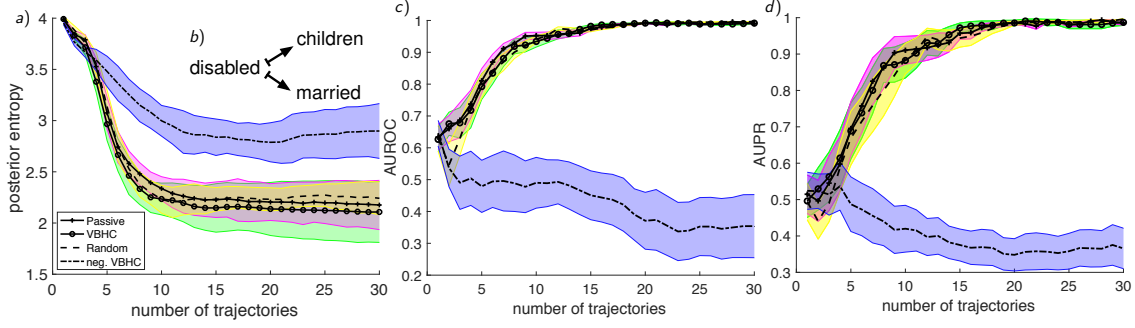


Figure 6. a) Mean and variance (area) of the evolution of the posterior entropy in BHPS data-set for 100 repetitions. b) Sketch of the underlying network. c) AUROC and d) AUPR converge equally fast to the inferred network b) for all criteria but negative VBHC.

on the history of observations $\mathcal{H} \xrightarrow{\text{design}} i$. We can make this formal by the following non-parametric dependency check:

The recommended intervention i is dependent on experimental sequence \mathcal{H} if they share high mutual information $I(i, \mathcal{H})$. We stress, that this does not reflect the quality of recommended interventions! We calculate the MI for random graphs of size $L = 3$ for different sample sizes N_S for random histories \mathcal{H} consisting of 30 trajectories drawn from our synthetic network. The results are displayed in figure 5. For all sample sizes considered, (V)BHC shares a much higher MI with their recommended interventions, than the sample estimate of the EIG.

C.2. Processing of British-Household Data-set

As mentioned in the main-text, the British-Household Data-set is incomplete, as no complete paths of variables are provided, but only their measurement at singular time-points $t_i \in \{1, \dots, 11\}$ (yearly for 11 years). In order to process, this kind of data, we employ a standard forward backward filter for continuous-time Markov jump processes, as in (Oppér & Winther, 2001; Oppér & Sanguinetti, 2008; Cohn et al., 2010; Linzner & Koepl, 2018). For this data $Y^{[0,T]} \equiv \{Y(t_i) \mid t_i \in \{1, \dots, 11\}\}$ and $Y(t_i) \sim p(Y(t_i) \mid S(t_i))$ some observation model, with measurements at singular time-points, posterior inference of the marginals $p(S(t) = s \mid Y^{[0,T]})$ is implemented by solving a time-dependent master-equation

$$\begin{aligned} \frac{d}{dt} p(S(t) = s \mid Y^{[0,T]}) = \\ \sum_{s' \neq s} \left[\hat{W}(s', s, t) p(S(t) = s \mid Y^{[0,T]}) - \hat{W}(s, s', t) p(S(t) = s' \mid Y^{[0,T]}) \right] \end{aligned}$$

with $\hat{W}(s, s', t) = W(s, s') \frac{\rho(s', t)}{\rho(s, t)}$ and

$$\begin{aligned} \frac{d}{dt} \rho(s', t) = - \sum_{s' \neq s} [W(s', s) \rho(s, t) - W(s, s') \rho(s', t)] \\ \text{subject to: } \lim_{t \rightarrow t_i^-} \rho(s, t) = \lim_{t \rightarrow t_i^+} \rho(s, t) \ln p(Y(t_i) \mid S(t_i) = s). \end{aligned}$$

This allows to calculate the marginal likelihood

$$p(Y^{[0,T]} \mid W) = \prod_{s, s' \neq s} W(s, s') \mathbb{E}^{[M(s, s') \mid Y^{[0,T]}]} \exp \left\{ W(s, s) \mathbb{E} [T(s) \mid Y^{[0,T]}] \right\},$$

with $\mathbb{E} [T(s) \mid Y^{[0,T]}] \equiv \int dt p(S(t) = s \mid Y^{[0,T]})$ and $\mathbb{E} [M(s, s') \mid Y^{[0,T]}] \equiv W(s, s') \mathbb{E} [T(s) \mid Y^{[0,T]}]$. By calculation of the corresponding moments of the CTBNs by appendix (17) and (16), we can also write this likelihood in terms of rates Λ and structure G

$$\begin{aligned} p(Y^{[0,T]} \mid \Lambda, G) = \\ \prod_{n, x, x' \neq x, u} \Lambda_n(x, x', u) \mathbb{E}^{[M_n(x, x', u) \mid Y^{[0,T]}]} \exp \left\{ \Lambda_n(x, x, u) \mathbb{E} [T_n(x, u) \mid Y^{[0,T]}] \right\}. \end{aligned}$$

As can be seen in (Linzner & Koeppl, 2018), this finally allows to form a posterior over parameters $p(\Lambda | Y^{[0,T]}) \propto p(Y^{[0,T]} | \Lambda)p(\Lambda)$, which is again a Gamma distribution, if $p(\Lambda)$ is gamma-distributed. Similarly, this holds for structures, by marginalization. Aside from this posterior calculation, everything about our method remains the same for incomplete data.

In Fig. 6 a), we track the evolution of the posterior entropy over structures for 100 independent runs. In Fig. 6 b) and c), we show that for all designs (except the "worst" design neg. VBHC) the inferred network converges against the one inferred using the full data-set (using AUROC and AUPR as metrics). We note that the effect of active learning can be expected to be small in this synthetic scenario, as we were only able to intervene on a single node.