

Active learning for optimal intervention design in causal models

Received: 25 November 2022

Accepted: 14 August 2023

Published online: 02 October 2023

 Check for updates

Jiaqi Zhang^{1,2}, Louis Cammarata^{2,3,4}, Chandler Squires^{1,2,4},
Themistoklis P. Sapsis¹ & Caroline Uhler^{1,2}

Sequential experimental design to discover interventions that achieve a desired outcome is a key problem in various domains including science, engineering and public policy. When the space of possible interventions is large, making an exhaustive search infeasible, experimental design strategies are needed. In this context, encoding the causal relationships between the variables, and thus the effect of interventions on the system, is critical for identifying desirable interventions more efficiently. Here we develop a causal active learning strategy to identify interventions that are optimal, as measured by the discrepancy between the post-interventional mean of the distribution and a desired target mean. The approach employs a Bayesian update for the causal model and prioritizes interventions using a carefully designed, causally informed acquisition function. This acquisition function is evaluated in closed form, allowing for fast optimization. The resulting algorithms are theoretically grounded with information-theoretic bounds and provable consistency results for linear causal models with known causal graph. We apply our approach to both synthetic data and single-cell transcriptomic data from Perturb-CITE-sequencing experiments to identify optimal perturbations that induce a specific cell-state transition. The causally informed acquisition function generally outperforms existing criteria, allowing for optimal intervention design with fewer but carefully selected samples.

An important problem across multiple disciplines, ranging from bioengineering to mechanical systems, operations research and environmental regulation, is the discovery of interventions on a system that can produce a desired outcome. With little prior knowledge of the outcome before performing the intervention, one can have a huge number of possible choices for the optimal design. In particular, the interventions in many applications are combinatorial, resulting in an exponential-size design space, making an exhaustive search infeasible. Examples include experimental design of genetic perturbations, such as those for cellular reprogramming in regenerative medicine¹, optimal feedback control in mechanical systems² as well as turbulent

flows³, dynamic pricing strategies in customer networks⁴ and iterative intervention research for climate change adaptation⁵.

In this context, ‘active learning’ has been proposed as a machine-learning strategy to efficiently explore the search space⁶. Such methods sequentially and strategically acquire new interventions, with the goal of discovering an optimal or a close-to-optimal intervention using the fewest number of samples. Although it can be of interest to identify optimal interventions for ‘estimating’ particular quantities in the model (for example, ref. 7), in this work, we consider the optimality of an intervention with respect to ‘optimizing’ its effect. Specifically, this is done by successively (1) updating the model belief using samples

¹Massachusetts Institute of Technology, Cambridge, MA, USA. ²Broad Institute of MIT and Harvard, Cambridge, MA, USA. ³Harvard University, Cambridge, MA, USA. ⁴These authors contributed equally: Louis Cammarata, Chandler Squires. ✉e-mail: sapsis@mit.edu; culer@mit.edu

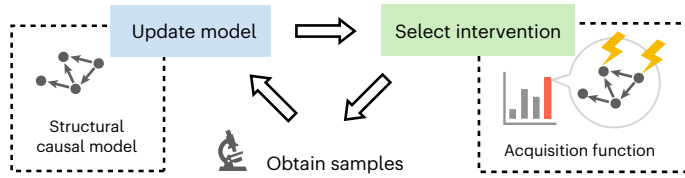


Fig. 1 Overview schematic of the active learning framework for optimal intervention design in causal models. Iterative process of active learning for intervention design in causal models, in which the main design steps are to update the structural causal model using the obtained samples and to select the next intervention based on an acquisition function.

acquired so far from different interventions and (2) selecting the next intervention to obtain samples from by constructing and optimizing an acquisition function, which prioritizes interventions that are more informative for the desired outcomes (Fig. 1).

Standard approaches towards this problem are correlation based. More precisely, the idea is to estimate associational relations between intervention and outcome to update the model belief and make decisions about which samples to acquire next. The two main approaches use either statistical theory, by minimizing the posterior variance of the outcome estimate⁸, or information theory, by maximizing the mutual information between its samples and the quantity of interest⁹. Although these methods are being widely applied, correlation-based approaches are not optimal when the underlying model is causal, as they do not take into account the structural information that can reduce the number of feasible models. Many systems that are relevant for applications are causal, in which an intervention can only have an impact on downstream variables. It is therefore of interest to develop methods that learn what is necessary about the underlying causal mechanisms to identify optimal interventions more quickly.

These limitations of correlation-based approaches have been noted in the related bandit setting, in which the goal is to minimize the cumulative regret by selecting arms iteratively, and causal relations have been used to improve over standard regret bounds^{10,11}. Recent works in the related field of Bayesian optimization have also considered exploiting causal structure^{12–14}. We provide a detailed review of these works in Supplementary Information I. Our work extends these works in two ways: (1) rather than optimizing a single target node, we optimize the entire distribution mean, and (2) rather than considering discrete or finite interventions, we consider continuous-valued interventions. This is important for various applications, such as optimizing drug dosages or product prices. As a concrete example, consider cellular reprogramming in genomics, a problem of great interest for regenerative medicine^{1,15}. The aim in this field is to reprogramme easily accessible cell types into a desired cell type via continuous-valued interventions such as the overexpression of particular transcription factors and genes. Because such interventions act on genes, which regulate each other through different pathways¹⁶, this problem can be formulated as optimal intervention design in a causal model represented by a directed network on genes.

In particular, we model the underlying causal model in a Bayesian way using a structurally informed prior. We then construct an acquisition function based on this model and show how to efficiently evaluate and optimize it. The acquisition function both enjoys an information-theoretic bound and provably recovers the optimal intervention in the appropriate limit. We demonstrate experimentally that our algorithms outperform baselines on both synthetic data and the design of genetic perturbations in the context of single-cell gene expression data. Finally, we conclude with an outlook to future research directions and discuss other potential applications.

Problem set-up

The state of the system of interest is described by a p -dimensional random variable $\mathbf{x} = (x_1, \dots, x_p) \in \mathbb{R}^p$ sampled from a structural causal

model (SCM)^{17,18}. Precisely, the ‘causal structure’ of the system is represented by a directed acyclic graph (DAG), and we assume that the joint distribution P of \mathbf{x} factorizes with respect to the DAG; that is, $P(\mathbf{x}) = \prod_{i=1}^p P(x_i | x_{\text{pa}(i)})$, where $\text{pa}(i) = \{j \in [p] : j \rightarrow i\}$ denotes the parents of node i in the DAG. In this formulation, the conditional distributions are the ‘causal mechanisms’ that generate the variable x_i from its parents $x_{\text{pa}(i)}$. Assuming a linear Gaussian model, then

$$x_i = \sum_{k=1}^p B_{ik} x_k + \epsilon_i, \quad \forall i \in [p], \tag{1}$$

where the real-valued coefficients $B_{ik} = 0$ if $k \notin \text{pa}(i)$ and the exogenous noise variables $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$ with variance $\sigma_i^2 > 0$ are mutually independent. An example of this model is given in Fig. 2. For simplicity, we assume that the system is centred to be mean zero, but intercepts can be easily added to equation (1) and the following approaches still apply.

An intervention, denoted by a vector $\mathbf{a} \in \mathbb{R}^p$, modifies the conditional distribution $P(x_i | x_{\text{pa}(i)})$ into a new conditional distribution, $P^{\mathbf{a}}(x_i | x_{\text{pa}(i)})$. Every $i \in [p]$ for which $a_i \neq 0$ is called an ‘intervention target’. In this work, we consider shift interventions^{19,20} (which are a special class of soft interventions²¹), in which the interventional distribution $P^{\mathbf{a}}(\mathbf{x})$ of the modified system under intervention is given by

$$x_i = \sum_{k=1}^p B_{ik} x_k + a_i + \epsilon_i, \quad \forall i \in [p], \tag{2}$$

where $a_i = 0$ for every i that is not an intervention target. This can be written in matrix form as $\mathbf{x} = B\mathbf{x} + \mathbf{a} + \boldsymbol{\epsilon}$, or $\mathbf{x} = (I - B)^{-1}(\mathbf{a} + \boldsymbol{\epsilon})$, where I is the p th order identity matrix and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \Sigma)$ with Σ being the diagonal matrix with $\sigma_1^2, \dots, \sigma_p^2$ on its diagonal. These interventions can be used to model a broad class of genetic perturbations including clustered regularly interspaced short palindromic repeat (CRISPR) interference and activation²², and transcription factor overexpression²³. Although we consider the case in which all variables can be intervened on, as for example in genetic perturbation experiments²⁴, for other applications, it may be of interest to consider extensions in which only a subset of variables can be intervened on. We also note that an extension to hard interventions can be obtained by setting B_{ik} to 0 for every i that is an intervention target and subsequently defining the post-interventional mean.

Whether or not an intervention induces the desired outcome is decided through samples obtained in the intervened environment, that is, $\mathbf{x} \sim P^{\mathbf{a}}$. Since \mathbf{x} is random, it is intuitive to use the average over multiple samples^{20,25}, that is, the empirical estimate of the distribution mean, and compare this with the ‘target mean’ $\boldsymbol{\mu}^*$, which is user specified and describes the desired outcome. Formally, we seek an intervention $\mathbf{a} \in \mathbb{R}^p$ such that, after obtaining n samples $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)} \in \mathbb{R}^p$, which we denote by $\mathbf{x}_{[n]}$, the squared distance between the empirical mean and the target mean, that is, the quantity $\| \frac{1}{n} \sum_{m=1}^n \mathbf{x}^{(m)} - \boldsymbol{\mu}^* \|_2^2$, is minimized. Denoting by $\boldsymbol{\epsilon}^{(1)}, \dots, \boldsymbol{\epsilon}^{(n)}$ (or short $\boldsymbol{\epsilon}_{[n]}$) the exogenous noise vectors independently sampled from $\mathcal{N}(0, \Sigma)$, then the squared distance between the empirical and target mean can be written as

$$\left\| (I - B)^{-1} \left(\mathbf{a} + \frac{1}{n} \sum_{m=1}^n \boldsymbol{\epsilon}^{(m)} \right) - \boldsymbol{\mu}^* \right\|_2^2 \tag{3}$$

using the matrix form. Here $\frac{1}{n} \sum_{m=1}^n \boldsymbol{\epsilon}^{(m)}$ accounts for the finite number n of interventional samples, where n is user specified based on the available budget. However, as $\boldsymbol{\epsilon}_{[n]}$ does not depend on the choice of \mathbf{a} , we can discard this term for the minimization of equation (3) with respect to \mathbf{a} by considering its infinite-sample version. When the sample size goes to infinity, by the law of large numbers, $\frac{1}{n} \sum_{m=1}^n \boldsymbol{\epsilon}^{(m)} = 0$ almost surely, and thus, the optimal intervention \mathbf{a}^* achieves the minimum value of zero in equation (3) and has an explicit form $\mathbf{a}^* = (I - B)\boldsymbol{\mu}^*$, which, as expected, depends on the unknown parameter B .

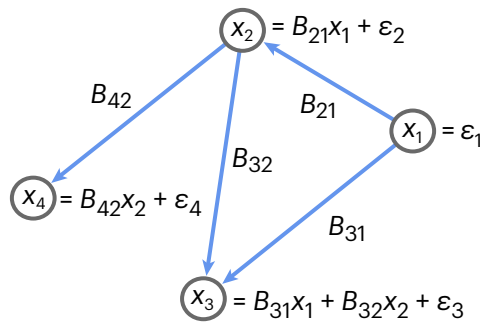


Fig. 2 | Example causal model. An example of a linear SCM with Gaussian noise on a four-node DAG, in which nodes and edges are labelled with variables and coefficients, respectively.

In what follows, we assume that the DAG structure is given; that is, the sparsity pattern of B is known but not the edge weights B_{ik} for $B_{ik} \neq 0$. This assumption is natural in many applications including fluid mechanics²⁶ and optimal pricing⁴, in which there are pre-specified networks given by either known laws or prior information. For applications in which the network cannot be assumed to be known such as problems in biology, the common approach is to use existing data from various sources to learn the DAG first^{15,27}. Further discussions on the implications of an unknown DAG structure are provided in Supplementary Information D.

Let $\mathcal{D}_t = \{(\mathbf{x}_{[n_1]}, \mathbf{a}^{(1)}), \dots, (\mathbf{x}_{[n_t]}, \mathbf{a}^{(t)})\}$ be the ‘current’ dataset, consisting of all samples obtained so far by performing interventions $\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(t)}$; here, n_1, \dots, n_t denotes the number of samples obtained for each of the t interventions. To simplify notation, we will assume $n_1 = \dots = n_t = n$ in the following, but all results still hold when using different sample sizes for each intervention. Given this dataset, the goal is to select the next intervention $\mathbf{a}^{(t+1)} \in \mathbb{R}^p$ such that the resulting dataset $\mathcal{D}_{t+1} = \mathcal{D}_t \cup (\mathbf{x}_{[n]}, \mathbf{a}^{(t+1)})$ contains as much information as possible about the underlying optimal intervention \mathbf{a}^* . The overall aim of iteratively picking interventions is to find the optimal intervention \mathbf{a}^* with a minimum number of samples.

Design of the acquisition function

To sequentially select the next best intervention, there are two important steps (Fig. 1): (1) updating the posterior of the edge weights in the causal model based on the samples in \mathcal{D}_t collected so far and (2) constructing an ‘acquisition function’, $h(\mathbf{a}, \mathcal{D}_t)$, such that the new dataset \mathcal{D}_{t+1} after adding samples from $\mathbf{a}^{(t+1)} = \operatorname{argmin}_{\mathbf{a}} h(\mathbf{a}, \mathcal{D}_t)$ is most informative of the optimal intervention \mathbf{a}^* and can be evaluated and optimized efficiently.

For the first step, we generalize the DAG–Wishart distribution^{28–30} to define a Bayesian model on the parameters B , which can be updated efficiently given a dataset \mathcal{D}_t (Methods). For the second step, we first characterize the uncertainty in estimating the optimality of an arbitrary intervention \mathbf{a} . Recall that the optimality of \mathbf{a} is given by the square distance in equation (3), which measures how close the intervention is to achieving the target mean. As B is unknown, we can only estimate this square distance based on the current collected samples in \mathcal{D}_t . The uncertainty of the estimation can be characterized by its variance $\operatorname{Var}(\| (I - B)^{-1}(\mathbf{a} + \frac{1}{n} \sum_{m=1}^n \epsilon_m) - \boldsymbol{\mu}^* \|_2^2 | \mathcal{D}_t)$. However, this quantity is typically hard to evaluate as it involves $(I - B)^{-1}$, whose posterior does not have a closed form. We instead multiply the term inside the variance by $I - B$ and characterize the following variance:

$$\sigma_{g(\mathbf{a})|\mathcal{D}_t}^2 := \operatorname{Var}(g(\mathbf{a}) | \mathcal{D}_t), \tag{4}$$

where we define

$$g(\mathbf{a}) := \left\| \left(\mathbf{a} + \frac{1}{n} \sum_{m=1}^n \epsilon^{(m)} \right) - (I - B)\boldsymbol{\mu}^* \right\|_2^2. \tag{5}$$

Interestingly, $g(\mathbf{a})$ can be interpreted as a noisy version of the ‘optimality gap’ by noting that the gap between an arbitrary intervention \mathbf{a} and the optimal intervention $\mathbf{a}^* = (I - B)\boldsymbol{\mu}^*$ can be written as $\| \mathbf{a} - \mathbf{a}^* \|_2^2 = \| \mathbf{a} - (I - B)\boldsymbol{\mu}^* \|_2^2$, which is a version of $g(\mathbf{a})$ without noise terms.

Building upon equation (4), the next acquired intervention should be such that after adding its samples to \mathcal{D}_t to obtain \mathcal{D}_{t+1} , the uncertainty $\sigma_{g(\mathbf{a})|\mathcal{D}_{t+1}}^2$ conditioned on \mathcal{D}_{t+1} is minimized. However, as the samples are unobserved before performing the intervention, we do not have access to \mathcal{D}_{t+1} yet. Therefore, when deciding which intervention to perform, we consider the \mathbf{a}' -augmented dataset $\mathcal{D}_t(\mathbf{a}') = \mathcal{D}_t \cup (\mathbf{x}'_{[n]}, \mathbf{a}')$ with ‘hypothetical samples’ $\mathbf{x}'_{[n]}$ of \mathbf{a}' , which are n repetitions of the plug-in estimator defined as $\mathbf{x}' := (I - \mathbb{E}(B|\mathcal{D}_t))^{-1}\mathbf{a}'$. This estimator is obtained through $\mathbf{x}' = (I - B)^{-1}(\mathbf{a}' + \epsilon)$ by using the maximum a posteriori estimate $\mathbb{E}(B|\mathcal{D}_t)$, which concentrates to B as \mathcal{D}_t grows³¹, and replacing ϵ with its mean 0.

By denoting the feasible set of interventions by \mathcal{A} , a reasonable choice is to select $\mathbf{a}^{(t+1)}$ based on integrating the uncertainty $\sigma_{g(\mathbf{a})|\mathcal{D}_t(\mathbf{a}')}^2$ over all $\mathbf{a} \in \mathcal{A}$. Formally, we let

$$\mathbf{a}^{(t+1)} = \operatorname{argmin}_{\mathbf{a}' \in \mathcal{A}} h(\mathbf{a}', \mathcal{D}_t), \tag{6}$$

and define the causal integrated variance (CIV) acquisition function h as follows.

Definition 1. The CIV acquisition function evaluated at \mathbf{a}' with current dataset \mathcal{D}_t is

$$h(\mathbf{a}', \mathcal{D}_t) = \int_{\mathcal{A}} \sigma_{g(\mathbf{a})|\mathcal{D}_t(\mathbf{a}')}^2 \, d\nu(\mathbf{a}), \tag{7}$$

where ν is a non-negative measure on \mathcal{A} .

Intuitively, this acquisition function provides a one-step lookahead of the overall uncertainty after acquiring intervention \mathbf{a}' . Minimizing it will prioritize interventions that are most informative towards estimating the optimal intervention. This acquisition function also automatically accounts for the causal model by using a posterior on B .

Note that in this formulation, we can choose the measure ν . For example, a uniform measure treats each intervention $\mathbf{a} \in \mathcal{A}$ equally and the resulting uncertainty captures how well we can estimate the entire landscape of the optimality gap g . In most cases, an overly concentrated measure (for example, a Dirac measure at a single point) is not preferred, as it can lead to the erroneous estimation of $g(\mathbf{a})$ for most $\mathbf{a} \in \mathcal{A}$, which makes minimizing g hard. Inspired by a recent line of work on output-weighted acquisition functions^{32,33}, we describe how to choose a non-uniform measure ν in the next section.

We discuss how to optimize CIV to solve for $\mathbf{a}^{(t+1)}$ in Methods, in which we show that the variance $\sigma_{g(\mathbf{a})|\mathcal{D}_t(\mathbf{a}')}^2$ can be computed in closed form. Considering \mathcal{A} to be the unit hypersphere and ν to be the uniform measure on it, this then leads to an explicit formula for CIV, which enables fast gradient-based optimizers to be used.

Making the acquisition function output weighted

Although the use of a uniform measure ν places an equal weight on reducing the variance of estimating the optimality gap for all \mathbf{a} in \mathcal{A} , as our goal is to identify the optimal intervention (that is, which minimizes the optimality gap), it is desirable to place more weight on interventions \mathbf{a} in \mathcal{A} with a smaller optimality gap. Note that as the ambient dimension grows, the volume (and thus the probability) of interventions with the optimality gap under a certain threshold shrinks (Supplementary Information C). This motivates the following measure, which uses the inverse of the optimality gap probability to up-weight interventions \mathbf{a} in \mathcal{A} that are closer to the optimal intervention:

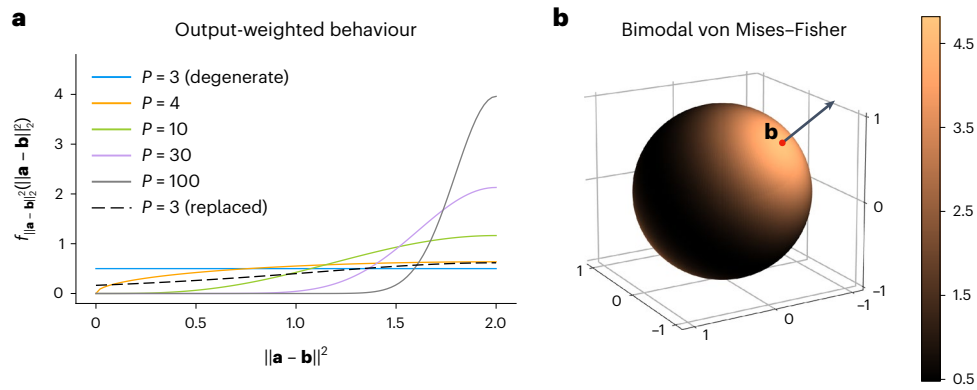


Fig. 3 | Illustration of output-weighted non-negative measure $v(\mathbf{a})$ on the space of all possible interventions \mathcal{A} . **a**, Values of $dv/d\mathbf{a}$ (up to constant multipliers) plotted against values of $\|\mathbf{a} - \mathbf{b}\|_2^2$ on the half sphere with $\|\mathbf{a} - \mathbf{b}\|_2^2 \leq 2$. Solid lines correspond to $f_{\|\mathbf{a} - \mathbf{b}\|_2^2}$ in different dimensions; the dashed line is the non-degenerate replacement for $p = 3$ using the bimodal von

Mises–Fisher distribution. **b**, Visualization of $v(\mathbf{a})$ in three dimensions using the bimodal von Mises–Fisher distribution. Here $\mathbf{b} \in \mathbb{S}^2$ was randomly generated and the points on the sphere are coloured corresponding to the value of $dv(\mathbf{a})/d\mathbf{a}$, which is larger for directions that are more aligned with \mathbf{b} .

$$dv(\mathbf{a}) = \frac{f_{\mathbf{a}}(\mathbf{a})}{f_{\|\mathbf{a} - \mathbf{b}\|_2^2}(\|\mathbf{a} - \mathbf{b}\|_2^2)} d\mathbf{a}, \quad (8)$$

where $\mathbf{b} = (I - \mathbb{E}(B|\mathcal{D}_t))\boldsymbol{\mu}^*$ is the estimated optimal intervention, and the probability density function $f_{\|\mathbf{a} - \mathbf{b}\|_2^2}$ is the distribution on the optimality gap $\|\mathbf{a} - \mathbf{b}\|_2^2$ induced by the uniform distribution $f_{\mathbf{a}}$ on \mathcal{A} .

Figure 3a shows that, as desired, the proposed weighting puts more mass on interventions that are closer to \mathbf{b} as the dimension increases. When $p = 3$, however, v degenerates to a uniform weighting (Supplementary Information C). In this case, we can use the bimodal von Mises–Fisher distribution, which behaves similar to equation (8) in higher dimensions (Fig. 3b and Supplementary Information C). Note that the weighting proposed here is symmetric and also puts more mass on interventions that are closer to $-\mathbf{b}$. The reason for this is that the optimal intervention can also be recovered by maximizing the optimality gap $\|\mathbf{a} - \mathbf{a}^*\|_2$ which gives $-\mathbf{a}^*$ (Supplementary Information C).

The corresponding causal output-weighted integrated variance (CIV-OW) acquisition function is given by

$$h_{\text{OW}}(\mathbf{a}', \mathcal{D}_t) = \int_{\mathcal{A}} \sigma_{g(\mathbf{a})|\mathcal{D}_t(\mathbf{a}')}^2 \cdot \frac{f_{\mathbf{a}}(\mathbf{a})}{f_{\|\mathbf{a} - \mathbf{b}\|_2^2}(\|\mathbf{a} - \mathbf{b}\|_2^2)} d\mathbf{a}. \quad (9)$$

Methods for evaluating and optimizing CIV-OW are given in Supplementary Information C and Supplementary Figs. 1 and 2.

Theoretical results

We provide two interpretations of the introduced CIV acquisition function. First, taking an information-theoretic perspective, we show that the proposed uncertainty measure can be lower bounded by the ‘negative’ mutual information between the variables of interest and the newly acquired samples. Thus, minimizing the uncertainty corresponds to maximizing a lower bound to the mutual information, which means that CIV approximately prioritizes the most informative interventions. Second, taking a graphical perspective, we illustrate how the causal structure is utilized by the CIV acquisition function to identify an intervention that is asymptotically consistent with the optimal intervention. The details underlying these theoretical results are provided in Methods, and Fig. 4 illustrates our consistency results experimentally, showing that a gradient-based optimizer with initialization close to the optimal intervention \mathbf{a}^* converges to \mathbf{a}^* as t increases.

Applications

Experiments on the synthetic dataset

To create a generative model following the linear SCM in equation (1), we first generate a DAG with p nodes (Supplementary Information G and Supplementary Fig. 3). Then, we randomly draw edge weights, namely, B_{ik} with $k \in \text{pa}(i)$, from a uniform distribution bounded away from zero. Next, we generate a sparse set of intervention targets and a randomly sampled optimal intervention \mathbf{a}^* over these targets. Finally, we calculate the target mean $\boldsymbol{\mu}^*$ using \mathbf{a}^* and the ground-truth causal model. These steps construct a synthetic instance of a causal system and optimal intervention \mathbf{a}^* . A more detailed description of the above procedure is given in Methods and Supplementary Information G.

We compare our two acquisition functions, CIV in equation (7) and CIV-OW in equation (9), against four relevant baselines. The ‘random’ baseline corresponds to a passive setting, in which each intervention is selected at random and no information from the collected samples is used. We also compare against three other active methods. The ‘greedy’ baseline selects the next intervention $\mathbf{a}^{(t)} = (I - \mathbb{E}(B|\mathcal{D}_t))\boldsymbol{\mu}^*$ purely based on the current estimate of \mathbf{a}^* (which is given by $\mathbb{E}(B|\mathcal{D}_t)$, where B is estimated from \mathcal{D}_t). The MaxV baseline seeks to select interventions that minimize the posterior variance of the estimate of the model parameters B . It uses as acquisition function a scalar version of this variance, namely, $h_{\text{MaxV}}(\mathbf{a}', \mathcal{D}_t) = \max_{i \in [p]} \|\text{Var}(B_{i,\text{pa}(i)}|\mathbf{a}', \mathcal{D}_t)\|_2$, where $\|\cdot\|_2$ denotes the spectral norm of the covariance matrix of $B_{i,\text{pa}(i)}$. As different rows of B (for example, $B_{i,\text{pa}(i)}$ and $B_{j,\text{pa}(j)}$ for $i \neq j$) are independent (Definition 2), we use the maximum of the spectral norms over $i \in [p]$. Whereas our acquisition functions concentrate on estimating the model parameters that are relevant for \mathbf{a}^* , this baseline estimates the entire model. Finally, the CV baseline seeks to minimize the posterior variance of estimating \mathbf{a}^* and uses the spectral norm of its covariance matrix as acquisition function; that is, $h_{\text{CV}}(\mathbf{a}', \mathcal{D}_t) = \|\text{Var}((I - B)\boldsymbol{\mu}^*|\mathbf{a}', \mathcal{D}_t)\|_2$. This is in contrast to our approach that integrates the estimation uncertainty over the entire feasible set of interventions.

To reduce evaluation noise effects, we run each method 20 times over 10 instances of a fixed DAG and optimal intervention \mathbf{a}^* . Figure 5 shows our results for 30-node DAGs with 10 intervention targets and sample size $n = 1$ per time step. Denoting the obtained samples at time step t with \mathcal{D}_t , we use the unbiased estimate $\mathbf{a}'_t = (I - \mathbb{E}(B|\mathcal{D}_t))\boldsymbol{\mu}^*$ of \mathbf{a}^* to obtain the current estimate of the target mean using the true model parameters, that is, $\boldsymbol{\mu}'_t = (I - B)^{-1}\mathbf{a}'_t$. Figure 5a shows the relative distance between current and target mean $\|\boldsymbol{\mu}'_t - \boldsymbol{\mu}^*\|_2 / \|\boldsymbol{\mu}^*\|_2$ across time steps. Figure 5b highlights the statistics of the last time

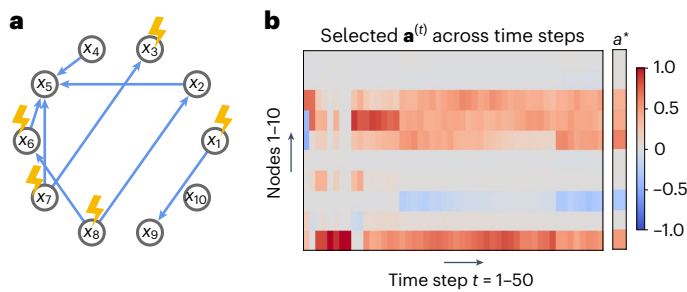


Fig. 4 | Convergence of the selected intervention $a^{(t)}$ to the optimal intervention a^* . **a**, DAG showing the targets of the optimal intervention a^* . **b**, Minimizing the CIV acquisition function using a sample size of $n = 1$ at each time step yields interventions $a^{(t)}$ for $t = 1, \dots, 50$ that converge to a^* .

step. As is apparent from these experiments, our proposed acquisition functions consistently outperform all baselines, with CIV-OW improving upon CIV by using an output-weighted measure. In Supplementary Information G, we examine the effect of varying different parameters including graph size, number of intervention targets and graph structure, demonstrating the robustness of our results (Extended Data Figs. 1–4). We also include three additional baselines from prior works^{9,12,34,35} (Supplementary Information G and I) on a ten-node DAG (Extended Data Fig. 5) and an experiment to examine the effect of misspecifying the underlying causal structure (Supplementary Fig. 4 and Extended Data Fig. 6). These results further demonstrate the effectiveness of our proposed method in terms of both accuracy and efficiency.

Experiments on the biological dataset

We next study the performance of our method to identify the optimal intervention for inducing a desired cell-state change in human melanoma cells, thereby mimicking a cellular reprogramming experiment. For this, we use Perturb-CITE-sequencing (Perturb-CITE-seq) data from ref. 36 consisting of single-cell transcriptomic readouts for a large collection of patient-derived melanoma cells. Here an intervention is a genetic perturbation that targets one or multiple genes and drives the expression of these genes towards zero (more precisely, these are knockout interventions). Cell states are measured by the joint distribution of the expression of a collection of genes. Samples from these distributions correspond to gene expression vectors of individual melanoma cells. To avoid dealing with batch effects, we use only one of the screens from ref. 36 (namely, the control screen with no additional treatment), which contains 5,039 cells with no perturbation and 30,486 cells with interventions on subsets of 248 genes associated with immunotherapy resistance (Supplementary Fig. 5). Gene expression of each sampled cell is captured as a vector of log-transformed transcripts per million.

For our optimal intervention design task, we focused on a particular functional context, namely, the $p = 36$ genes among the 248 interventional targets that are involved in interferon- γ signalling and immune response³⁶ (Supplementary Information H). As multi-target interventional samples are extremely scarce in this dataset, with most such interventions having no more than one sample (Supplementary Fig. 6), we only consider single-target interventions. We use the observational samples, that is, the cells with no perturbation, to learn a DAG over these 36 genes (Methods, Supplementary Fig. 7 and Extended Data Fig. 7). We model each of the 36 single-target interventions as a downshift of the target gene by its observational mean; that is, knocking out target gene $i \in [p]$ is modelled as a shift intervention \mathbf{a} with a_i being the negative of the observed mean of this gene, and all other entries being zero. This model assumes that the perturbations are effective at knocking out their target genes. This assumption is not always met by current technologies: we observe that some interventions do not

downregulate their target genes in a statistically significant manner (Supplementary Information H). However, this is only known after performing the experiments, and thus, we use the idealized model. If another dataset or prescreen (usually on a smaller scale) is available to help determine the effectiveness of an intervention to target a specific gene, then we can easily replace the idealized model with this estimate. As we will show, the idealized model still extracts enough signal for our method to aid in finding good interventions.

To evaluate our approach, we use the observational distribution as the source cell state and a particular single-node interventional distribution as the target cell state. The aim is to identify this single-node intervention or an intervention with similar effect using the least number of samples. For benchmarking purposes, we use a setting in which the optimal intervention is contained in the feasible set of interventions, but we note that this is not a requirement and the target cell state can be any desired distribution, the experimental design goal then being to identify an intervention that moves the distribution as close as possible to the target distribution.

Figure 6a shows three representative examples of interventional distributions (targeting *CDH19*, *SOX4* and *HLA-C*), visualized using a 2D projection along the most variable directions obtained using contrastive principal component analysis (PCA) (Methods). These examples highlight that the alterations induced by single-target interventions are subtle (Supplementary Fig. 8), with *SOX4* being more similar to *CDH19* than *HLA-C* as also corroborated by the squared distance between the interventional means (Fig. 6b). This observation is consistent with previous melanoma studies showing that *HLA-C* is associated with positive immune response³⁷, whereas *CDH19* and *SOX4* are associated with metastasis and immune evasion^{38,39}. Figure 6c shows the performance of our results with the optimal intervention targeting gene *CDH19*. More examples using other target genes are given in Supplementary Information H. Each method is run 50 times, in which each run starts with a warm-up set of 100 observational samples and then $n = 10$ interventional samples per time step. In Fig. 6c, we present the comparison between three acquisition functions: CIV and the two baselines random and greedy. More implementation details as well as the full results are given in Extended Data Figs. 8–10. Similar to what we observed in the synthetic data experiments, CIV outperforms the benchmarks in terms of distance to target mean across all time steps. This suggests that the proposed approach is beneficial for identifying perturbations to induce a desired cell-state change.

Discussion

In this work, we developed an active learning framework for optimal intervention design in causal models. Our method has two main ingredients: (1) modelling and updating the edge weights in the causal model using a Bayesian approach (using the DAG-Bayesian linear regression (BLR) distribution) and (2) optimizing the next intervention from which to obtain samples using a class of causally aware acquisition functions (CIV acquisition functions). The DAG-BLR distribution respects the underlying causal structure and allows for efficient posterior updates. The proposed CIV class of acquisition functions prioritizes the most informative intervention with respect to identifying the optimal intervention for moving the system towards a desired mean by minimizing an uncertainty quantity weighted based on the directions of interest. Importantly, the designed CIV acquisition function allows for efficient optimization by having tractable closed-form evaluations. In addition, we showed that the introduced acquisition function is characterized by attractive theoretical properties, such as by mutual information bounds and consistency. Finally, we demonstrated the developed active learning framework on both synthetic data and a biological dataset. In both cases, the designed acquisition function outperforms empirical analogues, allowing for accurate predictions with fewer experiments.

We made various assumptions that may be limiting for some applications and motivate future research directions. First, we focused on

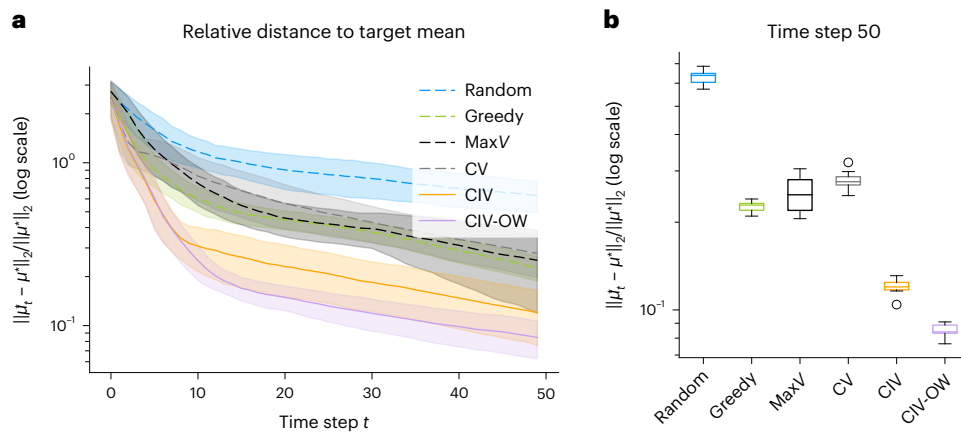


Fig. 5 | Comparison of different acquisition functions (random baseline, greedy baseline, MaxV baseline, CV baseline and our proposed CIV and CIV-OW) in a simulation study. a, b, Simulations on 10 instances of a 30-node DAG in which the optimal intervention a^* has 10 targets. Each method is run 20 times and averaged. **a,** Relative distance between the target mean μ^* and the best approximation μ_t^* across time step t . Lines denote the mean over the ten

instances; shading corresponds to 1 s.d. **b,** Relative distance statistics of each method averaged over ten instances at the last time step in **a**. Small circles: relative distance statistics of the corresponding method on a specific instance. Whiskers: first quartile minus 1.5× the inter-quartile range to third quartile plus 1.5× the inter-quartile range.

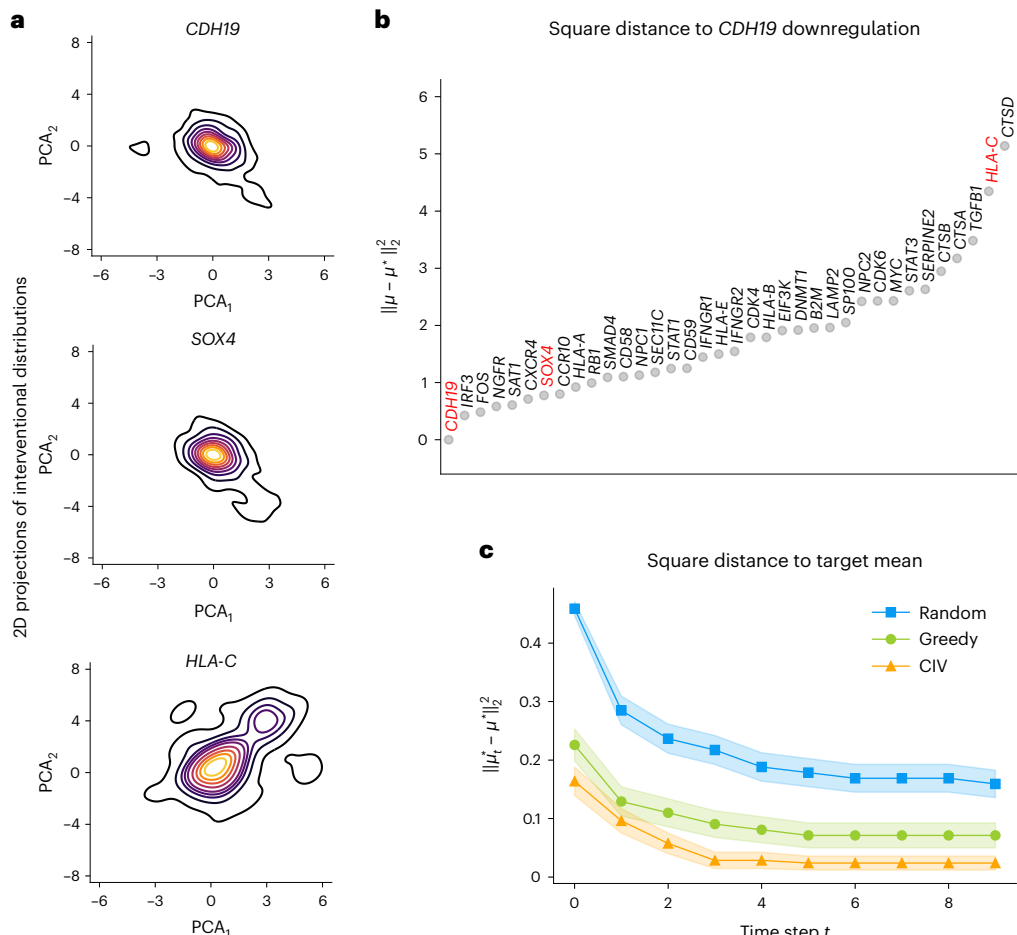


Fig. 6 | Results on perturbational single-cell gene expression dataset. a, Kernel density estimate plot of the two-dimensional projection of the interventional distributions obtained by targeting three representative genes: *CDH19*, *SOX4* and *HLA-C*. Colors indicate the density of the interventional

distribution targeting the corresponding gene. **b,** Squared distances between the target mean of the optimal intervention targeting *CDH19* and other interventions. **c,** Comparison of different acquisition functions for identifying interventions that match the target mean presented as mean \pm s.e.m.

the setting with known causal structure. An important future research direction is to consider the case in which the causal structure is partially or entirely unknown. We discuss potential avenues to approach this

problem in Supplementary Information D. Second, we considered linear SCMs with additive Gaussian noise. A potential extension to the nonlinear setting could be achieved using kernels with linearity

over the feature space. Similar derivations to our work could be used to evaluate or approximate the CIV acquisition functions in this setting. Third, in our model, we assumed ‘causal sufficiency’¹⁷, which excludes the existence of latent confounders as well as the possibility of performing interventions only on some nodes. This is violated when some system variables are unobserved and thus cannot be intervened on or specified in the desired state. A possible approach is to use a more agnostic model between interventions and their effects as proposed, for example, in ref. 12; however, this will generally lead to weaker results and a loss of structural information.

Although we discussed our work in the context of applications to cellular reprogramming, we envision our framework to be applicable broadly for sequential design problems arising in complex systems. In Supplementary Information J, we discuss several other applications and how they fit into our proposed framework.

Methods

Posterior update of edge weights

The current dataset \mathcal{D}_t induces a belief on the model parameters B . We consider the Bayesian setting in which the belief corresponds to a distribution. To account for the known causal structure, we assume a generalization of the DAG–Wishart distribution^{28,29}, which places a prior on B that respects the causal structure and allows for efficient posterior updates (Supplementary Information A). The prior is as follows.

Definition 2. The DAG–BLR prior models B and the noise variances Σ jointly as $\mathbb{P}(B, \Sigma) = \prod_{i=1}^p \mathbb{P}(B_{i,pa(i)}, \sigma_i^2)$, where each $\mathbb{P}(B_{i,pa(i)}, \sigma_i^2)$ is

$$\begin{aligned} \sigma_i^2 &\sim \mathcal{JG}(\alpha_i, \beta_i), \\ B_{i,pa(i)} | \sigma_i^2 &\sim \mathcal{N}(m_i, \sigma_i^2 M_i). \end{aligned} \quad (10)$$

Here \mathcal{JG} denotes the inverse-gamma distribution and $\{\alpha_i, \beta_i, m_i, M_i\}_{i=1}^p$ are hyperparameters satisfying certain constraints specified in Supplementary Information A.

This prior is consistent with the DAG structure as it sets $B_{ik} = 0$ for all $k \notin pa(i)$. Although it was developed in the observational setting⁴⁰, the following lemma shows that it can be extended to the interventional setting and is a conjugate prior for the model in equation (2). Thus, the posterior lies in the same family of distributions.

Lemma 1. The posterior corresponding to the DAG–BLR prior satisfies $\mathbb{P}(B, \Sigma | \mathcal{D}_t) = \prod_{i=1}^p \mathbb{P}(B_{i,pa(i)}, \sigma_i^2 | \mathcal{D}_t)$, where $\mathbb{P}(B_{i,pa(i)}, \sigma_i^2 | \mathcal{D}_t)$ is

$$\begin{aligned} \sigma_i^2 | \mathcal{D}_t &\sim \mathcal{JG}(\alpha_i(\mathcal{D}_t), \beta_i(\mathcal{D}_t)), \\ B_{i,pa(i)} | \sigma_i^2, \mathcal{D}_t &\sim \mathcal{N}(m_i(\mathcal{D}_t), \sigma_i^2 M_i(\mathcal{D}_t)). \end{aligned} \quad (11)$$

The hyperparameters $\{\alpha_i(\mathcal{D}_t), \beta_i(\mathcal{D}_t), m_i(\mathcal{D}_t), M_i(\mathcal{D}_t)\}_{i=1}^p$ are specified in Supplementary Information A.

The hyperparameters $\{\alpha_i(\mathcal{D}_t), \beta_i(\mathcal{D}_t), m_i(\mathcal{D}_t), M_i(\mathcal{D}_t)\}_{i=1}^p$ can be updated easily using the new samples obtained at step t ; thus, this choice of a ‘conjugate’ prior allows for efficient posterior updates.

We make one comment about the DAG–Wishart distribution. In previous literature^{28,29}, this distribution has been extended to model the posterior beliefs of both the edge weights and the DAG structure. The posterior for the DAG structure is known as the Bayesian Gaussian equivalence score. Thus, although we here assume that the DAG is known or prefixed and we only update the posterior on the non-zero entries of B , our framework can be extended to the unknown DAG setting by placing a probability on the DAG structure using the Bayesian Gaussian equivalence score, as discussed in Supplementary Information D.

Evaluation of CIV in closed form

We discuss how to optimize CIV to solve for $\mathbf{a}^{(t+1)}$. We start by providing a closed form of the variance $\sigma_{g(\mathbf{a})|\mathcal{D}_t, \mathbf{a}'}^2$. For ease of reading, we here

provide the formula for the special case in which the exogenous noise variances Σ are known. The general formula, which is similar in flavour but more complicated, including the proof is given in Supplementary Information B.

Proposition 1. Conditioning on Σ , we have

$$\sigma_{g(\mathbf{a})|\mathcal{D}_t, \mathbf{a}'}^2 = 2 \sum_{i=1}^p \left(v_i^2 + \frac{2}{n} v_i \sigma_i^2 + 2v_i(a_i - b_i)^2 \right) + c, \quad (12)$$

where only $v_i := \sigma_i^2 \mu_{pa(i)}^{*\top} M_i(\mathcal{D}_t(\mathbf{a}')) \mu_{pa(i)}^*$ depends on the augmented dataset $\mathcal{D}_t(\mathbf{a}')$, whereas $b_i := \mu_i^* - m_i(\mathcal{D}_t)^\top \mu_{pa(i)}^*$ and the constant c do not depend on \mathbf{a}' .

We note that the resulting expression is a function of b_i and v_i , which can be interpreted as follows: Lemma 1 can be used to rewrite $b_i = [(I - \mathbb{E}(B|\mathcal{D}_t))\mu^*]_i$ as the estimated shift of the optimal intervention at node i , and $v_i = \mu_{pa(i)}^{*\top} \text{Var}(B_{i,pa(i)} | \mathcal{D}_t(\mathbf{a}')) \mu_{pa(i)}^*$ as the covariance matrix of $B_{i,pa(i)}$ scaled in the direction of the target mean μ^* .

Next, we discuss how to integrate equation (12) over the space of possible interventions \mathcal{A} to evaluate $h(\mathbf{a}' | \mathcal{D}_t)$. Here we consider \mathcal{A} to be the hypersphere $\mathbb{S}^{p-1} = \{\mathbf{a} \in \mathbb{R}^p : \|\mathbf{a}\|_2 = 1\}$ and ν to be the uniform measure on the hypersphere (other types of feasible sets \mathcal{A} , for example, with sparsity constraints, can be considered, and similar derivations can be used to identify if a closed-form integration exists). This corresponds to fixing the magnitude of the intervention and only optimizing over its direction. This is suitable for various applications including the biological problem considered below, as the strength of the intervention is often designed separately or prefixed manually by the experimenter. We also note that for many linear problems, the uncertainty is decreased by making the magnitude of the selected point larger³². We show in Supplementary Information B that this is the case also for the problem considered in this paper. Considering the hypersphere allows us to avoid this ambiguity. The following proposition provides the resulting formula when Σ is known; see Supplementary Information B for the general case including the proof.

Proposition 2. For $\mathcal{A} = \mathbb{S}^{p-1}$ and ν being a uniform measure, the CIV acquisition function evaluated at \mathbf{a}' , conditioned on Σ , is

$$h(\mathbf{a}' | \mathcal{D}_t) = c_1 \cdot \sum_{i=1}^p \left(v_i^2 + 2v_i \left(\frac{\sigma_i^2}{n} + b_i^2 + \frac{1}{p} \right) \right) + c_2, \quad (13)$$

where $c_1 > 0$ and c_2 are constants that do not depend on \mathbf{a}' .

The gradient of this objective function can be calculated explicitly also in the general case when Σ is unknown, and thus, we can use gradient-based optimization methods with a ball constraint to solve for $\mathbf{a}^{(t+1)} = \text{argmin}_{\mathbf{a}' \in \mathbb{S}^{p-1}} \sum_{i=1}^p (v_i^2 + 2v_i(\frac{\sigma_i^2}{n} + b_i^2 + \frac{1}{p}))$ (see the section

Implementations below for details). For other types of feasible sets \mathcal{A} , for example, with sparsity constraints or consisting of hard interventions, the appropriate optimization method needs to be adjusted accordingly. We also note that the objective function is not necessarily convex (Supplementary Information B), and gradient-based optimizers may therefore only find a local minimum.

Mutual information bound

To provide an information-theoretic interpretation of the CIV acquisition function, we use the ‘relative decay’ of the uncertainty $\sigma_{g(\mathbf{a})|\mathcal{D}_t}^2$ measured by

$$\frac{\sigma_{g(\mathbf{a})|\mathcal{D}_t}^2 - \mathbb{E}_{\mathbf{x}'}(\sigma_{g(\mathbf{a})|\mathcal{D}_t \cup \{\mathbf{x}', \mathbf{a}'\}}^2)}{\sigma_{g(\mathbf{a})|\mathcal{D}_t}^2}, \quad (14)$$

where the expectation $\mathbb{E}_{\mathbf{x}'}$ is taken with respect to a new sample \mathbf{x}' from \mathbf{a}' whose distribution is given by $\mathbb{P}(\mathbf{x}' | \mathcal{D}_t, \mathbf{a}')$. To simplify the notation, we restrict the discussion in this section to the case of $n = 1$. Similar

results can easily be derived for sample sizes $n > 1$. The following theorem shows that the mutual information between $g(\mathbf{a})$ and a sample \mathbf{x}' from \mathbf{a}' can be lower bounded by this relative decay up to a multiplicative factor. The proof can be found in Supplementary Information E.

Theorem 1. Conditioning on \mathcal{D}_t , the following inequality holds for any $\mathbf{a} \in \mathcal{A}$ for the mutual information between $g(\mathbf{a})$ and a sample \mathbf{x}' from $\mathcal{P}^{\mathbf{a}}$:

$$I(g(\mathbf{a}); \mathbf{x}' | \mathbf{a}', \mathcal{D}_t) \geq \rho^2 \cdot \frac{\sigma_{g(\mathbf{a})|\mathcal{D}_t}^2 - \mathbb{E}_{\mathbf{x}'}(\sigma_{g(\mathbf{a})|\mathcal{D}_t \cup (\mathbf{x}', \mathbf{a}')}^2)}{\sigma_{g(\mathbf{a})|\mathcal{D}_t}^2}, \quad (15)$$

where ρ^2 is a constant that does not depend on \mathbf{a}' .

This result holds in general for any function $g(\mathbf{a})$ and its posterior modelling on \mathcal{D}_t , but the constant ρ in the bound depends on the specific choice of g and its posterior modelling. For example, for the prominent BLR task, Theorem 1 holds with $\rho^2 = 1/2$ (Supplementary Information E).

Note that as $\sigma_{g(\mathbf{a})|\mathcal{D}_t}^2$ does not depend on $\mathbf{a}' \in \mathcal{A}$, minimizing the uncertainty $\mathbb{E}_{\mathbf{x}'}(\sigma_{g(\mathbf{a})|\mathcal{D}_t \cup (\mathbf{x}', \mathbf{a}')}^2)$ is equivalent to maximizing its relative decay in equation (14). The uncertainty $\sigma_{g(\mathbf{a})|\mathcal{D}_t(\mathbf{a}')}^2$ used by the CIV acquisition function (equation 7) is a computationally efficient estimator of $\mathbb{E}_{\mathbf{x}'}(\sigma_{g(\mathbf{a})|\mathcal{D}_t \cup (\mathbf{x}', \mathbf{a}')}^2)$ in the relative decay (equation 14), where the expectation with respect to $\mathbf{x}' \sim \mathbb{P}(\mathbf{x}' | \mathcal{D}_t, \mathbf{a}')$ is replaced by a plug-in estimator $\bar{\mathbf{x}}' = \mathbb{E}(\mathbf{x}' | \mathcal{D}_t, \mathbf{a}')$. It follows from Theorem 1 that minimizing the integrated variance in equation (7) corresponds to selecting an intervention that maximizes a lower bound on the mutual information between the resulting samples and $g(\mathbf{a})$. This perspective shows how the CIV acquisition function connects to a prominent line of previous works, known as Bayesian active learning by disagreement (BALD)^{79,41}, in which the idea is to select interventions that maximize the information gain at each step. For example, if the quantity of interest is $g(\mathbf{a})$, then BALD seeks to find \mathbf{a}' such that the information gain measured by the entropy decay

$$H(g(\mathbf{a})|\mathcal{D}_t) - H(g(\mathbf{a})|\mathcal{D}_t \cup (\mathbf{x}', \mathbf{a}')) \quad (16)$$

is maximized. Note that this term equals $I(g(\mathbf{a}); \mathbf{x}' | \mathbf{a}', \mathcal{D}_t)$. As a consequence, BALD directly operates on the mutual information, which involves high-dimensional moments of the random variable $g(\mathbf{a})$. This highlights an important drawback of BALD, namely, that it is generally computationally difficult to evaluate and often requires approximation techniques such as Markov chain Monte Carlo⁹. By contrast, our CIV acquisition function only depends on the first and the second moments of $g(\mathbf{a})$, thereby allowing for direct computation and efficient optimization.

Consistency of CIV for identifying the optimal intervention

We now provide a graphical interpretation of the proposed CIV acquisition function. For simplicity, we restrict our discussion to the expression developed in Proposition 2 for the case of uniform measure and known variances. The general setting is discussed in Supplementary Information F. As c_1, c_2 in Proposition 2 are constants that do not depend on \mathbf{a}' , the new intervention is obtained by minimizing $\sum_{i=1}^p \left(v_i^2 + 2v_i \left(\frac{\sigma_i^2}{n} + b_i^2 + \frac{1}{p} \right) \right)$; with a slight abuse of notation, let

$$h(\mathbf{a}', \mathcal{D}_t) = \sum_{i=1}^p \left(v_i^2 + 2v_i \left(\frac{\sigma_i^2}{n} + b_i^2 + \frac{1}{p} \right) \right). \quad (17)$$

The term $v_i = \mu_{\text{pa}(i)}^{*\top} \text{Var}(B_{i,\text{pa}(i)} | \mathcal{D}_t(\mathbf{a}')) \mu_{\text{pa}(i)}^*$ can be understood as the epistemic uncertainty of node i in the direction of interest, namely, as $x_i = B_{i,\text{pa}(i)}^\top x_{\text{pa}(i)} + \epsilon_i$, the variance of $B_{i,\text{pa}(i)}$ characterizes the epistemic uncertainty of estimating x_i using its parent nodes; in v_i , this uncertainty, presented as a covariance matrix, is transformed into a scalar value by scaling it in the direction of the target mean μ^* . As the number

of time steps t increases, one can show (Supplementary Information E) using the Bernstein–von Mises theorem³¹ that $v_i = O(1/t)$, which implies that the second-order terms v_i^2 are dominated by the first-order terms in equation (17). Note that their coefficients $\sigma_i^2/n + b_i^2 + 1/p$ are larger for nodes i with larger noise variances and larger estimated optimal shift values $b_i = [(I - \mathbb{E}(B|\mathcal{D}_t))\mu^*]_i$. This scaling is intuitive, as it puts emphasis on reducing the variance of nodes in which there is still a high uncertainty (given by σ_i^2) or that require high shift values (estimated by b_i^2). As we show in the following theorem, the combined effect of these terms on acquiring new interventions is that \mathbf{a}^* will become an approximate local minimizer of equation (17) as t increases as shown in the following theorem, in which we use the conventional asymptotic notation, o , O and Θ , with respect to the time step t .

Theorem 2. For all $\mathbf{a}' \in \mathcal{A}$, the CIV acquisition function either decays linearly with respect to t , that is, $h(\mathbf{a}', \mathcal{D}_t) = \Theta(1/t)$, or degenerates to the constant zero, that is, $h(\mathbf{a}', \mathcal{D}_t) \equiv 0$. The optimal intervention \mathbf{a}^* satisfies $\|\nabla h(\mathbf{a}^*, \mathcal{D}_t)\|_2 = O(1/t^2)$ and $\nabla^2 h(\mathbf{a}^*, \mathcal{D}_t) \succeq -O(1/t^2)I$.

This theorem shows that as t increases, the gradient at \mathbf{a}^* vanishes to zero faster than the acquisition function and the Hessian becomes almost positive semi-definite. This suggests that gradient-based optimizers with an initialization close enough to \mathbf{a}^* converge to the optimal intervention \mathbf{a}^* as $t \rightarrow \infty$. In experiments, we observe this to hold even for a moderate number of time steps t as illustrated for a ten-node DAG example in Fig. 4. The proof of Theorem 2 can be found in Supplementary Information F.

Implementations

Optimization of the acquisition functions was performed using a gradient-based solver, more precisely, sequential least squares programming with the nonlinear constraint $\|\mathbf{a}^{(i)}\|_2 \leq 1$. As proven in Supplementary Information B, optimizing the proposed acquisition functions with this constraint always outputs a feasible $\mathbf{a}^{(i)}$ such that $\|\mathbf{a}^{(i)}\|_2 = 1$. For other types of feasible sets, for example, $\mathbf{a}^{(i)} \in [0, 1]^p$, one can use other solvers such as limited-memory Broyden–Fletcher–Goldfarb–Shanno with box constraints. Sequential least squares programming was implemented using the SciPy package⁴² and initialized in two ways: (1) using the intervention $\mathbf{a}^{(t-1)}$ that was selected in the previous time step and (2) using the current estimate $(I - \mathbb{E}(B|\mathcal{D}_t))\mu^*$ of the optimal intervention. Of the two resulting solutions, we used the one with the better acquisition function value as $\mathbf{a}^{(i)}$.

For the experiments on synthetic data, we generated the DAGs using the NetworkX package⁴³ and the CausalDAG package⁴⁴. We used different graph types and sizes, as detailed in Supplementary Information G. The edge weights were uniformly sampled from $[-1, -0.25] \cup [0.25, 1]$ to ensure that the parameters were bounded away from zero. The exogenous noise levels were set to 1, and the resulting linear Gaussian SCM was then rescaled with our implementation of the standardized model as described in Appendix F of ref. 45. For the DAG–BLR prior (equation 10), we set the edge-weight-related hyperparameters as follows: $m_i = 0$ and M_i equal to the identity matrix for all $i \in [p]$. We assumed the variances to be known and used the known-variance formula.

The transcriptomic dataset analysed in the biological application was processed using the Scanpy package⁴⁶. The kernel density estimate plots shown in Fig. 6a share the same coordinate axes and were produced by projecting the high-dimensional dataset along the first two principal components obtained using contrastive PCA⁴⁷, with the unperturbed samples as ‘background’ and all interventional samples as ‘foreground’. The purpose of contrastive PCA is to identify low-dimensional structures in a foreground dataset relative to a background dataset. The greedy sparsest permutation algorithm⁴⁸ was used to learn the DAG structure from the unperturbed data. The parameters used in the greedy sparsest permutation are given in Supplementary Information H and Fig. 7. For the DAG–BLR prior, we set the hyperparameters as follows: $m_i = 0$, M_i equal to the identity matrix,

$\alpha_i = 2$ and $\beta_i = 0$. These hyperparameters were then updated using posterior formula 1 based on 100 samples from the unperturbed data. Similar warm-up steps are used in the active learning literature^{7,32}. As we usually have access to some observational samples in applications of optimal intervention design, we recommend using these to obtain dataset-informed hyperparameters of the DAG–BLR prior. After this warm-up step using unperturbed data, interventions were acquired based on the known-variance formula, in which we used the estimated mean of the noise variances as a proxy.

The codebase for updating the model posterior and optimizing the causally aware acquisition functions proposed in this work can be obtained via refs. 49. We also provide a notebook to extract and process the Perturb–CITE-seq data from ref. 36. Our codebase can be used to replicate all the main results and figures as well as for other user-defined applications. All methods can be run efficiently on a CPU for a moderate number of variables.

Data availability

The Perturb–CITE-seq³⁶ data can be obtained from <https://doi.org/10.1038/s41588-021-00779-1>.

Code availability

All code has been deposited at ref. 49.

References

- Cherry, A. B. & Daley, G. Q. Reprogramming cellular identity for regenerative medicine. *Cell* **148**, 1110–1122 (2012).
- Todorov, E. & Jordan, M. I. Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.* **5**, 1226–1235 (2002).
- Blanchard, A. B. et al. Bayesian optimization for active flow control. *Acta Mech. Sin.* **37**, 1786–1798 (2021).
- Sunar, N., Birge, J. R. & Vitvasiri, S. Optimal dynamic product development and launch for a network of customers. *Oper. Res.* **67**, 770–790 (2019).
- Serrao-Neumann, S., Di Giulio, G. M., Ferreira, L. C. & Choy, D. L. Climate change adaptation: is there a role for intervention research? *Futures* **53**, 86–97 (2013).
- Fu, Y., Zhu, X. & Li, B. A survey on instance selection for active learning. *Knowl. Inf. Syst.* **35**, 249–283 (2013).
- Jesson, A. et al. Causal-BALD: deep Bayesian active learning of outcomes to infer treatment-effects from observational data. In *Adv. Neural Information Processing Systems* Vol. 34, 30465–30478 (NeurIPS, 2021).
- Cohn, D. A., Ghahramani, Z. & Jordan, M. I. Active learning with statistical models. *J. Artif. Intell. Res.* **4**, 129–145 (1996).
- Houlsby, N., Huszár, F., Ghahramani, Z. & Lengyel, M. Bayesian active learning for classification and preference learning. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.1112.5745> (2011).
- Lattimore, F., Lattimore, T. & Reid, M. D. Causal bandits: learning good interventions via causal inference. In *Adv. Neural Information Processing Systems* Vol. 29 (2016).
- Lee, S. & Bareinboim, E. Structural causal bandits: where to intervene? In *Adv. Neural Information Processing Systems* Vol. 31 (2018).
- Aglietti, V., Lu, X., Paley, A. & González, J. Causal Bayesian optimization. In *Int. Conf. Artificial Intelligence and Statistics* 3155–3164 (PMLR, 2020).
- Alabed, S. & Yoneki, E. BoGraph: structured Bayesian optimization from logs for expensive systems with many parameters. In *Proc. 2nd European Workshop on Machine Learning and Systems* 45–53 (2022).
- Branchini, N., Aglietti, V., Dhir, N. & Damoulas, T. Causal entropy optimization. In *Int. Conf. on Artificial Intelligence and Statistics* 8586–8605 (PMLR, 2023).
- Cahan, P. et al. CellNet: network biology applied to stem cell engineering. *Cell* **158**, 903–915 (2014).
- Kemmeren, P. et al. Large-scale genetic perturbations reveal regulatory networks and an abundance of gene-specific repressors. *Cell* **157**, 740–752 (2014).
- Spirtes, P., Glymour, C. N., Scheines, R. & Heckerman, D. *Causation, Prediction, and Search* (MIT Press, 2000).
- Pearl, J. *Causality* (Cambridge Univ. Press, 2009).
- Rothenhäusler, D., Heinze, C., Peters, J. & Meinshausen, N. Backshift: learning causal cyclic graphs from unknown shift interventions. In *Adv. Neural Information Processing Systems* Vol. 28 (2015).
- Zhang, J., Squires, C. & Uhler, C. Matching a desired causal state via shift interventions. In *Adv. Neural Information Processing Systems* Vol. 34 (2021).
- Eberhardt, F. & Scheines, R. Interventions and causal inference. *Philos. Sci.* **74**, 981–995 (2007).
- Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR–Cas9. *Nat. Rev. Genet.* **16**, 299–311 (2015).
- Joung, J. et al. A transcription factor atlas of directed differentiation. *Cell* **186**, 209–229 (2023).
- Replogle, J. M. et al. Mapping information-rich genotype–phenotype landscapes with genome-scale Perturb-seq. *Cell* **185**, 2559–2575 (2022).
- Sen, R., Shanmugam, K., Dimakis, A. G. & Shakkottai, S. Identifying best interventions through online importance sampling. In *Int. Conf. Machine Learning* 3057–3066 (PMLR, 2017).
- Koumoutsakos, P. & Leonard, A. High-resolution simulations of the flow around an impulsively started cylinder using vortex methods. *J. Fluid Mech.* **296**, 1–38 (1995).
- Rackham, O. J. et al. A predictive computational framework for direct reprogramming between human cell types. *Nat. Genet.* **48**, 331–335 (2016).
- Geiger, D. & Heckerman, D. Parameter priors for directed acyclic graphical models and the characterization of several probability distributions. *Ann. Stat.* **30**, 1412–1440 (2002).
- Kuipers, J. & Moffa, G. The interventional Bayesian Gaussian equivalent score for Bayesian causal inference with unknown soft interventions. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2205.02602> (2022).
- Kuipers, J., Moffa, G. & Heckerman, D. Addendum on the scoring of Gaussian directed acyclic graphical models. *Ann. Statist.* **42**, 1689–1691 (2014).
- Kleijn, B. J. & van der Vaart, A. W. The Bernstein–von-Mises theorem under misspecification. *Electron. J. Stat.* **6**, 354–381 (2012).
- Sapsis, T. P. Output-weighted optimal sampling for Bayesian regression and rare event statistics using few samples. *Proc. R. Soc. A* **476**, 20190834 (2020).
- Mohamad, M. A. & Sapsis, T. P. Sequential sampling strategy for extreme event statistics in nonlinear dynamical systems. *Proc. Natl Acad. Sci. USA* **115**, 11138–11143 (2018).
- Astudillo, R. & Frazier, P. Bayesian optimization of function networks. In *Adv. Neural Information Processing Systems* Vol. 34, 14463–14475 (NeurIPS, 2021).
- Bubeck, S. et al. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Found. Trends Mach. Learn.* **5**, 1–122 (2012).
- Frangieh, C. J. et al. Multimodal pooled Perturb–CITE-seq screens in patient models define mechanisms of cancer immune evasion. *Nat. Genet.* **53**, 332–341 (2021).
- Carretero, R. et al. Analysis of HLA class I expression in progressing and regressing metastatic melanoma lesions after immunotherapy. *Immunogenetics* **60**, 439–447 (2008).
- Jaeger, J. et al. Gene expression signatures for tumor progression, tumor subtype, and tumor thickness in laser-microdissected melanoma tissues. *Clin. Cancer Res.* **13**, 806–815 (2007).

39. Cheng, Q. et al. SOX4 promotes melanoma cell migration and invasion through the activation of the NF- κ B signaling pathway. *Int. J. Mol. Med.* **40**, 447–453 (2017).
40. Cao, X., Khare, K. & Ghosh, M. Posterior graph selection and estimation consistency for high-dimensional Bayesian DAG models. *Ann. Stat.* **47**, 319–348 (2019).
41. Kirsch, A., Van Amersfoort, J. & Gal, Y. BatchBALD: efficient and diverse batch acquisition for deep Bayesian active learning. In *Adv. Neural Information Processing Systems* Vol. 32 (2019).
42. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
43. Hagberg, A., Swart, P. & Schult, D. A. *Exploring Network Structure, Dynamics, and Function Using NetworkX* (Los Alamos National Lab, 2008).
44. Squires, C. CausalDAG: creation, manipulation, and learning of causal models. *GitHub* <https://github.com/uhlerlab/causal DAG> (2018).
45. Reisach, A., Seiler, C. & Weichwald, S. Beware of the simulated DAG! Causal discovery benchmarks may be easy to game. In *Adv. Neural Information Processing Systems* Vol. 34, 27772–27784 (NeurIPS, 2021).
46. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 1–5 (2018).
47. Abid, A., Zhang, M. J., Bagaria, V. K. & Zou, J. Exploring patterns enriched in a dataset with contrastive principal component analysis. *Nat. Commun.* **9**, 1–7 (2018).
48. Solus, L., Wang, Y. & Uhler, C. Consistency guarantees for greedy permutation-based causal inference algorithms. *Biometrika* **108**, 795–814 (2021).
49. Zhang, J. uhlerlab/actlearn_optint: v1, July. *Zenodo* <https://doi.org/10.5281/zenodo.8170179> (2023).

Acknowledgements

J.Z., C.S. and C.U. were partially supported by the National Center for Complementary and Integrative Health at the National Institutes of Health (NCCIH/NIH), the Office of Naval Research (N00014-22-1-2116), the National Science Foundation (DMS-1651995), the MIT-IBM Watson AI Lab, MIT J-Clinic for Machine Learning and Health, the Eric and Wendy Schmidt Center at the Broad Institute and a Simons Investigator Award to C.U. T.P.S. acknowledges support by the Office of Naval Research (N00014-21-1-2357) and the Air Force Office of

Scientific Research (MURI FA9550-21-1-0058). C.S. was partially supported by an NSF Graduate Fellowship.

Author contributions

J.Z., L.C., T.P.S. and C.U. conceived the research and designed the method. J.Z. derived the theoretical results and performed the numerical experiments. L.C. and J.Z. processed the biological data. C.S. and J.Z. derived the extension of the DAG–Wishart distribution. All authors interpreted the results and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s42256-023-00719-0>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42256-023-00719-0>.

Correspondence and requests for materials should be addressed to Themistoklis P. Sapsis or Caroline Uhler.

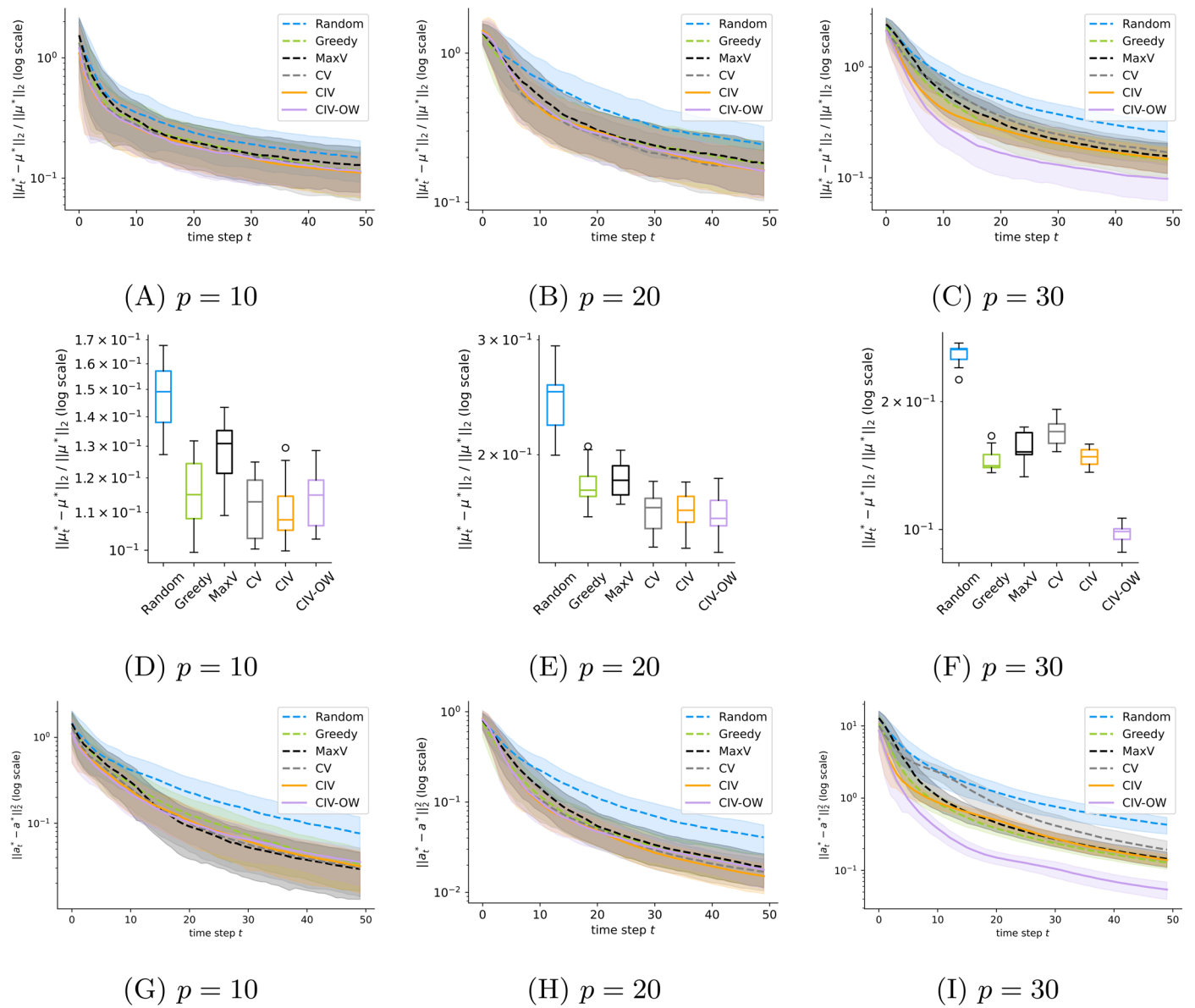
Peer review information *Nature Machine Intelligence* thanks Virginia Aglietti and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Mirko Pieropan, in collaboration with the *Nature Machine Intelligence* team.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

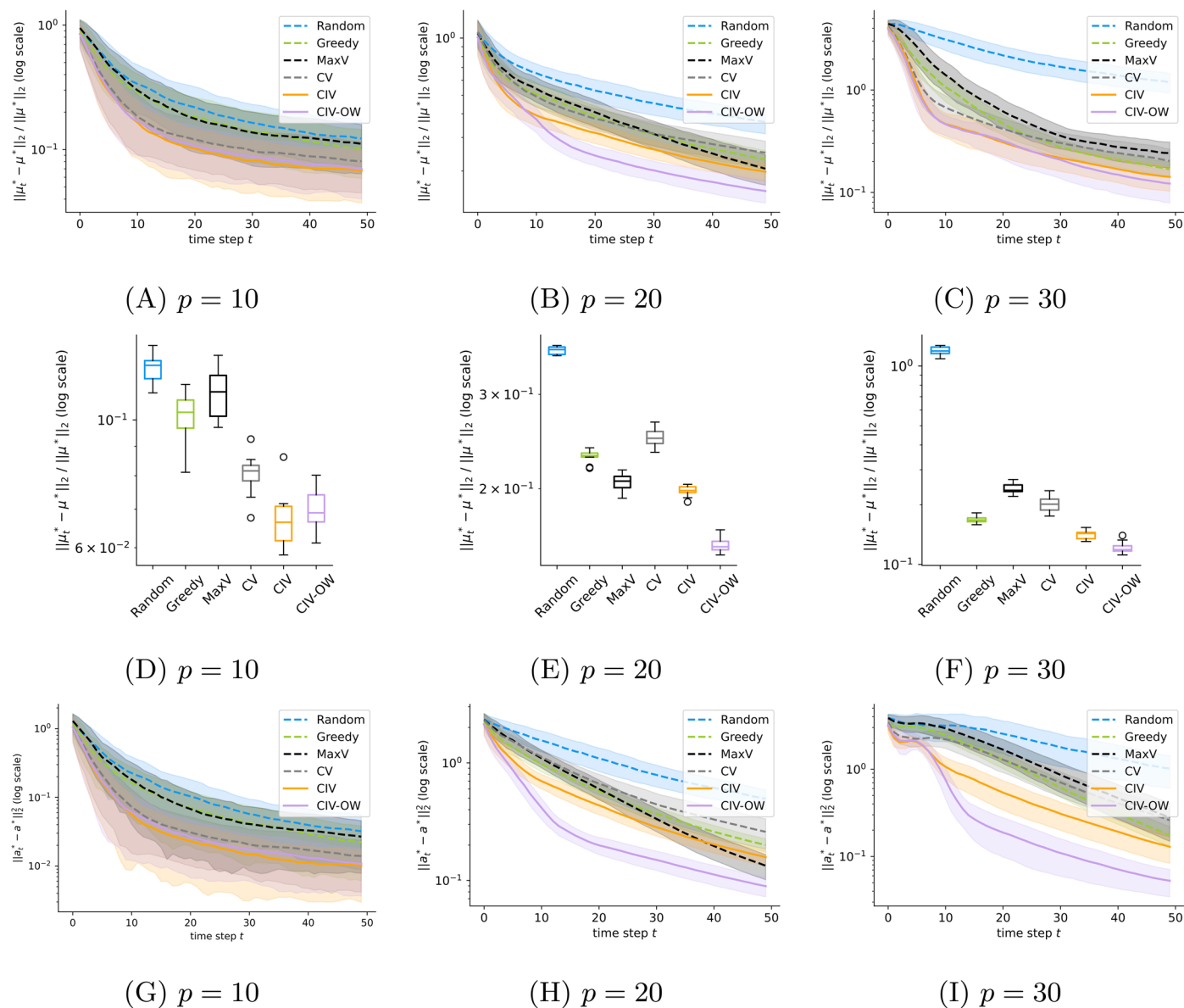
Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023



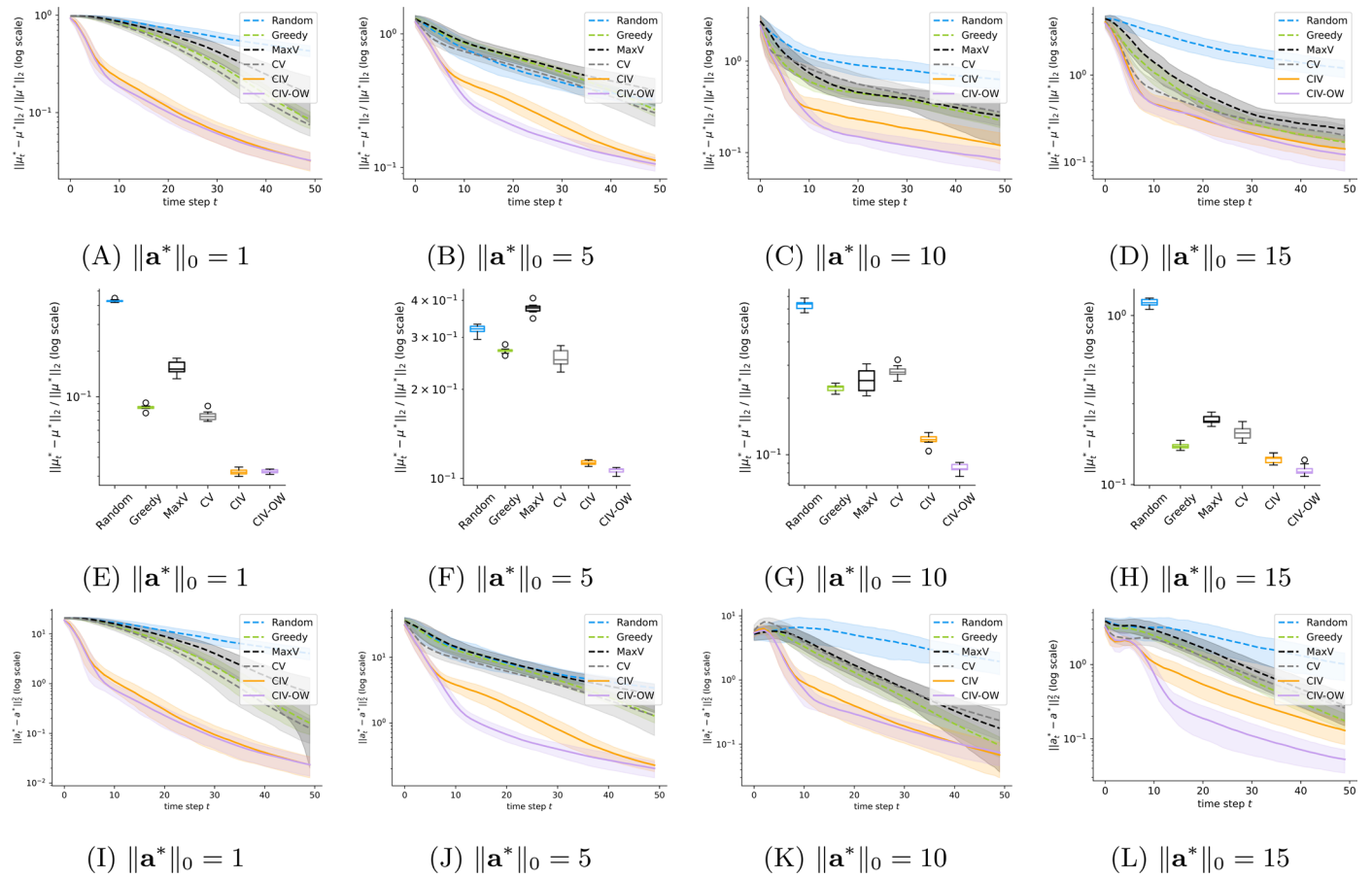
Extended Data Fig. 1 | Comparison of different acquisition functions in a simulation study where the underlying causal graph is the complete graph, half of the nodes are selected at random as intervention targets, and we vary the number of nodes p . Each plot corresponds to an average of 10 instances and each method is run 20 times and averaged. (A)-(C) Relative distance between the target mean μ and the best approximation μ_t^* (defined in Fig. 5A in the main text)

up to time step t . Lines denote the mean over 10 instances; the shading corresponds to one standard deviation. (D)-(F) Relative distance statistics of each method averaged over 10 instances at the last time step ($t = 50$). (G)-(I) Squared distance presented as mean value \pm SEM between the optimal intervention a^* and the best approximation a_t^* that is used to obtain μ_t^* up to time step t .



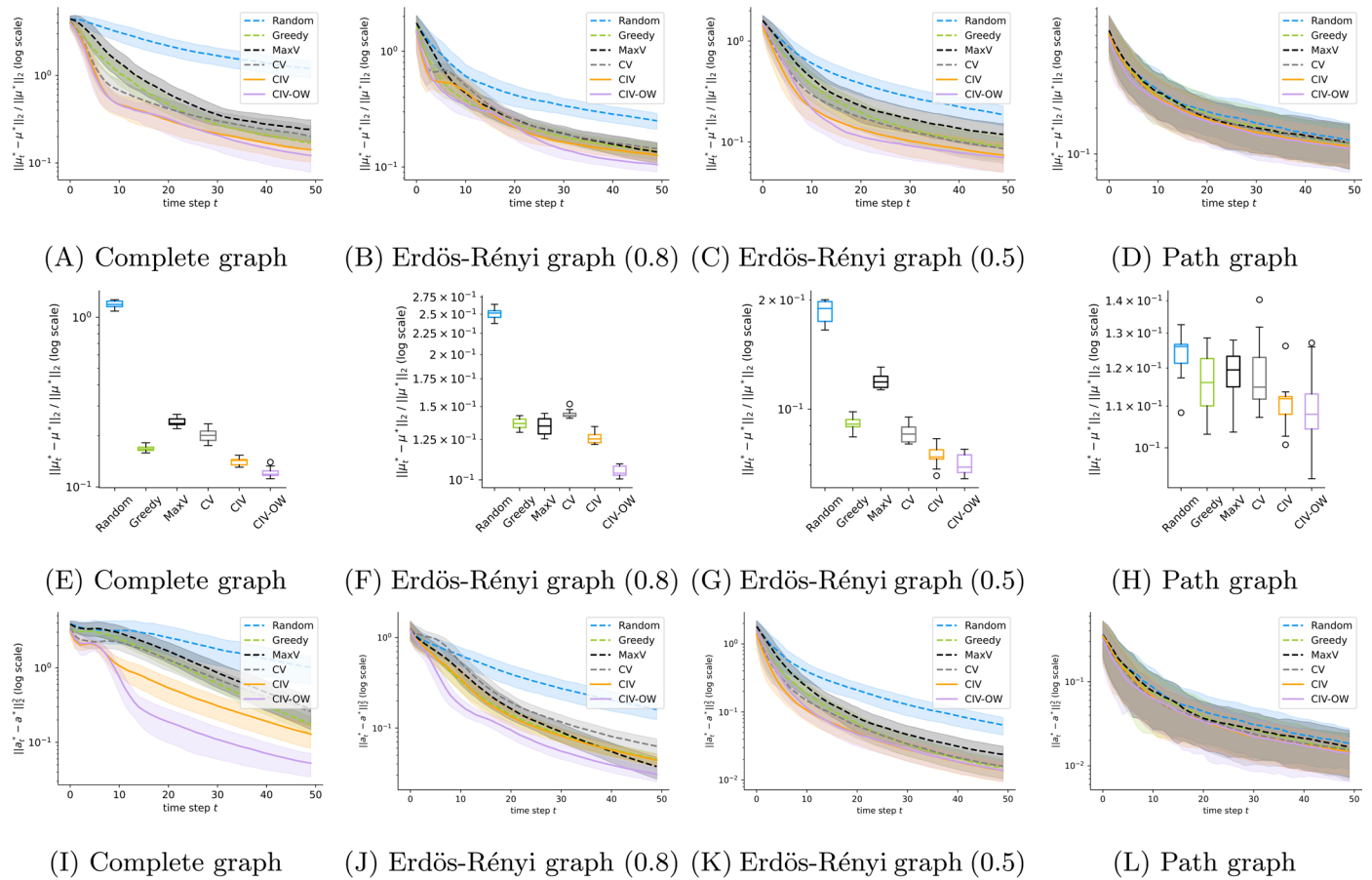
Extended Data Fig. 2 | Comparison of different acquisition functions in a simulation study where the underlying causal graph is the complete graph, the most downstream half of the nodes are fixed as intervention targets, and we vary the number of nodes p . Each plot corresponds to an average of 10 instances and each method is run 20 times and averaged. (A)-(C) Relative distance between the target mean μ and the best approximation μ_t^* (defined in

Fig. 5A in the main text) up to time step t . Lines denote the mean over 10 instances; the shading corresponds to one standard deviation. (D)-(F) Relative distance statistic of each method averaged over 10 instances at the last time step ($t=50$). (G)-(I) Squared distance presented as mean value \pm SEM between the optimal intervention a^* and the best approximation a_t^* that is used to obtain μ_t^* up to time step t .



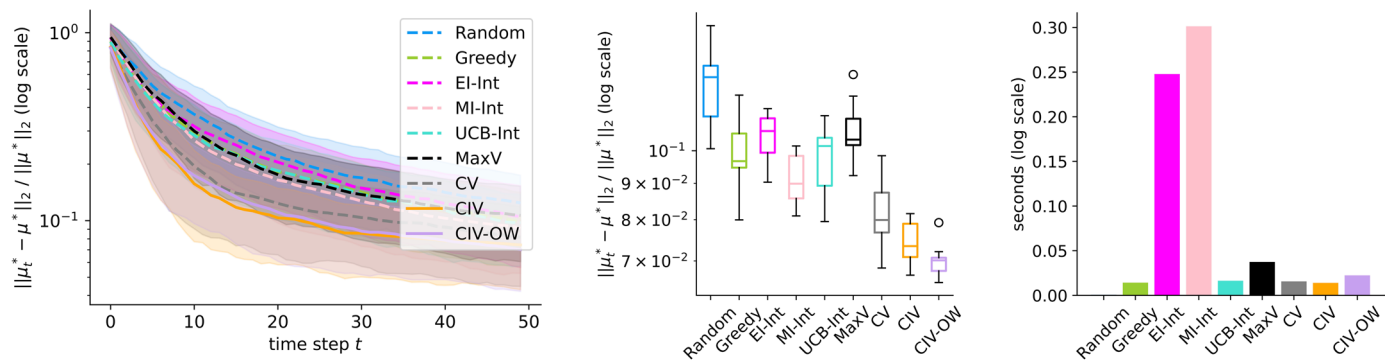
Extended Data Fig. 3 | Comparison of different acquisition functions in a simulation study where the underlying causal graph is the complete graph on 30 nodes, the most downstream nodes are fixed as intervention targets, and we vary the number of intervention targets. Each plot corresponds to an average of 10 instances and each method is run 20 times and averaged. (A)-(D) Relative distance between the target mean μ and the best approximation μ_t^*

(defined in Fig. 5A in the main text) up to time step t . Lines denote the mean over 10 instances; the shading corresponds to one standard deviation. (E)-(H) Relative distance statistic of each method averaged over 10 instances at the last time step ($t = 50$). (I)-(L) Squared distance presented as mean value \pm SEM between the optimal intervention a^* and the best approximation a_t^* that is used to obtain μ_t^* up to time step t .



Extended Data Fig. 4 | Comparison of different acquisition functions in a simulation study where we vary the underlying causal graph (complete graph, Erdős-Rényi graph with edge probability 0.8, Erdős-Rényi graph with edge probability 0.5, path graph) and the most downstream half of the nodes are fixed as intervention targets. Each plot corresponds to an average of 10 instances on a 30-node DAG with 15 perturbation targets. Each method is run 20 times and averaged. (A)-(D) Relative distance between the target mean μ^* and the

best approximation μ_t^* (defined in Fig. 5A in the main text) up to time step t . Lines denote the mean over 10 instances; the shading corresponds to one standard deviation. (E)-(H) Relative distance statistic of each method averaged over 10 instances at the last time step ($t = 50$). Note that the DAGs become sparser from left to right. (I)-(L) Squared distance presented as mean value +/- SEM between the optimal intervention a^* and the best approximation a_t^* that is used to obtain μ_t^* up to time step t .



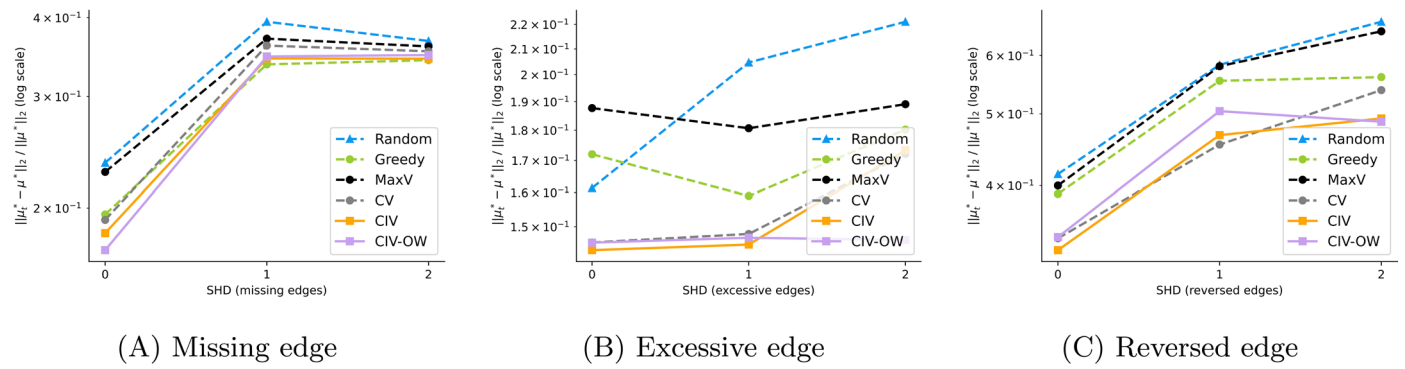
(A) Relative distance across time steps

(B) Last time step

(C) Runtime per iteration

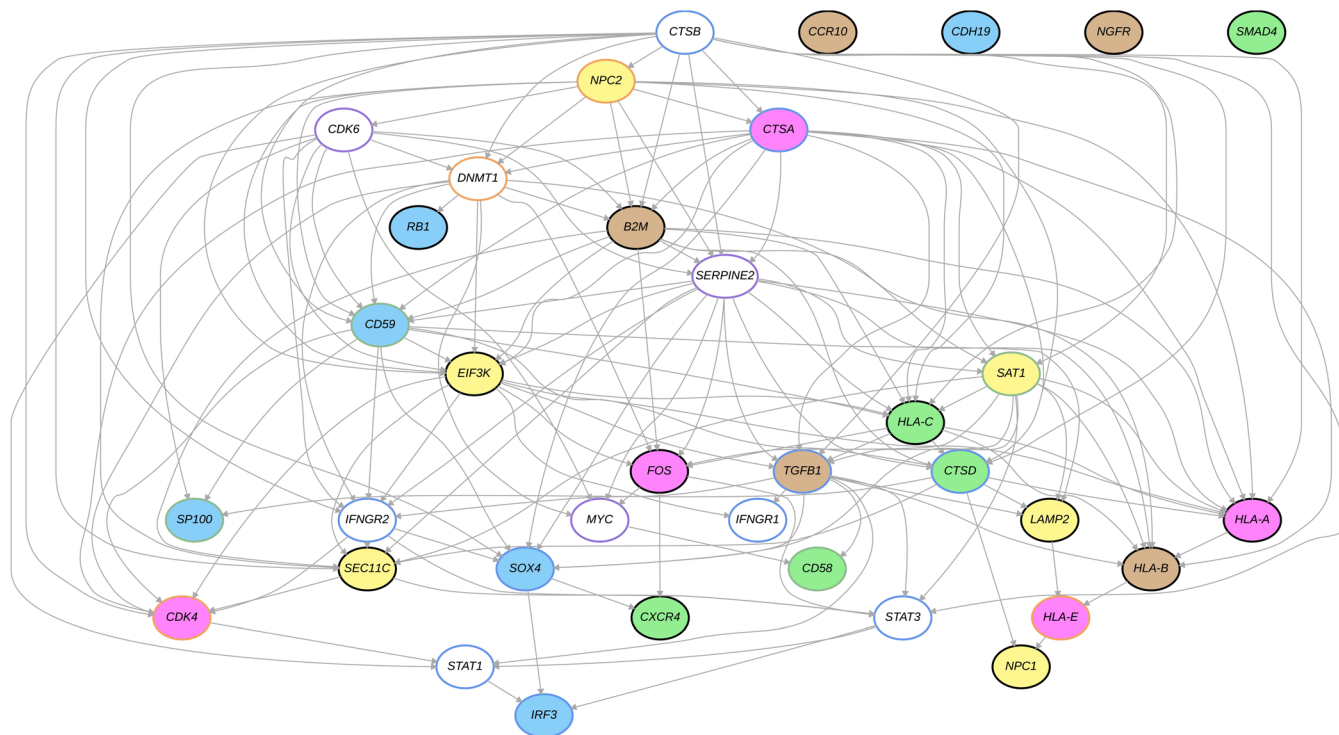
Extended Data Fig. 5 | Comparison of our acquisition functions to baseline acquisition functions adapted from prior works (EI-Int: based on Expected Improvement, MI-Int: based on Mutual Information, and UCB-Int: based on Upper Confidence Bound) in a simulation study where the underlying causal graph is the complete graph on 10 nodes and the most downstream 5 nodes are fixed as intervention targets. Each plot corresponds to an average of 10

instances and each method is run 20 times and averaged. (A) Relative distance between the target mean μ^* and the best approximation μ_t^* (defined in Fig. 5A in the main text) up to time step t . Lines denote the mean over 10 instances; the shading corresponds to one standard deviation. (B) Relative distance statistic of each method averaged over 10 instances at the last time step ($t = 50$). (C) Runtime per iteration of each method in seconds.



Extended Data Fig. 6 | Performance of the different acquisition functions under three types of DAG misspecifications where the underlying causal graph is a 5-node random Erdős-Rényi DAG with edge density 0.5 and 3 intervention targets. Each plot corresponds to an average of the relative

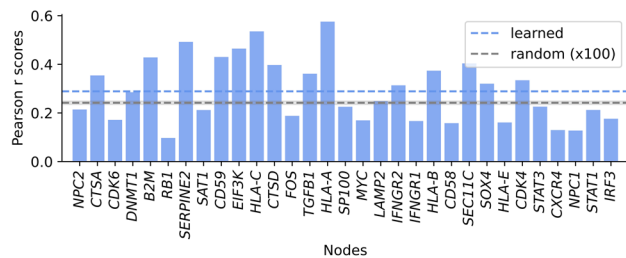
distance at time step 10 across 10 instances. Each method is run 10 times and averaged. SHD denotes the number of misspecified edges. (A)-(C) Three types of DAG misspecifications.



(A)

function/class	parameter	value
GSP	conditional independence tester	MemoizedCI_Tester
	number of runs	1000
CI Tester	conditional independence test	partial correlation test
	significance level ("α")	1E-03

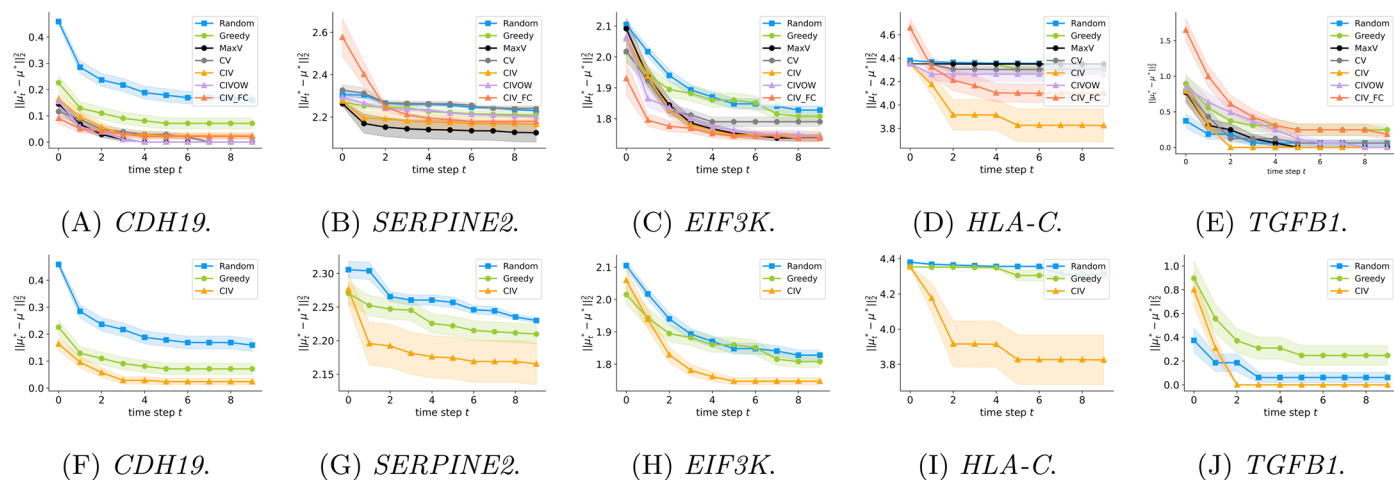
(B)



(C)

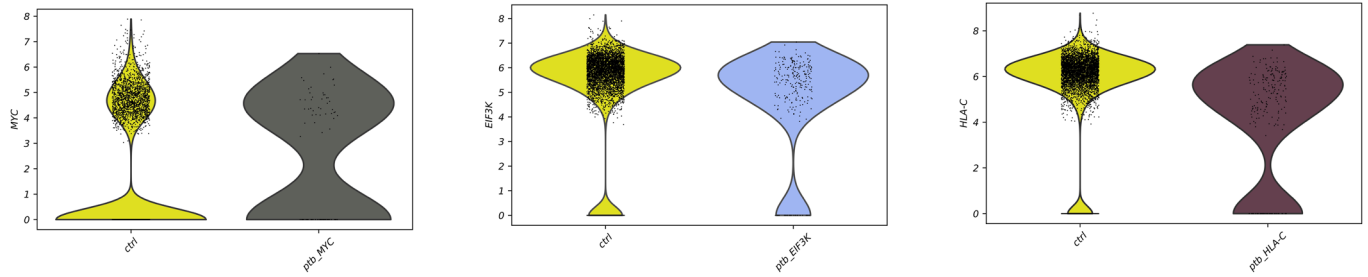
Extended Data Fig. 7 | Learned linear Gaussian SCM on the 36 genes of interest based on the control cells. (A) Learned DAG on the 36 considered genes. Nodes are oriented up-down by their topological order and colored by the module/program they belong to in Supplementary Fig. 7. (B) Parameters used in GSP⁴⁸

to learn the above DAG. (C) Pearson r scores of regressing each non-source gene against its parents. Blue: average scores in the learned DAG; grey (with errorbars): average of scores in a random graph (100 samples).



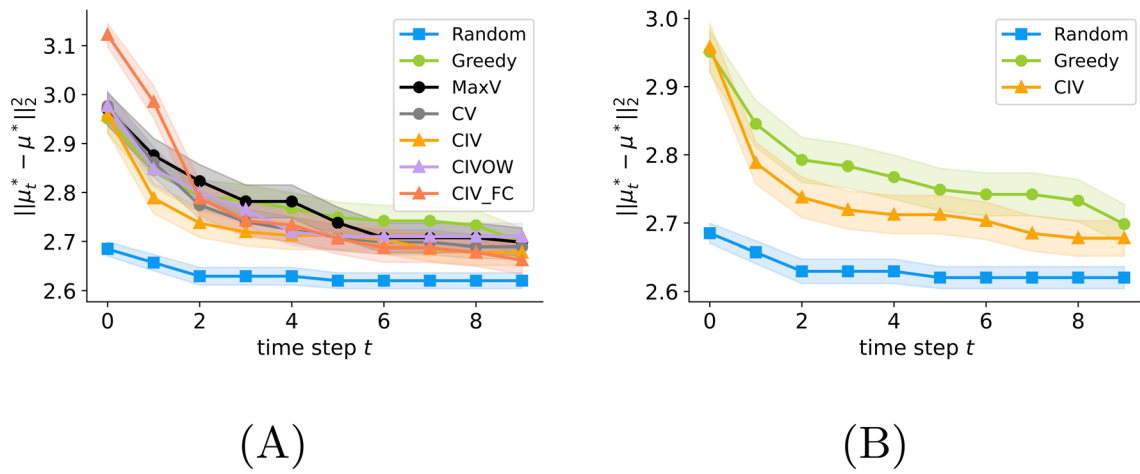
Extended Data Fig. 8 | Comparison of the different acquisition functions for identifying the intervention that matches the target mean for 5 different ground-truth target genes. Square distance presented as mean value \pm SEM between the target mean μ^* and the best approximation μ_t^* across time step t is

reported. (A)-(E) comparison of 6 acquisition functions. (F)-(J) same as top, showing only 3 methods to de-clutter the plots, comparing our CIV acquisition function against the random and greedy baseline. Each plot is captioned with its ground-truth target gene.



(A) *MYC* expression. Mean: 1.52 (ctrl) v.s. 2.58 (ptb_ *MYC*). (B) *EIF3K* expression. Mean: 5.19 (ctrl) v.s. 4.78 (ptb_ *EIF3K*). (C) *HLA-C* expression. Mean: 5.98 (ctrl) v.s. 4.16 (ptb_ *HLA-C*).

Extended Data Fig. 9 | Gene expression changes for three examples of different knock-out perturbations. Comparing target-gene expression in the control cell population and the perturbed cell population of the corresponding knock-out experiment. (A)-(C) Included target genes: *MYC*, *EIF3K*, and *HLA-C*. The mean expression of the target gene is given in each subcaption.



Extended Data Fig. 10 | Comparison of acquisition functions for identifying interventions that match the target mean of perturbing *MYC*. The reported metric is the square distance presented as mean value \pm SEM between the target mean μ^* and the best approximation μ_t^* across all time steps t . (A) All methods. (B) De-cluttered subset of methods.