Foundations of Testing for Finite-Sample Causal Discovery

Tom Yan [1] **Ziyu Xu** [2] **Zachary Lipton** [1]

Abstract

Discovery of causal relationships is a fundamental goal of science and vital for sound decision making. As such, there has been considerable interest in causal discovery methods with *provable* guarantees. Existing works have thus far largely focused on discovery under hard intervention and infinite-samples, in which intervening on a node readily reveals the orientation of every edge incident to the node. This setup however overlooks the stochasticity inherent in real-world, finite-sample settings. Our work takes a step towards studying finite-sample causal discovery, wherein multiple interventions on a node are now needed for edge orientation. In this work, we study the canonical setup in theoretical causal discovery literature, where one assumes causal sufficiency and access to the graph skeleton. Our key observation is that discovery may be viewed as structured, multiple testing, and we develop a novel testing framework to this end. Crucially, our framework allows for anytime valid testing as multiple tests are needed to conclude an edge orientation. It also allows for flexible combination of structured test-statistics (enabling one to use Meek rules to propagate edge orientation) as well as robust testing. Through empirical simulations, we confirm the usefulness of our framework. In closing, using this testing framework, we show how one may efficiently verify graph structure by drawing a connection to multi-constraint bandits and designing a novel algorithm to this end.

1. Introduction

Causal discovery is a fundamental goal of natural and social sciences, with widespread use across fields such as biology, physics and economics (Spirtes et al., 2000; Pearl, 2009). As a result, there has great interest in discovery methods with *provable guarantees*. In the task of causal discovery, one assumes access to the observational distribution, from which one can compute the undirected graph skeleton $\overline{G}$ with an unoriented edge between every cause and effect. Under specific functional assumptions on the graph, the underlying causal DAG can be identified from observational data alone. In more general settings, interventional data is needed for discovery. The goal of causal discovery is thus to minimize the amount of interventional data needed to identify the true causal graph. A typical discovery algorithm is outlined as in Algorithm 1, where the two key subroutines are the "query step" (adaptively determine which interventional data to collect next) and the "update step" (given the latest sample, orient edges in G using all the data collected so far).

The existing line of work on causal discovery with provable guarantees have largely focused on the query step; a non-exhaustive list of such papers include (Eberhardt, 2007; Hyttinen et al., 2013; Hu et al., 2014; Shanmugam et al., 2015; Kocaoglu et al., 2017; Ghassami et al., 2018; Lindgren et al., 2018; Choo et al., 2022). Key to the analysis is the assumption of hard intervention (under infinite samples), an idealized model of node intervention. That is, when node v is intervened on, the orientation of all edges in $\overline{G}$ incident to v is revealed. Thus, the update step can be easily implemented, and the algorithm performance be neatly defined in terms of the number of intervened nodes needed to fully orient the graph.

Importantly, this idealized model of node intervention overlooks the statistical complexity of orienting an edge in real world settings. If we view each edge orientation as a hypothesis, then almost always *multiple* samples are needed to reject with high probability (w.h.p.) an incorrect hypothesis (edge orientation), due to stochasticity in the data samples. Thus, towards studying finite-sample discovery, we consider the setup considered by Greenewald et al. (2019). An experiment with intervention v now provides *one sample* from v's interventional distribution, which by itself is may not be sufficient to orient the edge.

In this setting, it is no longer trivial to implement the update step. Thus, to *even begin* to study the finite-sample setting, we first need a framework that can implement the update

[1]Machine Learning Department, Carnegie Mellon University, Pittsburgh, USA [2]Department of Statistics and Data Science, Carnegie Mellon University, Pittsburgh, USA. Correspondence to: Tom Yan <tyyan@cmu.edu>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024.

step: given the interventional data obtained so far, decide which edges can be oriented. Put another way, a correct implementation of the update step is needed to *measure* algorithm performance. And only after we have this can we get to developing algorithms with provably good performance. Specifically, we note the following two properties are desirable for the framework to have:

1. **Anytime Valid Testing:** The most basic property required of any framework that implements the update step is correctness. That is, any edge that is oriented at any timestep should be correct w.h.p. In the finite-sample causal setting, this means that the testing framework has to have anytime validity.

   To see why, note that the number of samples needed for orientation varies depends on the unknown, underlying edge strength. For instance, many fewer samples are needed to orient $X_1 \xrightarrow{1000} X_2$ w.h.p. compared to that of $X_1 \xrightarrow{0.001} X_2$. And so, *anytime valid* testing is needed as hypotheses (corresponding to edge orientation e.g. of $X_1 \to X_2$) will be tested a number of times, where this number is unknown apriori.

2. **Encoding Propagation Implications:** Efficient discovery algorithms under hard intervention orient edges by considering the propagation implications of node interventions. Intervening on an "informative" node orients edges, whose orientations in turn propagate to many other edges via Meek rules (Meek, 2013).

   Thus in the finite-sample setting, a secondary, useful property for the framework to have is to be able to encode this structure, and relate hypotheses (edge orientations). We note that that this structure is useful for obtaining higher power tests. For a simple example, consider testing $X_1 \to X_2$ in $X_1 - X_2 - X_3$. Evidence against $X_2 \to X_3$ also serves as evidence against $X_1 \to X_2$, since by Meek rule $X_1 \to X_2 \Rightarrow X_2 \to X_3 \therefore \neg X_2 \to X_3 \Rightarrow \neg X_1 \to X_2$.

In this paper, we develop a framework that has both properties 1 and 2. To the best of our knowledge, our framework is the first that has these requisite properties. It perform anytime valid testing using the collected interventional data, with controlled error rate. That is, at any point in time (for however long it takes for the graph to be fully oriented), every oriented edge is correct w.h.p. This allows our framework to be paired with any causal discovery strategy (that implements the "query step") to perform finite-sample causal discovery.

The key observation used to develop the framework is that causal discovery can be viewed as structured, anytime hypothesis testing. The orientation of each edge in $\overline{G}$ corresponds to two hypotheses, one for each possible orientation.

Algorithm 1 Causal Discovery Algorithm Template

```
Input: Essential graph G, Query algorithm 𝒜
while |MEC(G)| > 1 do        ▷ multiple graphs in the Markov Equivalence Class
    𝒜(G) → X^t                                             ▷ query step
    Observe a sample from interventional distribution
    (x_1^t, ..., x_n^t) ~ X_1, ..., X_n | do(X^t)          ▷ collect new data
    Test orientation of each unoriented edge using data
    {(x_1^j, ..., x_n^j)}_{j=1}^t collected so far, and update G accordingly   ▷ update step
end while
Return G
```

There is structure among the hypotheses due to the Meek rules. Accordingly, our framework makes use of *e-processes* for testing. This is a type of test statistics that allows for both anytime valid testing and flexible combination of test statistics (Ramdas et al., 2023).

Our Contributions: First, in Section 3, we develop test statistics that is anytime valid (Property 1). In Section 4, we consider how one may combine test statistics to leverage graph structure (Property 2). In Section 5, we empirically verify the validity of our framework. Finally in Section 6, to make use of our testing framework, we develop a novel multi-constraint bandit algorithm for causal verification.

2. Problem Setup

We consider a linear graph with n nodes, where $X_i = \theta_i^T X_{\text{pa}(i)} + u_i$ with the set of exogenous noises U sub-Gaussian: $u_i \sim subG(\sigma^2)$. We note that our results generalize to additive graphs, provided knowledge of upper bounds on the variance of intervention distributions.

In line with the canonical setup in theoretical causal discovery literature, we assume causal sufficiency and access to the observational distribution $\mathcal{D}_0$ and graph skeleton $\overline{G}$ (Eberhardt, 2007; Hyttinen et al., 2013; Hu et al., 2014; Shanmugam et al., 2015; Kocaoglu et al., 2017; Ghassami et al., 2018; Lindgren et al., 2018; Choo et al., 2022). In certain settings, existing algorithms such as the PC algorithm (Spirtes et al., 2000) can orient additional edges on top of the graph skeleton. We note that our results are applicable to any essential graph returned by such algorithms. For most of our results, we are concerned with the worst-case setting wherein we only know the graph up to the graph skeleton. Still, our framework can be used to efficiently test and orient the *remaining* unoriented edges given any essential graph. Finally, with consideration for real-world robustness, we also consider the setting where the graph skeleton contains spurious edges in Section 3.4. We demonstrate that we can construct robust test statistics that do not propagate the error in the graph skeleton.

In addition to causal sufficiency, we also assume faithfulness: when a node is intervened upon, the expectation of each of its children nodes does change. Just as in the setup of (Greenewald et al., 2019), we consider a mildly stronger form of faithfulness where, for every cause-effect pair, there is a *minimal* causal effect b. That is, if we let the causal effect of i on j be $\mu_j(i) := \mathbb{E}[X_j|do(X_i)]$, then $\mu_j(i) \neq 0 \Rightarrow |\mu_j(i)| > b$.

For experimentation, we assume the scientist can perform sequential, single-node interventions with interventional value ν. In our setting, we focus on soft intervention, and we note that our testing framework is also applicable in the hard intervention setting, when mean-shift detection is used for edge orientation. Let I_t denote the node intervened on at time t. As in (Greenewald et al., 2019), following an intervention on node I_t, we observe one sample from the joint distribution, $\mathbf{X}_{\mathbf{I_t}} \sim \Pr(X_1, ..., X_n|do(X_{I_t}))$.

Our primary goal in this paper is to design a framework where we can use the data we collect from interventions to construct a sequence of partially oriented graphs $(\widehat{G}_t)$ such that *every* edge is oriented correctly at all time steps $t \in \mathbb{N}$ with high probability.

Definition 2.1 (Anytime-valid partially oriented graph)**.** Let $\widehat{G}_t$ denote the set of oriented edges after the first t interventions. A sequence of partially oriented graphs $(\widehat{G}_t)$ is an anytime-valid partially oriented graph if it satisfies:

$$\Pr(\exists t \in \mathbb{N} : \text{exists } \textbf{incorrect} \text{ edge in } \widehat{G}_t) \leq \alpha \tag{1}$$

for some predetermined error rate $\alpha \in [0, 1]$.

For further discussion on other relevant works and setups, please refer to Appendix A.

3. Anytime-valid testing via e-processes

First, in Section 3.1, we show that if we are able to construct an e-process for each edge orientation, then we can correctly implement the update step. This is because testing using e-processes guarantees that every edge that is oriented *at any point in time* is correct w.h.p. That is, the update step is correct across time w.h.p. With this motivation in mind, in Section 3.2, we construct e-processes that can be used for edge orientation. We begin with some definitions.

Definition 3.1 (Canonical Filtration)**.** A *filtration* $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$ is a sequence of nested sigma-algebras, i.e., $\mathcal{F}_t \subseteq \mathcal{F}_t$ for all $t \in \mathbb{N}$. We define the *canonical filtration* to have elements $\mathcal{F}_t := \sigma(\{\mathbf{X}_k\}_{k\in[t]} \cup \{U\})$ for each $t \in \mathbb{N}$ and let $\mathcal{F}_0 := \sigma(\{U\})$. $(\mathcal{F}_t)$ is essentially the sequence of variables observed after each intervention, and any internal randomness in the algorithm for selecting I_i for the first t interventions.

Definition 3.2 (Intervention-specific Filtration)**.** Define $(\mathcal{F}^i_t)$ as the filtration over data just from interventions on i: $\mathcal{F}^i_t := \sigma(\{\mathbf{X}_k\}_{k:k\in[t],I_t=i} \cup \{U\})$ for each $t \in \mathbb{N}$.

Definition 3.3. Define a *supermartingale* w.r.t. to filtration $(\mathcal{F}'_t)$ be any process $(M_t)_{t\in\mathbb{N}}$ s.t. $\mathbb{E}[M_t \mid \mathcal{F}'_{t-1}] \leq M_{t-1}$ and M_t is measurable w.r.t. $\mathcal{F}'_t$ for each $t \in \mathbb{N}$. For simplicity, we will always let *nonnegative supermartingale (NSM)* (M_t) satisfy $\mathbb{E}[M_1] \leq 1$.

Definition 3.4. Define an *e-process* $(E_t)_{t\in\mathbb{N}}$ w.r.t. to $(\mathcal{F}'_t)$ as a nonnegative process where there exists an NSM w.r.t. to $(\mathcal{F}'_t)$, (M_t), s.t. $E_t \leq M_t$ for all $t \in \mathbb{N}$ almost surely, and $\mathbb{E}[M_1] \leq 1$. Note that every NSM is an e-process. Equivalently, (E_t) is an e-process iff it satisfies $\mathbb{E}[E_\tau] \leq 1$ for any stopping time τ.

The only (key) property we use about e-processes is that it satisfies the following anytime guarantee, per Ville's inequality. At a high level, e-processes may be thought of something that satisfies the following crucial property, which is what enables sequential testing with provable error control.

Fact 1 (Ville's inequality (Ville, 1939)). For any *e-process* $(E_t)_{t\in\mathbb{N}}$: $\Pr(\exists t \in \mathbb{N} : E_t \geq 1/\alpha) \leq \alpha$.

3.1. A general approach for constructing anytime-valid partially oriented graphs

As mentioned previously, we may view each edge orientation as a hypothesis test. For an oriented edge $i \to j$, we may define the associated null hypothesis to be:

$$H_0^{i\to j} : \text{edge } (i, j) \text{ has orientation } i \to j \text{ in } G^*.$$

To test a hypothesis $H_0^{i\to j}$ with anytime validity, our testing framework simply requires the construction of an process $(E_t^{i\to j})$ that satisfies the following condition:

$$H_0^{i\to j} \text{ holds } \Rightarrow E_t^{i\to j} \text{ is an } \textit{e-process}$$

Note that this framework is general and one may design test statistics specific to the problem at hand, so long as the test statistic is an e-process under the null. Once we have such an $(E_t^{i\to j})$, our test is $\varphi_t^{i\to j}(\alpha) := \mathbb{1}\{E_t^{i\to j} \geq 1/\alpha\}$ and we may test as follows:

$$\begin{aligned} &\text{Reject } H_0^{i\to j} \text{ (i.e. claim } j \to i \text{ is correct)} \\ &\quad \text{if } \varphi_t^{i\to j}(\alpha) = 1 \text{ at any } t \in \mathbb{N}. \end{aligned} \tag{2}$$

Proposition 3.5. $(\varphi_t^{i\to j})$ *is an* anytime-valid test. *That is, the procedure in* (2) *ensures that for all error rates* $\alpha \in [0, 1]$*:*

$$\mathbb{P}(H_0^{i\to j} \textit{ is rejected } \mid H_0^{i\to j} \textit{ is true}) =$$
$$\mathbb{P}(\textit{exists } t \in \mathbb{N} : \varphi_t^{i\to j}(\alpha) = 1 \mid H_0^{i\to j} \textit{ is true}) \leq \alpha$$

Being able to construct anytime-valid test statistics is useful, because one can use it to produce anytime-valid partially-oriented graphs.

Using anytime-valid tests, we can construct an anytime-valid partially oriented graph by union bounding across the $|\overline{G}|$ tests.

Proposition 3.6. *Given an anytime-valid test $(\varphi_t^{i\to j})$, orient edge $i \to j$ in $\widehat{G}_t$ the first time $\varphi_t^{j\to i}(\alpha/|\overline{G}|) = 1$. Then, $(\widehat{G}_t)$ is an anytime-valid partially oriented graph.*

In summary, if we are able to construct anytime valid partially oriented graphs through anytime valid test statistics (such as e-processes), then we have in hand a testing framework that can correctly execute the update step w.h.p.

3.2. Construction of per-edge base e-processes

One way to construct e-processes is by combing a sequence of sequential e-values, defined as follows.

Definition 3.7. A sequence of *sequential e-values* (S_t) w.r.t. to a filtration $(\mathcal{F}'_t)$ under null hypothesis H_0 is defined as satisfying: $\mathbb{E}[S_t \mid \mathcal{F}'_{t-1}] \leq 1$ for all $t \in \mathbb{N}$ under H_0.

To develop a test statistic for testing hypothesis $i \to j$, we develop sequential e-values for testing $H_0^{i\to j}$.

It is natural to start by considering evidence from interventional data on node i and j. Both interventions provide evidence against $i \to j$ if the edge is actually $j \to i$. Below, we construct e-values under $do(i)$ and $do(j)$, which allows us to construct an e-process when we are given interventional data from i and j respectively.

Intervention on j: Suppose $I_t = j$, under $do(j)$, it natural to look at X_t^i. If $i \to j$, then X_t^i would still be mean 0, sub-Gaussian random variable, since the cause is not changed by changes in the effect. However, if $i \leftarrow j$, then X_t^i would have a shifted mean.

Thus, we define updates $S_t^{i\to j,+}(j), S_t^{i\to j,-}(j)$, which we show are sequential e-values:

$$S_t^{i\to j,+}(j) := \exp\left(\lambda_t X_t^i - \frac{\lambda_t^2\sigma_i^2}{2}\right)$$
$$S_t^{i\to j,-}(j) := \exp\left(\lambda_t(-X_t^i) - \frac{\lambda_t^2\sigma_i^2}{2}\right).$$

where (λ_t) is adapted to $(\mathcal{F}_t)$.

Proposition 3.8 (Effect on cause)**.** *For any sequence (λ_t) that is predictable w.r.t. $(\mathcal{F}_t^j)$, $S_t^{i\to j,+}(j)$ and $S_t^{i\to j,-}(j)$ are both sequential e-values under $H_0^{i\to j}$ w.r.t. filtration $(\mathcal{F}_t^j)$.*

Intervention on i: Suppose $I_t = i$, under $do(i)$, the assumption of minimal causal effect, b, allows us to include further evidence. We have that $H_0^{i\to j} = H_0^{i\to j,+} \cup H_0^{i\to j,-}$, where the two hypotheses are defined:

$$H_0^{i\to j,+} : H_0^{i\to j} \text{ is true and } \mu_i(j) \geq 0$$
$$H_0^{i\to j,-} : H_0^{i\to j} \text{ is true and } \mu_i(j) < 0$$

That is, if i causes j, then the casual effect of i on j is either positive or negative.

Since interventions result in a minimal shift of b in the mean, we can construct the following sequential e-values:

$$S_t^{i\to j,+}(i) := \exp\left(\lambda_t(b - X_t^j) - \lambda_t^2\sigma_j^2/2\right) \text{ if } \mu^j(i) > 0\ ,$$
$$S_t^{i\to j,-}(i) := \exp\left(\lambda_t(b + X_t^j) - \lambda_t^2\sigma_j^2/2\right) \text{ if } \mu^j(i) < 0$$

Proposition 3.9 (Cause on effect)**.** *Under the minimal causal effect condition, we have the following:*

Under $H_0^{i\to j,+}$, $S_t^{i\to j,+}(i)$ are sequential e-values w.r.t. filtration $(\mathcal{F}_t^i)$.

Under $H_0^{i\to j,-}$, $S_t^{i\to j,-}(i)$ are sequential e-values w.r.t. filtration $(\mathcal{F}_t^i)$.

With these e-values, we may construct aggregate test statistics under interventional data i and j, which we prove are e-processes.

Proposition 3.10. *Under $H_0^{i\to j}$, the following processes are e-processes w.r.t. filtrations $(\mathcal{F}_t^j)$, $(\mathcal{F}_t^i)$ respectively:*

$$E_t^{i\to j}(j) := \frac{1}{2}\left(\prod_{k:I_k=j}^{t} S_k^{i\to j,-}(j) + \prod_{k:I_k=j}^{t} S_k^{i\to j,+}(j)\right)$$
$$E_t^{i\to j}(i) := \min\left(\prod_{k:I_k=i}^{t} S_k^{i\to j,-}(i), \prod_{k:I_k=i}^{t} S_k^{i\to j,+}(i)\right)$$

3.3. Growth rate of e-processes

Suppose that it is the case that $j \to i$, we show that our test statistics in Proposition 3.10 are such the test *has power*. That is, it suffices to show that the test statistic will *increase* under the alternative, eventually exceed $1/\alpha$, and lead to the rejection of the null hypothesis $H_0^{i\to j}$.

Below, we derive the expected growth rate, which is a standard measure of the power of an e-process test. We note that the growth rate of (the log of) the e-values is edge-specific. It is a function of the edge's causal strength and variance. Also, we note that since the log of the e-values is sub-Gaussian, the test statistic concentrates quickly.

Proposition 3.11. *Suppose the true edge orientation is actually that $j \to i$ and WLOG $\mu^i(j) > 0$. By setting $\lambda_t = b/\sigma_i^2$ for $S_t^{i\to j}(i)$ and $\lambda_t = b/\sigma_j^2$ for $S_t^{i\to j}(i)$, we have the following growth rates:*

1. $\mathbb{E}[\log S_t^{i\to j,+}(j) \mid \mathcal{F}_{t-1}] = b(\mu_i(j) - b/2)/\sigma_i^2$
2. $\mathbb{E}[\log S_t^{i\to j,+}(i) \mid \mathcal{F}_{t-1}] = \mathbb{E}[\log S_t^{i\to j,-}(i) \mid \mathcal{F}_{t-1}] = b^2/(2\sigma_j^2)$

3.4. Robust Testing

In practical settings, the graph skeleton provided may contain mis-oriented edges. In what follows, we show that it is possible to detect and correct incorrect edges in the graph skeleton.

Specifically, we observe that by using only the test statistic $S_t^{i\to j}(j)$, our tests will be robust to spurious edges. The proof is simply that, if neither nodes have an effect on each other, the shift in mean is zero. Thus, both test statistics have expectation at most 1, and are thus e-processes. From Proposition 3.5, we then know that neither tests will reject w.h.p. And so, we *will not* mistakenly orient an edge w.h.p, when there is none there.

On top of this, we can then use the non-conclusiveness of both tests, after *sufficiently* many rounds, to correct an incorrectly specified edge. Indeed, when there is an edge, we should expect one of the two tests to reject within a bounded number of rounds with high probability. Thus, if we know a lower bound for the edge size, then we can use the non-rejection of both tests after sufficiently many rounds to determine that the edge is spurious. Indeed, if there is an edge, one of the two tests should have rejected w.h.p.

We now derive this bound as follows. For a sequential of sequential e-variables (S_t), define $\tau_\alpha := \min\{t \in \mathbb{N} \cup \{\infty\} : \prod_{k=1}^t S_k \geq \alpha^{-1}\}$ to be the first time $t \in \mathbb{N}$ where the product of S_t exceeds α^{-1} for any $\alpha \in [0, 1]$ (or ∞ if S_t never exceeds α^{-1}).

Proposition 3.12. *. If the edge $j \to i$ is the true orientation in G, then each of the the following statements hold true with probability $1 - \beta$ for each $\beta \in [0, 1]$:*

1. *For $(S_t^{i\to j,+}(j))$, we have that $\tau_\alpha \leq \frac{\sigma_i^2 \log(\alpha^{-1}\beta^{-1})}{b(\mu_i(j)-b)}$.*
2. *For $(S_t^{i\to j,\pm}(i))$, we have that $\tau_\alpha \leq \frac{\sigma_j^2 \log(\alpha^{-1}\beta^{-1})}{b^2}$*

Thus, these sample complexity results provide high probability upper bounds on the process corresponding to the product of sequential e-variables.

Please refer to Appendix B for the proofs of all results in this section and experiment plots.

4. Combining edge e-processes according to propagation rules

In this section, we study the theory of *combining* anytime valid e-processes, developed in the previous section. Recall, these test statistics (as in Proposition 3.10) were constructed for testing a single edge, in isolation. However, implications of Meek rules can allow us to propagate evidence from other edges to our edge of interest.

Importantly, this means that for testing $i \to j$, it is possible to make use of interventional data from *not just* nodes i, j. As we will show, e-processes can be flexibly combined and allow for propagation rules to be encoded into the test-statistic to take advantage of this structure.

Firstly, we observe that each Meek rule may be viewed as being one of two types of logical implications. Let $i_0 \to j_0, i_1 \to j_1, i_2 \to j_2$ be directed edges in the graph. Meek rules are of two forms:

$$i_1 \to j_1 \Rightarrow i_0 \to j_0 \text{ i.e., propagation of a single edge.} \quad (3)$$

$$(j_2 \to i_2 \wedge j_1 \to i_1) \Rightarrow j_0 \to i_0$$
$$\text{i.e., propagation of two edges to a single edge.} \quad (4)$$

Taking the contrapositive (CP) of Rule (4) results in the following rule: $i_0 \to j_0 \Rightarrow (i_2 \to j_2 \vee i_1 \to j_1)$.

Lemma 4.1 (Meek rules imply hypothesis conjunction/disjunction)**.** *For any edge orientation hypotheses $H_0^{i_0\to j_0}, H_0^{i_1\to j_1}, H_0^{i_2\to j_2}$, we have that*

$$H_0^{i_0\to j_0} = H_0^{i_0\to j_0} \cap H_0^{i_1\to j_1} \text{ by Rule 3}$$
$$H_0^{i_0\to j_0} = H_0^{i_0\to j_0} \cap (H_0^{i_1\to j_1} \cup H_0^{i_2\to j_2}) \text{ by CP of Rule 4}$$

This is useful, because under Rule 3 for example, testing $i_0 \to j_0$ is equivalent to testing $i_0 \to j_0$ *and* $i_1 \to j_1$. Thus, we can use evidence from $i_1 \to j_1$ to reject $i_0 \to j_0$, which increases the power of testing $i_0 \to j_0$.

In light of this observation, it is useful to enumerate $i_0 \to j_0$'s implications, to to obtain additional evidence for testing. Intuitively, the more implications an edge (hypothesis) has (due to propagation rules), the more ways there are to verify this hypothesis, since it only takes one false implication to reject a hypothesis. In the next subsection, we develop an algorithm that recursively enumerates these implications.

4.1. Enumeration of implications of an edge orientation

In this subsection, we develop an algorithm, Algorithm 2, for enumerating the "extended hypothesis" implied by the original hypothesis corresponding to the edge orientation of interest, $i \to j$. This algorithm allows us to operationalize the Meek rules and enumerate edges that are implied by the null hypothesis, $i \to j$.

In the algorithm, a tree of edges is recursively expanded to enumerate all the edges implied by the root edge. To emphasize, the tree we refer to in this section does not refer to the causal graph (which need not be a tree), but rather a

Algorithm 2 Enumerating edges implied by Meek rule for a given edge orientation

Require: Essential graph G, hypothesized orientation $i \to j$.

```
Initialize empty tree T, insert edge i → j as root.
while exists root to leaf path P such that the oriented edges in P imply new edge via a Meek rule in G do
    if Meek rule of the form (3) or (4) propagates a single new edge i′ → j′ not in P then
        Append i′ → j′ to the leaf node of P.
    end if
    if Meek rule of the form (4) propagates two new edges i1 → j1, i2 → j2 both not in P then
        Add i1 → j1 and i2 → j2 as children of the leaf node of P. ▷ do not include the pair if at least one of the edges is on the path
    end if
end while
Return tree T where each path from root to leaf P is a set of edges that are implied by i → j.
```

representation of the logical implications that are implied by the root edge.

Let $T^{i\to j}$ be the tree constructed by applying Algorithm 2. A *path* in a tree T is the set of edges encountered by traversing T from its root to a *leaf* node.

Definition 4.2. For a tree T, define the logical implications represented by T as follows:

$$H_0(T) := \bigcup_{P\in\mathcal{P}(T)} \bigcap_{i'\to j'\in P} H_0^{i'\to j'}.$$

Proposition 4.3. *Algorithm 2 satisfies the following properties:*

- *(Soundness) Algorithm 2 is sound and does terminate.*
- *(Correctness) Let $T^{i\to j}$ be the resultant tree of Algorithm 2, then:*

$$H_0^{i\to j} = \bigcup_{P\in\mathcal{P}(T^{i\to j})} \bigcap_{i'\to j'\in P} H_0^{i'\to j'}.$$

Remark 4.4. As we prove in Lemma C.2, we can stop the tree expansion in Algorithm 2 after any number of application of Meek rules. The corresponding tree T would still be valid ($H_0(T) = H_0^{i\to j}$). That is, we need not exhaust all implications based on the Meek rules. This is useful because one can trade off between the power of the test, and the time/space complexity of a more complicated tree/test statistic. Testing with fewer implications results in lower power, but has the benefit of being easier to track and evaluate.

4.2. Conversion of expanded hypothesis into an e-process

Having enumerated other edge orientations implied by the original edge orientation, in this subsection, we show how to convert these logical relationships into an "extended" e-process useful for testing.

Given a logical tree $T^{i\to j}$, we first design an e-process corresponding to a particular path $P \in T^{i\to j}$. Let V be the set of all nodes in the graph. Let Δ^d denote the probability simplex on d-dimensions. Let $P(i') := \{i \to j : i \to j \in P, i = i' \vee j = i'\}$ be the set of edges on path P with one its vertex node i'. We can now construct a corresponding e-process which is defined as follows.

Proposition 4.5. *Let $(E_t^{i\to j}(i))$ be an e-process w.r.t. $(\mathcal{F}_t^i)$ under $H_0^{i\to j}$. For a path P, define:*

$$E_t^P := \exp\left(\sum_{i'\in V} \max_{i\to j\in P(i')} \log E_t^{i\to j}(i') - \frac{|P(i')|-1}{2}\cdot \log(2|T_{i'}(t)|-2)\right),$$

Then (E_t^P) is an e-process w.r.t. $(\mathcal{F}_t)$ under H_0^P.

Having defined the e-process corresponding to some path $P \in T^{i\to j}$, we may now define the e-process corresponding the full tree $T^{i\to j}$ as follows.

Proposition 4.6 (Correctness of combined e-process)**.** *Define:*

$$E_t^{i\to j} := \min_{P\in\mathcal{P}(T^{i\to j})} E_t^P.$$

Then, $(E_t^{i\to j})$ is an e-process when $H_0^{i\to j}$ is true.

Theorem 4.7. *For any sequence of interventions (I_t) predictable w.r.t. $(\mathcal{F}_t)$, let $\widehat{G}_t$ be the partially oriented DAG where the test for each orientation is defined as follows:*

$$\varphi_t^{i\to j} = \mathbb{1}\{E_t^{i\to j} \geq |\overline{G}|/\alpha\}.$$

Then, $(\widehat{G}_t)$ is anytime-valid orientation (as defined in (11)).

4.3. Additional power in combined test statistics

We note that Proposition 4.6 applies to any expanded $T^{i\to j}$ tree, which includes the tree without any expansion i.e. $T^{i\to j} = (i \to j)$. So does an expanded tree lead to higher power? We note that an increase in power depends on the graph: for instance, if a graph comprises of only isolated edges, no additional power can be gained from propagation. Below, we present one instance where we can prove that the power of the test is sizably larger, thus providing a concrete example showing the value of combining evidence.

Proposition 4.8. *Consider an uniform intervention policy over nodes $[n]$. There exists a graph and edge $i \to j$, such that the expected growth rate (i.e. power) of $\log E_t^{i\to j}$ under the fully expanded tree $T^{i\to j}$ is $\Omega(|\overline{G}|)$ times that of $\log E_t^{i\to j}$ under the non-expanded tree (i.e. just the single edge $i \to j$).*

Please see Figure 15 of Appendix F for an illustration of a simple n-node chain graph, wherein additional power can be obtained due to Meek rules. Please refer to Appendix C for the proofs of all results in this section as well as time complexity analysis of the proposed algorithms.

In closing, we note that this approach to test statistic combination can apply more broadly to other structured hypothesis testing settings, wherein there are logical relationships (Meek rules in this case) relating the hypotheses.

5. Experiments on fixed-time versus anytime methods

To illustrate the usefulness of anytime valid tests, we compare our anytime-valid test statistic (as in Section 3.2) against a fixed-time test statistic across a variety of graphs.

Graph Setups: We consider two classes of graphs. (1) Erdos-Renyi graphs with varying number of nodes and density $(n, p) \in \{10, 20, 30\} \times \{0.3, 0.5\}$ (2) tree graphs with $n \in \{10, 20, 50, 100\}$. These are used to generate the graph skeleton. The SCM of the graphs is linear Gaussian; the edge strengths are randomly sampled from $\max(U[0, k], b)$, where k is the upper bound.

Fixed-time Baseline: We consider the following p-value that corresponds to the two-sided z-test for edge $i \to j$. The hypothesis test involves checking if the test statistic is below the acceptance threshold $\frac{\alpha}{2|\overline{G}|}$ (from union bound).

Let $\widehat{\mu}_t^{j|\mathrm{do}(i)} := \sum_{k\in T_i(t)} X_k^j, \widehat{\mu}_t^{i|\mathrm{do}(j)} := \sum_{k\in T_j(t)} X_k^i$. Let $T_i(t)$ be the number of times we have intervened on i at time t. Define the fixed-time p-value baseline as:

$$P_t^{i\to j} = 2\left(1 - \Phi\left(\frac{b \cdot |T_i(t)| - |\widehat{\mu}_t^{j|\mathrm{do}(i)}| + |\widehat{\mu}_t^{i|\mathrm{do}(j)}|)}{\sqrt{|T_j(t)| \operatorname{var} X_i + |T_i(t)| \operatorname{var} X_j}}\right)\right)$$

where Φ is the Gaussian CDF function.

Proposition 5.1 ($P_t^{i\to j}$ is a p-value)**.** *$P_t^{i\to j}$ satisfies $\mathbb{P}(P_t^{i\to j} \le s) \le s$ for all $s \in [0, 1]$ and $t \in \mathbb{N}$ under $H_0^{i\to j}$.*

Experiment Configurations: In the experiment, we fix $b = 0.1$, variance 1 and the interventional value $\nu = 1$. We vary the number of interventional samples $\in \{100, 500, 1000, 5000, 10000\}$, tolerated error rate $\alpha \in \{0.1, 0.2\}$ and edge strength $k \in \{0.1, 0.2, 1, 2, 10\}$, all of which affect hypotheses testing (i.e. number of orientations). Fixing a particular setting, we simulate 20 trials to compute the mean and standard deviation.

We plot two metrics. The most important is the miscoverage rate, which is defined to be the number of trials wherein the test statistic returns at least one falsely oriented edge. That is, the percentage of time that an update step that uses this test statistic is wrong. Alongside miscoverage, we also plot the number of oriented edges. This indicates the informativeness of a test statistic, as indeed a test that never rejects can trivially achieve 0 miscoverage rate.

Comparing anytime vs fixed-time: In the interest of space, we present results under the ER graph with $(n, p, \alpha) = (30, 0.5, 0.2)$ in Figure 1. Overall, we observe the following trends in our experiments.

Miscoverage: In every setting, we find that our testing framework achieves miscoverage rate below α (line in green), thus validating our theoretical anytime guarantee. On the other hand, in a number of settings, we observe that the fixed-time statistic leads to high miscoverage rate.

Number of Orientations: The reason for the high miscoverage seems to be that the fixed time test statistic is not conservative enough to control the error rate. The anytime test is more conservative in orienting fewer number of edges, so as to attain error control. Note that this control is important in preventing spurious edge orientations, which would then be fed back into the query step as an erroneous representation of the partially oriented graph.

Comparing combined e-values vs base e-values: We also conduct an experiment comparing the combined e-values (Section 4) against the base e-values (Section 3) in a chain graph, where we expect the combined e-values to be helpful. We find that combining e-values is more useful in large data/graph regimes, while the light-weight, base e-values are more effective in small data/graph regimes.

Please refer to Appendix D for all experimental results.

6. Optimizing test statistic for causal verification

Once we have an anytime valid test framework that correctly implements the update step, we can turn to designing query strategies that minimize sample complexity under this framework. Towards this goal, we consider the task of causal verification, which acts as a stepping-stone towards causal discovery. Knowing how to optimally intervene to verify a known graph is an useful building block for understanding how to optimally intervene to learn an unknown graph. In this section, we develop a novel querying algorithm with provable guarantees that we believe can be be a stepping stone to more practical algorithms. To do so, we

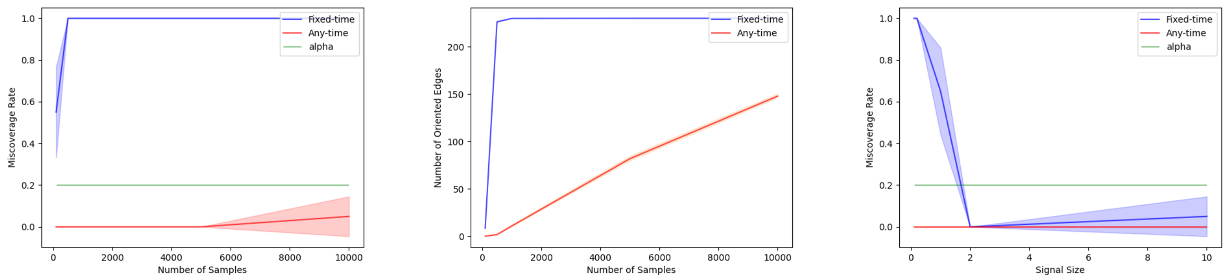


Figure 1. (Left) Number of samples vs Miscoverage rate (Middle) Number of samples vs Number of oriented edges (Right) Edge signalsize k vs Miscoverage rate.

highlight a connection between finite-sample causal verification and the structured bandit literature (Badanidiyuru et al., 2018), by demonstrating that causal verification reduces to multi-constraint bandit optimization.

To recap, the goal of active verification is: given knowledge of the true graph, verify the edge orientations, while minimizing the *expected* number of samples needed to conclude that each edge orientation is oriented as in the graph w.h.p.

Problem Setup: Formally, construct an intervention policy that (adaptively) intervenes on nodes $I_1, ..., I_\tau$ such that the the *expected* stopping time $\mathbb{E}[\tau]$ is minimized, where τ is defined as the earliest time step such that every hypothesis corresponding to incorrect orientation $j \to i$ is rejected. That is, $\forall j \to i$, $E_\tau^{j \to i} \geq |\overline{G}|/\alpha$.

6.1. Construction of test statistic for causal verification

Since the SCM is known in verification, all edge strengths are known. This allows us to construct a more simplified test-statistic than that of Proposition 4.5.

Consider some incorrect orientation $j \to i$. Let its logical tree be $T^{j \to i}$. With full information, it is natural to construct a test statistic for $j \to i$ by including only the e-value with the *highest expected growth* rate. Define $S_t^*(P, I_t) = S_t^{e^*, s^*}(I_t)$ for $e^*, s^* = \arg\max_{e \in P(I_t), s \in \{\pm\}} \mathbb{E}[S^{e,s}(I_t)]$. This represents the edge and sign e-value with the largest expected growth-rate under intervention I_t, out of all the possible e-values of edges in $P(I_t)$.

With this, we may define a test statistic with the highest expected growth rate under intervention i' as $E_t^{*j \to i}(i') = \prod_{k: I_k = i'}^t S_k^*(P, i')$, path test statistic as $E_t^{*P} := \exp\left(\sum_{i' \in V} \log E_t^{*j \to i}(i')\right)$ and full test statistic as $E_t^{*j \to i} = \min_{P \in \mathcal{P}(T^{j \to i})} E_t^{*P}$. In what follows, we will make the assumption that X_i is a bounded r.v. with b, ν such that $\log S_k^*(P, I)$ is positive (as arm rewards are usually assumed to be positive in bandits literature).

6.2. Reduction to multi-constraint bandit optimization

Having defined E_t^{*P}, causal verification then corresponds to choosing an apt intervention policy that jointly optimizes $E_t^{*j \to i}$ for every incorrect orientation $j \to i$, and only insofar as to have $E_t^{*j \to i}$ *exceed a threshold*, $|\overline{G}|/\alpha$. To solve this problem, we observe that causal verification reduces to multi-constraint bandit optimization, defined as follows.

Multi-constraint bandit optimization: An instance is parameterized by n arms, m constraints and budget b:

- There are T rounds for T a specified time horizon.
- At round i, the algorithm may pull an arm x_i, yielding a "gain" *vector*, where $r_{x_i} \sim D_{x_i}$ for $r_{x_i} \in [0, M]^m$.
- There is a known threshold $b \in \mathbb{R}^+$ on the aggregate gain of each constraint.
- The interaction terminates at the earliest round τ, when $\sum_{t=1}^\tau r_{x_t} \geq b \cdot 1$ (aggregate gain of every constraint exceeds b), or at the end of the Tth round.

The goal of the algorithm is to minimize the total *expected* cost $\sum_{i=1}^\tau c_{x_i}$ (node intervention cost c_{x_i} is set to 1).

Reduction to multi-constraint bandits: We observe that the test statistic for each path $P \in \mathcal{P}(T^{j \to i})$ grows *additively* in the log of e-values $\log S_k^*(P, I_k)$:

$$E_t^{*j \to i} \geq |\overline{G}|/\alpha \Leftrightarrow \forall P \in \mathcal{P}(T^{j \to i}), E_t^{*P} \geq |\overline{G}|/\alpha$$

$$\Leftrightarrow \forall P \in \mathcal{P}(T^{j \to i}), \sum_{k=1}^t \log S_k^*(P, I_k) \geq \log(|\overline{G}|/\alpha)$$

Thus, given a causal verification instance, we may instantiate a multi-constraint bandit instance as follows.

1. Arms: define $n = |V|$ arms, each corresponding to a node intervention in the graph.
2. Constraint: define a constraint corresponding to every (P, i') pair, for path $P \in T^{j \to i}$ and intervention $i' \in V$.

Algorithm 3 Causal Verification as multi-constraint bandits

Require: threshold b; time horizon T; for each node x, known expected gain vector $\bar{r}_x \in [0, M]^m$ ▷ *for node x, this vector's entries are the expected growth rates under intervention on node x ($\mathbb{E}[\log S^*(P, x)]$ of every path P of every logical tree $T^{j \to i}$)*

In the first n rounds, intervene on each node once
Initialize $v_1 = 1 \in [0, 1]^m$
Set $\epsilon = \sqrt{\frac{M \ln m}{b+M}}$
while $\sum_{i=1}^{t} r_{x_i} < b \cdot 1$ and $t < T$ **do** ▷ *not all tests have concluded, since not all test statistics have exceeded b*
  for node $x \in [n]$ **do**
    Set weighted total gain $g_x = \bar{r}_x \cdot v_t$
  end for
  Intervene on node $x_t = \arg\max_{x \in X} g_x$ with the highest weighted gain
  Receive vector r_x, whose entries are realizations of random variables $\log S^*(P, x)$ of every path $P \in T^{j \to i}$ of every tree $T^{j \to i}$
  Update v_t entry-wise with normalized r_x, where its ith entry changes as follows:

$$v_{t+1}(i) = v_t(i)(1 - \epsilon)^{\ell}, \ell = r_x(i)/M$$

end while

Thus, the gain of pulling arm $i' \in V$ corresponds to a vector of realizations of random variable $\log S^*(P, i')$ of every path $P \in T^{j \to i}$ of every tree $T^{j \to i}$.

3. Set the threshold $b = \log(|\overline{G}|/\alpha)$.

Guarantee: We develop Algorithm 3 that attains provable guarantees in the multi-constraint bandit setting, which applies immediately to the causal verification setting via the reduction. Let OPT be the expected total number of interventions needed by the optimal dynamic policy. Let REW_{tot} be the algorithm performance of Algorithm 3, which is the expected number of interventions such that every incorrect orientation test statistic exceeds b. Then, we have that:

Theorem 6.1. *The regret of Algorithm 3 is:*

$$\text{REW}_{tot} - \text{OPT} \leq \tilde{O}\left(\frac{M}{b} + \sqrt{\frac{(b+M)M}{b}}\right) \text{OPT} + \tilde{O}\left(\frac{M\sqrt{T}}{b} + nM\right).$$

To provide some intuition, in Algorithm 3, one may view v as a varying, weighting over each constraint. Each round, the algorithm greedily pulls the arm whose sum of weighted expected gain is the largest. After a round, if a constraint has seen a sizable increase, then its weighting in v is reduced. This adaptive re-balancing then allows for an arm selection that focuses more on increasing other constraints, which are further away from exceeding the threshold b. Please refer to Appendix E for the proofs of all results in this section.

7. Conclusion

We develop a general, anytime valid testing framework that can correctly implement the "update step" needed in finite-sample causal discovery. Using this framework, we develop a multi-constraint bandit algorithm for causal verification. Overall, we believe our results serve as an useful stepping stone towards making further progress on causal discovery and more broadly structure learning, in the real world.

Impact Statement

This paper presents work whose goal is to advance the topic of causality, primarily geared towards capturing the statistical sample complexity of causal discovery. To the best of our understanding, there are no societal consequences of our work which we feel must be specifically highlighted here.

References

Acharya, J., Bhadane, S., Bhattacharyya, A., Kandasamy, S., and Sun, Z. Sample complexity of distinguishing cause from effect. In *International Conference on Artificial Intelligence and Statistics*, pp. 10487–10504. PMLR, 2023.

Agrawal, R., Squires, C., Yang, K., Shanmugam, K., and Uhler, C. Abcd-strategy: Budgeted experimental design for targeted causal structure discovery. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 3400–3409. PMLR, 2019.

Badanidiyuru, A., Kleinberg, R., and Slivkins, A. Bandits with knapsacks. *Journal of the ACM (JACM)*, 65(3):1–55, 2018.

Choo, D., Shiragur, K., and Bhattacharyya, A. Verification and search algorithms for causal dags. *Advances in Neural Information Processing Systems*, 35:12787–12799, 2022.

Cover, T. and Ordentlich, E. Universal portfolios with side information. *IEEE Transactions on Information Theory*, 42(2):348–363, 1996.

Eberhardt, F. Causation and intervention. *Unpublished doctoral dissertation, Carnegie Mellon University*, pp. 93, 2007.

Ghassami, A., Salehkaleybar, S., Kiyavash, N., and Bareinboim, E. Budgeted experiment design for causal structure learning. In *International Conference on Machine Learning*, pp. 1724–1733. PMLR, 2018.

Glymour, C., Zhang, K., and Spirtes, P. Review of causal discovery methods based on graphical models. *Frontiers in genetics*, 10:524, 2019.

Greenewald, K., Katz, D., Shanmugam, K., Magliacane, S., Kocaoglu, M., Boix Adsera, E., and Bresler, G. Sample efficient active learning of causal trees. *Advances in Neural Information Processing Systems*, 32, 2019.

Hoyer, P., Janzing, D., Mooij, J. M., Peters, J., and Schölkopf, B. Nonlinear causal discovery with additive noise models. *Advances in neural information processing systems*, 21, 2008.

Hu, H., Li, Z., and Vetta, A. R. Randomized experimental design for causal graph discovery. *Advances in neural information processing systems*, 27, 2014.

Hyttinen, A., Eberhardt, F., and Hoyer, P. O. Experiment selection for causal discovery. *Journal of Machine Learning Research*, 14:3041–3071, 2013.

Jamieson, K. G. and Jain, L. A bandit approach to sequential experimental design with false discovery control. In Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems*, volume 31, 2018.

Kocaoglu, M., Dimakis, A., and Vishwanath, S. Cost-optimal learning of causal graphs. In *International Conference on Machine Learning*, pp. 1875–1884. PMLR, 2017.

Lindgren, E., Kocaoglu, M., Dimakis, A. G., and Vishwanath, S. Experimental design for cost-aware learning of causal graphs. *Advances in Neural Information Processing Systems*, 31, 2018.

Meek, C. Causal inference and causal explanation with background knowledge. *arXiv preprint arXiv:1302.4972*, 2013.

Pearl, J. *Causality*. Cambridge university press, 2009.

Porwal, V., Srivastava, P., and Sinha, G. Almost optimal universal lower bound for learning causal dags with atomic interventions. In *International Conference on Artificial Intelligence and Statistics*, pp. 5583–5603. PMLR, 2022.

Ramdas, A., Ruf, J., Larsson, M., and Koolen, W. Admissible anytime-valid sequential inference must rely on nonnegative martingales. *arXiv preprint arXiv:2009.03167*, 2020.

Ramdas, A., Grünwald, P., Vovk, V., and Shafer, G. Game-theoretic statistics and safe anytime-valid inference. *Statistical Science (forthcoming)*, 2023.

Shanmugam, K., Kocaoglu, M., Dimakis, A. G., and Vishwanath, S. Learning causal graphs with small interventions. *Advances in Neural Information Processing Systems*, 28, 2015.

Shimizu, S., Hoyer, P. O., Hyvärinen, A., Kerminen, A., and Jordan, M. A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, 7(10), 2006.

Spirtes, P., Glymour, C. N., and Scheines, R. *Causation, prediction, and search*. MIT press, 2000.

Squires, C., Magliacane, S., Greenewald, K., Katz, D., Kocaoglu, M., and Shanmugam, K. Active structure learning of causal dags via directed clique trees. *Advances in Neural Information Processing Systems*, 33:21500–21511, 2020.

Tigas, P., Annadani, Y., Jesson, A., Schölkopf, B., Gal, Y., and Bauer, S. Interventions, where and how? experimental design for causal models at scale. *Advances in Neural Information Processing Systems*, 35:24130–24143, 2022.

Toth, C., Lorch, L., Knoll, C., Krause, A., Pernkopf, F., Peharz, R., and Von Kügelgen, J. Active bayesian causal inference. *Advances in Neural Information Processing Systems*, 35:16261–16275, 2022.

Ville, J. *Étude critique de la notion de collectif*. 1939.

Wadhwa, S. and Dong, R. On the sample complexity of causal discovery and the value of domain expertise. *arXiv preprint arXiv:2102.03274*, 2021.

Xu, Z., Wang, R., and Ramdas, A. A unified framework for bandit multiple testing. In *Neural Information Processing Systems*, 2021.

Zhang, K. and Hyvarinen, A. On the identifiability of the post-nonlinear causal model. *arXiv preprint arXiv:1205.2599*, 2012.

A. More Related Works

Finite-Sample Considerations in Causality: The three papers most similar in motivation to that of ours are: Greenewald et al. (2019), Wadhwa & Dong (2021) and Acharya et al. (2023). Like Greenewald et al. (2019), our paper is similarly motivated by finite-sample considerations that exist in real-world settings, where the collection of interventional data (e.g. RCTs) is much more difficult and costly than that the collection of observational data. As such, we also assume that infinitely many observational samples are available, while only finitely many interventional samples can be obtained. (Wadhwa & Dong, 2021) is concerned with the sample complexity of causal discovery, albeit that of learning the equivalence class, and not the actual graph, given only observational data only. Finally, (Acharya et al., 2023) is also concerned with finite-sample causal discovery via testing. They study the two node setting, and assume both finite interventional and observational data, which are contrasted in the paper.

While our paper's goal of studying finite-sample causal discovery is the same as those of Greenewald et al. (2019); Acharya et al. (2023), our paper differs in focusing primarily on the update step. Additionally, our testing framework is applicable in general graphs, going beyond the two-node or tree settings. Different from (Greenewald et al., 2019), we study soft interventions instead of hard interventions, thus introducing the need to consider the strength of edges, as edges with weak causal strength require more samples to orient. Different from (Acharya et al., 2023), we study how to propagate edge orientations in hypothesis testing, which is needed when the graph comprises of more than two nodes and a single edge.

Causal Verification: Causal Verification is a well-known task in causal discovery. Besides having practical applications (e.g. verifying a scientific conjecture corresponding to some causal graph structure), it has the theoretical benefit of better understanding the lower bound that underlies any active causal discovery algorithm (Squires et al., 2020; Porwal et al., 2022; Choo et al., 2022).

Bayesian causal discovery: There has also been a line of work in Bayesian causal discovery, wherein one uses interventional data to update the posterior over all graphs (Agrawal et al., 2019; Toth et al., 2022; Tigas et al., 2022). Since the set of all graphs in the MEC may be prohibitively large, approximation methods are used to sample from the posterior, making less clear what provable guarantees one may be able to provide about such methods.

Functional Causal Discovery: Further afield, there has been a sizable number of paper that leverage specific functional forms of graphs for orientation, using observational data only. Examples of such methods include (Shimizu et al., 2006; Hoyer et al., 2008; Zhang & Hyvarinen, 2012). Interested readers may refer to for example (Glymour et al., 2019) for a more complete survey of this line of work.

Bandit Multiple-Testing: The closest type of methods in the bandit literature are those dealing with multiple testing (Jamieson & Jain, 2018; Xu et al., 2021). Current work on bandit multiple testing differs from the methods in this paper in two significant ways: (1) bandit multiple testing is primarily focused on controlling the false discovery rate (FDR) and (2) methods lie in the typical hypothesis testing problem setting where one can only reject a hypothesis — in the causal discovery setting, each unoriented edge will be one of two directions, and the negation of one implies the other — hence the relationships between the hypotheses require methods that will derive a certain conclusion for each unoriented edge.

Necessity of Non-Negative Martingales: (Ramdas et al., 2020) proves that under a suitable definition of admissibility, all admissible constructions of test statistics for any-time sequential inference must necessarily utilize nonnegative martingales. This shows that the martingale test statistic we construct is in some sense of the "right form".

B. Deferred Proofs from Section 3

Lemma B.1. *Let* $M_t := \prod_{k=1}^{t} S_k$. *Then,* (M_t) *is an NSM w.r.t. filtration* $(\mathcal{F}_t)$ *under* H_0.

Proof. The conditional expectation of M_t is as follows:

$$\mathbb{E}[M_t \mid \mathcal{F}_{t-1}] = \mathbb{E}[S_t \mid \mathcal{F}_{t-1}] \cdot \prod_{k=1}^{t-1} E_k = \mathbb{E}[S_t \mid \mathcal{F}_{t-1}] \cdot M_{t-1} \leq M_{t-1},$$

where the inequality is by definition of S_t being a sequential e-value for each $t \in \mathbb{N}$. □

B.1. Deferred Proofs from Section 3.1

Proposition B.2. $(\varphi_t^{i\to j})$ *is an* anytime-valid test*, that is, the procedure in* (2) *ensures that*

$$\mathbb{P}(H_0^{i\to j} \textit{ is rejected }) = \mathbb{P}(\textit{exists } t \in \mathbb{N} : \varphi_t^{i\to j}(\alpha) = 1) \le \alpha \textit{ when } H_0^{i\to j} \textit{ is true for all } \alpha \in [0,1].$$

Proof.

$$\begin{aligned}\mathbb{P}(\text{exists } t \in \mathbb{N} : \varphi_t^{i\to j}(\alpha) = 1 | H_0^{i\to j}) &= \Pr(\text{exists } t \in \mathbb{N} : M_t^{i\to j} \ge 1/\alpha | H_0^{i\to j}) && \text{(by definition of } \varphi_t^{i\to j}(\alpha)) \\ &\le \alpha && \text{(Ville's inequality, because under } H_0^{i\to j} \text{ is true} \Rightarrow (M_t^{i\to j}) \text{ is an e-process)}\end{aligned}$$

□

Proposition B.3. *Given an anytime-valid test* $(\varphi_t^{i\to j})$*, orient edge* $i \to j$ *in* $\widehat{G}_t$ *the first time* $\varphi_t^{j\to i}(\alpha/|\overline{G}|) = 1$*. Then,* $(\widehat{G}_t)$ *is an anytime-valid partially oriented graph.*

Proof. Let the final oriented graph be $\hat{G}$.

$$\begin{aligned}\mathbb{P}\Big(\text{exists } t \in \mathbb{N} : \text{exists oriented edge in } \widehat{G}_t \text{ not in } G^*\Big) &\le \sum_{i\to j \in \hat{G}} \mathbb{P}\left(\text{exists } t \in \mathbb{N} : \text{orient edge } i \to j \wedge j \to i \text{ in } G^*\right) \\ &= \sum_{i\to j \in \hat{G}} \Pr(\text{exists } t \in \mathbb{N} : \phi_t^{j\to i}(\alpha/|\overline{G}|) = 1 \wedge j \to i \text{ in } G^*) \\ &\le \sum_{i\to j \in \hat{G}} \alpha/|\bar{G}| && \text{(by Proposition 3.5)} \\ &= \alpha && (5)\end{aligned}$$

□

B.2. Deferred Results from Section 3.2

Proposition B.4. *For any sequence* (λ_t) *that is predictable w.r.t.* $(\mathcal{F}_t^j)$*,* $S_t^{i\to j,+}(j)$ *and* $S_t^{i\to j,-}(j)$ *are both sequential e-values under* $H_0^{i\to j}$ *w.r.t. filtration* $(\mathcal{F}_t^j)$*.*

Proof. At time t with $I_t = j$, under $H_0^{i\to j}$, we have that $\pm X_t^i \mid \mathcal{F}_{t-1}^j$ is a mean 0, σ_i^2-sub-Gaussian random variable. We work through the X_t^i case, and the $-X_t^i$ case follows analogously. From definition, its MGF is such that:

$$\mathbb{E}[\exp(\lambda X_t^i)] \le \exp\left(\frac{\lambda^2\sigma_i^2}{2}\right) \Leftrightarrow \mathbb{E}[S_t^{i\to j,+}(j) \mid \mathcal{F}_{t-1}^j] = \mathbb{E}[S_t^{i\to j,+}(j)] = \exp\left(\lambda X_t^i - \frac{\lambda^2\sigma_i^2}{2}\right) \le 1$$

□

Proposition B.5. *Under the minimal causal effect condition, we have the following:*

Under $H_0^{i\to j,+}$*,* $S_t^{i\to j,+}(i)$ *are sequential e-values w.r.t. filtration* $(\mathcal{F}_t^i)$*.*

Under $H_0^{i\to j,-}$*,* $S_t^{i\to j,-}(i)$ *are sequential e-values w.r.t. filtration* $(\mathcal{F}_t^i)$*.*

Proof. We prove the first statement, and the second follows analogously. WLOG $\mu_j(i) = \mathbb{E}[X_t^j(i)] \ge b$. We have that:

$$\mathbb{E}\left[\exp\left(\lambda(b - X_t^j(i)) - \lambda^2\sigma_j^2/2\right) \mid \mathcal{F}_{t-1}\right] \le 1 \Leftrightarrow \mathbb{E}[S_t^{i\to j,+}(i) \mid \mathcal{F}_{t-1}^i] = \mathbb{E}[S_t^{i\to j,+}(i)] \le 1$$

since $b - X_t^j(i) \mid \mathcal{F}_{t-1}$ is a σ_i^2-sub-Gaussian with nonpositive mean $b - \mathbb{E}[X_t^j(i)]$.

□

Proposition B.6. *Under $H_0^{i\to j}$, the following processes are e-processes w.r.t. to filtration $(\mathcal{F}_t^j)$, filtration $(\mathcal{F}_t^i)$ respectively:*

$$E_t^{i\to j}(j) := \frac{1}{2}\left(\prod_{k:I_k=j}^{t} S_k^{i\to j,-}(j) + \prod_{k:I_k=j}^{t} S_k^{i\to j,+}(j)\right), \qquad E_t^{i\to j}(i) \quad := \min\left(\prod_{k:I_k=i}^{t} S_k^{i\to j,-}(i), \prod_{k:I_k=i}^{t} S_k^{i\to j,+}(i)\right)$$

Proof.

- $(E_t^{i\to j}(j))$ is the average of two processes

$$M_t^{+}(j) = \prod_{k\in T_j(t)} S_k^{i\to j,+}(j), \qquad M_t^{-}(j) = \prod_{k\in T_j(t)} S_k^{i\to j,-}(j).$$

By Proposition 3.8, each of these processes are the product of sequences of sequential e-values (w.r.t. to filtration $(\mathcal{F}_t^j)$) under $H_0^{i\to j}$, i.e., $(S_k^{i\to j,-})$ and $(S_k^{i\to j,+})$. This implies that they are NSMs by Lemma B.1, and hence also e-processes w.r.t. to filtration $(\mathcal{F}_t^j)$.

To show that the average of these two e-processes is an e-process, we introduce the notion of a stopping time, and note the following e-process equivalence.

Definition B.7. A *stopping time* $\tau \in \mathbb{N}$ w.r.t. a filtration $(\mathcal{F}'_t)_{t\in\mathbb{N}}$ is a random variable that where $\mathbb{1}\{\tau = t\}$ is measurable w.r.t. $\mathcal{F}_t$.

Further, we use the following fact about e-processes from (Ramdas et al., 2020).

Fact 2 (Item (vi) from Lemma 6 of Ramdas et al. 2020). (E_t) is a an e-process w.r.t. to a filtration $(\mathcal{F}'_t)$ iff it is nonnegative and $\mathbb{E}[E_\tau] \leq 1$ for all stopping times τ defined w.r.t. $(\mathcal{F}'_t)$.

Now, we get that, for any stopping time τ defined w.r.t. to filtration $(\mathcal{F}_t^j)$:

$$\mathbb{E}[E_\tau^{i\to j}(j)] = \frac{1}{2}(\mathbb{E}[M_\tau^{+}(j)] + \mathbb{E}[M_\tau^{-}(j)]) \leq 1,$$

where the last inequality is by $(M_t^{+}(j)), (M_t^{-}(j))$ being NSMs defined w.r.t. to filtration $(\mathcal{F}_t^j)$.

- Now, we will prove $(E_t^{i\to j}(i))$ is also an e-process. Since $H_0^{i\to j} \Rightarrow H_0^{i\to j,+} \cup H_0^{i\to j,-}$, if $H_0^{i\to j}$ is true, one of $H_0^{i\to j,+}$ or $H_0^{i\to j,-}$ holds. Without loss of generality, let $H_0^{i\to j,+}$ be true.

Here, the processes under consideration are now:

$$M_t^{+}(i) = \prod_{k\in T_j(t)} S_k^{i\to j,+}(i), \qquad M_t^{-}(i) = \prod_{k\in T_j(t)} S_k^{i\to j,-}(i).$$

We will show that $M_t^{+}(i)$ is an NSM w.r.t. to filtration $(\mathcal{F}_t^i)$, which implies that $M_t^{i\to j}$ is an e-process since $M_t^{i\to j} \leq M_t^{+}(i)$ for all $t \in \mathbb{N}$ almost surely.

When $I_t = i$, by Proposition 3.9, $S_t^{i\to j,+}(i)$ is an e-value and so:

$$\mathbb{E}[M_t^{+}(i)|\mathcal{F}_{t-1}] = \mathbb{E}[S_t^{i\to j,+}(i)|\mathcal{F}_{t-1}] \cdot M_{t-1}^{+} \leq M_{t-1}^{+}.$$

Finally, we check that when $I_t \neq i$, we have that:

$$\mathbb{E}[M_t^{+}(i)|\mathcal{F}_{t-1}] = M_{t-1}^{+} \leq M_{t-1}^{+}.$$

And we note that at the base case $t = 1$, for the NSM, we have that:

$$\mathbb{E}[M_t^{+}(i)] = \mathbb{E}[S_1^{i\to j,+}(i)] \leq 1 \text{ or } \mathbb{E}[M_t^{+}(i)] = 1$$

□

B.3. Deferred Results from Section 3.3

Proposition B.8. *Suppose the true edge orientation is actually that $j \to i$ and WLOG $\mu^i(j) > 0$. By setting $\lambda = b/\sigma_i^2$ for $S_t^{i\to j}(i)$ and $\lambda = b/\sigma_j^2$ for $S_t^{i\to j}(i)$, we have the following growth rates:*

1. $\mathbb{E}[\log S_t^{i\to j,+}(j) \mid \mathcal{F}_{t-1}] = b(\mu_i(j) - b/2)/\sigma_i^2$
2. $\mathbb{E}[\log S_t^{i\to j,+}(i) \mid \mathcal{F}_{t-1}] = \mathbb{E}[\log S_t^{i\to j,-}(i) \mid \mathcal{F}_{t-1}] = b^2/(2\sigma_j^2)$

Proof. We analyze the growth rates of each case separately:

1.

$$\begin{aligned}\mathbb{E}[\log S_t^{i\to j}(j) \mid \mathcal{F}_{t-1}] &= \lambda\left(\mathbb{E}\left[X_t^i \mid \mathcal{F}_{t-1}\right]\right) - \frac{\lambda^2\sigma_i^2}{2}\\ &= \frac{b}{\sigma_i^2}\mu_i(j) - \frac{b^2}{2\sigma_i^2}\\ &= \frac{b(\mu_i(j) - b/2)}{\sigma_i^2}\end{aligned}$$

2. We have that:

$$\begin{aligned}\mathbb{E}[\log S_t^{i\to j,\pm}(i) \mid \mathcal{F}_{t-1}] &= \lambda(b \pm \mathbb{E}[X_t^j \mid \mathcal{F}_{t-1}]) - \lambda^2\sigma_j^2/2\\ &= \lambda b - \lambda^2\sigma_j^2/2\\ &= \frac{b^2}{2\sigma_j^2}.\end{aligned}$$

□

Remark B.9. We note that $\mathrm{var}(X_i)$ in any interventional distribution is identified, and the same as $\mathrm{var}_{\mathcal{D}_0}(X_i)$. This allows us to put in the exact multiplier for $\lambda^2/2$ in the the NSM.

From Linear Graphs to Additive Graphs: We note that our setting may be generalized to additive graphs, when given an upper bound on the variance of variables in the interventional.

This is because, to set the appropriate λ for sequential e-values, we only need to have knowledge of b and an upper bound on the variance interventional distribution. With this, we could set a rate such that the growth rate is positive as in the power analysis above.

Proposition B.10. *If the edge $j \to i$ is the true orientation in G, then each of the the following statements hold true with probability $1-\beta$ for each $\beta \in [0,1]$:*

1. *For $(S_t^{i\to j,+}(j))$, we have that $\tau_\alpha \leq \frac{\sigma_i^2\log(\alpha^{-1}\beta^{-1})}{b(\mu_i(j)-b)}$.*
2. *For $(S_t^{i\to j,\pm}(i))$, we have that $\tau_\alpha \leq \frac{\sigma_j^2\log(\alpha^{-1}\beta^{-1})}{b^2}$*

Proof. We prove this explicitly for $S_t^{i\to j,+}(j)$ and other results for $(S_t^{i\to j,\pm}(i))$ follow similarly.

Let $M_t := \prod_{k=1}^{t} S_k^{i\to j,+}(j)$ as follows.

$$\begin{aligned}M_t &= \exp\left(\sum_{k=1}^{t}\lambda X_t^i - \frac{\lambda^2\sigma_i^2}{2}\right)\\ &= \exp(t(\lambda\mu_i(j) - \lambda^2\sigma_i^2))\cdot\exp\left(\sum_{k=1}^{t}\lambda(X_t^i - \mu_i(j)) + \frac{\lambda^2\sigma_i^2}{2}\right).\end{aligned}$$

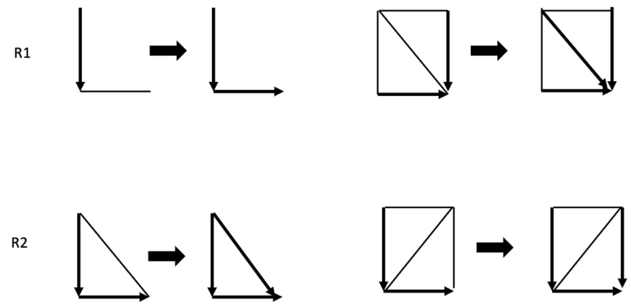


Figure 2. The four meek rules for propagating oriented edges.

Now, we note that $\exp\left(\sum_{k=1}^{t} -\lambda(X_t^i - \mu_i(j)) - \frac{\lambda^2\sigma_i^2}{2}\right)$ is a nonnegative supermartingale since $X_t^i - \mu_i(j)$ are i.i.d. σ_i^2-sub-Gaussian random variables with mean 0. As a result, we know that

$$M_t \geq \exp(t(\lambda\mu_i(j) - \lambda^2\sigma_i^2)) \cdot \beta$$

for all $t \in \mathbb{N}$ with probability $1-\beta$ by Ville's inequality. If we set $\lambda = b/\sigma_i^2$. We get that

$$\frac{\sigma_i^2 \log(\alpha^{-1}\beta^{-1})}{b(\mu_i(j) - b)} \leq t.$$

implies $M_t \geq \alpha^{-1}$ with probability $1-\beta$. This concludes our desired result. □

C. Deferred Proofs from Section 4

Lemma C.1 (Meek rules imply hypothesis conjunction/disjunction (general))**.** *For any edge orientation hypotheses $H_0^{i\to j}, H_0^{i_1\to j_1}, H_0^{i_2\to j_2}$, we have that*

$$\begin{aligned}
H_0^{i\to j} &= H_0^{i\to j} \cap H_0^{i_1\to j_1} && \text{if } i\to j \Rightarrow i_1 \to j_1 \\
H_0^{i\to j} \cap H_0^{i_1\to j_1} &= H_0^{i\to j} \cap H_0^{i_1\to j_1} \cap H_0^{i_2\to j_2} && \text{if } i\to j \wedge i_1 \to j_1 \Rightarrow i_2 \to j_2 \\
H_0^{i\to j} &= H_0^{i\to j} \cap (H_0^{i_1\to j_1} \cup H_0^{i_2\to j_2}) && \text{if } i\to j \Rightarrow i_1 \to j_1 \vee i_2 \to j_2
\end{aligned}$$

Proof. The results follow from an application of the logical rule that if $A \Rightarrow B$, then $A = A \cap B$.

For the first rule, we get the following implications:

$$H_0^{i\to j} \Leftrightarrow i \to j \text{ in } G^* \Rightarrow i_1 \to j_1 \text{ in } G^* \Leftrightarrow H_0^{i_1\to j_1}.$$

For the second rule, we can show its true by the following derivation.

$$H_0^{i\to j} \cap H_0^{i\to j} \Leftrightarrow i \to j \text{ and } i_1 \to j_1 \text{ in } G^* \Rightarrow i_2 \to j_2 \text{ in } G^* \Leftrightarrow H_0^{i_2\to j_2}.$$

For the last rule, we can derive the implication as follows:

$$\begin{aligned}
H_0^{i\to j} &\Leftrightarrow i \to j \text{ in } G^* \\
&\Rightarrow i_1 \to j_1 \text{ in } G^* \vee i_2 \to j_2 \text{ in } G^* \\
&\Leftrightarrow H_0^{i_1\to j_1} \cup H_0^{i_2\to j_2}.
\end{aligned}$$

□

C.1. Deferred Results from Section 4.1

Let $\mathcal{P}(T)$ denote the set of paths in T. The following lemma proves the correctness of the "extended hypothesis" generated by Algorithm 2.

Lemma C.2. *Given some tree T, let T' be the tree that results from applying a single Meek rule to T, i.e., through either Algorithm 2 or Algorithm 2 in Algorithm 2. Then, $H_0(T) = H_0(T')$.*

Proof. We perform a case analysis depending on the Meek rule (as defined in (3), (4), (4)) that is applied to T. Let the path in T that is expanded be $\hat{P}$.

1. In the case of (3) or (4), there exists a single path $P' = \hat{P} \cup \{i' \to j'\} \in \mathcal{P}(T')$ such that

$$\mathcal{P}(T') = \mathcal{P}(T) \setminus \{\hat{P}\} \cup \{P'\}\,,$$

   i.e., the only difference between T and T' is that path $\hat{P}$ gained a child $i' \to j'$ to become P'. We have that:

$$\bigcap_{i \to j \in \hat{P}} H_0^{i \to j} = \left(\bigcap_{i \to j \in \hat{P}} H_0^{i \to j} \right) \cap H_0^{i' \to j'} = \bigcap_{i \to j \in P'} H_0^{i \to j}. \tag{6}$$

   where the first equality is from Lemma 4.1. Hence, we get

$$\begin{aligned} H_0(T') &= \bigcup_{P \in \mathcal{P}(T')} \bigcap_{i \to j \in P} H_0^{i \to j} = \left(\bigcup_{P \in \mathcal{P}(T') \setminus \{P'\}} \bigcap_{i \to j \in P} H_0^{i \to j} \right) \cup \left(\bigcap_{i \to j \in P'} H_0^{i \to j} \right) \\ &\overset{\text{(a)}}{=} \left(\bigcup_{P \in \mathcal{P}(T) \setminus \{\hat{P}\}} \bigcap_{i \to j \in P} H_0^{i \to j} \right) \cup \left(\bigcap_{i \to j \in \hat{P}} H_0^{i \to j} \right) \\ &= H_0(T). \end{aligned}$$

   where equality (a) is by $\mathcal{P}(T') \setminus \{P'\} = \mathcal{P}(T) \setminus \{\hat{P}\}$ and (6).

2. In the case of (4), we know that there exist two paths $P_1', P_2' \in \mathcal{P}(T')$ such that $P_1' = \hat{P} \cup \{i_1' \to j_1'\}$ and $P_2' = \hat{P} \cup \{i_2' \to j_2'\}$, where $\hat{P} \in \mathcal{P}(T)$ and $\mathcal{P}(T') = \mathcal{P}(T) \setminus \{\hat{P}\} \cup \{P_1', P_2'\}$.

   Further, by Lemma 4.1, we know that:

$$\bigcap_{i \to j \in \hat{P}} H^{i \to j} \Rightarrow H^{i_i' \to j_1'} \cup H^{i_2' \to j_2'}.$$

   Hence, we get the following equality (using the logical relation $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$):

$$\bigcap_{i \to j \in \hat{P}} H_0^{i \to j} = \left(\bigcap_{i \to j \in \hat{P}} H_0^{i \to j} \right) \cap \left(H_0^{i_1' \to j_1'} \cup H_0^{i_2' \to j_2'} \right) = \left(\bigcap_{i \to j \in P_1'} H_0^{i \to j} \right) \cup \left(\bigcap_{i \to j \in P_2'} H_0^{i \to j} \right). \tag{7}$$

   From this, we obtain

$$\begin{aligned} H_0(T') &= \left(\bigcup_{P \in \mathcal{P}(T) \setminus \{P_1', P_2'\}} \bigcap_{i \to j \in P} H^{i \to j} \right) \cup \left(\bigcap_{i \to j \in P_1'} H^{i \to j} \right) \cup \left(\bigcap_{i \to j \in P_2'} H^{i \to j} \right) \\ &= \left(\bigcup_{P \in \mathcal{P}(T) \setminus \hat{P}} \bigcap_{i \to j \in \hat{P}} H^{i \to j} \right) \cup \left(\bigcap_{i \to j \in P_1'} H^{i \to j} \right) \cup \left(\bigcap_{i \to j \in P_2'} H^{i \to j} \right) \\ &\qquad \text{(since } \mathcal{P}(T') \setminus \{P_1', P_2'\} = \mathcal{P}(T) \setminus \{\hat{P}\}\text{)} \\ &= \left(\bigcup_{P \in \mathcal{P}(T) \setminus \hat{P}} \bigcap_{i \to j \in \hat{P}} H^{i \to j} \right) \cup \left(\bigcap_{i \to j \in \hat{P}} H^{i \to j} \right) \qquad \text{(by (7))} \\ &= H_0(T). \end{aligned}$$

□

Proposition C.3. *Algorithm 2 satisfies the following properties.*

- *(Soundness) Algorithm 2 is sound and does terminate.*
- *(Correctness) Let $T^{i\to j}$ be the resultant tree of Algorithm 2, then:*

$$H_0^{i\to j} = \bigcup_{P\in\mathcal{P}(T^{i\to j})} \bigcap_{i'\to j'\in P} H_0^{i'\to j'}.$$

Proof. **Soundness:** We note that a Meek rule cannot introduce a novel edge to a path in the tree if the path is of length $|E|$ and already contains an orientation for each possible edge in G. And so, the algorithm must terminate since each root to leaf path's length is bounded. This in turn means that so is the depth of the final tree $T^{i\to j}$.

Correctness: This follows from Lemma C.2 that each added edge(s) maintains the invariant that the logical expression corresponding to the tree is equal to $H_0^{i\to j}$.

□

C.2. Deferred Proofs from Section 4.2

Proposition C.4. *Let $(E_t^{i\to j}(i))$ be an e-process w.r.t. $(\mathcal{F}_t^i)$ under $H_0^{i\to j}$. For any path $P \subseteq \mathfrak{E}$ define*

$$E_t^P := \exp\left(\sum_{i'\in V} \max_{i\to j\in P(i')} \log E_t^{i\to j}(i') - \frac{|P(i')|-1}{2}\cdot \log(2|T_{i'}(t)|-2)\right),$$

Then (E_t^P) is an e-process w.r.t. $(\mathcal{F}_t)$ under H_0^P.

We are able to justify that (E_t^P) is an e-process by constructing an NSM that upper bounds E_t^P. We begin with the following fact.

Fact 3 (Theorem 2 of Cover & Ordentlich (1996)). Define $\mathbf{x}_t \in \mathbb{R}_+^d$ to be a d-dimensional nonnegative real vector for each $t\in\mathbb{N}$. Then, there exists a sequence of weight vectors, $(\mathbf{w}_t)$, where $\mathbf{w}_t \in \Delta^d$ and $\mathbf{w}_t$ is solely a function of $(\mathbf{x}_k)_{k\in[t-1]}$ for each $t \in \mathbb{N}$, such that

$$\log\left(\prod_{k=1}^t \mathbf{w}_k^\top \mathbf{x}_t\right) \geq \max_{\mathbf{w}\in\Delta^d} \log\left(\prod_{k=1}^t \mathbf{w}^\top \mathbf{x}_t\right) - \frac{d-1}{2}\cdot\log(2(t+1)) \text{ for all } t\in\mathbb{N}.$$

Proof. For each e-process $E_t^{i\to j}(i')$, let $M_t^{i\to j}(i')$ be the corresponding $(\mathcal{F}_t^{i'})$-NSM (under $H_0^{i\to j}$) such that $E_t^{i\to j}(i') \leq M_t^{i\to j}(i')$ for all $t\in\mathbb{N}$ almost surely. Now, define

$$\Delta M_t^{i\to j}(i') := \begin{cases} M_t^{i\to j}(i') & \text{if } t=1 \\ 1 & \text{if } M_{t-1}^{i\to j}(i') = 0 \\ \frac{M_t^{i\to j}(i')}{M_{t-1}^{i\to j}(i')} & \text{otherwise} \end{cases}.$$

For $t \geq 2$, we have that:

$$\mathbb{E}[\Delta M_t^{i\to j}(i')|\mathcal{F}_{t-1}^{i'}] = \frac{\mathbb{E}[M_t^{i\to j}(i')|\mathcal{F}_{t-1}^{i'}]}{M_{t-1}^{i\to j}(i')} \leq 1$$

as a result of $M_t^{i\to j}(i')$ being an NSM. And so, $(\Delta M_t^{i\to j}(i'))$ is a sequence of sequential e-values with respect to the filtration $\mathcal{F}_t^{i'}$. And $\mathbb{E}[M_1^{i\to j}(i')] \le 1$ by Definition 3.3.

Furthermore, we use the following lemma.

Lemma C.5. *$M_t^{i\to j}(i')$ is a NSM and $(\Delta M_t^{i\to j}(i'))$ is a sequence of sequential e-values under $(\mathcal{F}_t)$ as well.*

Proof. The filtration $(\mathcal{F}_t^{i'})$ is important here, since this implies that

$$M_t^{i\to j}(i') = M_{t-1}^{i\to j}(i') \text{ and } \Delta M_t^{i\to j}(i') = 1 \text{ if } I_t \neq i', \tag{8}$$

as $M_t^{i\to j}(i')$ is $\mathcal{F}_t^{i'}$-measurable (i.e., a function of samples from i') for each $t \in \mathbb{N}$.

Note that for each $t \in \mathbb{N}$,

$$X_t \perp\!\!\!\perp \mathcal{F}_{t-1} \mid I_t. \tag{9}$$

We will now show that $\mathbb{E}[\Delta M_t^{i\to j} \mid \mathcal{F}_{t-1}] \le 1$, i.e., is a sequential e-value under $(\mathcal{F}_t)$. This is trivially true if $I_t \neq i'$, so we consider the case where $I_t = i'$.

$$\begin{aligned}\mathbb{E}[\Delta M_t^{i\to j} \mid I_t = i', \mathcal{F}_{t-1}] &= \mathbb{E}[\Delta M_t^{i\to j} \mid \mathcal{F}_{t-1}^{i'}, I_t = i, \bigcup_{j\in V, j\neq i'} \mathcal{F}_{t-1}^{j}] \\ &= \mathbb{E}[\Delta M_t^{i\to j} \mid \mathcal{F}_{t-1}^{i'}, I_t = i'] \le 1.\end{aligned}$$

The first equality is because $\mathcal{F}_t = \mathcal{F}_{t-1}^{i'} \cup \bigcup_{j\in V, j\neq i'} \mathcal{F}_{t-1}^{j}$. The last line is by (9) and $\Delta M_t^{i\to j}$ being a sequential e-value under $\mathcal{F}_{t-1}^{i'}$.

□

Let $\Delta \mathbf{M}_t(i')$ be the vector of $\Delta M_t^{i\to j}(i')$ indexed for each $i \to j \in P(i')$. Now, we utilize the following regret bound from Fact 3, which implies that there exists a sequence of weights $(\mathbf{w}_t)$ predictable w.r.t. $(\mathcal{F}_t)$ such that we can define the following process:

$$\begin{aligned}
M_t^P &:= \prod_{k=1}^{t} \mathbf{w}_k^\top \Delta \mathbf{M}_k(I_k) = \exp\left(\sum_{k=1}^{t} \log(\mathbf{w}_k^\top \Delta \mathbf{M}_k(I_k))\right) \\
&= \exp\left(\sum_{i'\in V} \log\left(\prod_{k\in T_{i'}(t)} \mathbf{w}_k^\top \Delta \mathbf{M}_k(i')\right)\right) && \text{(collecting terms across } I_k \in V) \\
&\overset{(a)}{\ge} \exp\left(\sum_{i'\in V} \max_{\mathbf{w}\in\Delta^{|P(i')|}} \log\left(\prod_{k\in T_{i'}(t)} \mathbf{w}^\top \Delta \mathbf{M}_k(I_k)\right) - \frac{|P(i')|-1}{2}\cdot \log(2(|T_{i'}(t)|-1))\right) \\
&\overset{(b)}{=} \exp\left(\sum_{i'\in V} \max_{\mathbf{w}\in\Delta^{|P(i')|}} \log\left(\prod_{k\in [t]} \mathbf{w}^\top \Delta \mathbf{M}_k(I_k)\right) - \frac{|P(i')|-1}{2}\cdot \log(2(|T_{i'}(t)|-1))\right) \\
&\overset{(c)}{\ge} \exp\left(\sum_{i'\in V} \max_{i\to j\in P(i')} \log M_t^{i\to j}(i') - \frac{|P(i')|-1}{2}\cdot \log(2(|T_{i'}(t)|-1))\right) \\
&\ge \exp\left(\sum_{i'\in V} \max_{i\to j\in P(i')} \log E_t^{i\to j}(i') - \frac{|P(i')|-1}{2}\cdot \log(2(|T_{i'}(t)|-1))\right)
\end{aligned}$$

Inequality (a) is a result of Fact 3. For equality (b), we note that $\Delta \mathbf{M}_k(i') = \mathbf{1}$ (i.e., the vector of ones) for each $k \notin T_{i'}(t)$, as a result of (8). Consequently, is we can change the index of the product from $T_{i'}(t)$ to $[t]$, since multiplying by $\mathbf{w}^\top \mathbf{1} = 1$ does not change the product. Inequality (c) is because the elementary bases is a subset of $\Delta^{|A_{i'}|}$ and $\prod_{k \in T_{i'}(t)} \Delta M_k^{i \to j}(i') = M_t^{i \to j}(i')$ due to telescoping product. The last inequality is by definition of $M_t^{i \to j}(i') \geq E_t^{i \to j}(i')$ for all $t \in \mathbb{N}$.

Now, we only need to show that M_t^P is an NSM w.r.t. $(\mathcal{F}_t)$. Recall $M_t^P = \prod_{k \in T_{i'}(t)} \mathbf{w}_k^\top \Delta \mathbf{M}_k(I_k)$. We have that:

$$\mathbb{E}[M_t^P | \mathcal{F}_{t-1}] = \mathbb{E}[\mathbf{w}_t^\top \Delta \mathbf{M}_t(I_t) \mid \mathcal{F}_{t-1}] M_{t-1}^P$$

Thus, it suffices to show the following:

$$\mathbb{E}[\mathbf{w}_t^\top \Delta \mathbf{M}_t(I_t) \mid \mathcal{F}_{t-1}] \leq 1 \text{ under } H_0^P. \tag{10}$$

We know the following is true under H_0^P:

$$\mathbb{E}[\mathbf{w}_t^\top \Delta \mathbf{M}_t(I_t) \mid \mathcal{F}_{t-1}] = \sum_{i \to j \in P_{I_t}} w_t^{i \to j} \mathbb{E}[\Delta M_t^{i \to j}(I_t) \mid \mathcal{F}_{t-1}] \leq \sum_{i \to j \in P_{I_t}} w_t^{i \to j} = 1.$$

The inequality is by definition of $\Delta M_t^{i \to j}$ of being a sequential e-value (under $(\mathcal{F}_t)$) under $H_0^{i \to j}$, which holds as $i \to j \in P$ and H_0^P holds by assumption. The last equality is by $\mathbf{w}_t \in \Delta^{|A_{I_t}|}$.

Thus, we have shown (10) and proven our desired result.

□

Proposition C.6 (Correctness of combined e-process)**.** *Define*

$$E_t^{i \to j} := \min_{P \in \mathcal{P}(T^{i \to j})} E_t^P.$$

Then, $(E_t^{i \to j})$ *is an e-process when* $H_0^{i \to j}$ *is true.*

Proof. By the definition of $\mathcal{P}(T^{i \to j})$:

$$H_0^{i \to j} = H_0(T^{i \to j}) = \bigcup_{P \in \mathcal{P}(T^{i \to j})} H_0^P$$

Thus, if $H_0^{i \to j}$ is true, then there exists $P \in \mathcal{P}(T^{i \to j})$ such that H_0^P is true.

(E_t^P) is an e-process by Proposition 4.5. Since, $E_t^{i \to j} \leq E_t^P$ for all $t \in \mathbb{N}$ almost surely, $(E_t^{i \to j})$ is an e-process, and we have shown our desired result. □

Theorem C.7. *For any sequence of interventions* (I_t) *predictable w.r.t.* $(\mathcal{F}_t)$*, let* $\widehat{G}_t$ *be the partially oriented DAG where the test for each orientation is defined as follows:*

$$\varphi_t^{i \to j} = \mathbb{1}\{E_t^{i \to j} \geq |\overline{G}|/\alpha\}.$$

Then, $(\widehat{G}_t)$ *is anytime-valid orientation (as defined in (11)).*

Proof. Let the final oriented graph be $\hat{G}$.

$$\begin{aligned}
\mathbb{P}\left(\text{exists } t \in \mathbb{N} : \text{exists oriented edge in } \widehat{G}_t \text{ not in } G^*\right) &\leq \sum_{i \to j \in \hat{G}} \mathbb{P}\left(\text{exists } t \in \mathbb{N} : \text{orient edge } i \to j \wedge j \to i \text{ in } G^*\right) \\
&= \sum_{i \to j \in \hat{G}} \Pr(\text{exists } t \in \mathbb{N} : E_t^{j \to i} \geq |\overline{G}|/\alpha \wedge j \to i \text{ in } G^*) \\
&\leq \sum_{i \to j \in \hat{G}} \alpha/|\bar{G}| \qquad \text{(by Proposition 4.6)} \\
&= \alpha
\end{aligned} \tag{11}$$

□

Proposition C.8 (Additional power of combining test statistics)**.** *Consider an uniform set of interventions over nodes* $[n]$. *There exists a graph and edge* $i \to j$*, such that the expected growth rate (i.e. power) of* $\log E_t^{i\to j}$ *under the fully expanded tree* $T^{i\to j}$*, is* $\Omega(|\bar{G}|)$ *times that of* $\log E_t^{i\to j}$ *under the non-expanded tree (i.e. just the single edge* $i \to j$*).*

Proof. Consider a chain graph $\bar{G} = X_1 - X_2 - ... - X_n$ (generalizable to trees where the root has only one child), where the underlying graph is such that $X_1 \leftarrow X_2 ... \leftarrow X_n$. Such a setting allows for a simple, closed-form expression for the test statistic.

Suppose we are interested in testing $H_0^{1\to 2}$. Suppose there are m interventions, which means m/n interventions of each node.

In this setting, we assume that edge causal effects and variance are equal for fair comparisons. Thus $\forall i, j, \mathbb{E}[\log E_t^{i\to i+1}(i+1)] = \mathbb{E}[\log E_t^{j\to j+1}(j+1)]$. Certainly, if $E_t^{j\to j+1}(j+1)] > E_t^{1\to 2}(2)]$, then gain in power will be even more pronounced.

Using Proposition 4.5, we have that:

$$
\begin{aligned}
\mathbb{E}[\log E_t^{1\to 2}] &= \mathbb{E}\left[\log E_t^{1\to 2}(1) + \sum_{i=2}^{n-1}[\max(\log E_t^{i-1\to i}(i), \log E_t^{i\to i+1}(i)) - 1/2(\log(2|T_i(t)| - 2))] + \log E_t^{n-1\to n}(n)\right] \\
&\geq \mathbb{E}\left[\log E_t^{1\to 2}(1) + \sum_{i=2}^{n-1} \log E_t^{i-1\to i}(i) + \log E_t^{n-1\to n}(n) - 1/2(n-2)\log(2m/n - 2)\right] \\
&\geq \sum_{i=2}^{n-1} \mathbb{E}[\log E_t^{i-1\to i}(i)] - 1/2(n-2)\log(2m/n-2) \\
&= (n-2)\cdot \mathbb{E}[\log E_t^{1\to 2}(2)] - \tilde{O}(n) \\
&\geq (n-2)/2 \cdot \mathbb{E}[\log E_t^{1\to 2}(2) + \log E_t^{1\to 2}(1)] - \tilde{O}(n) \\
&\qquad \text{(for any edge } i \to j, \mathbb{E}[\log S_t^{i\to j,+}(j) \mid \mathcal{F}_{t-1}] \geq \mathbb{E}[\log S_t^{i\to j,+}(i) \mid \mathcal{F}_{t-1}]) \\
&\geq C \cdot \mathbb{E}[\log E_t^{1\to 2}(2) + \log E_t^{1\to 2}(1)] \qquad \text{(we assume that } \log E_t^{1\to 2}(2) = \Omega(m) >> \tilde{O}(n))
\end{aligned}
$$

for constant $C = \Omega(n)$.

□

Remark C.9. Note that the combination of evidence is such that we need not reject any of $X_i \to X_{i+1}$ to reject $X_1 \to X_2$. The cumulative evidence is enough, despite the data not being conclusive for any of the downstream edges!

C.3. Time complexity analysis of algorithms

Algorithm 2: Each path in the loop contains at most $|E|$ edges. Each round in the while loop requires examining at most $|E|$ to see if there is a new edge that is implied via Meek's rule. Thus, if the algorithm is run for T rounds, the time complexity is $T \cdot |E|$. We note that how much the tree is expanded out, as a function of T, is an user choice.

Algorithm 4: First, we consider the time complexity of updating test-statistic given a new intervention i at time t. It suffices to update E_t^P for every path $P \in T^{i\to j}$, which is pre-computed. Using the definition of E^P, it suffices to just re-compute $\log E_t^{i\to j}(i')$ to incorporate the new interventional data, and then take the minimum.

If one re-computation is taken to require one unit of computation, then there are $|P(i')|$ many $\log E_t^{i\to j}(i')$ re-computations. Using the definition of $|P(i')|$, we know it is upper bounded by the degree of i'. Thus, if the max degree of the graph is $deg(G)$, then the update to each edge test statistic requires at most $deg(G) \cdot |P \in P(T^{i\to j})|$ updates. In total, updating this for all edge hypotheses is upper bounded by: $2|E| \cdot deg(G) \cdot |P \in P(T^{i\to j})| = O(|V||E| \max_{i\to j} |P \in P(T^{i\to j})|)$. Finally, if there are T rounds with T interventions, the total number of updates comes out to: $O(T \cdot |V||E| \max_{i\to j} |P \in P(T^{i\to j})|)$.

Note that this characterizes the time-complexity of the e-process updates in as being polynomial (more precisely linear) in terms of the graph parameters and the size of the implication trees. As we previously note, the size of this implication tree (that is pre-computed) is an user-based choice. The more implications that are enumerated in the tree, the higher the power of the test. However, this in turn increases the time-complexity (and memory), which we can observe above.

Algorithm 4 Anytime Testing for Updates in Finite-Sample Causal Discovery

Require: Input: pre-compute logical tree $T^{i\to j}$ for each hypothesized edge orientation for edge $i - j$ in skeleton (via Algorithm 2)
Require: Sample from intervention distribution $(x_1^t, ..., x_n^t) \sim X_1, ..., X_n | do(X^t)$

for node X_i adjacent to X_{I_t} **do**
  if edge $X_i - X_{I_t}$ unoriented **then**
    Update $E_t^{i\leftarrow I_t}$, $E_t^{i\to I_t}$
    Test $E_t^{i\leftarrow I_t} \geq |\overline{G}|/\alpha, E_t^{i\to I_t} \geq |\overline{G}|/\alpha$ ▷ *Test if we can conclude $i \not\leftarrow I_t$ or $i \not\to I_t$ w.h.p.*
  end if
end for
for hypothesized orientation edge $i' \to j'$; $i' - j'$ unoriented, $i', j' \neq I_t$ **do** ▷ *Propagation*
  if exists edge $i \leftarrow I_t$ or $i \to I_t$ in $T^{i'\to j'}$ **then**
    Update $E_t^{i'\to j'}$ using updated $E_t^{i\leftarrow X^t}$ or $E_t^{i\to X^t}$
    Test $E_t^{i'\to j'} \geq |\overline{G}|/\alpha$ ▷ *Test if we can conclude $i' \not\to j'$ w.h.p.*
  end if
end for

D. Experiments

D.1. Fixed-time test statistic construction

Proposition D.1. *$P_t^{i\to j}$ satisfies $\mathbb{P}(P_t^{i\to j} \leq s) \leq s$ for all $s \in [0,1]$ and $t \in \mathbb{N}$ under $H_0^{i\to j}$.*

Proof. For $c \in \{\pm 1\}$, define

$$P_t^+(c) := 1 - \Phi\left(\frac{b \cdot T_i(t) - \widehat{\mu}_t^{j|\text{do}(i)} + c \cdot \widehat{\mu}_t^{i|\text{do}(j)}}{\sqrt{|T_j(t)| \operatorname{var} X_i + |T_i(t)| \operatorname{var} X_j}}\right),$$

$$P_t^-(c) := 1 - \Phi\left(\frac{\widehat{\mu}_t^{j|\text{do}(i)} + b \cdot T_i(t) + c \cdot \widehat{\mu}_t^{i|\text{do}(j)}}{\sqrt{|T_j(t)| \operatorname{var} X_i + |T_i(t)| \operatorname{var} X_j}}\right).$$

Note that, for any choice of c, $P_t^+(c)$ and $P_t^-(c)$ are z-test p-values under $H_0^{i\to j,+}$ and $H_0^{i\to j,-}$ respectively. As a result, $\max(P_t^+, P_t^-)$ is a p-value under $H_0^{i\to j}$.

Now, we can see that the following is true:

$$P_t^{i\to j} = 2\min(\max(P_t^+(1), P_t^-(1)), \max(P_t^+(-1), P_t^-(-1))).$$

Since taking double the minimum of any two p-values is still a valid p-value by union bound, we get our desired result that $P_t^{i\to j}$ is a p-value.

□

Remark D.2. Note the difference in qualifiers from the anytime guarantee, wherein correctness is guaranteed across time t, and not only at some fixed point t in time.

D.2. Comparing fixed-time vs anytime test statistics

Experiment Configurations: In experiments, we fix $b = 0.1$, variance as 1 and interventional value $\nu = 1$. Each setting is run for 20 trials to evaluate the mean and standard deviation.

We plot two metrics: (1) the mis-coverage rate (number of trials wherein the test statistic returns at least one falsely oriented edge) (2) the number of oriented edges (indeed an uninformative test that never rejects can trivially achieve 0 miscoverage rate).

To assess the guarantee of anytime approaches across a number of settings, we have the following experiments:

- Figures 3 and 5: Varying the number of interventional samples $\{100, 500, 1000, 5000, 10000\}$ (fixing $k = 0.2$) in ER graphs with number of nodes $\in \{10, 20, 30\}$, $\alpha \in \{0.1, 0.2\}$ and $p = 0.3$.
- Figures 4 and 6: Varying the number of interventional samples $\{100, 500, 1000, 5000, 10000\}$ (fixing $k = 0.2$) in ER graphs with number of nodes $\in \{10, 20, 30\}$, $\alpha \in \{0.1, 0.2\}$ and $p = 0.5$.
- Figures 7 to 10: Varying the number of interventional samples $\{100, 500, 1000, 5000, 10000\}$ (fixing $k = 0.2$) in tree graphs with number of nodes $\in \{10, 20, 50, 100\}$ and $\alpha \in \{0.1, 0.2\}$.
- Figures 11 and 12: Varying edge causal strength $k \in \{0.1, 0.2, 1, 2, 10\}$ (fixing number of samples at 1000) in ER graphs $(n, p) \in \{10, 20, 30\} \times (0.3, 0.5)$ and $\alpha = 0.2$.

In all these settings, in terms of miscoverage, we find that the anytime approach has controlled error rate below that of α (in green), although the miscoverage rate is not always 0. On the other hand, the fixed time approach can attain sizable error rate and introduce spuriously oriented edges. This trend seems consistent across two classes of graphs (ER and trees), as well as in ER graphs with varying SCM parameters edge strength k.

In terms of number of orientations, we observe that the number of orientations increases with sample complexity (as expected). However, the anytime test statistic orients conservatively at a (much) lower pace than does the fixed time approach. In exchange, this provides the error control and guarantees a high probability of only correct orientations.

D.3. Understanding the effectiveness of combining test statistics

To examine the effectiveness of propagating evidence from test statistics, we have the following experiment:

- Figure 13: Varying the number of interventional samples $\{100, 500, 1000, 5000, 10000\}$ (fixing $k = 0.2$) and plot the number of oriented edges in a chain graph with the number of nodes in $\{5, 10, 20, 50\}$ and edges have alternating causal strength in $\{0.1, 10\}$.

Chain graphs are an example where edges may benefit from propagation effects. We set up the causal strength to vary such that some edge orientations (those with low edge strength) will benefit from other edges (those with high edge strength). In the experiment, we compare the number of oriented edges at a fixed sample size by base e-values (as in Section 3) against the combined e-values (as in Section 4). Note that we also check for miscoverage rate to ensure correctness (in order to have a fair comparison); we do find that the miscoverage rate under both are 0.

In the plot, we observe that combining test statistics *may* help. It orients more edges than base e-values, when one has a sizable number of samples. Interestingly, we find that the base e-values is better in lower sample regimes. Moreover, the number of data points after which the combined test statistic is more effective increases with graph size.

We believe that this happens, because the combined test statistic, while having higher mean, also has higher variance. Thus, it is most effective when there are more samples. Overall, this suggests we should favor base e-values in smaller data and/or graph regimes, and combined test statistics in big data/graph regimes. The lighter-weight base e-values can be surprisingly effective. Verily, an interesting future work would be to develop testing methods that adapt the test statistic to the (unknown) SCM parameters and the test parameters (e.g. number of budgeted samples).

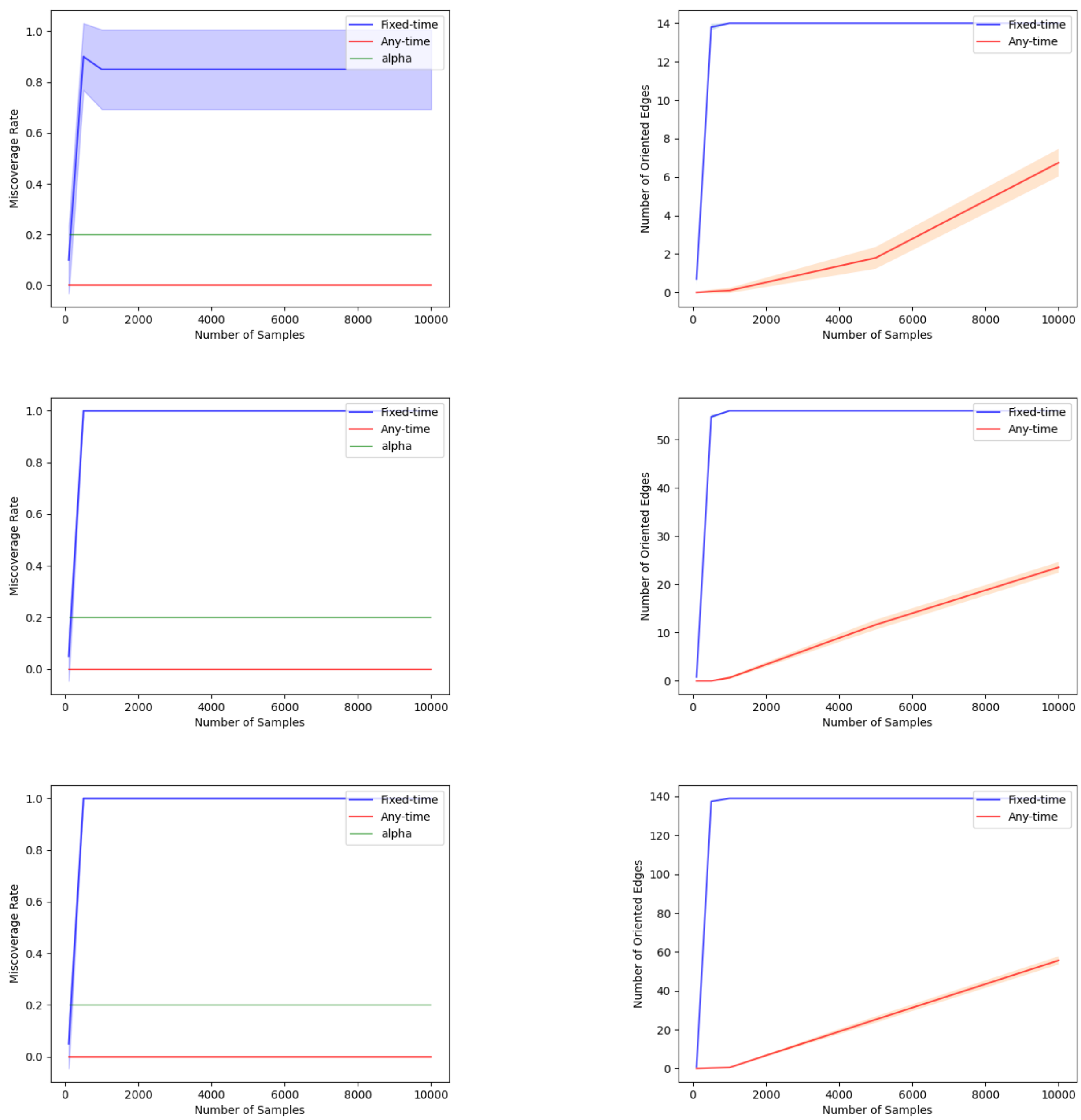


Figure 3. Plotting miscoverage rate and number of orientations in Erdos-Renyi Graphs with $\alpha = 0.2, p = 0.3$. First Row: $(n, p) = (10, 0.3)$; Second Row: $(n, p) = (20, 0.3)$; Third Row: $(n, p) = (30, 0.3)$.

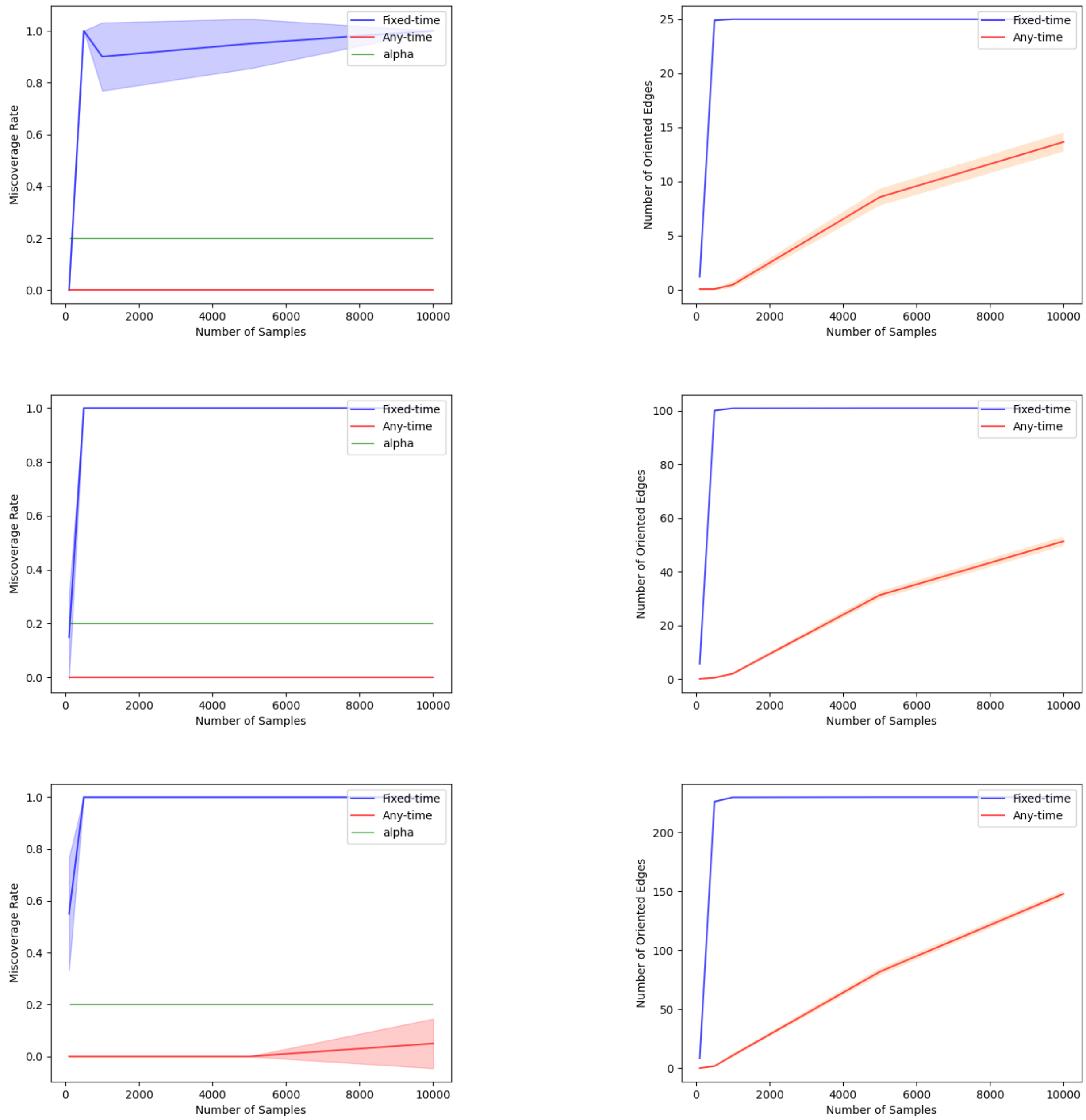


Figure 4. Plotting miscoverage rate and number of orientations in Erdos-Renyi Graphs with $\alpha = 0.2, p = 0.5$. First Row: $(n, p) = (10, 0.5)$; Second Row: $(n, p) = (20, 0.5)$; Third Row: $(n, p) = (30, 0.5)$.

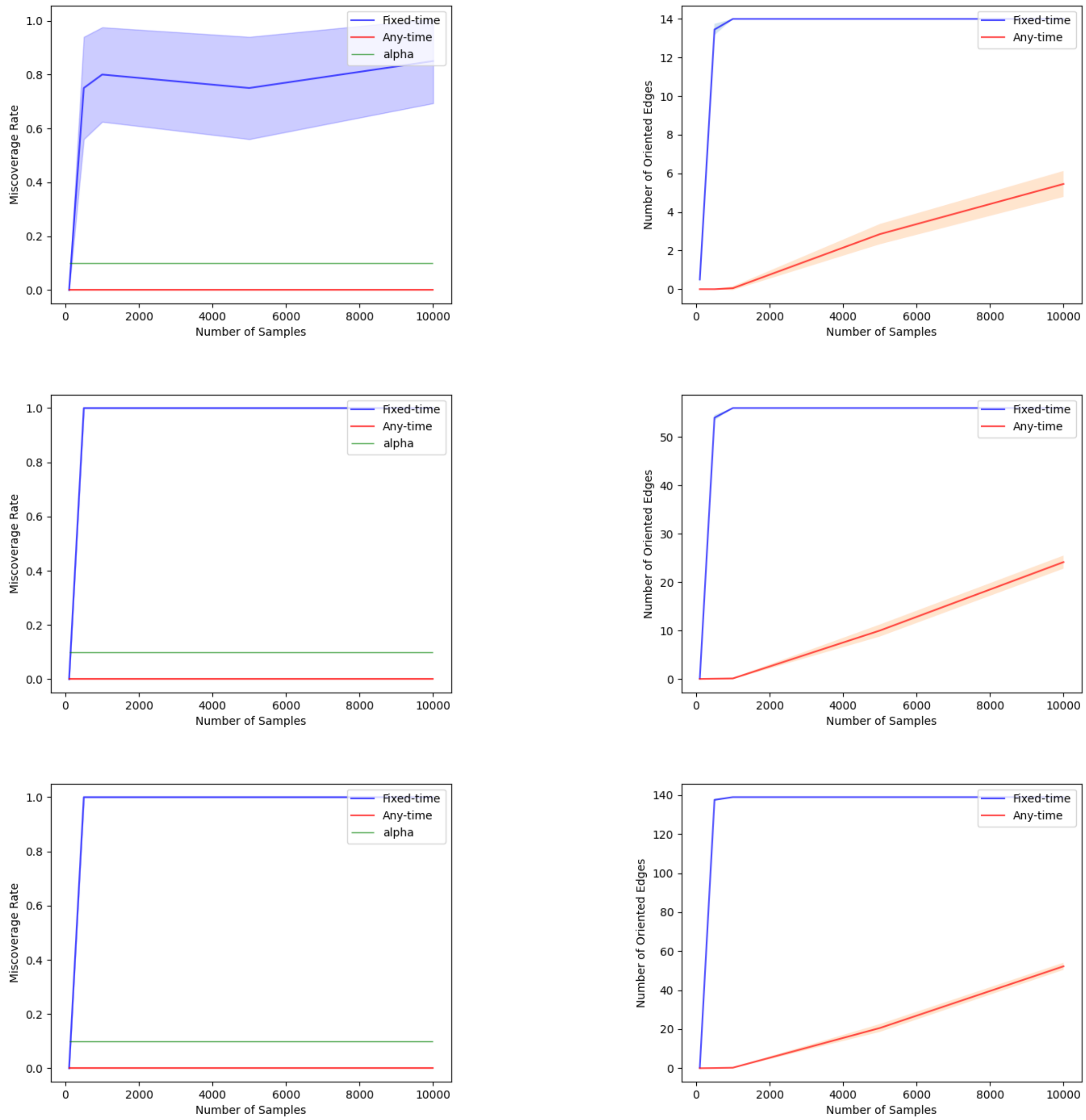


Figure 5. Plotting miscoverage rate and number of orientations in Erdos-Renyi Graphs with $\alpha = 0.1, p = 0.3$. First Row: $(n, p) = (10, 0.3)$; Second Row: $(n, p) = (20, 0.3)$; Third Row: $(n, p) = (30, 0.3)$.

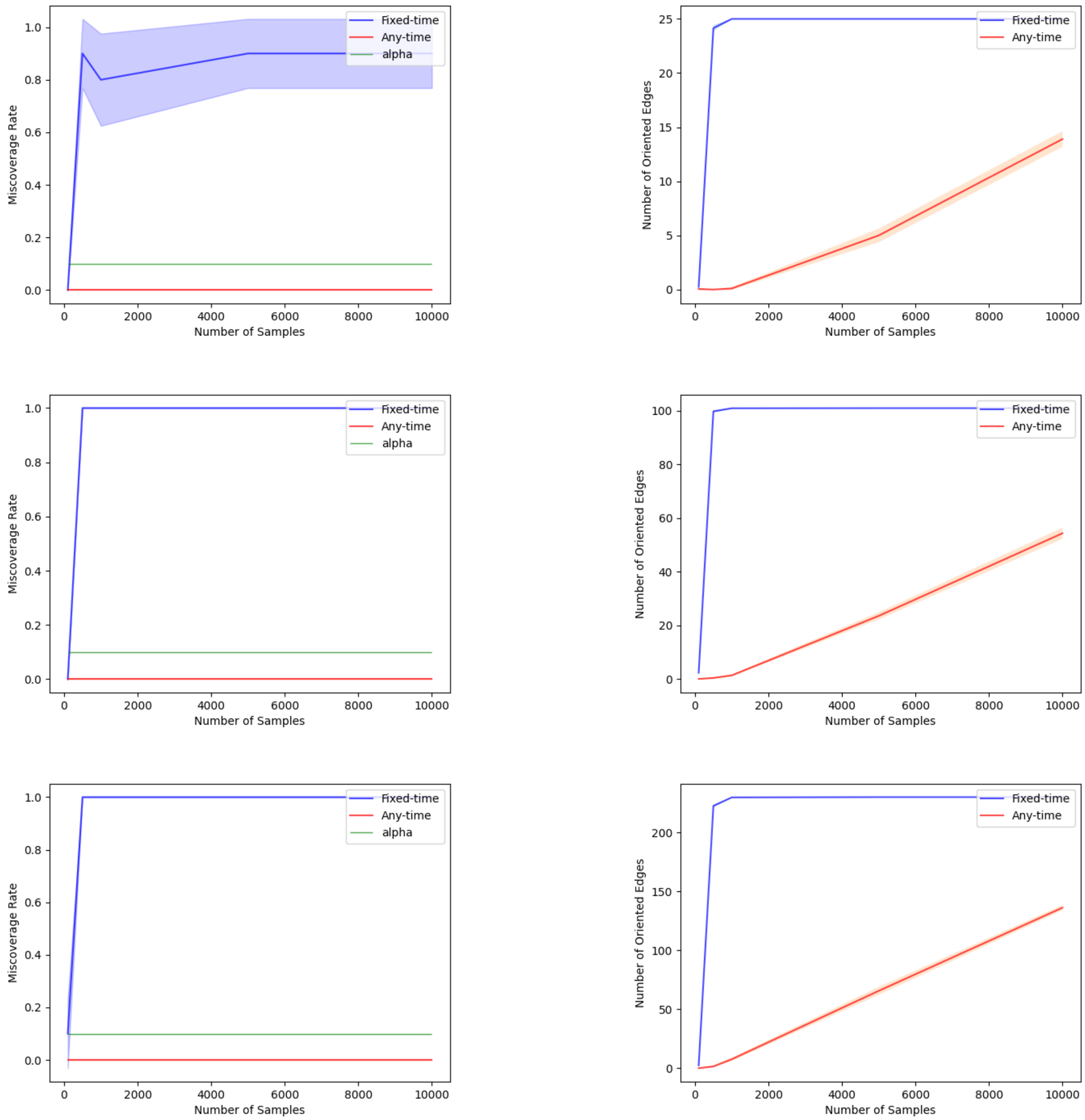


Figure 6. Plotting miscoverage rate and number of orientations in Erdos-Renyi Graphs with $\alpha = 0.1, p = 0.5$. First Row: $(n, p) = (10, 0.5)$; Second Row: $(n, p) = (20, 0.5)$; Third Row: $(n, p) = (30, 0.5)$.

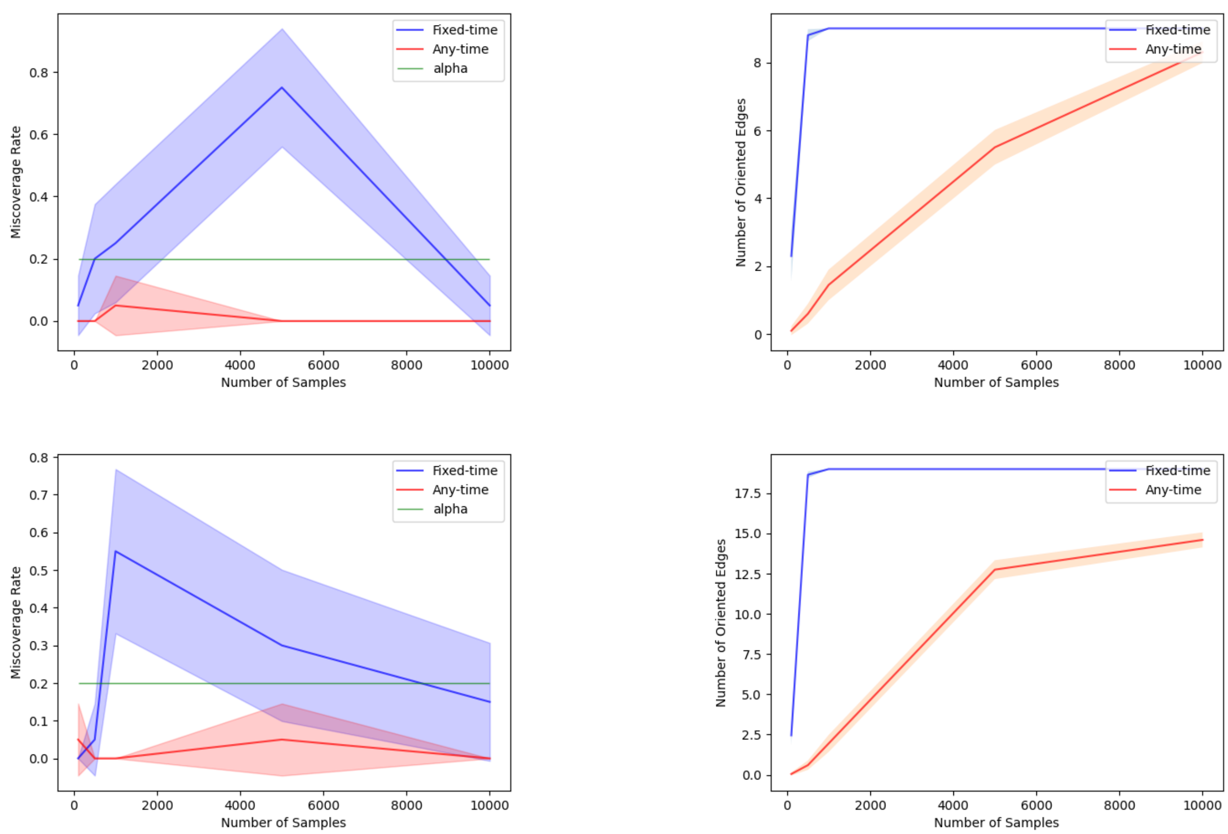


Figure 7. Plotting miscoverage rate and number of orientations in tree graphs with $\alpha = 0.2, n \in \{10, 20\}$. First Row: $n = 10$; Second Row: $n = 20$.

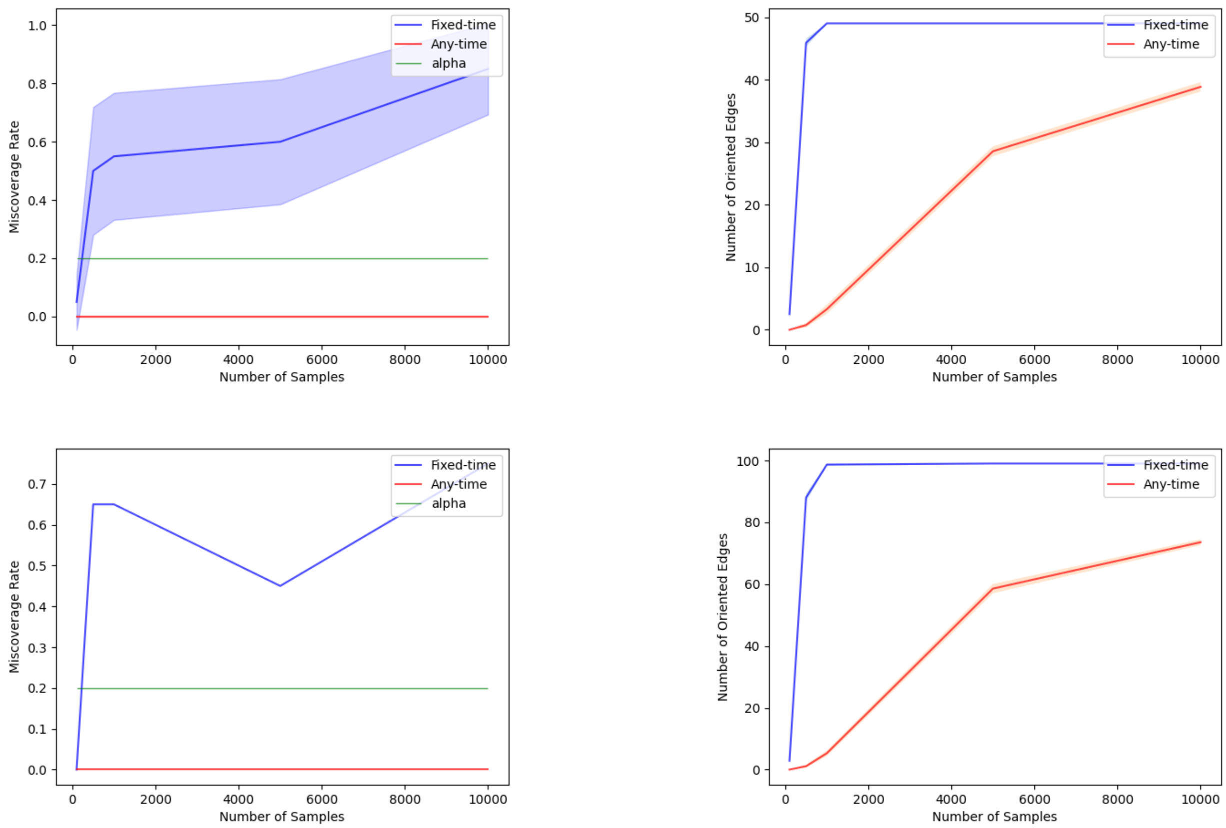


Figure 8. Plotting miscoverage rate and number of orientations in tree graphs with $\alpha = 0.2, n \in \{50, 100\}$. First Row: $n = 50$; Second Row: $n = 100$.

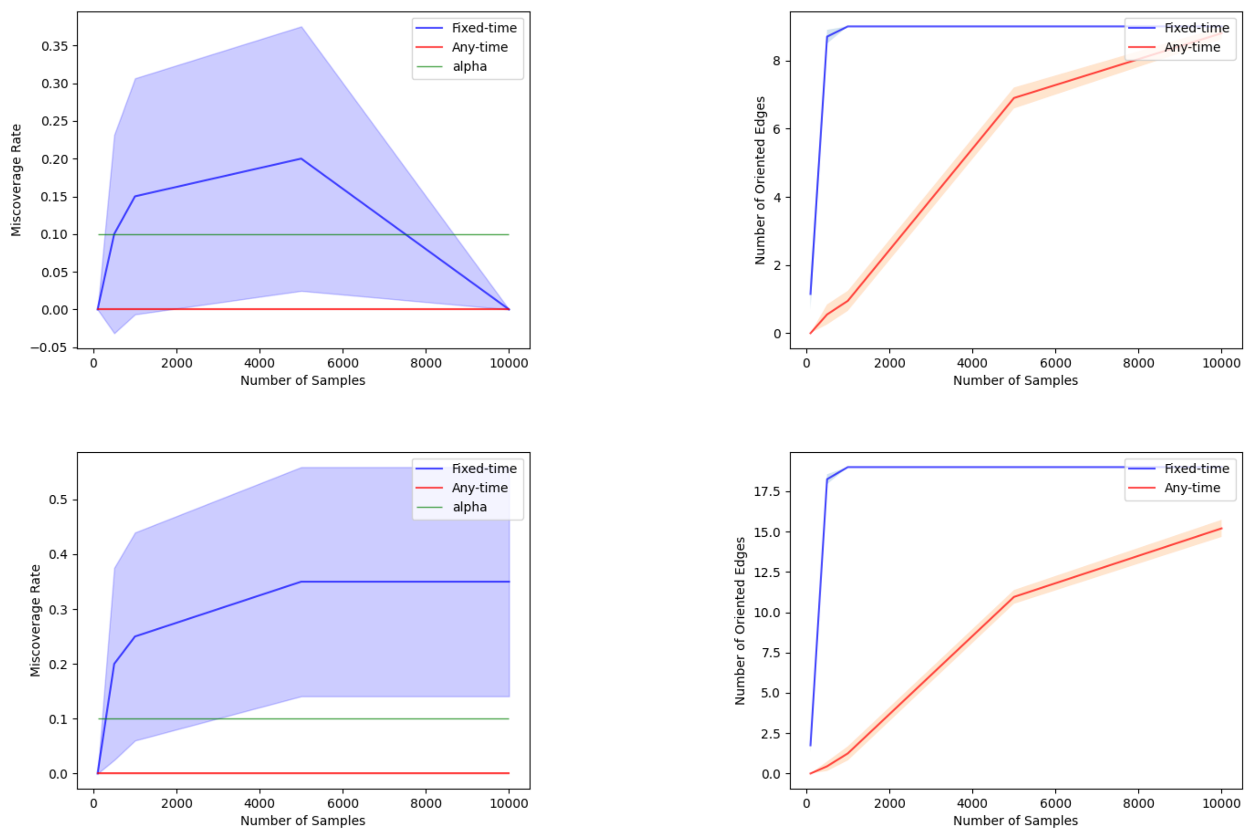


Figure 9. Plotting miscoverage rate and number of orientations in tree graphs with $\alpha = 0.1, n \in \{10, 20\}$. First Row: $n = 10$; Second Row: $n = 20$.

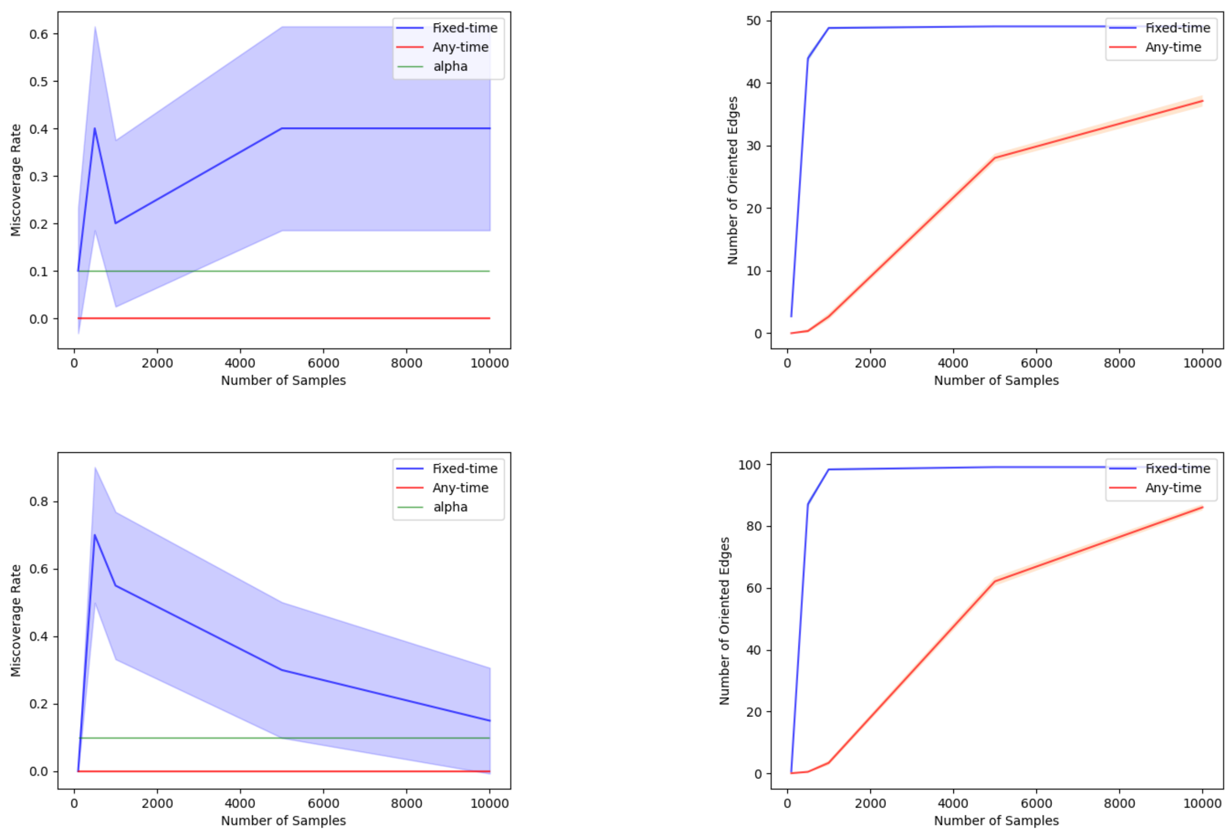


Figure 10. Plotting miscoverage rate and number of orientations in tree graphs with $\alpha = 0.1, n \in \{50, 100\}$. First Row: $n = 50$; Second Row: $n = 100$.

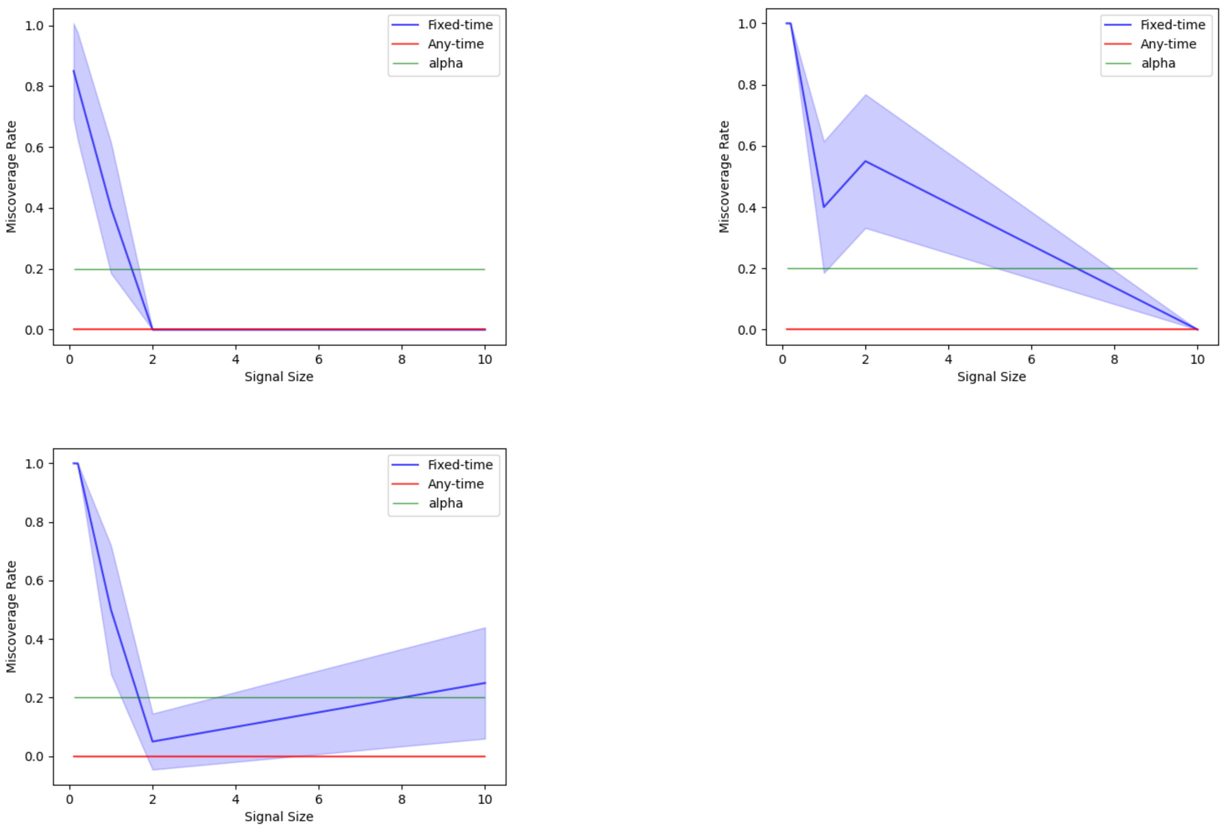


Figure 11. Plotting SCM parameter (edge strength k) vs miscoverage rate in ER graphs with $\alpha = 0.2, p = 0.3$. First Row: $n = 10$ (left) and $n = 20$ (right); Second Row: $n = 30$.

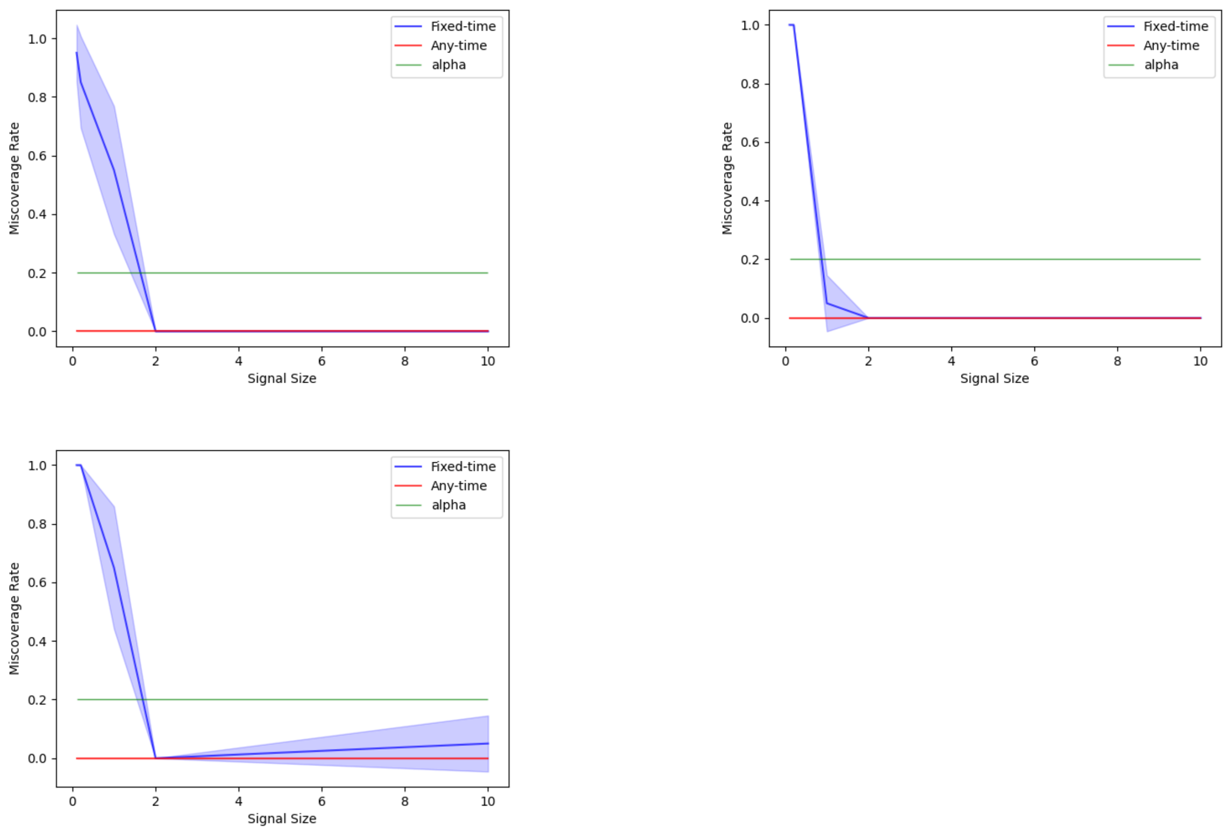


Figure 12. Plotting SCM parameter (edge strength k) vs miscoverage rate in ER graphs with $\alpha = 0.2, p = 0.5$. First Row: $n = 10$ (left) and $n = 20$ (right); Second Row: $n = 30$.

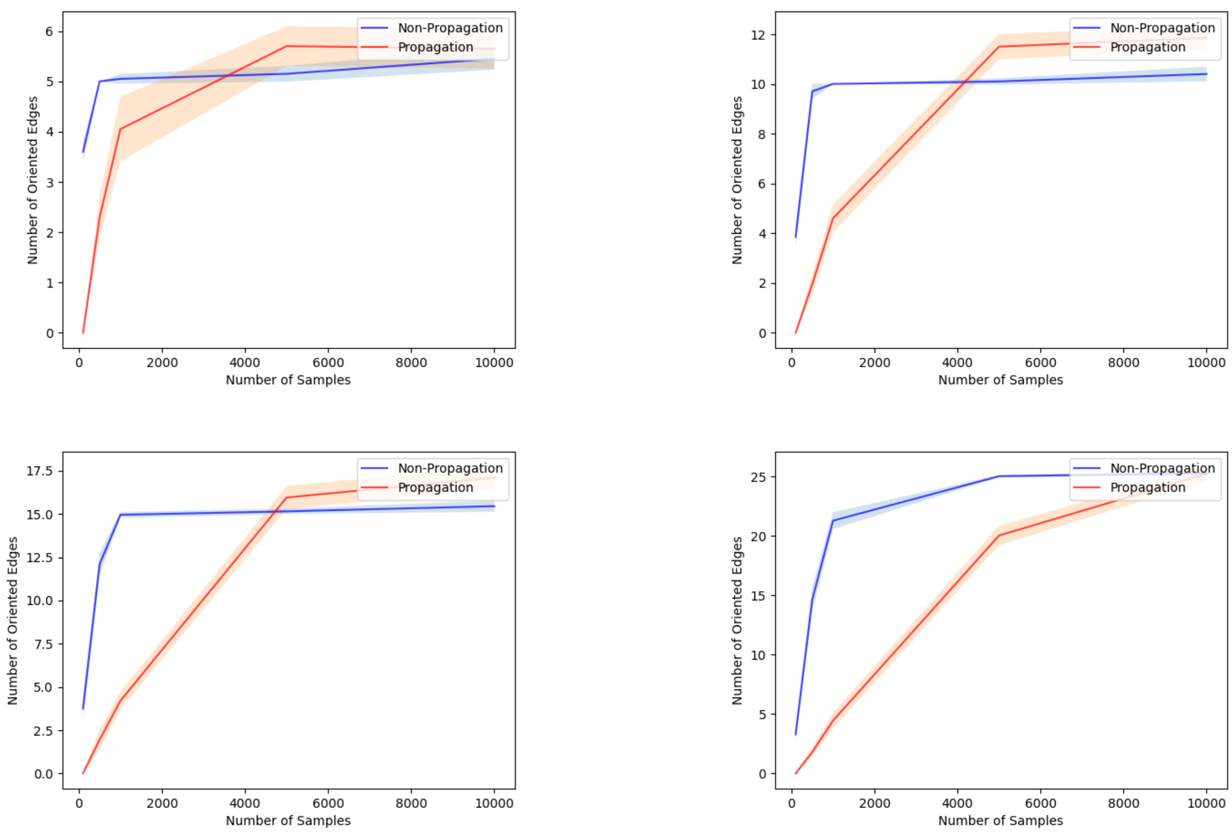


Figure 13. Comparing number of orientations of combined e-values vs those of base e-values in chain graphs with $\alpha = 0.2$. First Row: $n = 10$ (left) and $n = 20$ (right); Second Row: $n = 30$ (left) and $n = 50$ (right).

D.4. Evaluating Derived Upper Bounds on Stopping Time useful for Robust Testing

In Subsection 3.4, we derive a set of upper bounds on the number of samples needed for testing. One important implication is that this allows one to have an upper bound estimate on the amount of interventional data that one needs to collect to do the test. The other implication of this, useful for robust testing, is that one can use the non-conclusiveness of the test after this number of samples to detect spurious, non-edges.

In this subsection, we empirically verify this claim by evaluating the sample complexity needed for testing the orientation of some edge. Please refer to Figure 14 for a plot of the results.

- Firstly, we verify that with high probability, the bounds derived in Proposition 3.12 holds.

  In the experiments, we vary one parameter and fix the rest, checking the number of times the number of samples needed for testing is *below* the derived upper bound, out of 100 trials for each setting.

  We vary $\alpha = \{1\text{e-}6, 1\text{e-}5, 1\text{e-}4, 1\text{e-}3, 1\text{e-}2, 1\text{e-}1, 5\text{e-}1\}$, while setting $\mu_i(j) = 1.0, \sigma = 1.0, b = 0.1, \beta = 0.1$.

  We vary $\sigma = \{1\text{e-}6, 1\text{e-}3, 1\text{e-}2, 1\text{e-}1, 1, 10\}$, while setting $\alpha = 0.01, \mu_i(j) = 1.0, b = 0.1, \beta = 0.1$.

  We vary $\mu_i(j) = \{5\text{e-}1, 1, 5, 10, 100, 1000\}$, while setting $\alpha = 0.01, \sigma = 1.0, b = 0.1, \beta = 0.1$.

- Secondly, we verify that when there is no edge between the two edges, then with high probability the test statistic *does not reject* before the derived number of samples, thus allowing us to use the contrapositive of Proposition 3.12 to detect spuriously oriented edges.

  In this set of experiments, we use the same parameter setting as above, with the only difference that there is no causal effect from node i to node j. We check the number of times the number of samples needed for test conclusion is *below* the derived upper bound, out of 100 trials for each setting.

  We vary $\alpha = \{1\text{e-}6, 1\text{e-}5, 1\text{e-}4, 1\text{e-}3, 1\text{e-}2, 1\text{e-}1, 5\text{e-}1\}$, while setting $\sigma = 1.0, b = 0.1, \beta = 0.1$. Here, $\mu_i(j) = 0.0$.

  We vary $\sigma = \{1\text{e-}6, 1\text{e-}3, 1\text{e-}2, 1\text{e-}1, 1, 10\}$, while setting $\alpha = 0.01, b = 0.1, \beta = 0.1$. Here, $\mu_i(j) = 0.0$.

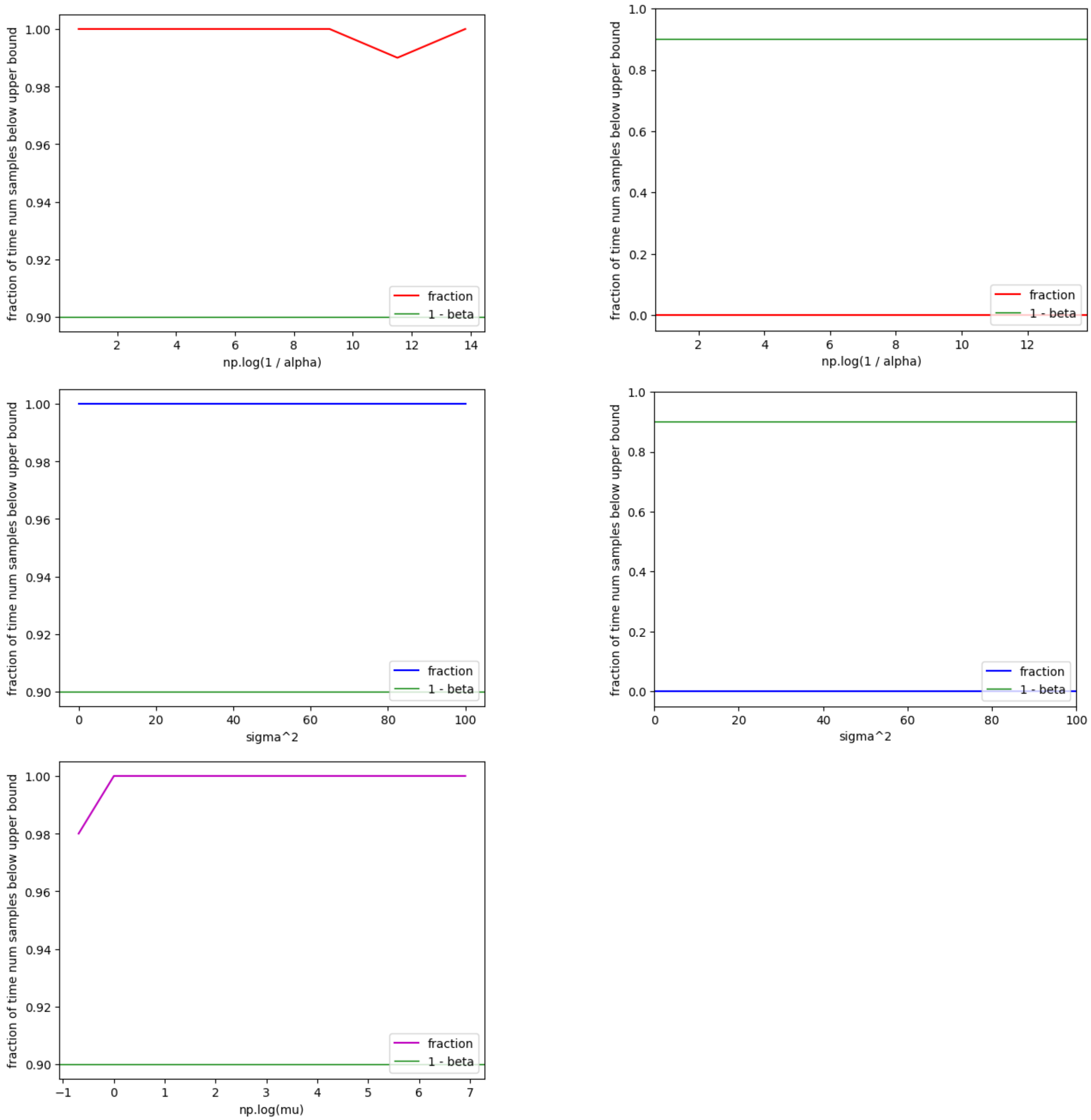


Figure 14. (Left Column) the fraction of 100 trials where the needed number of samples to conclude the test is below that of the derived upper bound (Right Column) the fraction of 100 trials where the needed number of sample complexity is below that of the derived upper bound, when there is no edge between the two nodes (i.e. $\mu_j(i) = 0$).

E. Multi-constraint Bandit Optimization

Causal verification is a well-known and important task in causal discovery (Squires et al., 2020; Porwal et al., 2022; Choo et al., 2022). Besides having practical applications (e.g. verifying a scientific conjecture corresponding to a causal graph structure), it has the theoretical benefit of understanding the lower bound that underlies any causal discovery algorithm.

The key challenge that arises in this problem is that an intervention policy needs to choose nodes I_t in order to grow the test statistic of all n edges *simultaneously*. Moreover, it only needs to optimize each test statistic only to the extent that the test statistic exceeds a threshold. In this section, we develop a novel, *multi-constraint* bandit algorithm needed for verification. Our key observation is that the causal verification setting reduces to the dual of the Bandits with Knapsack (BwK) setup (Badanidiyuru et al., 2018).

To this end, we develop Algorithm 3 that attains provable guarantees in the multi-constraint bandit setting, and applies immediately to the causal verification setting using the reduction. As observed in (Badanidiyuru et al., 2018), OPT, the expected total number of interventions needed by the optimal dynamic policy is difficult to characterize. In fact, even evaluating the expected number of interventions needed by *a given* time invariant, intervention policy is difficult. This is due to the difficulty of characterizing the *random stopping time* τ, when every test statistic exceeds the threshold. Thus, an algorithmic approach is taken where the proposed algorithm is shown to attain provable guarantees with respect to OPT.

E.1. Problem Statement

Setup: An instance of multi-constraint bandit optimization is parameterized by n arms, m constraints (henceforth "resources"), gain vector distributions $\mathcal{D}_i$ for arm i and budget b:

- There are n arms and m resources.
- Time proceeds in T rounds, where T is a finite time horizon given as input into the algorithm.
- Each round t, the learning algorithm picks some arm $x_t \in X$.
- Pulling arm x incurs deterministic cost c_x.
- The algorithm receives a gain vector, $R_{x_t} \in [0, M]^m$ where $R_{x_t} \sim D_{x_t}$ (some known distribution).
- There is a threshold $b \in \mathbb{R}^+$ on the total gain of each resource.
- The interaction terminates the first time $\sum_{t=1}^{\tau} R_{x_t} \geq b \cdot 1$.
- The goal of the algorithm is to minimize the total *expected* cost $\sum_{i=1}^{\tau} c_{x_i}$.

E.2. Reducing causal verification to multi-constraint bandits

We show that the causal verification problem corresponds to an instance of the multi-constraint bandit optimization problem. This algorithm is needed as our choice of intervention affects the e-processes of all edges.

We observe that bandit optimization is possible, because for any $j \to i$, the test statistic grows *additively* in the log of e-values:

$$
\begin{aligned}
E_t^{*j\to i} \geq 1/\alpha &\Leftrightarrow \forall P \in \mathcal{P}(T^{j\to i}), E_t^{*P} \geq 1/\alpha \\
&\Leftrightarrow \forall P \in \mathcal{P}(T^{j\to i}), \sum_{i' \in V} \log E_t^{*j\to i}(i') \geq \log(1/\alpha) \\
&\Leftrightarrow \forall P \in \mathcal{P}(T^{j\to i}), \sum_{k=1}^{t} \log S_k^*(P, I_k) \geq \log(1/\alpha)
\end{aligned}
$$

Reduction: Thus, given a causal verification instance, we may reduce to a multi-constraint bandit instance as follows:

1. Arms: define $n = |V|$ arms, each corresponding to a node intervention in the graph.

   Set $c_i = 1$ for all $i \in V$, as we only care about the total number of interventions. However, we note that our algorithm can handle differing node intervention costs.

2. Resources: define a resource corresponding to each (P, i') pair each path $P \in T^{j \to i}$ and intervention $i' \in V$.

   Accordingly, define the gain of pulling arm $i' \in V$ (i.e. intervening on node i') as a random draw of $\log S^*(P, i')$.

3. Define budget $b = \log 1/\alpha$.

In the causal verification setting, node intervention cost is set as $c_i = 1$ for all $i \in V$, since the objective of interest is the total number of interventions. Note however that Algorithm 3 can handle varying node intervention costs.

In the analysis below, we assume that X_i is a bounded r.v. with b, ν such that $\log S^*_k(P, I)$ is positive (as arm rewards are usually assumed to be positive in bandits literature). While this represents a subset of all SCM instances, as we will see, the query strategy design already results in solving an involved and novel multi-constraint bandit problem.

Finally, we note that one needs to manually specified the horizon T as in the multi-constraint bandit setting. This is a common assumption in BwK literature (Badanidiyuru et al., 2018), which has proven to be difficult to remove. T in the causal verification setting may be viewed as the maximum number of experiments a scientist can run, or a known upper bound on the number of experiments needed to verify the graph. Verily, an exciting future direction is understanding how to remove the need to specify T, and develop a

E.3. Algorithm Guarantee:

The goal is to compete with the optimal dynamic policy given all the latent information. That is, OPT is the expected total number of steps of the optimal dynamic policy, given foreknowledge of the distribution of outcome vectors.

Since OPT is difficult to analyze, consider the fractional relaxation of this problem in which the number of rounds in which a given arm is selected (and also the total number of rounds) can be fractional, and the reward and resource consumption per unit time are deterministically equal to the corresponding expected values in the original instance.

$$\begin{aligned} \min_{k_1,\ldots,k_n} \quad & c_1 k_1 + \ldots + c_n k_n \\ \text{s.t.} \quad & \sum_{j=1}^{n} r_{ji} k_i \geq b \text{ for each resource } i \in [m] \\ & k_i \geq 0 \end{aligned}$$

where k_i is the the fractional relaxation for the number of rounds in which a given arm i is selected.

This is a bounded LP, because $\sum_{i=1}^{n} k_i \leq T$ by definition. The optimal value of this LP is denoted by $\mathrm{OPT_{LP}}$. We may construct the dual program:

$$\begin{aligned} \max_{v_1,\ldots,v_m} \quad & b(v_1 + \ldots + v_m) \\ \text{s.t.} \quad & \sum_{i=1}^{m} r_{ji} v_i \leq c_j \text{ for each arm } j \in [n] \\ & v_j \geq 0 \end{aligned}$$

The dual variables v_i can be interpreted as a unit gain for the corresponding resource i.

Lemma E.1. $\mathrm{OPT_{LP}}$ *is a lower bound on the value of the optimal dynamic policy:* $\mathrm{OPT_{LP}} \leq \mathrm{OPT}$.

Proof. Let v^* be the optimal solution to the dual program. We note that by strong duality, $b \sum_{i=1}^{m} v_i^* = \mathrm{OPT_{LP}} = \sum_{j=1}^{n} c_i k_i^*$.

Algorithm 5

In the first n rounds, pull each arm once
For each arm x, define known expected gain vector $R_x \in [0, M]^m$
$v_1 = 1 \in [0,1]^m$ ▷ *$v_t \in [0,1]^m$ is the round-t estimate of the optimal solution to the dual v^**
Set $\epsilon = \sqrt{\frac{M \ln m}{b+M}}$
for rounds $t = n+1, ..., \tau$ **do**
  for arm $x \in X$ **do**
    Set expected gain $g_x = R_x \cdot v_t$
  end for
  Pull arm $x = x_t \in X$ that maximizes g_x / c_x
  Observe realized reward for each resource $r_x \in [0, M]$
  Update estimated unit gain for each resource i with normalized gain $r_x(i)/M$: ▷ *Cost-based MWU*

$$v_{t+1}(i) = v_t(i)(1-\epsilon)^{\ell}, \ell = r_x(i)/M$$

end for

Let Z_t denote the potential function: sum of costs incurred in optimal dynamic policy, plus total gain of the remaining resource endowment after round t.

At the start, the total gain of the remaining (all the) resource endowment is $Z_0 = b \sum_{i=1}^{m} v_i^*$.

We have that $Z_t = Z_{t-1} + c_{x_i t} - \sum_{i=1}^{m} r_{x_t i} v_i$ from arm pull x_t at time t.

From dual feasibility, we have that $c_j - \sum_{i=1}^{m} r_{ji} v_i \leq 0$. Then, it follows that the stochastic process $Z_0, Z_1, ..., Z_T$ is a submartingale.

Let τ be the stopping time of the optimal dynamic algorithm, i.e. the total number of rounds.

Thus, $Z_{\tau-1}$ equals the algorithm's total cost, plus the gain of the remaining (non-negative) resource supply at the start of round τ.

By Doob's optional stopping theorem, we have that $Z_0 \leq \mathbb{E}[Z_{\tau-1}] \leq \mathrm{OPT}$. □

Let us $\mathrm{REW}_{tot} = \sum_{t=1}^{\tau} c_t$.

The algorithmic approach will make use of dual vectors, computed as follows.

Learning the dual variable: We use the multiplicative weights update method to learn the optimal dual vector. This method raises the cost of a resource exponentially as it is consumed, which ensures that heavily demanded resources become costly, and thereby promotes balanced resource consumption.

Scaled-Hedge: This update scheme is such that for any τ and asequence of vectors $\pi_1, ..., \pi_\tau \in [0, M]^m$, feed in normalized $\pi_1/M, ..., \pi_\tau/M$ vectors into the hedge algorithm and obtain guarantee:

$$\forall y \in \Delta[m], \sum_{t=1}^{\tau} y_t^T \pi_t \leq (1+\epsilon) \sum_{t=1}^{\tau} y^T \pi_t + \frac{M \ln m}{\epsilon}$$

E.4. Algorithm Analysis under Known Arm Means

Let $\hat{R}_t \in [0, M]^{m \times n}$ be the actual gain matrix for round t. The (i, x) entry is the realized gain of resource i in round t if arm x were chosen in this round.

Suppose it holds with probability at least $1 - 1/T$ that the confidence interval for every latent parameter, in every round of execution, contains the true value of that latent parameter. We call this high-probability event a "clean execution".

The regret guarantee will hold almost surely assuming that a clean execution takes place. The regret can be at most $T \cdot M$ when a clean execution does not take place, and since this event has probability at most $1/T$ it contributes only $O(M)$ to the regret. We will henceforth assume a clean execution.

Claim 1. The Algorithm total cost is such that:

$$\text{REW}_{tot} - \text{OPT}_{\text{LP}} \leq \left[\tilde{O}\left(\frac{M}{b} + \sqrt{\frac{(b+M)M}{b}}\right)\text{OPT}_{\text{LP}} + nM\right] + \tilde{O}\left(\frac{1}{b}\right)\text{OPT}_{\text{LP}} \| \sum_{t=n+1}^{\tau} E_t z_t \|_\infty$$

where $E_t = R - \hat{R}_t$ under Algorithm 3.

Proof. Let k^* be the optimal solution to the LP-Primal with $\text{OPT}_{\text{LP}} = \sum_{j=1}^{n} c_j k_j^*$. For any realized gains by the algorithm policy, we have the following analysis.

Let $\hat{y} = e_i$, where resource i is (one of) the last resources, whose gain exceeds b:

$$\hat{y}^T \left(\sum_{t=1}^{\tau-1} \hat{R}_t z_t\right) \leq b \Rightarrow \hat{y}^T \left(\sum_{t=n+1}^{\tau-1} \hat{R}_t z_t\right) \leq b$$

Let the total cost after exploration be $\text{REW} = \sum_{t=n+1}^{\tau} c_t$ and define:

$$\bar{y} = \frac{1}{\text{REW}} \sum_{t=n+1}^{\tau-1} c_t y_t$$

Under Algorithm 3, we have at time t, by our choice of x_t, the corresponding z_t must be such that:

$$z_t \in \arg\max_{z \in \Delta(X)} \frac{y_t^T R z}{c^T z}$$

$$\begin{aligned}
b &\leq \bar{y}^T R k^* && \text{(from primal feasibility, } Rk^* \geq b1) \\
&= \frac{1}{\text{REW}} \sum_{t=n+1}^{\tau-1} c_t y_t^T R k^* && \text{(plug in definition of } \bar{y}) \\
&= \frac{c^T k^*}{\text{REW}} \sum_{t=n+1}^{\tau-1} c_t y_t^T R \frac{k^*}{c^T k^*} \\
&\leq \frac{\text{OPT}_{\text{LP}}}{\text{REW}} \sum_{t=n+1}^{\tau-1} y_t^T R z_t && \text{(since } \tfrac{y_t^T R z_t}{c^T z_t} \geq \tfrac{y_t^T R k^*}{c^T k^*}) \\
&\leq \min_y \frac{\text{OPT}_{\text{LP}}}{\text{REW}} \left[(1+\epsilon)\sum_{t=n+1}^{\tau-1} y^T R z_t + M \ln m/\epsilon\right] && \text{(since this holds for all } y \in \Delta[m] \text{ using hedge)} \\
&< (1+\epsilon)\frac{\text{OPT}_{\text{LP}}}{\text{REW}} \min_y \left[y^T \sum_{t=n+1}^{\tau-1} \hat{R}_t z_t + y^T \sum_{t=n+1}^{\tau-1} E_t z_t + M \ln m/\epsilon\right] && \text{(pull out } (1+\epsilon)) \\
&\leq (1+\epsilon)\frac{\text{OPT}_{\text{LP}}}{\text{REW}} \left[\hat{y}^T \sum_{t=n+1}^{\tau-1} \hat{R}_t z_t + \hat{y}^T \sum_{t=n+1}^{\tau-1} E_t z_t + M \ln m/\epsilon\right] && \text{(choose } y = \hat{y}) \\
&\leq (1+\epsilon)\frac{\text{OPT}_{\text{LP}}}{\text{REW}} \left[b + \hat{y}^T \sum_{t=n+1}^{\tau-1} E_t z_t + M \ln m/\epsilon\right] && \text{(since } \hat{y}^T \left(\textstyle\sum_{t=n+1}^{\tau-1} \hat{R}_t z_t\right) \leq b)
\end{aligned}$$

From this we get that by setting $\epsilon = \sqrt{\frac{M \ln m}{b+M}}$:

$$\begin{aligned}
&\text{REW} \leq \text{OPT}_{\text{LP}} \left((1+\epsilon) + \frac{1+\epsilon}{b} \left[\hat{y}^T \sum_{t=n+1}^{\tau-1} E_t z_t \right] + \frac{1+\epsilon}{b} \frac{M \ln m}{\epsilon} \right) \\
&\Leftrightarrow \text{REW} - \text{OPT}_{\text{LP}} \leq \text{OPT}_{\text{LP}} \left(\epsilon + \frac{1+\epsilon}{b} \frac{M \ln m}{\epsilon} + \frac{1+\epsilon}{b} \left[\hat{y}^T \sum_{t=n+1}^{\tau-1} E_t z_t \right] \right) \\
&\Leftrightarrow \text{REW}_{tot} - \text{OPT}_{\text{LP}} \leq \left(\epsilon + \frac{1+\epsilon}{b} \frac{M \ln m}{\epsilon} \right) \text{OPT}_{\text{LP}} + \sum_{t=1}^{n} c_t + \frac{1+\epsilon}{b} \text{OPT}_{\text{LP}} \left[\hat{y}^T \sum_{t=n+1}^{\tau-1} E_t z_t \right] \\
&\Leftrightarrow \text{REW}_{tot} - \text{OPT}_{\text{LP}} \leq \left(\sqrt{\frac{M \ln m}{b+M}} + \frac{M \ln m}{b} + \frac{\sqrt{(b+M) M \ln m}}{b} \right) \text{OPT}_{\text{LP}} \\
&\qquad\qquad + nM + \frac{1+\epsilon}{b} \text{OPT}_{\text{LP}} \left[\hat{y}^T \sum_{t=n+1}^{\tau-1} E_t z_t \right] \\
&\Leftrightarrow \text{REW}_{tot} - \text{OPT}_{\text{LP}} \leq \tilde{O} \left(\frac{M}{b} + \sqrt{\frac{(b+M)M}{b}} \right) \text{OPT}_{\text{LP}} + nM + \tilde{O} \left(\frac{1}{b} \right) \text{OPT}_{\text{LP}} \left[\hat{y}^T \sum_{t=n+1}^{\tau-1} E_t z_t \right]
\end{aligned}$$

□

Remark E.2. This roughly leads to a M factor larger than when $M = 1$, which should yield a $O(\sqrt{\frac{\ln m}{b}} + \frac{\ln m}{b})$ multiplier.

E.4.1. KNOWN R CONCENTRATION

For Error Analysis, it remains to bound the error term $\| \sum_{t=n+1}^{\tau} E_t z_t \|_\infty$.

In this case, we observe that each entry of E_t is a mean-zero random variable bounded in $[0, M]$. We may then use Hoeffding and union bound across all m resources to get that:

$$\Pr \left(\| \sum_{t=n+1}^{\tau} E_t z_t \|_\infty \leq (\tau - n - 1)\kappa \right) \geq 1 - m \cdot \left(2 \exp(-2(\tau - n - 1)\kappa^2 / M^2) \right)$$

Setting $1/T = m \cdot \left(2 \exp(-2(\tau - n - 1)\kappa^2 / M^2) \right)$, we obtain that:

$$\kappa = \sqrt{\frac{M^2 \log 2mT}{2(\tau - n - 1)}} \Rightarrow (\tau - n - 1)\kappa = O(M \sqrt{T \log 2mT}).$$

E.4.2. REGRET GUARANTEE

Theorem E.3. *Algorithm 3 with parameter $\epsilon = \sqrt{\frac{M \ln m}{b+M}}$ attains total regret:*

$$\text{REW}_{tot} - \text{OPT}_{\text{LP}} \leq \tilde{O} \left(\frac{M}{b} + \sqrt{\frac{(b+M)M}{b}} \right) \text{OPT}_{\text{LP}} + \tilde{O} \left(\frac{M\sqrt{T}}{b} + nM \right)$$

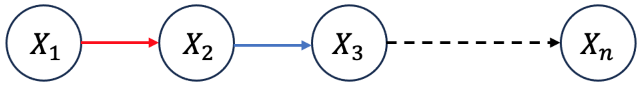
Proof. We have that:

$$\begin{aligned}
\mathrm{REW}_{tot} - \mathrm{OPT}_{\mathrm{LP}} &\leq \tilde{O}\left(\frac{M}{b} + \sqrt{\frac{(b+M)M}{b}}\right)\mathrm{OPT}_{\mathrm{LP}} + nM + \tilde{O}\left(\frac{1}{b}\right)\left[\hat{y}^T \sum_{t=n+1}^{\tau-1} E_t z_t\right] \\
&\leq \tilde{O}\left(\frac{M}{b} + \sqrt{\frac{(b+M)M}{b}}\right)\mathrm{OPT}_{\mathrm{LP}} + \tilde{O}\left(\frac{\|\sum_{t=n+1}^{\tau} E_t z_t\|_\infty}{b} + Mn\right) \\
&\leq \tilde{O}\left(\frac{M}{b} + \sqrt{\frac{(b+M)M}{b}}\right)\mathrm{OPT}_{\mathrm{LP}} + \tilde{O}\left(\frac{M\sqrt{T}}{b} + Mn\right)
\end{aligned}$$

□

Remark E.4. The regret dependence on the number of resources m is $O(\ln m)$.

We note that Theorem 6.1 follows from that $\mathrm{OPT}_{\mathrm{LP}} \leq \mathrm{OPT}$.



$$T^{1\to 2} = (X_1 \to X_2) \wedge (X_2 \to X_3) ... \wedge (X_{n-1} \to X_n)$$

$$E_t^{1\to 2} = \exp\left(\log E_t^{1\to 2}(1) + \sum_{i=2}^{n-1} \max(\log E_t^{i-1\to i1}(i), \log E_t^{i\to i+1}(i)) - 1/2(\log(2|T_i(t)| - 2)) + \log E_t^{n-1\to n}(n) \right)$$

Figure 15. Consider testing $X_1 \to X_2$ in the n-node chain graph $X_1 - X_2 - ... - X_n$. This is a graph, where the propagation of edge orientation is crucial for minimizing interventional complexity. We have that $X_1 \to X_2 \Rightarrow X_i \to X_{i+1}$, with which we can derive $T^{1\to 2}$ and $E_t^{1\to 2}$ explicitly. We note that, asymptotically (ignoring log factors in t), $E_t^{1\to 2}$ has much higher power than the e-process of $\exp(\log E_t^{1\to 2}(1) + \log E_t^{1\to 2}(2))$ (from Section 3), under non-expanded tree $(1 \to 2)$. $E_t^{1\to 2}$ leverages evidence from for example hypothesis $X_2 \to X_3$ (blue).

F. Worked through Examples

We work out in close form the test statistic of simple graphs to illustrate our test statistic construction and illustrate how it draws on power from its implications. Having already seen the chain graph example, we turn to the triangle example.

Three-node Triangle Graph: Consider testing $X_1 \to X_3$ in triangle graph $X_1 - X_2 - X_3$.

Since $X_3 \to X_2 \wedge X_2 \to X_1 \Rightarrow X_3 \to X_1, \therefore X_1 \to X_3 \Rightarrow X_2 \to X_3 \vee X_2 \to X_1$.

We have that $T^{1\to 3} = (X_2 \to X_3 \wedge X_1 \to X_3) \vee (X_1 \to X_2 \wedge X_1 \to X_3)$.

For path $P = X_2 \to X_3 \wedge X_1 \to X_3$, we have that:

$$E_t^P = \exp\left(\log E_t^{1\to 3}(1) + \log E_t^{2\to 3}(2) + \max(\log E_t^{1\to 3}(3), \log E_t^{2\to 3}(3)) - 1/2(\log(2|T_3(t)| - 2))\right)$$

For path $P' = X_1 \to X_2 \wedge X_1 \to X_3$, we have that:

$$E_t^{P'} = \exp\left(\max(\log E_t^{1\to 2}(1), \log E_t^{1\to 3}(1)) - 1/2(\log(2|T_1(t)| - 2)) + \log E_t^{1\to 2}(2) + \log E_t^{1\to 3}(3)\right)$$

Thus we see that, asymptotically (ignoring log factors in t), $E_t^{1\to 3} = \min(E_t^P, E_t^{P'})$ has *strictly* higher power. This is because both E_t^P and $E_t^{P'}$ have higher power the single-edge e-process corresponding $E_t^{1\to 3}(1)E_t^{1\to 3}(3) = \exp(\log E_t^{1\to 3}(1) + \log E_t^{1\to 3}(3))$.

Here, we can also observe that there are two possible updates when node X_3 is intervened upon, which corresponds to a choice when optimizing the test statistic. This naturally later motivates the use of bandit optimization.

Remark F.1. In Figure 15, another interesting observation of note is that, due to the combination of evidence, we need not reject any of $X_i \to X_{i+1}$ to reject $X_1 \to X_2$. The cumulative evidence across all $n-1$ edges may be enough for $E_t^{1\to 2}$ to exceed $1/\alpha$, and lead to the rejection of the null. This is despite the data being inconclusive for any of the downstream edges (i.e. $E_t^{i\to i+1}$ need not exceed $1/\alpha$).