A Meek Orientation Rules

In in Figure 3, we provide the four Meek orientation rules that are used in the definition of the essential graph.

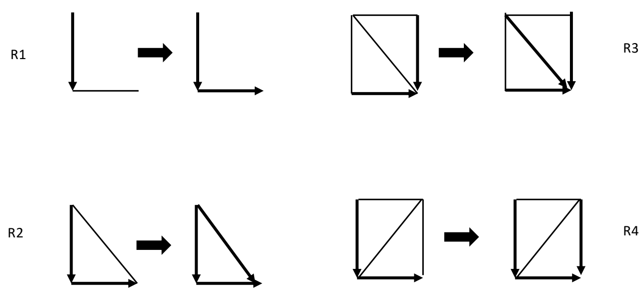


Figure 3: Meek orientation rules used to direct edges in the interventional essential graph representing the I-MEC.

The following two observable properties play an important role in various results in the main paper.

Property 1. *If a node v is involved in any of the four Meek rules and if the node v does not have an outgoing edge in the original causal DAG, then the oriented edge (in the right hand side motif of any of the four rules in Figure 3) is incident to v.*

Property 2. *If a node v is involved in a motif for any of the four rules, then either v has an outgoing edge or it has an adjacent undirected edge (on the left hand side motif appearing in that rule).*

B Additional Proofs

B.1 Proof of Lemma 1

Observe that all edges between G_n and v_{n+1} are directed to v_{n+1} and that v_{n+1} does not have any outgoing edges. Suppose that v_{n+1} is involved in one of the four Meek rules in Appendix A. Then by Property 1 in Appendix A, the discovered edge has to be incident to v_{n+1}. On the other hand, if v_{n+1} is not part of any Meek rule, then the rules must have already been applied in G_n to orient edges maximally, which completes the proof.

B.2 Proof of Lemma 2

The proof is similar to the proof of Lemma 1, i.e. it follows from Property 1 in Appendix A.

B.3 Proof of Equation 1

We can simplify this sum as follows:

$$\begin{aligned}\sum_{i\geq n} ix^{i-1} &= \frac{d}{dx}(\sum_{i\geq n} x^i)\\ &= \frac{d}{dx}\frac{x^n}{1-x}\\ &= \frac{x^n}{(1-x)^2} + nx^{n-1}\frac{1}{1-x}\end{aligned}$$

Substituting $(1-\rho(1-\rho))$ for x, we obtain

$$\begin{aligned}\sum_{i=n}^{\infty} \rho i * (1-\rho(1-\rho))^{i-1} &= \rho\left[\frac{(1-\rho(1-\rho))^n}{\rho^2(1-\rho)^2} + \right.\\ &\left. n\frac{(1-\rho(1-\rho))^{n-1}}{\rho(1-\rho)}\right]\\ &= \frac{(1-\rho(1-\rho))^n}{\rho(1-\rho)^2} + \\ &n\frac{(1-\rho(1-\rho))^{n-1}}{(1-\rho)}\end{aligned}$$

B.4 Proof of Lemma 3

Suppose $v_n \notin J$, then by Property 2, it cannot be part of any of the Meek rules in Appendix A . Therefore, it cannot aid in any of the rule applications after new edges have been discovered by interventions in J. This means that removing it before or after applying the Meek rules is irrelevant. Hence $J(G\backslash v_n) = J(G)\backslash v_n$. If $v_n \in J$, then the intervention on v_n gives no additional information as all its adjacent edges have already been discovered and hence it is equivalent to using $J\backslash v_n$. Hence, this reduces to the previous case with J replaced by $J\backslash v_n$ and thereby completes the proof.

B.5 Proof of Theorem 2

a) This follows directly from Lemma 1.
b) Suppose the MEC of G_n is given by the DAG set $\{H_1, H_2, \ldots, H_k\}$. For each i, let H_i' be the DAG H_i extended by adding the vertex v_{n+1}, with the same incoming edges as it has in G_{n+1}. From Lemma 1, it follows that the DAGs $\{H_1', H_2', \ldots, H_k'\}$ are contained in the MEC of G_{n+1}.
c) Due to the coupling, if a set of interventions orients G_{n+1}, it also orients G_n. The result follows.

The results for the expected values follow from the almost sure results.

B.6 Proof of Theorem 3

Let R be the set of interventions that achieves the minimum number $X_{n+1}(r)$ of unoriented edges in G_{n+1}

and let $|R| = r$. We apply the same set of interventions to G_n barring the possible intervention on node v_{n+1}. By Lemma 2, all edges unorientable in G_n after these 'copied' interventions are also unorientable in G_{n+1} even with/without the possible additional intervention on v_{n+1}. Since we can (possibly) add the extra intervention in G_n to bring the total number to r, this means that after these r interventions on G_n we have at most $X_{n+1}(r)$ unorientable edges in G_n, which completes the proof.

B.7 Proof of Theorem 5

Let R be the set of optimal set of r interventions on G_{n+1} that orient the maximum number of edges in the $\mathrm{Ess}(G_{n+1}, R)$. Therefore, $|\mathrm{Ess}(G_{n+1}, R)| = L_{n+1}(r)$. Let $R' = R \backslash v_{n+1}$. Suppose the $\mathrm{Ess}(G_n, R')$ is given by the DAG set $\{H_1, H_2, \ldots, H_k\}$. For each i, let H_i' be the DAG H_i extended by adding the vertex v_{n+1}, with the same incoming edges as in G_{n+1}. From Lemma 2, it follows that the DAGs $\{H_1', H_2', \ldots, H_k'\}$ are contained in $\mathrm{Ess}(G_{n+1}, R)$. This means, that, $L_n(r) \le |\mathrm{Ess}(G_n, R')| \le |\mathrm{Ess}(G_{n+1}, R)| = L_{n+1}(r)$

B.8 Proof of Theorem 6

For all monotonic sequences $\{x_n\}$,

$$\liminf\{x_n\} = \limsup\{x_n\} = \lim\{x_n\}.$$

Further, when a sequence of measurable functions $\{f_n\}$ converges pointwise to a function f, then f is also measurable. Here, X_n is a measurable function of the random variables G_n. Hence, $\mathbb{E}[X_\infty] = \lim_{n\to\infty} \mathbb{E}[X_n]$ follows from the Lebesgue Monotone Convergence Theorem.

B.9 Proof of Theorem 7

We first prove the following Lemma. The theorem follows from the Lemma.

Lemma 5. $\mathbb{E}(X_{n+1}) - \mathbb{E}(X_n) \le \rho n * (1-\rho(1-\rho))^{n-1}$.

Proof. Observe that the left hand side equals the expected number of unorientable edges incident to v_{n+1} in G_{n+1} by Lemma 1. We will upper bound this number as follows:

For each vertex i the edge $(i, n+1)$ is unoriented if it is present (probability ρ), and not part of an uncovered collider. The probability that $(i, n+1)$ and $(j, n+1)$ form an uncovered collider given $(i, n+1)$ is present is $\rho(1-\rho)$, and such probabilities for different j are independent. Thus the probability that $(i, n+1)$ is not part of an uncovered collider given that it is present is $(1-\rho(1-\rho))^{n-1}$. Multiplying by ρ for the probability that $(i, n+1)$ is present, and by n for the total number of such potential edges leads to the desired bound. ☐

B.10 Proof of Theorem 8

Since the unoriented edges incident to v_{n+1} can be oriented with at most one intervention each, we have that $I_{n+1} - I_n \le X_{n+1} - X_n$. This, combined with Theorem 7 results in the desired bound.

B.11 Proof of Theorem 11

If $A_{i,n} = 1$, and for all $j \neq \{i, n\}$ it holds that $A_{i,j} = Aj,n = 0$, then the edge (i, n) is isolated and thus unorientable. That happens with probability $\rho(1-\rho)^{n-2}(1-\rho)^{n-2}$ for each i. Note that there are $n-1$ such potential edges that are adjacent to vertex v_n and therefore figure into $\mathbb{E}(X_n) - \mathbb{E}(X_{n-1})$, which completes the proof.

B.12 Proof of Theorem 9

We provide a lemma and its proof regarding successive differences of the interventional metric $X_n(r)$. The result in the theorem follows immediately from this.

Lemma 6. $\mathbb{E}(X_{n+1}(r)) - \mathbb{E}(X_n(r)) \le \rho n * (1-\rho(1-\rho))^{n-1}$.

Proof. Let $X'_{n+1}(r)$ be the number of unoriented edges in G_{n+1} after we apply r interventions that achieve $X_n(r)$ unoriented edges in G_n. Observe that $X_{n+1}(r) \le X'_{n+1}(r)$. Since we can show that $\mathbb{E}(X'_{n+1}(r)) - \mathbb{E}(X_n(r)) \le \rho n * (1-\rho(1-\rho))^{n-1}$ by following the proof of Lemma 5 in the proof of Theorem 7 , this completes the proof. ☐

B.13 Proof of Theorem 10

Observe that $isuEss_n(r) - isuEss_{n+1}(r) \le X_{n+1}(r) - X_n(r)$, because $isuEss_n(r) - isuEss_{n+1}(r)$ is either 0 or 1 (it cannot be -1 by Theorem 4), and can only be 1 if $X_{n+1}(r) > X_n(r)$. This, combined with Lemma 6 means that $\mathbb{E}(isuEss_n(r)) - \mathbb{E}(isuEss_{n+1}(r)) \le \rho i * (1-\rho(1-\rho))^{n-1}$, which provides the necessary bound.

B.14 Proof of Theorem 12

It follows from the proof of Lemma 5 that the probability that v_{n+1} has an undirected edge is less than $\mathrm{RHS}(\rho, n)$. If v_{n+1} does not have an undirected edge, then, by the fact that A is a downstream independent algorithm it follows that $E(Y(r, A)_{n+1} = \mathbb{E}(Y(r, A)_n)$, and therefore that $\mathbb{E}(Y(r, A)_{n+1} - \mathbb{E}(Y(r, A)_n) = 0$. If v_n is adjacent to undirected edges, then

$\mathbb{E}(Y(r,A)_{n+1} - \mathbb{E}(Y(r,A)_n) < n(n+1)/2$, the number of possible edges in G_{n+1}. The bound follows since this happens with probability $< \rho n * (1-\rho(1-\rho))^{n-1})$.

B.15 Proof of Theorem 14

We consider the following bound for $\rho = 0.5$ from Theorem 1:

$$E[L_\infty] \leq E[X_\infty] \leq E[X_{30}] + \epsilon_{30}. \tag{3}$$

$\epsilon_{30} < 0.02$ (by direct calculation) based on Lemma **??**.

Now, we apply the following theorem from (quoted with appropriate modifications for a random variable taking values in $[0, B]$).

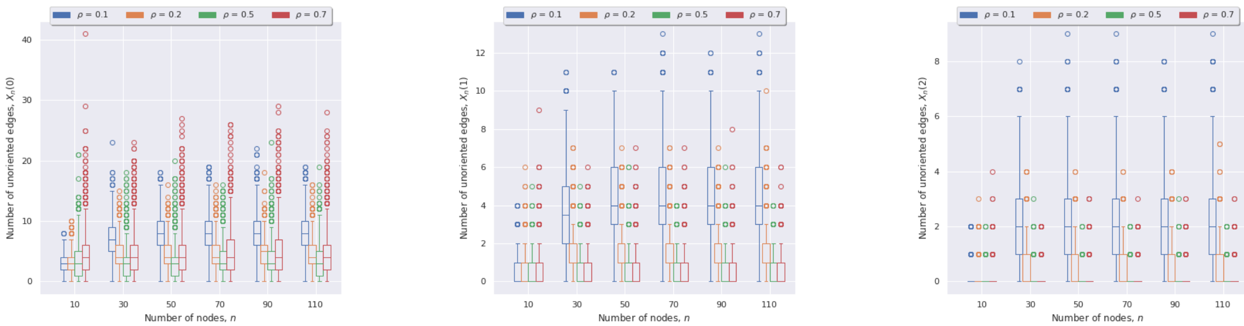
Theorem 15. *(Maurer and Pontil, 2009) If $Z_1, Z_2 \ldots Z_s$ are i.i.d random variables each bounded in $[0, B]$. Let M be the empirical mean and V be the empirical variance of the samples. Then, with probability $1 - \delta$,*

$$E[Z] \leq M + \sqrt{\frac{2V \log(2/\delta)}{s}} + B\frac{7\log(2/\delta)}{3(s-1)} \tag{4}$$

Now, for X_{30}, $B = 450$. Substituting $V = 7.054, M = 3.394, \delta = 0.01$ in the above theorem and using (3) we have the bound in the theorem.

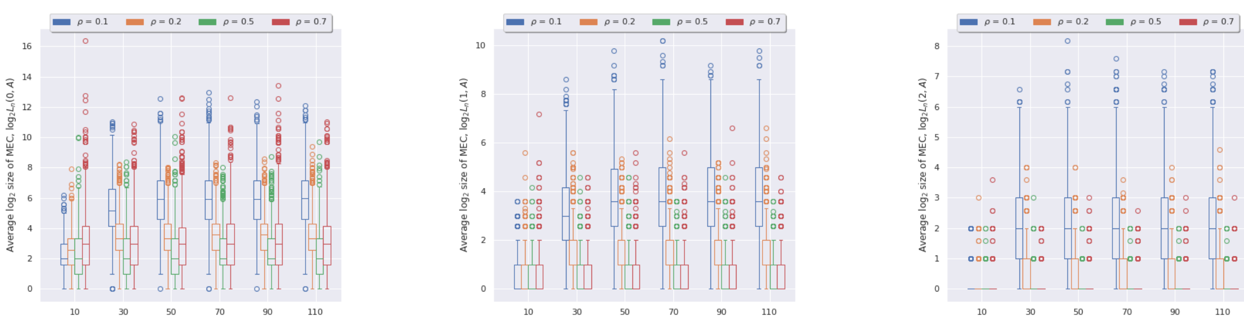
C Additional Figures

The additional figures regarding numerical simulations are given on the next page.



(a) (b) (c)

Figure 4: Number of unoriented edges of the 2,000 orderDAG samples. The middle line of the box is the median, the upper and lower edges are the upper and lower quartiles, and the circles are outliers.



(a) (b) (c)

Figure 5: $\log_2$ MEC sizes of the 2,000 orderDAG samples. The middle line of the box is the median, the upper and lower edges are the upper and lower quartiles, and the circles are outliers.