Appendices

A. Example of Comparison with the Influence Maximization Problem

Suppose $k = 1$. Figure 3 depicts a graph for which the optimal solution to the influence maximization problem is different from the optimal solution to the budgeted experiment design problems. Clearly, influencing vertex $v_1$ leads to influencing all the vertices in the graph, and hence, this vertex is the solution to the influence maximization problem. But, intervening on $v_1$ leads to discovering the orientation of only 3 edges, while intervening on, say $v_2$, leads to discovering the orientation of 5 edges.

B. Proof of Lemma 1

First we show that for a given directed graph $G_i \in \mathcal{G}$ the function $D(\mathcal{I}, G_i)$ is a monotonically increasing function of $\mathcal{I}$. In the proposed method, intervening on elements of $\mathcal{I}$, we first discover the orientation of the edges in $A_{G_i}^{(2)}$, and then applying the Meek rules, we possibly learn the orientation of some extra edges. Having $\mathcal{I}_1 \subseteq \mathcal{I}_2$, we have more information about the direction of edges. Hence, in the step of applying Meek rules, by soundness and order-independence of Meek algorithm, we recover the direction of more extra edges, i.e., $R(A_{G_i}^{(2)}, G_i) \subseteq R(A_{G_i}^{(2)}, G_i)$, which in turn implies that $D(\mathcal{I}_1, G_i) \leq D(\mathcal{I}_2, G_i)$. Finally, from the relation $D(\mathcal{I}) = \frac{1}{|\mathcal{G}|} \sum_{G_i \in \mathcal{G}} D(\mathcal{I}, G_i)$, the desired result is immediate.

C. Proof of Lemma 2

The direction $R(A_{G_i}^{(2)}, G^*) \leq R(A_{G_i}^{(2)}, G^*) \subseteq R(A_{G_i}^{(2)}, G^*) \cup R(A_{G_i}^{(2)}, G^*)$ is proved in the proof of Lemma 1. Also, as observed in the proof of Lemma 1, we have $R(A_{G_i}^{(2)}, G^*) \subseteq R(A_{G_i}^{(2)}, G^*) \cup R(A_{G_i}^{(2)}, G^*)$. Therefore, in order to prove that $R(A_{G_i}^{(2)}, G^*) \subseteq R(A_{G_i}^{(2)}, G^*) \cup R(A_{G_i}^{(2)}, G^*)$, for which it suffices to show that $R(A_{G_i}^{(2)}, G^*) \cup R(A_{G_i}^{(2)}, G^*) \subseteq R(A_{G_i}^{(2)}, G^*)$, we need to show that if $e \notin R(A_{G_i}^{(2)}, G^*)$ and $e \notin R(A_{G_i}^{(2)}, G^*)$, then $e \notin R(A_{G_i}^{(2)}, G^*)$. Proof by contradiction. Let $e \notin R(A_{G_i}^{(2)}, G^*)$ and $e \notin R(A_{G_i}^{(2)}, G^*)$, but its orientation is learned in the first iteration of applying Meek rules to $R(A_{G_i}^{(2)}, G^*) \cup R(A_{G_i}^{(2)}, G^*) \cup A(\text{Ess}(G^*))$. Then, we have learned the orientation of $e$ due to one of Meek rules (Verma & Pearl, 1992):

- Rule 1. $e = \{a, b\}$ is oriented as $(a, b)$ if $\exists c$ s.t. $e_1 = (c, a) \in R(A_{G_i}^{(2)}, G^*) \cup R(A_{G_i}^{(2)}, G^*) \cup A(\text{Ess}(G^*))$, and $\{c, b\} \notin \text{skeleton of } G^*$.
- Rule 2. $e = \{a, b\}$ is oriented as $(a, b)$ if $\exists c$ s.t. $e_1 = (c, a) \in R(A_{G_i}^{(2)}, G^*) \cup R(A_{G_i}^{(2)}, G^*) \cup A(\text{Ess}(G^*))$, and $e_2 = (c, b) \in R(A_{G_i}^{(2)}, G^*) \cup R(A_{G_i}^{(2)}, G^*) \cup A(\text{Ess}(G^*))$.
- Rule 3. $e = \{a, b\}$ is oriented as $(a, b)$ if $\exists c, d$ s.t. $e_1 = (c, a) \in R(A_{G_i}^{(2)}, G^*) \cup R(A_{G_i}^{(2)}, G^*) \cup A(\text{Ess}(G^*))$, $e_2 = (d, b) \in R(A_{G_i}^{(2)}, G^*) \cup R(A_{G_i}^{(2)}, G^*) \cup A(\text{Ess}(G^*))$, and $\{c, d\} \in \text{skeleton of } G^*$.
- Rule 4. $e = \{a, b\}$ is oriented as $(a, b)$ and $e = \{b, c\}$ is oriented as $(c, b)$ if $\exists d$ s.t. $e_1 = (d, c) \in R(A_{G_i}^{(2)}, G^*) \cup R(A_{G_i}^{(2)}, G^*) \cup A(\text{Ess}(G^*))$, and $\{a, d\} \in \text{skeleton of } G^*$.

In what follows, we show that the orientation of $e$ cannot be learned due to any of the Meek rules unless it belongs to $R(A_{G_i}^{(2)}, G^*)$ or $R(A_{G_i}^{(2)}, G^*)$.

Rule 1.

Without loss of generality, assume $e_1 \in R(A_{G_i}^{(2)}, G^*) \cup A(\text{Ess}(G^*))$. Therefore, we should have the condition of rule 1 satisfied when only intervening on $\mathcal{I}_1$ as well, which implies that $e \in R(A_{G_i}^{(2)}, G^*)$, which is a contradiction.

Rule 2.

If both $e_1$ and $e_2$ belong to $R(A_{G_i}^{(2)}, G^*) \cup A(\text{Ess}(G^*))$ (or $R(A_{G_i}^{(2)}, G^*) \cup A(\text{Ess}(G^*))$), then we should have the condition of rule 2 satisfied when only intervening on $\mathcal{I}_1$ (or $\mathcal{I}_2$) as well, which implies that $e \in R(A_{G_i}^{(2)}, G^*)$ (or $e \in R(A_{G_i}^{(2)}, G^*)$), which is a contradiction. Therefore, it suffices to show that the case that $e_1$ belongs to exactly one of
$R(A^{(Z)}_{G^*}, G^*) \cup A(\text{Ess}(G^*))$ or $R(A^{(Z)}_{G^*}, G^*) \cup A(\text{Ess}(G^*))$ and $e_2$ belongs only to the other one, does not happen. To this end, it suffices to show that there does not exist intervention target set $I$ such that $e_1 \in R(A^{(Z)}_{G^*}, G^*) \cup A(\text{Ess}(G^*))$ and $e_1, e_2 \notin R(A^{(Z)}_{G^*}, G^*) \cup A(\text{Ess}(G^*))$, i.e., there does not exist intervention target set $I$ that has structure $S_0$, depicted in Figure 4, as a subgraph of $\text{Ess}(G^*)$ after applying the orientations learned from $R(A^{(Z)}_{G^*}, G^*)$.

![Figure 4. Structure $S_0$](image)

If $e_1 \in A^{(Z)}_{G^*}$, then $a \in I$ or $c \in I$, which implies $e \in A^{(Z)}_{G^*}$ or $e_2 \in A^{(Z)}_{G^*}$, respectively, and hence, $e \in R(A^{(Z)}_{G^*}, G^*)$ or $e_2 \in R(A^{(Z)}_{G^*}, G^*)$, respectively. Therefore, in either case, $e \in R(A^{(Z)}_{G^*}, G^*)$, and $S_0$ will not be a subgraph. Therefore, $e_1 \notin A^{(Z)}_{G^*}$, and hence, $e_1$ was learned by applying one of the Meek rules. We consider each or the rules in the following:

- If we have learned the orientation of $e_1$ from rule 1, then we should have had one of the structures in Figure 5 as a subgraph of $\text{Ess}(G^*)$ after applying the orientations learned from $R(A^{(Z)}_{G^*}, G^*)$. In case of structure $S_1$, using rule 1 on subgraph induced on vertices $\{v_1, a, b\}$, we will also learn $(a, b)$. In case of structure $S_2$, using rule 4, we will also learn $(b, c)$. Therefore, we cannot learn only the direction of $e_1$ and hence, $S_0$ will not be a subgraph.

![Figure 5. Rule 1](image)

- If we have learned the orientation of $e_1$ from rule 3, then we have had one of the structures in Figure 6 as a subgraph of $\text{Ess}(G^*)$ after applying the orientations learned from $R(A^{(Z)}_{G^*}, G^*)$. In case of structures $S_3$ and $S_4$, using rule 1 on subgraph induced on vertices $\{v_2, c, b\}$, we will also learn $(c, b)$. In case of structure $S_5$, using rule 3 on subgraph induced on vertices $\{b, v_2, c, v_1\}$, we will also learn $(b, c)$. Therefore, we cannot learn only the direction of $e_1$ and hence, $S_0$ will not be a subgraph.

![Figure 6. Rule 3](image)

- If we have learned the orientation of $e_1$ from rule 4, then we have had one of the structures in Figure 7 as a subgraph of $\text{Ess}(G^*)$ after applying the orientations learned from $R(A^{(Z)}_{G^*}, G^*)$. In case of structures $S_6$, using rule 1 on subgraph induced on vertices $\{v_1, c, b\}$, we will also learn $(c, b)$. In case of structure $S_7$, using rule 4 on subgraph induced on vertices $\{b, v_2, v_1\}$, we will also learn $(a, b)$. In case of structures $S_8$, using rule 4 on subgraph induced on vertices $\{b, v_2, v_1, c\}$, we will also learn $(b, c)$. Therefore, we cannot learn only the direction of $e_1$ and hence, $S_0$ will not be a subgraph.

![Figure 7. Rule 4](image)

- If we have learned the orientation of $e_1$ from rule 2, then we should have had one of the structures in Figure 8 as a subgraph of $\text{Ess}(G^*)$ after applying the orientations learned from $R(A^{(Z)}_{G^*}, G^*)$. In case of structure $S_9$, using rule 1 on subgraph induced on vertices $\{v_1, c, b\}$, we will also learn $(c, b)$ and hence, $S_0$ will not be a subgraph. In case of structure $S_{10}$ if $v_1 \in I$, then the direction of the edge $\{v_1, b\}$ will be also known. If the direction of this edge is $(v_1, b)$, then using rule 2 on subgraph induced on vertices $\{a, v_1, b\}$, we will also learn $(a, b)$; otherwise, using rule 2 on subgraph induced on vertices $\{b, v_1, c\}$, we will also learn $(c, b)$. Therefore, $v_1 \notin I$. Also, as mentioned earlier, $a \notin I$. Therefore, we have learned the orientation of $(a, v_1)$ from applying Meek rules.

In the triangle induced on vertices $\{v_1, b, a\}$, we have learned only the orientation of one edge, which is $(a, v_1)$. But as seen in structures $S_1$ to $S_9$, all of them lead to learning the orientation of at least 2 edges of a triangle. In the following, we will show that a structure
of form S_{10}, does not lead to learning the orientation of only (a, v_1) and making S_{10} a subgraph either.

\[ v_1 \quad c \quad e_1 \]
\[ a \quad b \]

Figure 8. Rule 2

Suppose we had learned (a, v_1) via a structure of form S_{10}, as depicted in Figure 9(a). Using rule 4 on subgraph induced on vertices \( \{v_2, v_1, c, b\} \), we will also learn \((b, c)\). Therefore, we should have the edge \( \{v_2, c\} \) too. Also, using rule 2 on triangle induced on vertices \( \{v_2, v_1, c\} \), the orientation of this edges should be \((v_2, c)\). Therefore, in order to have the structure depicted in Figure 9(b) as a subgraph. As seen in Figure 9(b), we again have a structure similar to \( S_{10} \): a complete skeleton \( K_5 \), which contains \( \{v_j, c, \{a, v_j\}, \{v_j, b\}\} \) for \( j \in \{1, 2\} \) and \( \{v_2, v_1\} \), with a triangle on vertices \( \{v_2, b, a\} \), in which we have learned only the orientation of \((a, v_2)\).

\[ v_2 \quad c \quad e_1 \]
\[ a \quad b \]

Figure 9. Step of the induction.

We claim that this procedure always repeats, i.e., at step \( i \), we end up with skeleton \( K_i \), which contains \( \{v_j, c, \{a, v_j\}, \{v_j, b\}\} \), for \( j \in \{1, ..., i\} \) and \( \{v_k, v_j\} \), for \( 1 \leq j < k \leq i \), with a triangle on vertices \( \{v_1, b, a\} \), in which we have learned only the orientation of \((a, v_i)\). We prove this claim by induction. We have already proved the base of the induction above. For the step of the induction, suppose the hypothesis is true for \( i - 1 \). Add vertex \( v_i \) to form a structure of form \( S_{10} \) for \( (a, v_{i-1}) \). \( v_i \) should be adjacent to \( v_j \), for \( j \in \{1, ..., i - 2\} \); otherwise, using rule 4 on subgraph induced on vertices \( \{v_i, v_{i-1}, v_j, b\} \), we will also learn \((b, v_j)\). Moreover, using rule 2 on triangle induced on vertices \( \{v_i, v_{i-1}, v_j\} \), the direction of \((v_i, v_j)\) should be \((v_i, v_j)\). Also, using rule 4 on subgraph induced on vertices \( \{v_i, v_{i-1}, c, b\} \), we will also learn \((b, c)\). Therefore, we should have the edge \( \{v_i, c\} \) too.

We showed that \( S_0 \) is a subgraph only if \( S_{10} \) is a subgraph, and \( S_{10} \) is a subgraph only if the structure in Figure 9(b) is a subgraph, and this chain of required subgraphs continue. Therefore, since the order of the graph is finite, there exist a step where since we cannot add a new vertex, it is not possible to have one of the required subgraphs, and hence we conclude that \( S_0 \) is not a subgraph.

**Rule 3.**

Since edges \( e_1 \) and \( e_2 \) form a V-structure, they should appear in \( A(\text{Ess}(G^*)) \) as well. Therefore, we should have the condition of rule 3 satisfied when only intervening on \( I_1 \) as well, which implies that \( e \in R(A_{G^*}^{[I_1]}), G^* \), which is a contradiction.

**Rule 4.**

Without loss of generality, assume \( e_1 \in R(A_{G^*}^{[I_1]}), G^* \) and \( e \), which is a contradiction.

The argument above proves that there is no edge \( e \) such that \( e \notin R(A_{G^*}^{[I_1]}), G^* \) but \( e \in R(R(A_{G^*}^{[I_1]}), G^* \) or \( R(A_{G^*}^{[I_2]}), G^* \).

### D. Proof of Theorem 3

Let \( I^* = \{v_1^*, ..., v_k^*\} \in \arg\max_{l \subseteq \mathcal{V}} \cup \{I \in \mathbb{P}(\mathcal{D}(I)) \}

\[
\begin{align*}
\mathcal{D}(I^*) & \stackrel{(a)}{=} \mathcal{D}(I^* \cup I_1) \\
& + \sum_{j=1}^{k} [\mathcal{D}(I_i \cup \{v_1^*, ..., v_j\}) - \mathcal{D}(I_i \cup \{v_1^*, ..., v_{j-1}\})] \\
& \leq \mathcal{D}(I_i) + \sum_{j=1}^{k} [\mathcal{D}(I_i \cup \{v_j\}) - \mathcal{D}(I_i)],
\end{align*}
\]

(3)

where \( (a) \) follows from Lemma 1, and \( (b) \) follows from Theorem 1. Define \( \hat{\mathcal{D}}_{i,v}, \mathcal{D}_{i,v} \) as the first and second calls of subroutine in \( i \)-th step for variable \( v \); respectively. By the assumption of the theorem we have

\[
\mathcal{D}(I_i \cup \{v_j\}) - \mathcal{D}(I_i \cup \{v_j\}) < \hat{\mathcal{D}}_{i,v_j}(I_i \cup \{v_j\}),
\]

with probability larger than \( 1 - \delta \). Therefore

\[
\mathcal{D}(I_i \cup \{v_j\}) < \hat{\mathcal{D}}_{i,v_j}(I_i \cup \{v_j\}) + \mathcal{E}(I_i),
\]

with probability larger than \( 1 - \delta \). Similarly

\[
\Rightarrow -\mathcal{D}(I_i) < -\hat{\mathcal{D}}_{i,v_j}(I_i) + \mathcal{E}(I_i) \quad w.p. \quad > 1 - \delta,
\]

where \( \mathcal{E}(I_i) \) is the error term.
We run the algorithm for \( k \) iterations. In each iteration, we execute the function \( \hat{D}(\cdot) \) using Subroutine 1 for at most \( n \) vertices. Furthermore, in this subroutine, we generate \( N \) random DAGs by calling the function \textsc{RandEdge}, where in \cite{ghassami2018} it is shown that the complexity of each call is \( O(n^2) \). Hence, the computational complexity of the algorithm is \( O(knN \times n^2) \).

### F. Proof of Lemma 3

We require the following lemma for the proof:

**Lemma 4.** A chordal graph has a directed cycle only if it has a directed cycle of size 3.

**Proof.** If the directed cycle is of size 3 itself, the claim is trivial. Suppose the cycle \( C_n \) is of size \( n > 3 \). Relabel the vertices of \( C_n \) to have \( C_n = (v_1, \ldots, v_n, v_1) \). Since the graph is chordal, \( C_n \) has a chord and hence we have a triangle on vertices \( \{v_i, v_{i+1}, v_{i+2}\} \) for some \( i \). If the direction of \( \{v_i, v_{i+2}\} \) is \( \{v_i, v_{i+2}\} \), we have the directed cycle \( (v_i, v_{i+1}, v_{i+2}, v_i) \) which is of size 3. Otherwise, we have the directed cycle \( C_{n-1} = (v_1, \ldots, v_{i+2}, v_{i+3}, \ldots, v_n, v_1) \) on \( n-1 \) vertices. Relabeling the vertices from 1 to \( n-1 \) and repeating the above reasoning concludes the lemma.

**Proof of Lemma 3.** All the components in the undirected subgraph of \( \text{Ess}(G^*) \) are chordal \cite{hauser2012}. Therefore, by Lemma 4, to insure that a generated directed graph is a DAG, it suffices to make sure that it does not have any directed cycles of length 3, which is one of the checks that we do in the proposed procedure. For checking if the generated DAG is in the same Markov equivalence class as \( G^* \), it suffices to check if they have the same set of v-structures \cite{verma1991}, which is the other check that we do in the proposed procedure.

### E. Proof of Theorem 4

We run the algorithm for \( k \) iterations. In each iteration, we execute the function \( \hat{D}(\cdot) \) using Subroutine 1 for at most \( n \) vertices. Furthermore, in this subroutine, we generate \( N \) random DAGs by calling the function \textsc{RandEdge}, where in \cite{ghassami2018} it is shown that the complexity of each call is \( O(n^2) \). Hence, the computational complexity of the algorithm is \( O(knN \times n^2) \).