Do-calculus when the True Graph Is Unknown

Antti Hyttinen
HIIT, Dept. Computer Science
University of Helsinki
Finland

Frederick Eberhardt
Humanities and Social Sciences
California Institute of Technology
Pasadena, CA, USA

Matti Järvisalo
HIIT, Dept. Computer Science
University of Helsinki
Finland

Abstract

One of the basic tasks of causal discovery is to estimate the causal effect of some set of variables on another given a statistical data set. In this article we bridge the gap between causal structure discovery and the *do*-calculus by proposing a method for the identification of causal effects on the basis of arbitrary (equivalence) classes of semi-Markovian causal models. The approach uses a general logical representation of the equivalence class of graphs obtained from a causal structure discovery algorithm, the properties of which can then be queried by procedures implementing the *do*-calculus inference for causal effects. We show that the method is more efficient than determining causal effects using a naive enumeration of graphs in the equivalence class. Moreover, the method is complete with respect to the identifiability of causal effects for settings, in which extant methods that do not require knowledge of the true graph, offer only incomplete results. The method is entirely modular and easily adapted for different background settings.

1 INTRODUCTION

The theory of causal learning is aimed at finding ways to estimate causal effects in a variety of different settings. In the most basic setting the starting point is a statistical data set of measurements over the variables of interest. In this article we explore how quantitative causal effects can be estimated from such data alone, that is, without the additional knowledge of the causal structure.

When the true causal structure (the causal graph) *is* known, the well-known *do*-calculus enables the complete inference (i.e., the identification) of causal effects from the passive observational distribution over the variables (Pearl, 2000; Shpitser and Pearl, 2006b). However, full knowledge of the true graph requires a rather extensive understanding of the system under investigation. Data alone is in general insufficient to uniquely determine the true causal graph. Even complete discovery methods will usually leave the graph underdetermined (Spirtes et al., 1993).

Here we develop a general method for the combined task of causal structure discovery and the inference about causal effects.[1] Leveraging a constraint satisfaction approach to connect the output of causal discovery algorithms to the *do*-calculus, our method enables the identification of causal effects for arbitrary (equivalence) classes of semi-Markovian causal models (DAGs with latent variables). The primary advantages of the approach are that (i) it does not assume a unique true causal structure, (ii) it is not restricted to particular types of equivalence classes of causal structures, such as partial ancestral graphs (PAGs), (iii) it provides an algorithm that outputs (when possible) at least one estimator of the causal effect, rather than only specifying rules of a calculus, (iv) it considers all *do*-calculus inferences, not just e.g. the so-called backdoor conditions, and (v) it gives the user flexibility e.g. in how statistical conflicts in the data are handled and how the possibility of multiple estimators of a causal effect is addressed. To simplify notation we will, throughout this article, describe our method using a single observational data set as input. However, we emphasize that the method is extremely general, so, where relevant, we will indicate how the approach is extended or adapted to other scenarios.

Figure 1 gives an overview of the computational flow of the proposed method. The method determines whether a causal effect of the form $P(y \mid do(x), w)$ is identified given a data set as input, and if so, provides a numerical estimate. We use a (complete) causal discovery method to extract from the data as much information as possible about the true causal graph in terms of so-called *d-separation/connection constraints*. We encode these constraints in the language of propositional logic for the constraint solving component, thereby *implicitly* representing the equivalence class

[1] See Section 5 for a discussion of related approaches.

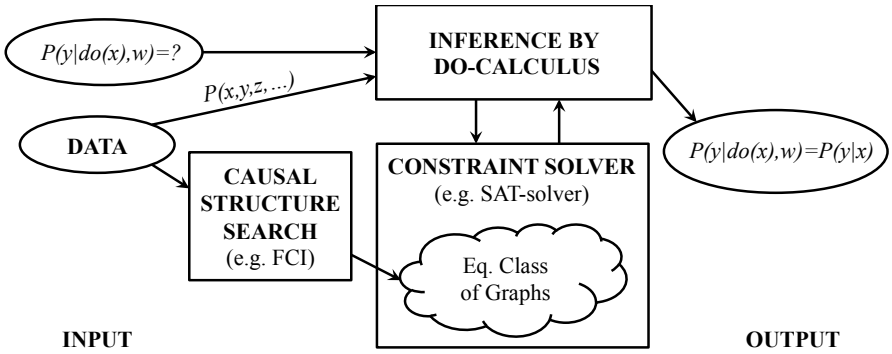

Figure 1: Overall structure of the system.

of causal structures (Hyttinen et al., 2013, 2014). This enables the application of modern constraint solving techniques, such as Boolean satisfiability (SAT) solvers (Biere et al., 2009). We can then use the do-calculus to determine whether the causal effect $P(y \mid do(x), w)$ is identified by alternating between do-calculus inferences and constraining the equivalence class to graphs for which no estimator of the causal effect has been found yet. If the causal effect is identifiable, we obtain a formula and a numerical estimate of the causal effect from the joint probability distribution over the variables.

Our method enables considerable flexibility in addressing the identification problem: The representation of the candidate causal structures in terms of a logical formula frees us from the restriction to settings where standard graphical representations of equivalence classes of causal graphs apply. We can include a wide variety of background constraints or additional knowledge, e.g., from experiments. We can leverage the full inferential power of the do-calculus without having to explicitly enumerate every causal structure consistent with the data. Nevertheless, we can (and do) instantiate an implicit exhaustive search that ensures that we preserve the completeness guarantees of *both* the causal discovery procedure and the do-calculus.

The paper is structured as follows: In Section 2 we describe the model space and assumptions used and give a concise problem statement. Section 3 explains the inference algorithms, whose completeness properties are discussed in Section 4. In Section 5 we describe known results and approaches that are closely related to ours or that provide context. Section 6 provides simulated results illustrating our main points.

2 PROBLEM SETUP

Following the standard set-up of the do-calculus, we assume that the causal system can be represented by a semi-Markovian causal model (SMCM). In other words, the underlying causal structure over a set of causal variables $\mathbf{V}$ is described by a directed acyclic graph G, in which the directed edges correspond to direct causal relations between the variables (relative to $\mathbf{V}$), and confounding of any two observed variables by some unobserved common cause U is represented by a bi-directed edge between the variables (thereby omitting U for simplicity in the graph; see Figure 2). The causal structure gives rise to a probability distribution that is assumed to be Markov and faithful to the graph. No further parametric assumption about the distribution is made. Importantly, Markov and faithfulness ensure that the probabilistic (in)dependencies of the distribution correspond to *d-separation/connection* relations in the causal graph.[2]

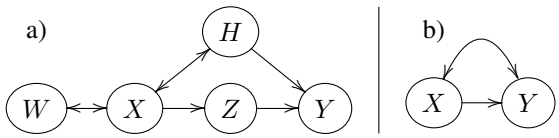

Figure 2: a) Example of a SMCM graph for which the causal effect $P(y|do(x))$ is identifiable even when the graph is unknown. b) Example of a SMCM graph for which the causal effect $P(y|do(x))$ is not identifiable even when the graph is known (since the graph includes a *hedge*).

Under specific d-separation conditions on the underlying causal structure, the rules of the do-calculus (see Figure 3) license inferences between the passive observational distribution $P(\mathbf{V})$ and the corresponding (conditional) interventional distributions $P(y \mid w, do(x))$, where one or more variables $x \subset \mathbf{V}$ have been subject to intervention and variables $w \subset \mathbf{V}$ are conditioned on. By using an additional exogenous intervention variable I_X with $I_X \to X$ for each variable X, the d-separation conditions of the do-calculus can be stated as in Figure 3 (see Pearl (1995, p. 686) for the

[2]See Spirtes et al. (1993) for a precise statement of the assumptions and for a definition of d-separation.

Rule 1 (Insertion/deletion of observations):

$$P(y|do(x), z, w) = P(y|do(x), w) \text{ if } Y \perp\!\!\!\perp Z|X, W||X$$

Rule 2 (Action/observation exchange):

$$P(y|do(x), do(z), w) = P(y|do(x), z, w) \text{ if } Y \perp\!\!\!\perp I_Z|X, Z, W||X$$

Rule 3 (Insertion/deletion of actions):

$$P(y|do(x), do(z), w) = P(y|do(x), w) \text{ if } Y \perp\!\!\!\perp I_Z|X, W||X$$

Rule 4 (Marginalization/sum-rule):

$$P(y|do(x), w) = \sum_z P(y, z|do(x), w)$$

Rule 5 (Conditioning):

$$P(y|do(x), z, w) = \frac{P(y, z|do(x), w)}{\sum_y P(y, z|do(x), w)}$$

Rule 6 (product/chain-rule):

$$P(y, z|do(x), w) = P(y|do(x), w, z)P(z|do(x), w)$$

Figure 3: Rules of the do-calculus.

proof of equivalence to the standard conditions; see also Spirtes et al. (1993, p. 79)). The d-separation conditions of each rule have the general form of '$Y \perp\!\!\!\perp Z \mid X, W \mid\mid X$', where W, X, Y, Z are disjoint sets of variables in the graph (including intervention variables), and '$\mid\mid X$' denotes an intervention on X: any edges with arrowheads into the variables in X are cut.

Given a graph G, the identifiability of a causal effect is now defined as follows (see Pearl (2000), Def. 3.2.4, p. 77): a causal effect is identifiable if and only if it can be uniquely computed from G and any positive input distribution $P()$ that is Markov to G, i.e., there are no two causal models with structure G that are Markov to $P()$ but have different numerical values for the causal effect.

An algorithm to apply the *do*-calculus when the true graph is known was developed by Tian and Pearl (2002), which with some modifications was shown to be complete for the identification of (conditional) causal effects by Shpitser and Pearl (2006b) (see also Huang and Valtorta (2006)). The *Shpitser algorithm* provides (given the graph and the observational distribution) one estimator of the causal effect, if such an estimator exists. When the causal effect is non-identifiable, it returns a feature of the graph, known as a *hedge*, that proves non-identifiability (see Figure 2b for an example; see Shpitser and Pearl (2006b) for the exact definition).

Since we do not assume that the true graph G is known, we take a causal effect to be identifiable given the equivalence class of causal structures deemed consistent with the input data if and only if the causal effect is (Pearl-) identifiable by the same estimator for each member of the equivalence class.

As shown in the simulations in Section 6, the causal effect is very often not uniquely identifiable from data when the true graph is unknown. So, instead of outputting only whether an effect is identifiable, and the estimate if it is, we follow Maathuis et al. (2009), who output a (multi)set of causal effect estimates that in some cases can be used to obtain bounds on the true causal effect. This leads to the following problem statement.

Problem Statement

INPUT: Data set D generated from an SMCM over variables $\mathbf{V}$ and a query about a causal effect $P(y \mid do(x), w)$.

TASK: Output a set of causal effect estimates S such that it includes an estimate for $P(y \mid do(x), w)$ for any causal structure that is consistent with D. Include 'NA' in S, if the causal effect is not identifiable for some causal structure consistent with D.

3 THE APPROACH

We proceed by describing the main contribution of this work: a general method for the estimation of causal effects. In the following, we will specify the main components (recall Figure 1) of our approach. First, we give details on how we connect the causal structure discovery algorithm of choice with the constraint solving component that maintains a logical representation of the equivalence class of models under consideration. Then, we describe the do-calculus inference component and its iterative interactions with the constraint solver.

3.1 Querying the Equivalence Class

We use (for purposes of illustration) the FCI-algorithm (Spirtes et al., 1993) to determine the equivalence class of candidate causal structures from the data set. The FCI-algorithm considers the same class of causal models as the *do*-calculus: acyclic causal structures with latent variables. It is complete with respect to knowledge about the underlying causal structure that can be obtained from conditional independence tests. Most importantly, FCI achieves this d-separation completeness while performing very few redundant tests. It thus also lends itself to the efficient characterization of the equivalence class in terms of a small set of d-separation constraints that can be fed to the constraint solver.

Algorithm 1 Do-calculus Inference.

Input: $P(y|do(x), w)$, an equivalence class E of SMCM graphs.

Initialize the set S of causal effect estimates as empty.

While E is nonempty:

Find a graph G from the equivalence class E.

Find a formula F by calling Shpitser's algorithm for graph G. If the algorithm does not find a formula but returns a hedge H, restrict eq. class E not to include H, add NA to S and continue the loop from the beginning.

Find a derivation for F by calling Algorithm 2 for graph G and the input distributions used in F.

Using F and the (estimated) $P(\mathbf{V})$, compute the estimate for the causal effect and add it to S.

Restrict the eq. class E to not satisfy at least one of the required d-separations in derivation D.

Return a set S of numerical causal effect estimates.

Algorithm 2 Do-calculus Derivation.

Input: $P(y|do(x), w)$, a SMCM graph G, and a set of distributions $P = \{P_1, \ldots\}$.

For each P_i in P:

Derive the distributions computable from P_i using the rules of the do-calculus such that:

- The required d-separation conditions are satisfied by G.
- All variables appearing in the derived distributions are ancestors of $Y \cup W$ (see Shpitser and Pearl (2006a)).
- For an application of the product rule, both required distributions are in P.

Add the new distributions to P and record the used rules and the required d-separations.

If $P(y|do(x), w)$ was derived, return the formula, the rules, and the d-separations used on the way.

Return "the effect is not identifiable".

We translate the d-separation constraints of the equivalence class into a logical representation using the ASP-encoding of Hyttinen et al. (2014). We can then query the constraint solver to obtain graphs from the equivalence class or to check whether any graphical conditions, such as d-separations or ancestral relations, apply to all, some or none of the members in the equivalence class, and we can further restrict the equivalence class with additional constraints.

For other settings or different background assumptions, the FCI algorithm can be substituted with any other structure discovery method. If one has reason to think that there are no latent variables, we would recommend using an exact Bayesian search algorithm, or the PC- or GES-algorithms if something more scalable is required (Spirtes et al., 1993; Chickering, 2002). If the causal constraints are to be obtained from heterogeneous data sets, possibly including experimental data or background knowledge, then a search algorithm such as GIES, IOD or a SAT-based procedure may be better (Hauser and Bühlmann, 2012; Tillman and Spirtes, 2011; Triantafillou and Tsamardinos, 2014; Hyttinen et al., 2014). The overall completeness of our method depends in part, of course, on whether the causal discovery method is complete.

3.2 Identifying Causal Effects

Algorithm 1 instantiates the *do*-calculus inference on a given equivalence class. It queries the constraint solver for a graph G in the equivalence class (one truth-value assignment to the logical formula) and calls Shpitser's algorithm on G to identify the desired causal effect.

If Shpitser's algorithm fails to identify the effect and returns a hedge H, then the causal effect is non-identifiable for G. Consequently, Algorithm 1 adds 'NA' to the set S of causal effect estimates to mark the non-identifiability. In addition, we restrict the equivalence class to not include any graphs that have the hedge H, as the causal effect is unidentifiable for such graphs as well.

If Shpitser's algorithm returns a formula F, then the causal effect is identifiable on the basis of the (marginal conditional) distributions $\{P_1, \ldots\}$ used in F. Algorithm 2 is then called with the causal effect query, the graph G and the list of distributions $\{P_1, \ldots\}$ to obtain a 'derivation' for the formula. This derivation specifies the rules of the do-calculus used to derive the formula F and consequently the set of d-separation constraints C that warrant the use of this estimate. The numerical estimate is added to S. This estimate is now valid for all graphs that satisfy the d-separation constraints in C. Then, the equivalence class is again restricted to disregard such graphs by ensuring that at least one of the constraints in C is no longer satisfied, i.e., we add the negation of the conjunction of constraints in C to the constraint solver. We repeatedly solve for a new graph G from the restricted equivalence class until the class becomes empty, at which time we have the solution to the Problem Statement. Note that the repeated restrictions of the equivalence class avoid an explicit enumeration of all the members of the equivalence class, allowing for faster operation.

Algorithm 2 implements the search for a valid do-calculus derivation for a formula. We need such a derivation as Shpitser's algorithm does not output the set of d-separations needed for the validity of the formula. Algorithm 2 does an exhaustive breadth-first-search, producing computable distributions that are warranted by the input graph and do-calculus. It stops when a derivation for the causal effect is found. The algorithm can be made sufficiently efficient because 1) we only input the distributions that are used in

the formula given by Shpitser's algorithm, 2) we use the fact that only variables that are ancestors of $Y \cup W$ can be helpful in determining the causal effect (Shpitser and Pearl, 2006b), and 3) causal structures usually permit the identification of fairly few distributions.

Figure 9 at the end of the paper shows an example run of Algorithm 1 for one particular equivalence class.

4 COMPLETENESS RESULTS

The completeness properties of our method are derivative of the completeness properties of the causal discovery algorithm of choice and the *do*-calculus. Given a complete structure search algorithm, such as FCI, we obtain in the large sample limit the Markov equivalence class of the true causal structure. If the causal effect in this equivalence class is non-identifiable, it is either because one graph in the equivalence class contains a *hedge*, which proves non-identifiability, or because there are two graphs which have different estimators for the causal effect. Algorithm 1 repeatedly performs the (complete) Shpitser algorithm on members of the equivalence class, each time restricting the equivalence class to graphs for which the discovered derivations of an estimator do not hold. Consequently, its output must eventually identify a graph with a hedge if there is one, since implicitly the entire set of graphs in the equivalence class is enumerated. If no graph with a hedge is found, then Algorithm 1 only terminates once there is a derivation of an estimator of the causal effect for each graph in the equivalence class. A check whether these formulas are the same determines the identifiability. (In the presented version of the algorithms we only output the numerical estimates, but the formulas could easily be added to avoid any formal concern that there could be coincidentally identical numerical values of the causal effect derived from two different estimators.)

We note that the FCI algorithm is not complete with regard to so-called *Verma constraints* that can further restrict the equivalence class of causal structures (Shpitser and Pearl, 2008). We are not aware of any search algorithm that is complete in this regard. However, any specific Verma constraint that may be established for a particular case, can easily be included and the set of estimators our method returns will be complete with regard to that additional constraint.

For settings involving multiple data sets or experimental data sets, there exist d-separation complete structure search algorithms, but it is not known whether the *do*-calculus is complete for these settings. Certainly, Shpitser's algorithm is restricted to the passive observational distribution. For the restricted experimental settings described in Bareinboim and Pearl (2012), we could replace Shpitser's algorithm with Bareinboim's method in Algorithm 1 and retain completeness, since Bareinboim shows completeness for the identification problem in these so-called "surrogate experiments". In the future, we hope our approach can aid the identification of causal effects from multiple data sets of non-identical populations (Bareinboim and Pearl, 2013).

5 RELATED WORK

Building the connection between causal structure discovery and causal effect inference seems essential if one wants to complete the aim of causal learning from data sets to causal effects. We consider it all the more important given that significant parts of the causal literature regard the problem of identifying the causal effect *given* the causal structure as entirely separate from the problem of discovering the causal structure in the first place. For example, the entire literature on algorithms applying the *do*-calculus assumes — generally without further discussion — that the causal graph is known (Tian and Pearl, 2002; Huang and Valtorta, 2006; Shpitser and Pearl, 2006b; Bareinboim and Pearl, 2012). In the general model space that the *do*-calculus allows for, the causal structure can hardly ever be uniquely determined from the passive observational distribution or even from the experimental distributions that Bareinboim and Pearl (2012) consider. Still, the algorithms rely on being able to check complicated features of the causal structure. Similarly, the methods of causal structure discovery often remain silent on how exactly one should determine the causal effects given their output equivalence class.[3] There need not be any harm in this division of labor if there is an obvious and satisfactory answer of how to connect the two. We assume that the standard proposal would be to take the equivalence class of causal structures output by a search algorithm, enumerate each member of the equivalence class, and perform the *do*-calculus algorithm on each member to determine (the identifiability of) the causal effect. Effectively, this is what is done in the IDA-algorithm which returns (multi-sets of) estimates of the causal effects of variables from an equivalence class of causal structures under the assumption that there are no latent variables (Maathuis et al., 2009). However, such an explicit enumeration of the members of an equivalence class can very quickly become unwieldy (see also (Malinsky, 2015)). In our simulations (Section 6) we show that our approach is more efficient than a naive enumeration combined with the *do*-calculus inference.

Zhang (2008), instead, developed a *do*-calculus directly for the equivalence classes represented by partial ancestral graphs (PAGs; see also Richardson and Spirtes (2003)). As he explains, the calculus is not complete and no inference algorithm to apply the calculus is given (although Zhang notes that his results could be used to improve on the ear-

[3]The same is not true for methods of causal discovery that include a parametric assumption (e.g. linearity, additive noise, non-Gaussianity, etc), since in these cases the identification of the qualitative causal effect generally corresponds to providing a quantitative estimate of it.

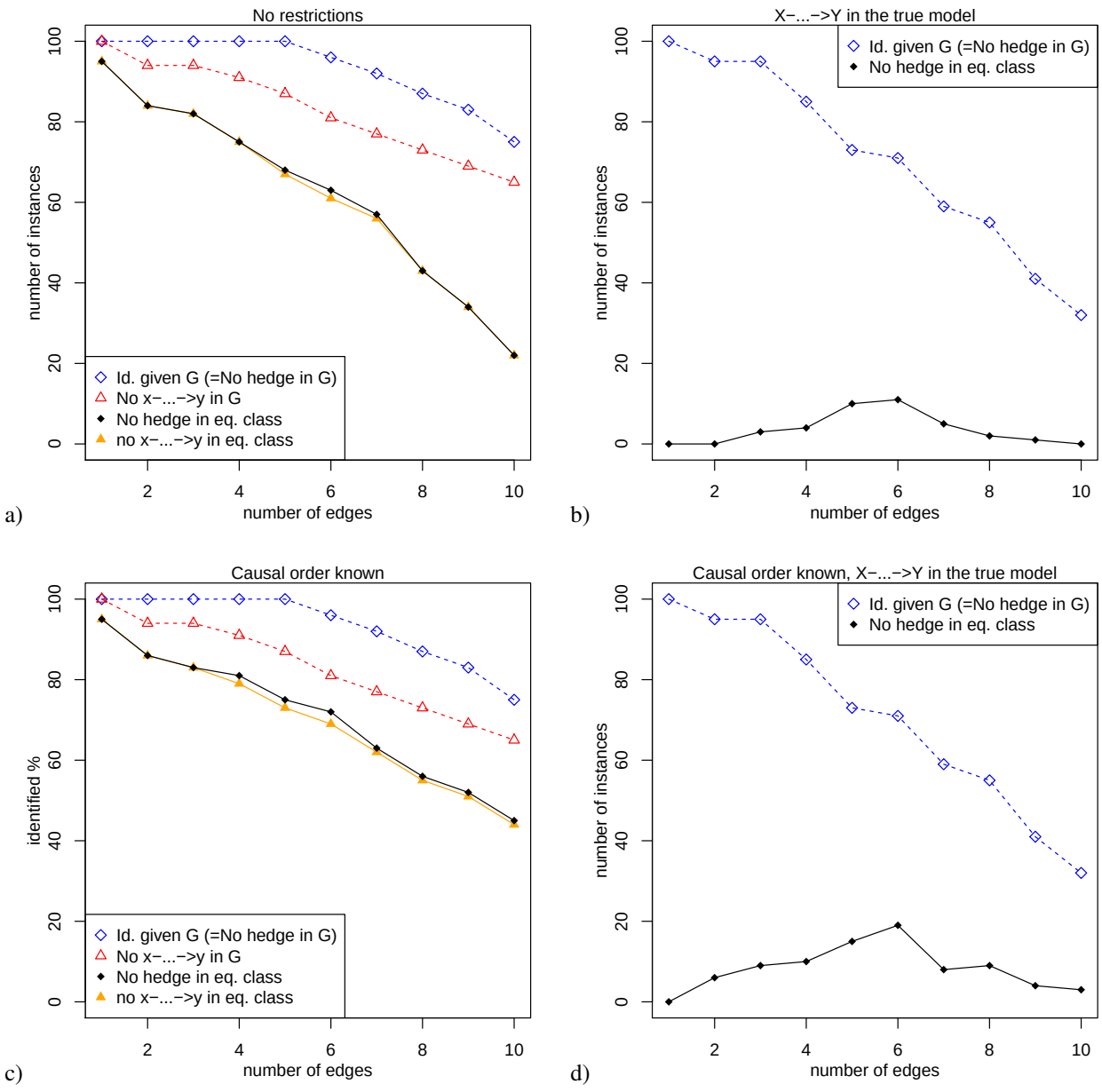

Figure 4: Identifiability of the causal effect $P(y|do(x))$ over random 5-variable graphs.

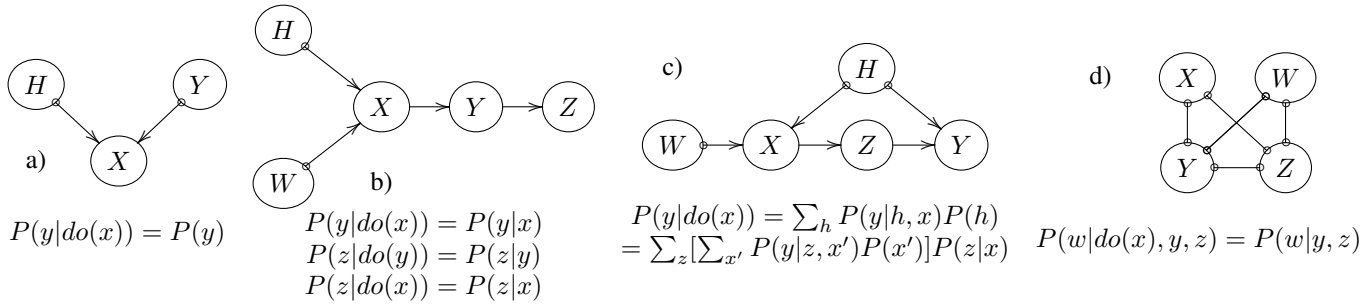

Figure 5: Some examples of graph equivalence classes where causal effects can be identified. The circles on the edges can be edge heads or tails (or both), as long as no new unshielded colliders are formed.

lier Prediction Algorithm by Spirtes et al. (1993)). His approach is closely related to ours, with the main difference that we do not restrict ourselves to equivalence classes of SMCMs that are PAGs and that our procedure is complete for his setting.

Also relying on PAGs, Maathuis and Colombo (2015) focus on a subset of Zhang's invariance principles (which themselves are implied by the *do*-calculus) in order to specify conditions on a PAG that allow for the identification of an *adjustment set*, i.e., a set of variables W that block all so-called "backdoor paths" between X and Y, and allow for the causal effect $P(y \mid do(x))$ to be estimated using $P(y \mid w, x)$. In our approach their "generalized backdoor criterion" corresponds to the simple and intuitive d-separation conditions that are required for the estimation of the causal effect ($w \perp\!\!\!\perp I_x$ and $y \perp\!\!\!\perp I_x \mid x, w$). These we can directly query on *any* class of SMCMs, not just on PAGs (or MAGs or DAGs).

More data-driven methods to find adjustment sets have also been developed (De Luna et al., 2011; Entner et al., 2013; VanderWeele and Shpitser, 2011). These do not rely explicitly on a graphical representation of the causal knowledge, but specify independence conditions that can be directly checked in the data. This is a very attractive direction of research, since it cuts out the graph from the inference procedure altogether. However, extant methods rely on a variety of general background assumptions about how the causal variables may be related (e.g. order assumptions) that we do not require. More generally, methods for the identification of adjustment sets obviously do not exhaust the identifiability conditions for causal effects for which the *do*-calculus was shown to be complete. Most prominently, the so-called "front-door" criterion for identifiability is not considered. Thus, one of the contributions of our method is to enable the full identification power of the *do*-calculus in settings when the causal structure is underdetermined.

6 SIMULATIONS

These simulations explore the identifiability of the causal effects when the true causal graph is unknown, the scalability of the methods presented in this paper, and, finally, the accuracy of different causal effect estimates when multiple estimators can be calculated. We implemented the algorithms using `R` with various packages (Tikka, 2014; Kalisch et al., 2012). Following Hyttinen et al. (2014) for the implementation of the constraint solving component, we employed the off-the-shelf state-of-the-art answer set programming (ASP) solver Clingo version 4.4.0, which at its core uses modern SAT solving techniques to perform a complete search for solutions, and at the same time allows for a natural high-level representation of the structural constraints in logical form (Gebser et al., 2011).

Figure 4 compares the identifiability of the basic causal effect $P(y|do(x))$ when the graph is known vs. when only its equivalence class is known. Figure 4a shows the number of identified causal effects in random 5-variable graphs without any restrictions, as density increases. When only the equivalence class is known, the causal effect is almost always either trivially identified as X is discovered not to be an ancestor of Y in any member of the equivalence class, or trivially unidentifiable due to a possible hedge in the equivalence class. When the graph is known, a significant number of causal effects are identified even when X is an ancestor of Y. This is further highlighted in Figure 4b where X is required to be an ancestor of Y in the true graph. Almost no causal effects are identified when only the equivalence class is known. Only in very rare cases can one orient enough edges to deduce the absence of hedges from the equivalence class. This seems to happen when 4-7 edges are present; additional edges often prevent the determination of the orientation. Figure 4c and 4d consider the same comparison with the modification that the true causal order is known by the causal discovery algorithm, and thus fixed in the equivalence class. A few more effects are identified, but since the fixed causal order does not prevent bidirected edges, the improvement on non-trivial instances is limited in Figure 4d.

Figure 5 shows equivalence classes of graphs for which the causal effect is identified even when the true graph is unknown. In each case, we are able to deduce enough edge orientations to prevent the presence of a hedge, and to fix the orientation of the paths from X to Y. Note that Figure 5d shows an example for which the *do*-calculus formulation of Zhang (2008) over PAGs is incomplete. Our approach is complete here and can hence identify the conditional causal effect.

We also compared the running times of Algorithm 1 against

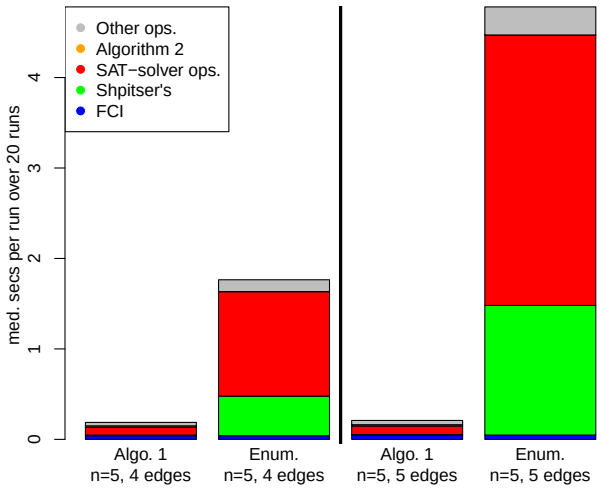

Figure 6: Algorithm 1 vs. Enumeration approach.

trivially enumerating all graphs in the equivalence class and running Shpitser's algorithm on all of them. Figure 6 shows the median times spent by the different parts of the algorithms. Algorithm 1 is much faster. In addition to the time spent on enumerating the whole equivalence class by the ASP constraint solver used here, running Shpitser's algorithm on so many graphs also takes a considerable amount of time.

Figure 7 shows the median of the time spent during the different operations of Algorithm 1 on larger instances. FCI was run using an independence oracle. Algorithm 1 spends the majority of its time finding graphs for which the formulas obtained in previous iteration rounds are not warranted. For some outlier instances not visible in the median here, Algorithm 2 also needs a considerable amount of time when finding the derivation for a particularly complicated formula. Note that we undergo here a rather heavy task of finding estimates for all graphs in the equivalence class. If we were content to just decide whether the effect is identifiable, the total running times would be considerably lower. However, as shown in Figure 4, the results of that kind of an algorithm would be quite uninformative.

Finally, we examined the benefits of finding more estimators for the causal effect using Algorithm 2 in cases where multiple different estimators exist. We drew random parameters for a binary SMCM with the graph in Figure 2. Given the equivalence class (shown in Figure 5c), the causal effect $P(y|do(x = 0))$ can be calculated from the passively observed distribution $P(h, w, x, y, z)$ either by the backdoor formula adjusting for h, or by the front-door formula relative to z (see Figure 5c). We also estimated the causal effect by directly sampling from the model when x is surgically fixed to 0. Figure 8 shows the average KL-divergence of the different estimators. The distributions

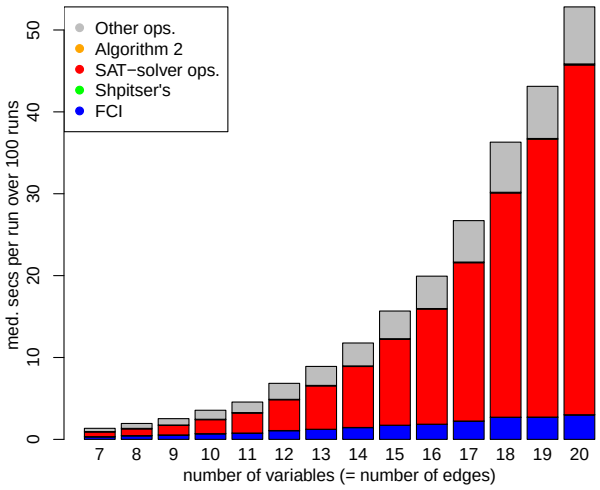

Figure 7: Time spent by Algorithm 1.

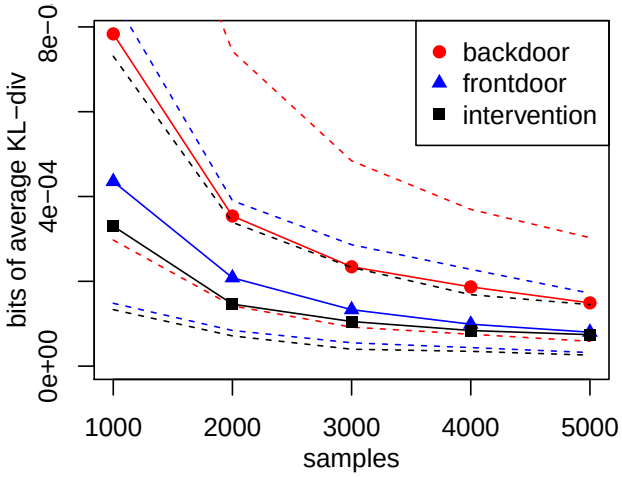

Figure 8: Average KL-divergence for different estimates of $P(y|do(x))$ for the equivalence class of Figure 2. Median, 33% and 66% quantiles are plotted.

needed for the estimators were estimated directly (with regularization to avoid zero probabilities). In these simulations the front-door estimator seems to offer better accuracy than the backdoor estimator. By intervention we can obtain still higher accuracy than either of the estimates obtained from passively observed data. Shpitser's algorithm gives here only the backdoor formula. This simulation shows that it may be beneficial to consider many estimators of a causal effect instead of using only a single consistent formula.

7 CONCLUSION

In this work we explored the possibilities of estimating causal effects from data. We have considerably relaxed the assumption of the known true graph, which has been standard in the literature on the *do*-calculus. Although causal effects are rarely identified when the true graph is unknown, our approach can still generate informative output in terms of a set of estimates. Unlike other approaches that perform only a limited set of causal effect inferences, our method retains the completeness properties of the used causal discovery algorithm and the *do*-calculus inference. We hope that the flexible machinery presented in this paper can be used to obtain further graphical criteria for identifiability, and will help in achieving more completeness results.

Acknowledgements This work was supported in part by the Finnish Foundation for Technology Promotion TES (A.H.), and by the Academy of Finland (grants 251170 COIN Centre of Excellence in Computational Inference Research, 276412 and 284591) and Research Funds of the University of Helsinki (M.J.).

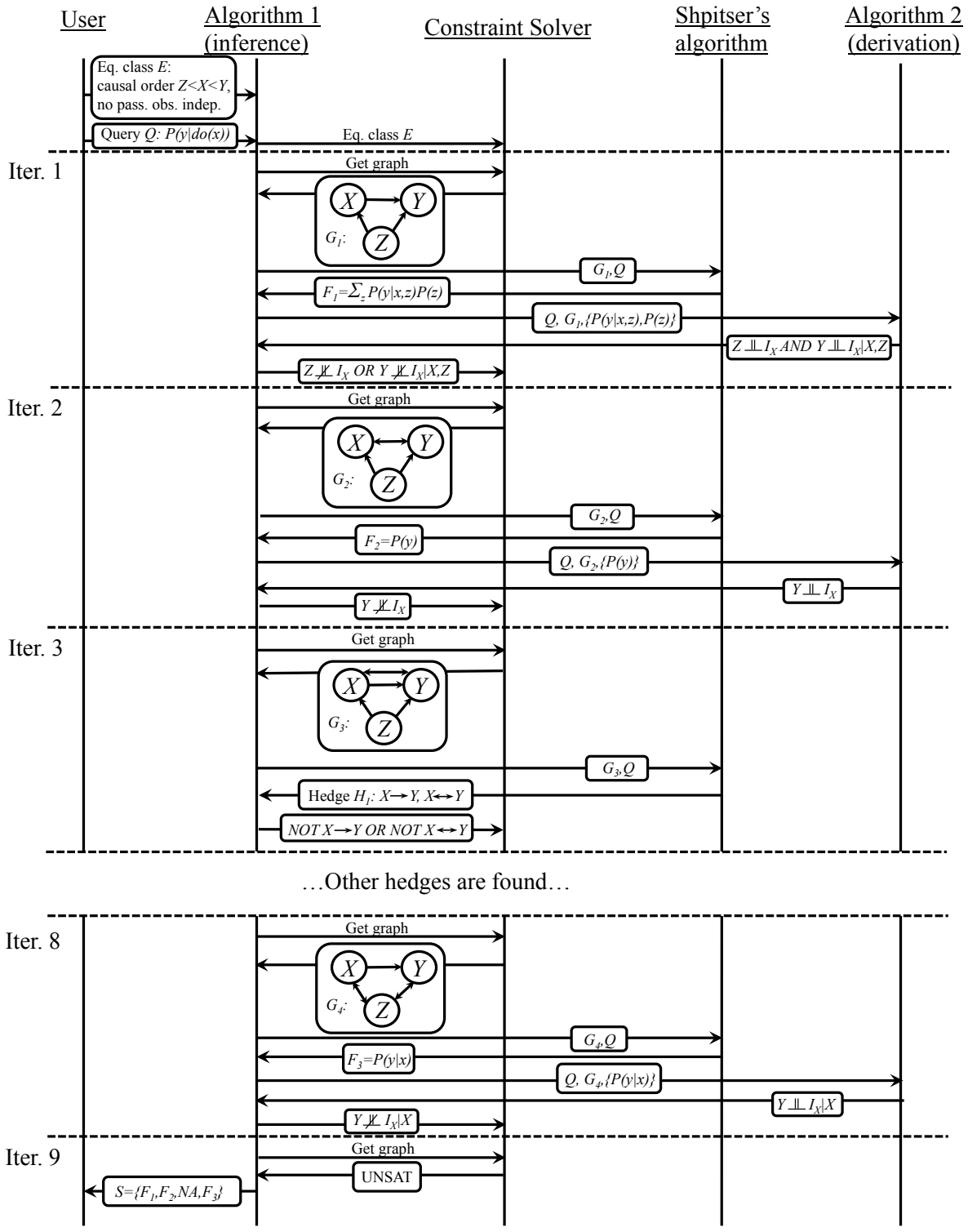

Figure 9: An example run of Algorithm 1 that estimates $P(y|do(x))$ given the input equivalence class consisting of all SMCM graphs with three variables, no passively observed independencies, and the causal order $Z < X < Y$.

References

Bareinboim, E. and Pearl, J. (2012). Causal inference by surrogate experiments: z-identifiability. In *Proc. UAI*, pages 113–120. AUAI Press.

Bareinboim, E. and Pearl, J. (2013). A general algorithm for deciding transportability of experimental results. *Journal of Causal Inference*, 1(1):107–134.

Biere, A., Heule, M., van Maaren, H., and Walsh, T., editors (2009). *Handbook of Satisfiability*, volume 185 of *Frontiers in Artificial Intelligence and Applications*. IOS Press.

Chickering, D. M. (2002). Optimal structure identification with greedy search. *Journal of Machine Learning Research*, 3:507–554.

De Luna, X., Waernbaum, I., and Richardson, T. S. (2011). Covariate selection for the nonparametric estimation of an average treatment effect. *Biometrika*, 98(4):861–875.

Entner, D., Hoyer, P., and Spirtes, P. (2013). Data-driven covariate selection for nonparametric estimation of causal effects. In *Proc. AISTATS*, volume 31 of *JMLR Workshop and Conference Proceedings*, pages 256–264. JMLR.org.

Gebser, M., Kaufmann, B., Kaminski, R., Ostrowski, M., Schaub, T., and Schneider, M. T. (2011). Potassco: The Potsdam answer set solving collection. *AI Communications*, 24(2):107–124.

Hauser, A. and Bühlmann, P. (2012). Characterization and greedy learning of interventional markov equivalence classes of directed acyclic graphs. *Journal of Machine Learning Research*, 13:2409–2464.

Huang, Y. and Valtorta, M. (2006). Identifiability in causal Bayesian networks: a sound and complete algorithm. In *Proc. AAAI*, pages 1149–1154. AAAI Press.

Hyttinen, A., Eberhardt, F., and Järvisalo, M. (2014). Constraint-based causal discovery: Conflict resolution with answer set programming. In *Proc. UAI*, pages 340–349. AUAI Press.

Hyttinen, A., Hoyer, P. O., Eberhardt, F., and Järvisalo, M. (2013). Discovering cyclic causal models with latent variables: A general SAT-based procedure. In *Proc. UAI*, pages 301–310. AUAI Press.

Kalisch, M., Mächler, M., Colombo, D., Maathuis, M. H., and Bühlmann, P. (2012). Causal inference using graphical models with the R package pcalg. *Journal of Statistical Software*, 47(11):1–26.

Maathuis, M. H. and Colombo, D. (2015). A generalized backdoor criterion. *Annals of Statistics*, pages 1060–1088.

Maathuis, M. H., Kalisch, M., Bühlmann, P., et al. (2009). Estimating high-dimensional intervention effects from observational data. *The Annals of Statistics*, 37(6A):3133–3164.

Malinsky, D. (2015). Estimating intervention effects in systems with unmeasured confounding. Technical report, Carnegie Mellon University.

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82(4):669–688.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press.

Richardson, T. and Spirtes, P. (2003). Causal inference via ancestral graph models. In Green, P., Hjort, N., and Richardson, S., editors, *Highly Structured Stochastic Systems*, pages 83–105. Oxford University Press.

Shpitser, I. and Pearl, J. (2006a). Identification of conditional interventional distributions. In *Proc. UAI*, pages 437–444. AUAI Press.

Shpitser, I. and Pearl, J. (2006b). Identification of joint interventional distributions in recursive semi-markovian causal models. In *Proc. AAAI*, pages 1219–1226. AAAI Press.

Shpitser, I. and Pearl, J. (2008). Dormant independence. In *Proc. AAAI*, pages 1081–1087. AAAI Press.

Spirtes, P., Glymour, C., and Scheines, R. (1993). *Causation, Prediction, and Search*. Springer-Verlag. (2nd ed. MIT Press 2000).

Tian, J. and Pearl, J. (2002). A general identification condition for causal effects. In *Proc. AAAI*, pages 567–573. AAAI Press.

Tikka, S. (2014). *Package 'causaleffect'*.

Tillman, R. E. and Spirtes, P. (2011). Learning equivalence classes of acyclic models with latent and selection variables from multiple datasets with overlapping variables. In *Proc. AISTATS*, volume 15 of *JMLR Proceedings*. JMLR.org.

Triantafillou, S. and Tsamardinos, I. (2014). Constraint-based causal discovery from multiple interventions over overlapping variable sets. arXiv:1403.2150.

VanderWeele, T. J. and Shpitser, I. (2011). A new criterion for confounder selection. *Biometrics*, 67(4):1406–1413.

Zhang, J. (2008). Causal reasoning with ancestral graphs. *The Journal of Machine Learning Research*, 9:1437–1474.