

## A General Method for Deriving Tight Symbolic Bounds on Causal Effects

Michael C. Sachs, Gustav Jonzon, Arvid Sjölander & Erin E. Gabriel

**To cite this article:** Michael C. Sachs, Gustav Jonzon, Arvid Sjölander & Erin E. Gabriel (2023) A General Method for Deriving Tight Symbolic Bounds on Causal Effects, Journal of Computational and Graphical Statistics, 32:2, 567-576, DOI: [10.1080/10618600.2022.2071905](https://doi.org/10.1080/10618600.2022.2071905)

**To link to this article:** <https://doi.org/10.1080/10618600.2022.2071905>



© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC.



[View supplementary material](#)



Published online: 23 May 2022.



[Submit your article to this journal](#)



Article views: 2683



[View related articles](#)



[View Crossmark data](#)



Citing articles: 9 [View citing articles](#)

# A General Method for Deriving Tight Symbolic Bounds on Causal Effects

Michael C. Sachs<sup>a,b</sup> , Gustav Jonzon<sup>a</sup>, Arvid Sjölander<sup>a</sup> , and Erin E. Gabriel<sup>a,b</sup> 

<sup>a</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; <sup>b</sup>Section of Biostatistics, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

## ABSTRACT

A causal query will commonly not be identifiable from observed data, in which case no estimator of the query can be contrived without further assumptions or measured variables, regardless of the amount or precision of the measurements of observed variables. However, it may still be possible to derive symbolic bounds on the query in terms of the distribution of observed variables. Bounds, numeric or symbolic, can often be more valuable than a statistical estimator derived under implausible assumptions. Symbolic bounds, however, provide a measure of uncertainty and information loss due to the lack of an identifiable estimand even in the absence of data. We develop and describe a general approach for computation of symbolic bounds and characterize a class of settings in which our method is guaranteed to provide tight valid bounds. This expands the known settings in which tight causal bounds are solutions to linear programs. We also prove that our method can provide valid and possibly informative symbolic bounds that are not guaranteed to be tight in a larger class of problems. We illustrate the use and interpretation of our algorithms in three examples in which we derive novel symbolic bounds. Supplementary materials for this article are available online.

## ARTICLE HISTORY

Received January 2022  
Accepted April 2022

## KEYWORDS

Causal bounds; Causal inference; Unmeasured confounding

## 1. Introduction

In many fields of research, a common goal is to determine causal relationships or mechanistic pathways. This investigation is often complicated by common causes of the outcome and either the exposure, or other variables of interest along the causal pathway from the exposure to the outcome, causing confounding. When common causes are unmeasured, the causal effect of interest is usually not identifiable. When the causal effect of interest, which we will refer to as a causal query, cannot be identified, one can derive bounds, that is, a range of possible values for this quantity in terms of the observed data distribution.

In general, arbitrarily wide bounds are trivial to derive, but not informative in the sense that they will not provide further insight into the magnitude of the effect. Deriving narrower bounds that are still valid, that is, containing all possible values of the true causal effect, can be a complicated task, and in particular, deriving tight bounds, that is, the narrowest possible given all and only explicit assumptions, may be highly nontrivial. An approach to deriving numeric tight bounds in quite general settings is given in Duarte et al. (2021). A drawback of the numeric approach, however, is the need for recomputation with each new dataset.

Computing bounds symbolically, that is, as closed form analytic expressions in terms of known observable quantities, rather than numerically, may provide useful information with which to draw conclusions about a study design or form of data collection in the absence of data, in addition to their transparent ease of

use in real data once derived. Symbolic tight bounds on a causal query thus, in many ways, provide us with an ideal summary of our effect of interest given our current state of knowledge and/or set of assumptions.

In 1994, in his Ph.D. dissertation, Alexander Balke gave a method for translating a certain type of causal theory, represented by a directed acyclic graph (DAG), and causal query into a constrained optimization problem in terms of unmeasured response function variables (Balke and Pearl 1994a, 1994b). The causal query is expressed in terms of the distribution of these variables and the DAG gives rise to linear relationships between this distribution and that of the observed variables. In conjunction with standard probabilistic constraints, this yields a bounded constrained optimization problem. If the problem is linear then a vertex enumeration algorithm can be used to find the global extrema of the causal query in terms of the true probability distribution of the observed variables (Dantzig 1963). Balke and Pearl (1994a) state that the resulting extrema give tight bounds for their causal query in the instrumental variable setting.

The linear programming method proposed by Balke in his thesis has been used to derive bounds for various causal parameters and scenarios. Balke and Pearl (1997) derived bounds for the causal risk difference in scenarios where the exposure and outcome are confounded and there is data on an unconfounded instrumental variable, and Cai, Kuroki, and Sato (2007) showed how these bounds can be made tighter by using information on measured covariates. Kaufman et al. (2005) used

linear programming to derive numeric bounds on the controlled direct effect in scenarios where the mediator is confounded but the exposure is not, and Cai et al. (2008) derived symbolic expressions for these bounds. Sjölander (2009) considered the same scenario, and derived bounds on the natural direct effects. Kaufman and MacLehose (2009) considered various causal effects in scenarios involving three measured variables, and showed how the linear programming bounds in these scenarios can be made tighter by certain monotonicity and no-interactions assumptions. Tian and Pearl (2000) derived bounds on the probability of causation, and Imai and Yamamoto (2010) derived bounds on the causal risk difference in the presence of differential measurement error. Sjölander et al. (2014b, 2014a) and derived bounds for causal interactions in scenarios with two unconfounded exposures. Gabriel, Sachs, and Sjölander (2020) derived bounds for the causal risk difference in observational studies with outcome dependent sampling, and Gabriel, Sjölander, and Sachs (2021) derived bounds for the causal risk difference in randomized controlled trials with both imperfect compliance and nonignorable missingness in the outcome. MacLehose et al. (2005) provided a gentle introduction to the linear programming method for epidemiologists.

Related theoretical results have been shown in specific settings that are extensions to the binary instrumental variable problem (Heckman and Vytlačil 2001; Ramsahai 2012; Bonet 2013). To the knowledge of the authors, there has been no attempt in the literature to characterize the set of causal problems that are always linear or an approach for determining whether a problem is linear, given its DAG and target query. In this article, we generalize and extend Balke and Pearl's approach for computation of bounds by characterizing a class of causal problems that always give rise to linear programs and describing a general algorithm for constructing the objective and constraints based on the DAG and query. The only input required is thus subject matter knowledge in the form of a causal DAG as well as an effect of interest.

In Section 2, we introduce the transformation of a causal DAG over categorical variables with unmeasured causal influences into an equivalent one where those influences have been discretized. In Section 3 we characterize a set of DAGs that have linear relations between the distributions of their observed variables and unobserved influences, along with an algorithm that extracts those relations from the DAG. Section 4 develops notation and requirements for general forms of causal queries that are linear in the distribution of the unmeasured discrete influences, and details an algorithm that constructs such relations from a complex causal query expressed in terms of potential outcomes and observable variables. Section 5 then states the final linear program, possible extensions of it and a suitable optimization method. Finally, Section 6 details a few interesting examples using this method. Proofs of the main propositions, notes on computational complexity of the algorithms, and the details of a simulation study are given in the [supplementary materials](#). The algorithms described herein are implemented in an R (R Core Team 2019) package called `causaloptim`, available on the Comprehensive R Archive Network (CRAN), with a user friendly interface.

## 2. Canonical Partitions

Let the set of observed variables be denoted  $\mathcal{W} = \{W_1, \dots, W_n\}$ , with corresponding vector  $\mathbf{W} = (W_1, \dots, W_n)$  and realized values represented by a vector  $\mathbf{w} = (w_1, \dots, w_n)$ . We assume that all of these variables are categorical, and constitute the vertices of a DAG, whose vertex set is thus  $\mathcal{W}$ . Each variable of interest  $W_i \in \mathcal{W}$  is affected by a set of unmeasured variables  $U_{W_i}$  as well as a subset of the remaining variables. Potential outcomes will be denoted using brackets; for example,  $W_1(W_2 = w_2)$ , and the probability that the variable  $W_1$  would have value  $w_1$ , if the variable  $W_2$  was intervened upon to have value  $w_2$  will be denoted as  $p\{W_1(W_2 = w_2) = w_1\}$ . For any given random variable  $X$ , we let  $v(X)$  denote its support, that is, for discrete variables, the set of all values it can take on with positive probability.

We are assuming the nonparametric structural equation framework, that is, for each  $W_i \in \mathcal{W}$ , we assume that there exists a function  $F_{W_i}$  such that  $w_i$ , the value of  $W_i$  is given by  $w_i = F_{W_i}(\mathbf{pa}_{W_i}, u_{W_i})$ , where  $\mathbf{pa}_{W_i}$  denotes the values of variables  $\mathbf{Pa}_{W_i}$  in  $\mathcal{W}$  that are parents of  $W_i$ , and  $u_{W_i}$  represents the values of  $U_{W_i}$ , the unmeasured causes of  $W_i$ . We make no assumptions about the unmeasured variables  $U_{W_i}$ ; they may, for example, be continuous or multivariate. Since all observed variables of interest in the graph are assumed to be categorical, we can, without loss of generality, recode the assumptions by defining a series of new categorical variables  $R_{W_i}$ , one for each variable  $W_i \in \mathcal{W}$ , which specifies how the value of  $W_i$  is determined from those of its parents.

For each  $W_i \in \mathcal{W}$ , we let  $R_{W_i}$  be the variable corresponding to the canonical partition of  $v(U_{W_i})$  into finite states with respect to the given causal DAG, as stated formally in [Proposition 1](#).

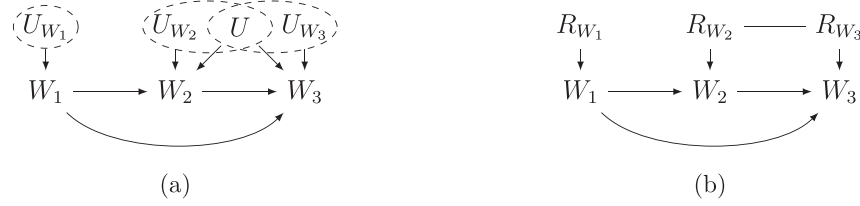
**Proposition 1 (Canonical partitions).** Let  $G$  be a causal DAG, let  $\mathcal{W} := V(G)$  be its vertices and suppose that  $\forall W \in \mathcal{W}, |v(W)| < \infty$  (i.e., each variable is categorical). Let  $\mathcal{D} := v(\mathbf{Pa}_W) \times v(R_W)$  if  $\mathbf{Pa}_W \neq \emptyset$  and  $v(R_W)$  otherwise. Then there exists a categorical variable  $R_W$  (so  $|v(R_W)| < \infty$ ) and a mapping  $f_W : \mathcal{D} \rightarrow v(W)$  such that for each value  $u_W \in v(U_W)$  there exists a unique value  $r_W \in v(R_W)$  for which  $F_W(\cdot, u_W) = f_W(\cdot, r_W)$ .

For proof, see Section S1 of the [supplementary materials](#). The function  $f_W : \mathcal{D} \rightarrow v(W)$  above is called the *response function* of  $W$ , and the variable  $R_W$  is called its *response function variable*.

Regardless of the characteristics of  $U_{W_i}$ , we have

$$|v(R_{W_i})| = |v(W_i)|^{|\mathbf{Pa}_{W_i}|} = |v(W_i)|^{\prod_{V \in \mathbf{Pa}_{W_i}} |v(V)|} < \infty,$$

since all variables in  $\mathcal{W}$  are assumed categorical. Let  $c_{W_i} := |v(W_i)|$ , so  $|v(R_{W_i})| = c_{W_i}^{|\mathbf{Pa}_{W_i}|}$ , where  $|\mathbf{Pa}_{W_i}| = \prod_{V \in \mathbf{Pa}_{W_i}} c_V$ , and note that without loss of generality we can assume that  $v(W_i) = \{0, \dots, c_{W_i} - 1\}$  and enumerate  $v(R_{W_i})$  as  $\{0, \dots, c_{W_i}^{|\mathbf{Pa}_{W_i}|} - 1\}$ . Let  $\mathbf{R} := (R_{W_1}, \dots, R_{W_n})$  and  $\aleph := |v(\mathbf{R})| = \prod_{i=1}^n |v(R_{W_i})| = \prod_{i=1}^n c_{W_i}^{|\mathbf{Pa}_{W_i}|}$ . The joint distribution of  $\mathbf{R}$  together with the response functions fully characterize the probabilistic causal model.



**Figure 1.** Example DAG to illustrate the concepts and notation. In this example, the measured variables are  $W_1$ ,  $W_2$ , and  $W_3$ , and the remaining are unmeasured. Since the measured variables are categorical, an equivalent representation of (a) is given in (b), where  $R_{W_1}$ ,  $R_{W_2}$ ,  $R_{W_3}$  are categorical response function variables.

For a given  $W_i \in \mathcal{W}$  and fixed  $\mathbf{r} \in \nu(\mathbf{R})$ , we define a procedure for determining its value  $w_i$  by recursively evaluating the corresponding functional expression. Using nested subscripts, we let  $W_{i1}, \dots, W_{ik_i}$  denote the parents of  $W_i$  that are in  $\mathcal{W}$ . Then  $w_i$ , the value of  $W_i$ , can be obtained by recursively evaluating

$$w_i = g_{W_i}^*(\mathbf{r}) := f_{W_i}(g_{W_{i1}}^*(\mathbf{r}), \dots, g_{W_{ik_i}}^*(\mathbf{r}), r_{W_i}).$$

Any set of observed probabilities can be related to the distribution of response function variables as follows:

$$\begin{aligned} p\{\mathbf{W} = \mathbf{w}\} &= p\{W_1 = w_1, \dots, W_n = w_n\} \\ &= \sum_{\mathbf{r} \in \nu(\mathbf{R}) : \forall i \in \{1, \dots, n\}, w_i = g_{W_i}^*(\mathbf{r})} p\{\mathbf{R} = \mathbf{r}\}. \end{aligned}$$

As an example, Figure 1 shows a simple setting with three binary variables of interest. Figure 1(a) shows the DAG for a model in which variables  $W_1$  and  $W_2$  both directly affect an outcome  $W_3$ , with  $W_1$  also directly affecting  $W_2$ . Figure 1(b) shows the equivalent DAG with response function variables in place of the original unmeasured variables. The variables that have an unmeasured common cause have response function variables that are dependent, as indicated by the dashed ellipse that outlines the unmeasured causal influences of  $W_2$  and  $W_3$ . Since they both contain  $U$ , the common cause, their response function variables are dependent as indicated by an undirected edge. We can encode  $R_{W_2}$  so the values 0, 1, 2, 3 of  $R_{W_2}$  correspond to the response patterns  $w_2 = f_{W_2}(pa_{W_2}=w_1, r_{W_2}=0) := 0$ ,  $w_2 = f_{W_2}(pa_{W_2}=w_1, r_{W_2}=1) := w_1$ ,  $w_2 = f_{W_2}(pa_{W_2}=w_1, r_{W_2}=2) := 1 - w_1$ ,  $w_2 = f_{W_2}(pa_{W_2}=w_1, r_{W_2}=3) := 1$ , respectively. We encode  $R_{W_1}$  taking values 0 and 1 according to  $w_1 = f_{W_1}(r_{W_1}=0) := 0$  and  $w_1 = f_{W_1}(r_{W_1}=1) := 1$ , respectively. Under the model shown in Figure 1(b), with for example,  $r_{W_1} = 0$ ,  $r_{W_2} = 1$ ,  $r_{W_3} = 3$ , we can evaluate the function to determine  $w_2$ :

$$\begin{aligned} w_2 &= g_{W_2}^*(\mathbf{r}=(0, 1, 3)) = f_{W_2}(f_{W_1}(r_{W_1}=0), r_{W_2}=1) \\ &= f_{W_2}(0, 1) = 0. \end{aligned}$$

For  $W_3$ , we need to enumerate the response patterns for each of the  $2^2$  possible combinations of values of  $(w_1, w_2)$ , that is,  $2^2 = 16$ . Then, to evaluate the probability  $p\{W_1 = 1, W_2 = 0, W_3 = 1\}$  in terms of  $\mathbf{R}$ , we can follow the same procedure as above for all  $2^1 \cdot 2^2 \cdot 2^4 = 128$  possible combinations of  $\mathbf{r}$ , keeping track of the resulting values  $\mathbf{w}$ . It can be shown that the variable value  $\mathbf{w} = (1, 0, 1)$  is consistent with 16 values of  $\mathbf{r}$ . Thus, the probability of this event is the sum over the set of these 16 values of the probability that  $\mathbf{R}$  equals them. See Balke

and Pearl (1994a) or Pearl (2009), Chapter 8 for another example and further interpretation.

Using this discretization, we can enumerate the relationships between the observable probabilities and the distribution of the response function variables. If those relationships are linear, then they define linear constraints in an optimization problem. Next, we describe a general class of DAGs having linear relationships between their distributions of response function variables and distributions of observable variables.

### 3. A Class of Linear DAGs

To characterize our class of linear problems, we first extract any measured variables that, akin to instrumental variables, are known to share no common causes with another set of variables. In Gabriel, Sachs, and Sjölander (2020) it was shown that the presence of such variables may tighten the bounds in several scenarios. The set  $\mathcal{W}$  is divided into two subsets  $\mathcal{W} = \{\mathcal{W}_{\mathcal{L}}, \mathcal{W}_{\mathcal{R}}\}$ , where  $\mathcal{W}_{\mathcal{L}}$  may be empty. We assume without loss of generality that the indices of the variables are ordered in such a way that  $\mathcal{L} = \{1, \dots, \mathfrak{J}\}$  and  $\mathcal{R} = \{\mathfrak{J} + 1, \dots, n\}$ , where  $\mathfrak{J}$  may be 0 in which case  $\mathcal{L}$  is the empty set. We will denote the corresponding subdivisions of the vectors  $\mathbf{W}$  and  $\mathbf{R}$  by  $(\mathbf{W}_{\mathcal{L}}, \mathbf{W}_{\mathcal{R}})$  and  $(\mathbf{R}_{\mathcal{L}}, \mathbf{R}_{\mathcal{R}})$ , respectively, and likewise for their lowercase value-vector counterparts.  $\mathcal{L}$  and  $\mathcal{R}$ , connote *left* and *right* sides, where the causal paths flow from left to right. We make this division because in our class of problems, the  $\mathcal{L}$ -side variables share no common causes with the  $\mathcal{R}$ -side variables.

Let  $B := |\nu(\mathbf{W})| = \prod_{i=1}^n |\nu(W_i)| = \prod_{i=1}^n c_{W_i}$ , and let  $\{1, \dots, B\} \ni b \mapsto \mathbf{w}_b \in \nu(\mathbf{W})$  be an enumeration of  $\nu(\mathbf{W})$  that preserves the ordering of the  $\mathcal{L}$ -indices before the  $\mathcal{R}$ -indices such that  $\forall b \in \{1, \dots, B\}$ ,  $\mathbf{w}_{b,\mathcal{L}} := (\mathbf{w}_b)_{\mathcal{L}}$  and  $\mathbf{w}_{b,\mathcal{R}} := (\mathbf{w}_b)_{\mathcal{R}}$ . Let  $\mathbf{p}^*, \mathbf{p} \in [0, 1]^B$  be given by  $\forall b \in \{1, \dots, B\}$ ,  $p_b^* := p\{\mathbf{W} = \mathbf{w}_b\} = p\{(\mathbf{W}_{\mathcal{L}}, \mathbf{W}_{\mathcal{R}}) = (\mathbf{w}_{b,\mathcal{L}}, \mathbf{w}_{b,\mathcal{R}})\}$  and  $p_b := p\{\mathbf{W}_{\mathcal{R}} = \mathbf{w}_{b,\mathcal{R}} | \mathbf{W}_{\mathcal{L}} = \mathbf{w}_{b,\mathcal{L}}\}$ . Thus, the vector  $\mathbf{p}^*$  represents the joint distribution of all observed variables and the vector  $\mathbf{p}$  contains the observed conditional distribution of all variables in  $\mathcal{W}_{\mathcal{R}}$  given all variables in  $\mathcal{W}_{\mathcal{L}}$ . As shown in Proposition 2, we will only need to observe the components of  $\mathbf{p}$ .

We will focus on the response function variables of the  $\mathcal{R}$ -side, and will provide them a dedicated enumeration. Let

$$\aleph_{\mathcal{R}} := |\nu(\mathbf{R}_{\mathcal{R}})| = \prod_{i=\mathfrak{J}+1}^n |\nu(R_{W_i})| = \prod_{j=\mathfrak{J}+1}^n c_{W_j}^{|\mathbf{pa}_{W_j}|}.$$

Let  $\{1, \dots, \aleph_{\mathcal{R}}\} \ni \gamma \mapsto \mathbf{r}_{\gamma} \in \nu(\mathbf{R}_{\mathcal{R}})$  enumerate  $\nu(\mathbf{R}_{\mathcal{R}})$  and  $\mathbf{q} \in [0, 1]^{\aleph_{\mathcal{R}}}$  be given by  $\forall \gamma \in \{1, \dots, \aleph_{\mathcal{R}}\}$ ,  $q_{\gamma} := p\{\mathbf{R}_{\mathcal{R}} = \mathbf{r}_{\gamma}\}$ . In particular, the vector  $\mathbf{q}$  contains the joint probability



distribution of the response function variables  $\mathbf{R}_{\mathcal{R}}$ . For  $i \in \mathcal{R}$  and a fixed value-vector  $\mathbf{w}_{\mathcal{L}} \in \nu(\mathbf{W}_{\mathcal{L}})$ , we let

$$g_{W_i}(\mathbf{w}_{\mathcal{L}}, \mathbf{r}_{\gamma}) := f_{W_i}(w_{i1}, \dots, w_{iI_i}, g_{W_{iI_i+1}}^*(\mathbf{r}_{\gamma}), \dots, g_{W_{iI_i}}^*(\mathbf{r}_{\gamma}), r_{W_i}) \\ = w_i,$$

where  $w_{i1}, \dots, w_{iI_i}$  are the values of the parents of  $W_i$  that are in  $\mathcal{W}_{\mathcal{L}}$ , and  $W_{iI_i+1}, \dots, W_{iI_i}$  are the parents of  $W_i$  that are in  $\mathcal{W}_{\mathcal{R}}$ .

**Proposition 2.** Let  $G$  be a causal DAG satisfying the following Conditions:

1. Any edge that connects two variables  $W_{\mathcal{L}} \in \mathcal{W}_{\mathcal{L}}$  and  $W_{\mathcal{R}} \in \mathcal{W}_{\mathcal{R}}$  must be directed from  $W_{\mathcal{L}}$  to  $W_{\mathcal{R}}$ .
2. There exists no unmeasured variable  $U$  that has children in both  $\mathcal{W}_{\mathcal{L}}$  and  $\mathcal{W}_{\mathcal{R}}$ . That is, the variables in  $\mathcal{W}_{\mathcal{L}}$  and  $\mathcal{W}_{\mathcal{R}}$  do not share a common cause.
3. There exists an unmeasured variable  $U_{\mathcal{L}}$  such that  $U_{\mathcal{L}}$  is a parent of  $W_i$  for all  $i \in \mathcal{L}$ . That is, all variables in  $\mathcal{L}$  share an unmeasured common cause.
4. There exists an unmeasured variable  $U_{\mathcal{R}}$  such that  $U_{\mathcal{R}}$  is a parent of  $W_i$  for all  $i \in \mathcal{R}$ . That is, all variables in  $\mathcal{R}$  share an unmeasured common cause,

Then there exist matrices  $P \in \{0, 1\}^{B \times \aleph_{\mathcal{R}}}$ ,  $P^* \in [0, 1]^{B \times \aleph_{\mathcal{R}}}$  and  $\Lambda \in [0, 1]^{B \times B}$  such that  $\mathbf{p} = P\mathbf{q}$ ,  $\mathbf{p}^* = P^*\mathbf{q}$ ,  $\Lambda$  is diagonal with nonzero diagonal entries,  $\Lambda P = P^*$ ,  $\mathbf{p}^* = \Lambda\mathbf{p}$ , and there are no other constraints on the distribution of response function variables that are not redundant with these.

See Section S1 of the [supplementary materials](#) for proof. **Conditions 1** and **2** imply that the matrices and linear relations exist; thus, a causal model satisfying those conditions implies those linear constraints. **Conditions 1** and **2** together with **Conditions 3** and **4** imply there are no other nonredundant constraints implied by the causal model. As a counterexample, removing for example, an arc from  $U_{\mathcal{R}}$  to a variable in  $\mathcal{W}_{\mathcal{R}}$  would violate **Condition 4** and potentially introduce a new, nonlinear constraint. A concrete example for why **Conditions 3** and **4** are needed is given in Gabriel, Sachs, and Sjölander (2020): **Figure 1(d)** (depicting unconfounded, outcome-dependent sampling). In that example, the lack of an arrow from  $U = U_{\mathcal{R}}$  to  $S \in \mathcal{W}_{\mathcal{R}}$  implies nonlinear constraints, requiring them to use other methods to derive valid but not necessarily tight bounds.

Though **Proposition 2** guarantees their existence, it may not be trivial to construct these linear relations. **Algorithm 1** details a method for constructing the matrices  $P$ ,  $P^*$ , and  $\Lambda$ .

#### 4. Functional Expressions Incorporating Interventions

In order to determine the values of variables of interest for potential outcomes that incorporate interventions, we must also define a procedure for evaluating a functional expression that allows for variables to be externally forced to certain values. To illustrate, we consider extended DAGs, which add additional nodes for potential outcomes of interest as in Balke and Pearl (1994b). These are called multi-networks in Shpitser and Pearl (2007). Two examples are shown in **Figure 2(a)** and **(b)**. For each potential outcome of interest, nodes are added such that

**Result:** Systems of linear equations relating  $\mathbf{p}^*$  and  $\mathbf{p}$  to  $\mathbf{q}$

Initialize  $P$  as a  $B \times \aleph_{\mathcal{R}}$  matrix of 0s;

Initialize  $P^*$  as a  $B \times \aleph_{\mathcal{R}}$  matrix of 0s;

Initialize  $\Lambda$  as a  $B \times B$  matrix of 0s;

**for**  $b \in 1, \dots, B$  **do**

**for**  $\gamma \in 1, \dots, \aleph_{\mathcal{R}}$  **do**

        Initialize  $\omega$  as an empty vector of length

$|\mathcal{R}| (= n - \mathbb{I})$ ;

**for**  $i \in \mathcal{R}$  **do**

            Set  $\omega_i := g_{W_i}(\mathbf{w}_{b,\mathcal{L}}, \mathbf{r}_{\gamma})$ ;

**end**

**if**  $\omega = \mathbf{w}_{b,\mathcal{R}}$  **then**

$P_{b,\gamma} := 1$ ;

$\Lambda_{b,b} := p\{\mathbf{W}_{\mathcal{L}} = \mathbf{w}_{b,\mathcal{L}}\}$ ;

$P_{b,\gamma}^* := p\{\mathbf{W}_{\mathcal{L}} = \mathbf{w}_{b,\mathcal{L}}\}$ ;

**end**

**end**

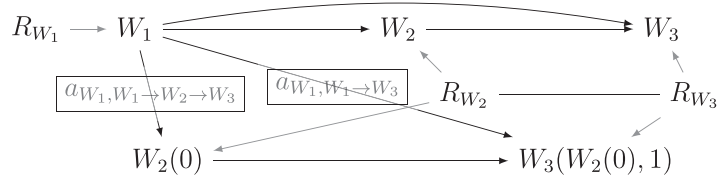
**end**

**Algorithm 1:** An algorithm to determine a system of linear equations relating  $\mathbf{p}$  and  $\mathbf{p}^*$  to  $\mathbf{q}$ .

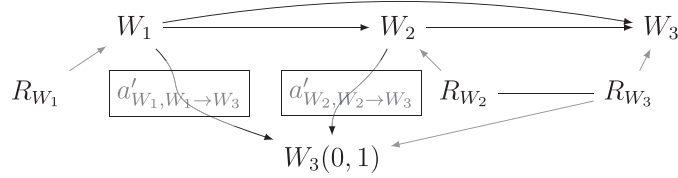
the corresponding factual and potential outcome nodes share the same response function variables. Edges that connect factual nodes to potential outcome nodes are labeled with letters that denote intervention sets indexed by the tail variable of that edge and the path to the head of that edge sequence. These sets define the variables being externally set, the values that they are being set to, and their indices indicate for which edge sequences they apply.

Balke and Pearl (1994a) considered cases where we externally force a single subset of the variables to some fixed values. This construction suffices for the examples they consider, but not for defining and bounding effects like the natural direct effect of  $W_1$  in the graph in **Figure 2(a)** whose first term is  $p\{W_3(W_2(W_1 = 0), W_1 = 1) = 1\}$ . In that expression, we see that the variable  $W_1$ , which is a parent of both  $W_3$  and  $W_2$ , is simultaneously being set to 0 and 1, the difference being which child is in question. As another example, the causal query  $p\{W_3(W_2(W_1 = 0)) = 1, W_2(W_1 = 1) = 1\}$  is a joint probability statement, and the two events in question are under interventions on  $W_1$ . Therefore, to be completely general, the variables that one assign to values cannot be a single set; the values that variables are being externally forced to may depend on which children are being considered and also on the term of the probability statement. Thus, we define an extended function expression, which “remembers” the path of edges taken to get the value that is being determined at each call.

For  $i \in \{\mathbb{I} + 1, \dots, n\}$ , let  $A_i$  be a matrix that encodes the interventions and variables on which to intervene, with rows indexed by  $l$  corresponding to the variables in  $\mathcal{W}$  and the columns indexed by  $j$  corresponding to all possible paths terminating at  $W_i$ ; the entries in row  $l$  are in  $\nu(W_l) \cup \{\emptyset\}$ . The desired interventions within the causal query then define the entries of  $A_i$  which are denoted  $a_{ij}$ . In our procedure for evaluating potential outcomes, there is a distinct interventional matrix  $A_i$  corresponding to each outcome variable  $W_i$  used in the causal query. We define the procedure for evaluating the



(a) Extended graph for evaluation of the potential outcome  $W_3(W_2(W_1 = 0), W_1 = 1)$ .



(b) Extended graph for evaluation of the potential outcome  $W_3(W_2 = 0, W_1 = 1)$ .

**Figure 2.** Extended DAGs to illustrate that multiple intervention sets are needed to define certain potential outcomes. In these two examples, the variables are binary.

interventional functional expression for an outcome variable  $W_i$  as

$$w_i = h_{W_i}^{A_i}(\mathbf{r}, W_i),$$

where for all  $l \in \{1, \dots, n\}$ , all  $\mathbf{r} \in \nu(\mathbf{R})$  and all strings  $j$  representing paths to  $W_i$ , we define  $h_{W_i}^{A_i}(\mathbf{r}, j)$  recursively by

$$h_{W_i}^{A_i}(\mathbf{r}, j) := \begin{cases} a_{lj} & \text{if } a_{lj} \neq \emptyset \\ f_{W_i}(r_{W_i}) & \text{if } a_{lj} = \emptyset \text{ and } \mathbf{Pa}_{W_i} = \emptyset \\ f_{W_i}(h_{W_{i_1}}^{A_i}(\mathbf{r}, W_{i_1} \rightarrow j), \\ \dots, \\ h_{W_{i_{k_i}}}^{A_i}(\mathbf{r}, W_{i_{k_i}} \rightarrow j), r_{W_i}) & \text{else,} \end{cases}$$

where  $k_i := |\mathbf{Pa}_{W_i}|$  and  $\{W_{i_1}, \dots, W_{i_{k_i}}\} := \mathbf{Pa}_{W_i}$ , and the notation  $i \rightarrow j$  means that  $i \rightarrow$  is prepended to  $j$ . This notation allows us to trace the full path taken from the outcome of interest to the variable being intervened upon.

For example, considering the DAG in Figure 2(a) and the causal query  $p\{W_3(W_2(W_1 = 0), W_1 = 1) = 1\}$ , we have the interventional matrix

$$A_3 = \left[ \begin{array}{c|cccc} & W_1 \rightarrow W_2 \rightarrow W_3 & W_1 \rightarrow W_3 & W_2 \rightarrow W_3 & W_3 \\ \hline W_1 & 0 & 1 & \emptyset & \emptyset \\ W_2 & \emptyset & \emptyset & \emptyset & \emptyset \\ W_3 & \emptyset & \emptyset & \emptyset & \emptyset \end{array} \right].$$

Thus, evaluating the functional expression  $w_3 = h_{W_3}^{A_3}(\mathbf{r}, W_3)$  results (since  $W_3$  is not intervened upon and  $\mathbf{Pa}_{W_3} = \{W_1, W_2\}$ ) in

$$\begin{aligned} w_3 &= h_{W_3}^{A_3}(\mathbf{r}, W_3) \\ &= f_{W_3}(w_1 = h_{W_1}^{A_3}(\mathbf{r}, W_1 \rightarrow W_3), w_2 = h_{W_2}^{A_3}(\mathbf{r}, W_2 \rightarrow W_3), r_{W_3}). \end{aligned}$$

For the first argument of that function call we have  $w_1 = h_{W_1}^{A_3}(\mathbf{r}, W_1 \rightarrow W_3) = a_{1, W_1 \rightarrow W_3} = 1$ . Then for the second argument,  $a_{2, W_2 \rightarrow W_3} = \emptyset$  and  $\mathbf{Pa}_{W_2} = \{W_1\}$ , so we recurse,

$$w_2 = h_{W_2}^{A_3}(\mathbf{r}, W_2 \rightarrow W_3)$$

$$= f_{W_2}(w_1 = h_{W_1}^{A_3}(\mathbf{r}, W_1 \rightarrow W_2 \rightarrow W_3), r_{W_2}).$$

Now,  $w_1 = h_{W_1}^{A_3}(\mathbf{r}, W_1 \rightarrow W_2 \rightarrow W_3) = a_{1, W_1 \rightarrow W_2 \rightarrow W_3} = 0$ , giving  $w_2 = f_{W_2}(w_1 = 0, r_{W_2})$ , so we get  $w_3 = f_{W_3}(w_1 = 1, w_2 = f_{W_2}(w_1 = 0, r_{W_2}), r_{W_3})$ .

For the DAG in Figure 2(b) and the first part of the causal query  $p\{W_3(W_2 = 0, W_1 = 1) = 1\}$ , we have

$$A_3 = \left[ \begin{array}{c|cccc} & W_1 \rightarrow W_2 \rightarrow W_3 & W_1 \rightarrow W_3 & W_2 \rightarrow W_3 & W_3 \\ \hline W_1 & \emptyset & 1 & \emptyset & \emptyset \\ W_2 & \emptyset & \emptyset & 0 & \emptyset \\ W_3 & \emptyset & \emptyset & \emptyset & \emptyset \end{array} \right].$$

Thus, evaluating the functional expression  $w_3 = h_{W_3}^{A_3}(\mathbf{r}, W_3)$  results in

$$\begin{aligned} w_3 &= h_{W_3}^{A_3}(\mathbf{r}, W_3) \\ &= f_{W_3}(w_1 = h_{W_1}^{A_3}(\mathbf{r}, W_1 \rightarrow W_3), w_2 = h_{W_2}^{A_3}(\mathbf{r}, W_2 \rightarrow W_3), r_{W_3}). \end{aligned}$$

For the first argument of that function call we have  $w_1 = h_{W_1}^{A_3}(\mathbf{r}, W_1 \rightarrow W_3) = a_{1, W_1 \rightarrow W_3} = 1$ . Then, for the second argument,  $w_2 = h_{W_2}^{A_3}(\mathbf{r}, W_2 \rightarrow W_3) = a_{2, W_2 \rightarrow W_3} = 0$ , giving the result  $w_3 = f_{W_3}(w_1 = 1, w_2 = 0, r_{W_3})$ .

The procedures for evaluating the functions  $g$  and  $h^{A_i}$  are sufficient to translate any combined factual and/or potential outcome joint probability statement into probability statements involving only the response function variables  $\mathbf{R}$ . Thus, using our response function formulation, any potential outcome or factual joint probability statement can be written

$$\begin{aligned} Q &:= p\{h_{W_{i_1}}^{A_{i_1}}(\mathbf{R}, W_{i_1}) = w_{i_1}, \dots, h_{W_{i_p}}^{A_{i_p}}(\mathbf{R}, W_{i_p}) = w_{i_p}, \\ &\quad g_{W_{j_1}}(\mathbf{R}) = w_{j_1}, \dots, g_{W_{j_o}}(\mathbf{R}) = w_{j_o}\}, \quad (1) \end{aligned}$$

where  $\mathcal{P} = \{i_1, \dots, i_p\}$  denote the indices of potential outcomes, and  $\mathcal{O} = \{j_1, \dots, j_o\}$  the indices of the factual outcomes (and these sets may be overlapping). Given the functional expressions we have defined and our procedures for evaluating them, we can therefore write

$$Q = \sum_{\mathbf{r} \in \Gamma(Q)} p\{\mathbf{R} = \mathbf{r}\}, \text{ where}$$

$$\Gamma(Q) := \{\mathbf{r} \in v(\mathbf{R}) : \forall i_p \in \mathcal{P}, w_{i_p} = h_{W_{i_p}}^{A_{i_p}}(\mathbf{r}, W_{i_p}) \text{ and} \\ \forall j_o \in \mathcal{O}, w_{j_o} = g_{W_{j_o}}(\mathbf{r})\}.$$

We will call an expression of this form an *atomic* query. Their form is completely general, and allows arbitrarily nested potential outcomes, and combinations with observational quantities. We will combine atomic queries to obtain causal contrasts of interest, such as the causal risk difference.

**Proposition 3.** Let  $G$  be a causal DAG satisfying [Conditions 1](#) and [2](#), and let  $Q$  be an atomic query satisfying the following Conditions:

5. Each atomic query is a probability as given in Equation (1) where  $i_1, \dots, i_p, j_1, \dots, j_o \in \mathcal{R}$  (i.e., all outcome variables must be in  $\mathcal{W}_{\mathcal{R}}$ ) and
6. if  $\mathcal{L} \neq \emptyset$  then: (i) none of the variables in  $\mathcal{W}_{\mathcal{L}}$  that are intervened upon can have any children in  $\mathcal{W}_{\mathcal{L}}$ , (ii) all paths from  $\mathcal{W}_{\mathcal{L}}$  to  $\mathcal{W}_{\mathcal{R}}$  must be blocked by the intervention set (here the intervention set refers to variables in the rows of the  $A$  matrices that are not  $\emptyset$ ), (iii) no observations are allowed, that is,  $\mathcal{O} = \emptyset$ .

Then there exists a constant binary vector  $\alpha \in \{0, 1\}^{\mathcal{N}_{\mathcal{R}}}$  such that  $Q = \alpha^\top \mathbf{q}$ .

See Section S1 of the [supplementary materials](#) for proof. Note that we allow for observations to appear in the query. Some examples of causal queries that involve observations in addition to potential outcomes are the treatment effect on the treated (Angrist and Imbens 1995), and the proportion factually helped by treatment:  $p\{Y = 1, Y(X = 0) = 0\}$ . A procedure for construction of this  $\alpha$  is detailed in [Algorithm 2](#) which converts the atomic query  $Q$  into a binary linear combination of probabilities of response function variables of the  $\mathcal{R}$ -side.

The following corollary, which specifies the general form of a causal query, follows immediately since linear combinations of linear combinations again are just linear combinations.

**Corollary 1.** Let  $Q^*$  be any real linear combination of atomic queries (in particular,  $Q$  may be a classic linear causal contrast such as a causal risk difference). Under [Conditions 1, 2, 5, and 6](#), there exists a constant vector  $\alpha^* \in \mathbb{R}^{\mathcal{N}_{\mathcal{R}}}$  such that  $Q^* = \alpha^{*\top} \mathbf{q}$ .

The algorithms are formulated so that bounds are derived in terms of the true probabilities of the observed variables in  $\mathcal{W}_{\mathcal{R}}$  conditional on the variables in  $\mathcal{W}_{\mathcal{L}}$ . Provided one is not intervening on any of the variables in  $\mathcal{W}_{\mathcal{L}}$ , [Conditions 1](#) and [2](#) imply that the directions of the edges within  $\mathcal{W}_{\mathcal{L}}$  cannot influence the bounds. That is, the bounds are tight for the equivalence class of DAGs that contains the set of DAGs for all possible directions of edges among variables in  $\mathcal{W}_{\mathcal{L}}$ . For example, the bounds computed for a query such as  $p\{Y(X = 1) = 1\}$  are tight and equal for both of the DAGs in [Figure 3\(a\)](#) and (b). In either case, the knowledge of whether  $Z$  causes  $Z_2$  or vice versa does not influence the bounds because both of those variables are conditioned upon in the algorithm.

Alternatively, if the desired query was  $p\{Y(X(Z = 1)) = 1\}$ , the DAGs in [Figure 3\(a\)](#) and (b) may not result in the same bounds, and in fact, the causal problem under [Figure 3\(a\)](#) may

**Result:**  $Q$  expressed as a simple sum of a subset of the components of  $\mathbf{q}$ .

Initialize  $\alpha \in \{0, 1\}^{\mathcal{N}_{\mathcal{R}}}$  by  $\forall \gamma \in \{1, \dots, \mathcal{N}_{\mathcal{R}}\}, \alpha_\gamma := 1$ ;

Let  $\mathcal{P}, \mathcal{O}$  be the index sets as defined above corresponding to  $Q$ ;

```

for  $\gamma \in 1, \dots, \mathcal{N}_{\mathcal{R}}$  do
  for  $l \in \mathcal{P}$  do
    Construct  $A_l$  according to  $l$ ;
    Compute  $\omega := h_{W_l}^{A_l}(\mathbf{r}_\gamma, W_l)$ ;
    if  $\omega \neq w_l$  then
      Set  $\alpha_\gamma := 0$ ;
      break;
    end
  end
  if  $\alpha_\gamma = 0$  then
    break;
  end
  for  $l \in \mathcal{O}$  do
    Compute  $\omega := g_{W_l}(\mathbf{r}_\gamma)$ ;
    if  $\omega \neq w_l$  then
      Set  $\alpha_\gamma := 0$ ;
      break;
    end
  end
end

```

**Algorithm 2:** Converting  $Q$  to a binary linear combination of  $\mathbf{q}$ .

not be linear. As required by [Conditions 5](#) and [6](#), if we intervene upon a variable in  $\mathcal{W}_{\mathcal{L}}$ , then the direction of edges within  $\mathcal{W}_{\mathcal{L}}$  matters, and in fact if the intervened upon variable has a child also in  $\mathcal{W}_{\mathcal{L}}$ , Condition 6 will not be met.

## 5. Optimization via Vertex Enumeration

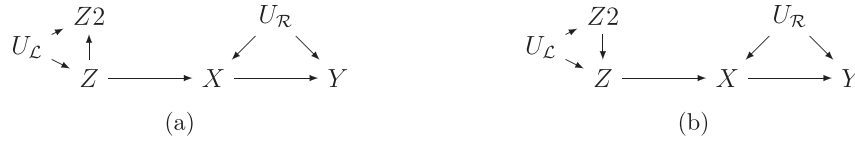
After applying [Algorithms 1](#) and [2](#), we have a linear objective and a system of linear constraints. We also have the probabilistic constraints:

$$\forall \mathbf{w}_{\mathcal{L}} \in v(\mathbf{W}_{\mathcal{L}}), \quad \sum_{\mathbf{w}_{\mathcal{R}} \in v(\mathbf{W}_{\mathcal{R}})} p\{\mathbf{W}_{\mathcal{R}} = \mathbf{w}_{\mathcal{R}} | \mathbf{W}_{\mathcal{L}} = \mathbf{w}_{\mathcal{L}}\} = 1$$

and

$$\sum_{\gamma=1}^{\mathcal{N}_{\mathcal{R}}} q_\gamma = \sum_{\gamma=1}^{\mathcal{N}_{\mathcal{R}}} p\{\mathbf{R}_{\mathcal{R}} = \mathbf{r}_\gamma\} = \sum_{\mathbf{r}_{\mathcal{R}} \in v(\mathbf{R}_{\mathcal{R}})} p\{\mathbf{R}_{\mathcal{R}} = \mathbf{r}_{\mathcal{R}}\} = 1.$$

Additional linear constraints on  $\mathbf{q}$  can be optionally given as  $B\mathbf{q} \geq \mathbf{d}$  where  $B$  and  $\mathbf{d}$  are respectively a matrix and vector of real constants. These constraints can be used to encode assumptions about the response functions that are not possible to encode in a DAG, for example, restricting the probabilities of implausible response patterns. We thus arrive at the following linear programming problem for the lower bound; the upper bound is given by the corresponding maximization problem.



**Figure 3.** An equivalence class of DAGs defined by arbitrary connections in  $\mathcal{W}_{\mathcal{L}}$ . Bounds for causal queries that involve intervening on  $X$  that meet our conditions are equivalent and tight for these two graphs in (a) and (b).

$$\begin{aligned} & \text{minimize } Q = \alpha^T \mathbf{q} \\ & \text{subject to } P\mathbf{q} = \mathbf{p}, \\ & \quad B\mathbf{q} \geq \mathbf{d}, \\ & \quad \mathbf{q} \geq \mathbf{0}, \text{ and } \mathbf{1}^T \mathbf{q} = 1. \end{aligned}$$

Note that the constraint space constitutes a bounded (due to the probabilistic constraints) convex polytope. By the fundamental theorem of linear programming, the global extrema must occur at one of the vertices of the polytope. We can thus solve this problem symbolically by applying an efficient vertex enumeration algorithm, such as the double description algorithm (Motzkin et al. 1953; Fukuda 2018) to enumerate the vertices of the polytope of the dual linear program. For instance, the dual of the minimization problem above is given by

$$\begin{aligned} & \text{maximize } (\mathbf{d}^T \mathbf{1} \mathbf{p}^T) \mathbf{y} \\ & \text{subject to } \begin{pmatrix} B^T & \mathbf{1} & P^T \\ I & 0 & 0 \end{pmatrix} \mathbf{y} \leq \begin{pmatrix} \alpha \\ 0 \end{pmatrix}. \end{aligned}$$

So by the strong duality theorem, the optimum of the dual, and thus also of the primal problem, is of the form  $(\mathbf{d}^T \mathbf{1} \mathbf{p}^T) \bar{\mathbf{y}}$  where  $\bar{\mathbf{y}}$  is a vertex of the polytope  $\{\mathbf{y} : \begin{pmatrix} B^T & \mathbf{1} & P^T \\ I & 0 & 0 \end{pmatrix} \mathbf{y} \leq \begin{pmatrix} \alpha \\ 0 \end{pmatrix}\}$ . This gives a lower bound on the causal effect of interest as the maximum of a set of expressions involving only observable probabilities. Similarly, the upper bound is given by reversing the dual inequality and minimizing over the corresponding polytope.

**Proposition 4.** Under Conditions 1–6 and subject to any additional linear constraints of the form  $B\mathbf{q} \geq \mathbf{d}$ , the procedure above yields valid and tight symbolic bounds for a causal query that is a linear combination of atomic queries.

**Corollary 2.** If Condition 4 does not hold, then the bounds derived using the above procedure are still valid.

See the [supplementary materials](#) for proof. The Conditions 3 and 4 represent a worst-case scenario of confounding and ensure that the decompositions giving rise to the linear constraints cannot be further factorized to yield more granular but nonlinear constraints. If however there is any known (partial) absence of such confounding, then these bounds are still valid, and may be narrow enough to be informative, while not necessarily tight. Such an absence of confounding on the  $\mathcal{R}$ -side implies some independence among the  $\mathbf{R}_{\mathcal{R}}$  variables, and hence additional constraints on their distribution. Thus, the true feasible space may be smaller than the one considered in our algorithm, but completely contained inside it.

## 6. Examples

The graphs in the following examples are divided into a left side, which corresponds to the  $\mathcal{W}_{\mathcal{L}}$  set, and a right side, which corresponds to the  $\mathcal{W}_{\mathcal{R}}$  set, as in Figure 4(a). The left side is displayed as a violet (dark gray) box, and the right side a yellow (light gray) box.

### 6.1. Confounded Exposure and Outcome

The basic DAG with two variables that are confounded as shown in Figure 4(a) conforms to our class of models. In this case, the variable  $X$  is the exposure of interest, and  $Y$  the outcome of interest.  $X$  and  $Y$  have a common, unmeasured cause  $U$  which we make no assumptions about. We specify  $X$  and  $Y$  to be ternary and binary, respectively, so  $X$  takes values in  $\{0, 1, 2\}$  and  $Y$  in  $\{0, 1\}$ . Our causal effects of interest are the risk differences  $p\{Y(X = 2) = 1\} - p\{Y(X = 0) = 1\}$ ,  $p\{Y(X = 2) = 1\} - p\{Y(X = 1) = 1\}$  and  $p\{Y(X = 1) = 1\} - p\{Y(X = 0) = 1\}$ , and we have no additional constraints to specify.

Here we have two variables and therefore two response function variables. The response function variable formulation of the graph in Figure 4(b) is an equivalent representation of the causal model. The following tables define the values of the response functions and variables:

		$y = f_Y(x, r_Y)$			
		$r_Y = 0$	$x = 0$	$x = 1$	$x = 2$
$x = f_X(r_X)$	$r_X = 0$	$r_Y = 0$	$y = 0$	$y = 0$	$y = 0$
	$r_X = 1$	$r_Y = 1$	$y = 1$	$y = 0$	$y = 0$
	$r_X = 2$	$r_Y = 2$	$y = 0$	$y = 1$	$y = 0$
	$x = 0$	$r_Y = 3$	$y = 1$	$y = 1$	$y = 0$
	$x = 1$	$r_Y = 4$	$y = 0$	$y = 0$	$y = 1$
	$x = 2$	$r_Y = 5$	$y = 1$	$y = 0$	$y = 1$
		$r_Y = 6$	$y = 0$	$y = 1$	$y = 1$
	$r_Y = 7$	$y = 1$	$y = 1$	$y = 1$	

$R_X$  is a random variable that can take on 3 possible values, and  $R_Y$  is a random variable that can take on  $2^3 = 8$  possible values. Thus, the joint distribution of  $(R_X, R_Y)$  is characterized by  $3 \cdot 8 = 24$  parameters, say  $q_{i,j}$ , where  $i \in \{0, 1, 2\}$  and  $j \in \{0, 1, 2, 3, 4, 5, 6, 7\}$ . Applying Algorithm 1, we can relate the  $3 \cdot 2 = 6$  observed probabilities to the parameters of the response function variable distribution as follows:

$$\begin{aligned} p_{0,0} &:= p\{X = 0, Y = 0\} = q_{0,0} + q_{0,2} + q_{0,4} + q_{0,6} \\ p_{1,0} &:= p\{X = 1, Y = 0\} = q_{1,0} + q_{1,1} + q_{1,4} + q_{1,5} \\ p_{2,0} &:= p\{X = 2, Y = 0\} = q_{2,0} + q_{2,1} + q_{2,2} + q_{2,3} \\ p_{0,1} &:= p\{X = 0, Y = 1\} = q_{0,1} + q_{0,3} + q_{0,5} + q_{0,7} \\ p_{1,1} &:= p\{X = 1, Y = 1\} = q_{1,2} + q_{1,3} + q_{1,6} + q_{1,7} \\ p_{2,1} &:= p\{X = 2, Y = 1\} = q_{2,4} + q_{2,5} + q_{2,6} + q_{2,7}. \end{aligned}$$



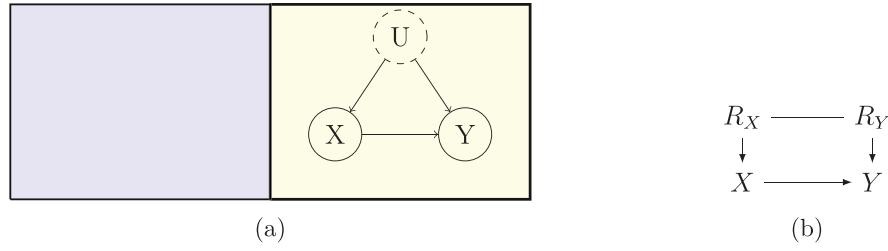


Figure 4. Simple confounded example and the equivalent response function variable graph.

We get

$$A = \begin{bmatrix} X \rightarrow Y & Y \\ X & a & \emptyset \\ Y & \emptyset & \emptyset \end{bmatrix}, \text{ for } p\{Y(X = a) = 1\}, a \in \{0, 1, 2\}$$

Applying Algorithm 2, we get

$$\begin{aligned} p\{Y(X = 0) = 1\} &= q_{0,1} + q_{0,3} + q_{0,5} + q_{0,7} + q_{1,1} \\ &\quad + q_{1,3} + q_{1,5} + q_{1,7} + q_{2,1} + q_{2,3} \\ &\quad + q_{2,5} + q_{2,7}, \\ p\{Y(X = 1) = 1\} &= q_{0,2} + q_{0,3} + q_{0,6} + q_{0,7} + q_{1,2} \\ &\quad + q_{1,3} + q_{1,6} + q_{1,7} + q_{2,2} + q_{2,3} \\ &\quad + q_{2,6} + q_{2,7}, \\ p\{Y(X = 2) = 1\} &= q_{0,4} + q_{0,5} + q_{0,6} + q_{0,7} + q_{1,4} \\ &\quad + q_{1,5} + q_{1,6} + q_{1,7} + q_{2,4} + q_{2,5} \\ &\quad + q_{2,6} + q_{2,7}, \end{aligned}$$

from which the contrasts of interest are easily derived.

Together with the probabilistic constraints, we then have the fully specified linear programming problem. The bounds are, after some algebra on the output from the program,

$$\begin{aligned} &p\{X = x_1, Y = 1\} + p\{X = x_2, Y = 0\} - 1 \\ &\leq p\{Y(X = x_1) = 1\} - p\{Y(X = x_2) = 1\} \leq \\ &1 - p\{X = x_1, Y = 0\} - p\{X = x_2, Y = 1\} \end{aligned}$$

for  $(x_1, x_2) \in \{(1, 0), (2, 0), (2, 1)\}$ .

Note that these expressions are not unique and may be reparameterized using that conditional probabilities sum to 1.

## 6.2. Two Instruments

Our next example is shown in the DAG in Figure 5. This extends the instrumental variable example to the case where there are two binary variables on the left side that may be associated with each other and that both have a direct effect on  $X$ , but no direct effect on  $Y$ . This situation may arise in Mendelian randomization studies, wherein multiple genes may be known to cause changes in an exposure but not directly on the outcome.

The bounds on risk difference  $p\{Y(X = 1)\} - p\{Y(X = 0)\}$  under this DAG can be computed using our method. In this problem, there are 16 constraints involving the conditional probabilities, the distribution of the response function variables of the  $\mathcal{R}$ -side has 64 parameters, and the causal query is a function of 32 of these parameters. The bounds are the extrema over 112 vertices, and are therefore too long to be presented simply, but code to reproduce them, as well as details about the simulation, is included in the [supplementary materials](#).

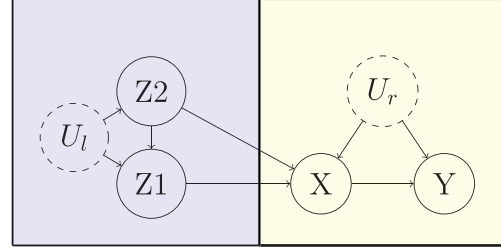


Figure 5. Two instrumental variables example with binary variables

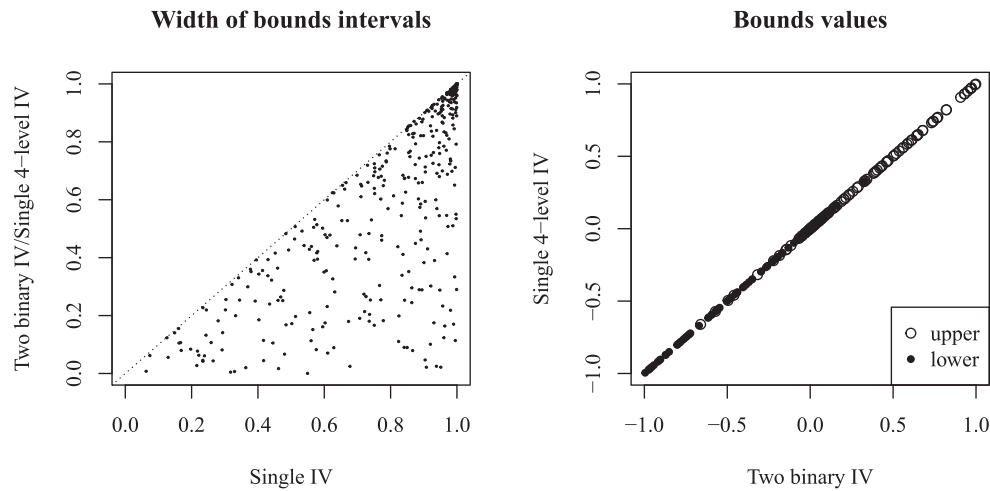
To illustrate these bounds, we computed them for specific values of observed probabilities generated from a model (described fully in Section S2 of the [supplementary materials](#)) which satisfies the DAG in Figure 5. Using these simulations, we compare our bounds to the classic IV bounds from Balke and Pearl (1997) for a single binary instrument and to bounds derived using our method for a single but 4-level categorical instrument.

The widths of the classic IV bounds and the dual binary instruments are compared for a subsample of the simulations in Figure 6. The bounds with two instruments are never wider than the classic IV bounds with a single binary instrument. The simulations also verify that a single four level instrument yields exactly the same bounds as two binary ones. Details and R code for these simulations are provided in the [supplementary materials](#).

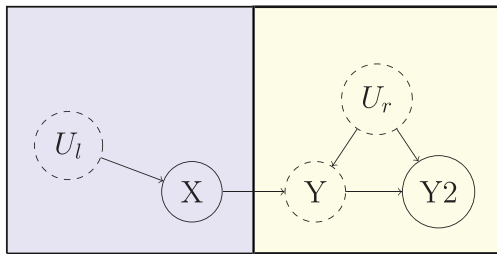
## 6.3. Measurement Error in the Outcome

Our final example illustrates some additional features of our method. In Figure 7, we have a binary variable  $X$  affecting a binary variable  $Y$ , but  $Y$  is not observed. Instead, the binary variable  $Y_2$  which is a child of  $Y$  is observed, and the effect of the true  $Y$  on the measured  $Y_2$  is confounded. Additionally, we would like to include a constraint that  $Y_2(Y = 1) \geq Y_2(Y = 0)$ , which is often called the monotonicity constraint. This constraint encodes the assumption that the outcome measured with error would not be equal to 0 unless the true unobserved outcome is also equal to 0. In terms of the response functions, this constraint removes the case where  $f_{Y_2}(y, r_{Y_2}) = 1 - y$ , thereby reducing the number of possible values that  $r_{Y_2}$  can take by 1.

The fact that  $Y$  is unobserved implies that we have four possible conditional probabilities to work with;  $p\{Y_2 = y_2 | X = x\}$ , for  $y_2, x \in \{0, 1\}$ . There are 12 parameters that characterize the distribution of the response function variables of the  $\mathcal{R}$ -side, and 4 constraints involving conditional probabilities. The



**Figure 6.** Under a DAG with two instruments, the left panel is a comparison of the width of the bounds intervals for the causal risk difference assuming only one of the instruments is observed to the width of the bounds assuming both are observed. The right panel compares the values of the upper and lower bounds for each replicate for two binary instruments versus a single 4-level instrument.



**Figure 7.** Example with measurement error in the outcome. Dashed circles indicate unobserved variables.

bounds for the risk difference  $p\{Y(X = 1) = 1\} - p\{Y(X = 0) = 1\}$  derived using our method are given by

$$\begin{aligned} & \max\{-1, 2p\{Y2 = 0|X = 0\} - 2p\{Y2 = 0|X = 1\} - 1\} \\ & \leq p\{Y(X = 1) = 1\} - p\{Y(X = 0) = 1\} \leq \\ & \min\{1, 2p\{Y2 = 0|X = 0\} - 2p\{Y2 = 0|X = 1\} + 1\}. \end{aligned}$$

Except in cases where  $p\{Y2 = 0|X = 0\} = p\{Y2 = 0|X = 1\}$ , these bounds are informative; meaning they give an interval that is shorter than the a priori interval  $[-1, 1]$ .

## 7. Conclusion and Discussion

We have described a general method for the symbolic computation of bounds on causal queries that are not identified from the true probability distribution of the observed variables. For this method, we give two algorithms for deriving the needed constraints and objective to construct such bounds. We describe a class of causal graphs and queries that will always define a linear program, for which we have shown the derived symbolic bounds will always be both valid and tight. We also show that under a broader class of problems our method will provide valid and possibly informative bounds that are not guaranteed to be tight.

Our approach is useful in several novel scenarios, as illustrated in the examples above. Other practical settings covered by our class of problems include Mendelian randomization, using categorical genotypes as instruments on the left side (Didelez

and Sheehan 2007), and many applied research problems that are given in terms of categorical measured variables (e.g., finitely many dosage levels for exposures and discrete health states for outcomes), but without making any assumptions about the unmeasured confounding. Additional applications of this method to unsolved problems in causal inference are now much more accessible to researchers as a class of problems for which linear programming can always be used is well-defined and clear algorithms exist for translating DAGs plus causal queries into linear programs. Our representation of causal estimands as arbitrarily nested counterfactuals and our procedure for translating them into functional expressions provides a significant advance over previous methods. This allows for bounding of cross-world counterfactual quantities which are highly relevant in mediation settings. The generality yet accessibility of the method all but guarantees that practitioners will find novel applications that we have not foreseen. We note that throughout this article and in the implementation we have used the graphical representations of causal models as suggested by Pearl (2009) and Shpitser and Pearl (2007), namely causal DAGs and multi-networks. However, as our method only relies on the nonparametric structural equations framework, it may be possible to use other graphical representations, such as single world interventional graphs (Richardson and Robins 2013).

Although our class of problems and method from deriving bounds puts no limit of the number of variables or categories for a given variable, in practice attention must be paid to computational complexity. Since we have  $|v(R_{W_i})| = \prod_{i=1}^n c_{W_i}^{|\text{Pa}_{W_i}|}$  for each variable  $W_i$ , the cardinalities of the domains of the response function variables grow exponentially with the those of other variables in the DAG. The exact growth pattern will of course depend on the DAG and its connectivity as well as the number of categorical levels of select influential variables. Thus, the number of variables or levels may be limited by computing power. More detail on the computational complexity is available in Section S3 of the [supplementary materials](#).

It should be noted that our conditions for a class of problems to be linear are sufficient, but not required. Thus, we cannot rule out that there exist problems outside of our class that can

be stated as linear. It may be possible to identify a broader class of problems or a different algorithm that may apply on a case-by-case basis. For example, nonlinear causal queries such as the relative risk or odds ratio yield nonlinear optimization problems yet in some cases it may be possible to translate them to equivalent linear problems.

We have shown that we can apply our methods in settings where all variables may have arbitrarily many categorical levels, expanding considerably the usually binary settings previously considered. This also provides opportunity for reasonable approximation via discretization of continuous variables, although the amount and impact of information lost due to such discretization on bounds has not been well studied. Extensions and insights into discretization would be useful and are areas of future research for the authors.

## Supplementary Materials

Supplementary material available online includes:

1. Supplementary Content—This includes proofs of the propositions, details on the simulation for the instrumental variable example, and discussion of computational complexity. (supplement.pdf, portable document format)
2. The R package `causaloptim`—An Interface to Specify Causal Graphs and Compute Bounds on Causal Effects, is also available from CRAN, and from Github at <https://sachsmc.github.io/causaloptim>, with additional documentation and examples. (`causaloptim_0.9.2.tar.gz`, R source package)
3. Reproducibility code—An R script that contains the R code used to run the examples and simulations presented in the main text. (`example-code.R`, R script)

## Disclosure Statement

The authors report there are no competing interests to declare.

## Funding

MCS and GJ are partially supported by Swedish Research Council grant 2019-00227, EEG by Swedish Research Council grant 2017-01898, and AS by Swedish Research Council grant 2016-01267.

## ORCID

Michael C. Sachs  <http://orcid.org/0000-0002-1279-8676>

Arvid Sjölander  <http://orcid.org/0000-0001-5226-6685>

Erin E. Gabriel  <http://orcid.org/0000-0002-0504-8404>

## References

- Angrist, J., and Imbens, G. (1995), "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–475. [572]
- Balke, A., and Pearl, J. (1994a), "Counterfactual Probabilities: Computational Methods, Bounds and Applications," in *Proceedings of the Tenth International Conference on Uncertainty in Artificial Intelligence*, pp. 46–54, Morgan Kaufmann Publishers Inc. [567,569,570]
- (1994b), "Probabilistic Evaluation of Counterfactual Queries," in *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pp. 230–237, Menlo Park, CA: The AAAI Press. [567,570]
- Balke, A., and Pearl, J. (1997), "Bounds on Treatment Effects from Studies with Imperfect Compliance," *Journal of the American Statistical Association*, 92, 1171–1176. [567,574]
- Bonet, B. (2013), "Instrumentality Tests Revisited," arXiv: 1301.2258, 2013. [568]
- Cai, Z., Kuroki, M., and Sato, T. (2007), "Non-parametric Bounds on Treatment Effects with Non-compliance by Ccovariate Adjustment," *Statistics in Medicine*, 26, 3188–3204. [567]
- Cai, Z., Kuroki, M., Pearl, J., and Tian, J. (2008), "Bounds on Direct Effects in the Presence of Confounded Intermediate Variables," *Biometrics*, 64, 695–701. [568]
- Dantzig, G. B. (1963), *Linear Programming and Extensions*, Princeton: Princeton University Press. [567]
- Didelez, V., and Sheehan, N. (2007), "Mendelian Randomization as an Instrumental Variable Approach to Causal Inference," *Statistical Methods in Medical Research*, 16, 309–330. [575]
- Duarte, G., Finkelstein, N., Knox, D., Mummolo, J., and Shpitser, I. (2021), "An Automated Approach to Causal Inference in Discrete Settings," arXiv: 2109.13471. [567]
- Fukuda, K. (2018), *cdd, cddplus and cddlib homepage*. Swiss Federal Institute of Technology, Zurich. Available at [https://people.inf.ethz.ch/fukudak/cdd\\_home/](https://people.inf.ethz.ch/fukudak/cdd_home/). [573]
- Gabriel, E. E., Sachs, M. C., and Sjölander, A. (2020), "Causal Bounds for Outcome-Dependent Sampling in Observational Studies," *Journal of the American Statistical Association*, pp. 1–12. [568,569,570]
- Gabriel, E. E., Sjölander, A., and Sachs, M. C. (2021), "Nonparametric Bounds for Causal Effects in Imperfect Randomized Experiments," *Journal of the American Statistical Association*, pp. 1–9. [568]
- Heckman, J. J., and Vytlacil, E. J. (2001), "Instrumental Variables, Selection Models, and Tight Bounds on the Average Treatment Effect," in *Econometric Evaluations of Active Labor Market Policies in Europe*, eds. M. Lechner and F. Pfeiffer, pp. 1–25, Berlin: Physica-Verlag. [568]
- Imai, K., and Yamamoto, T. (2010), "Causal Inference with Differential Measurement Error: Nonparametric Identification and Sensitivity Analysis," *American Journal of Political Science*, 54, 543–560, 2010. [568]
- Kaufman, J., Kaufman, S., and MacLehose, R. (2009), "Analytic Bounds on Causal Risk Differences in Directed Acyclic Graphs Involving Three Observed Binary Variables," *Journal of Statistical Planning and Inference*, 139, 3473–3487. [568]
- Kaufman, S., Kaufman, J., MacLehose, R., Greenland, S., and Poole, C. (2005), "Improved Estimation of Controlled Direct Effects in the Presence of Unmeasured Confounding of Intermediate Variables," *Statistics in Medicine*, 24, 1683–1702. [567]
- MacLehose, R., Kaufman, S., Kaufman, J., and Poole, C. (2005), "Bounding Causal Effects Under Uncontrolled Confounding Using Counterfactuals," *Epidemiology*, 16, 548–555. [568]
- Motzkin, T., Raiffa, H., Thompson, G., and Thrall, R. (1953), "The Double Description Method," *Contributions to Theory of Games*, 2, 51–74. [573]
- Pearl, J. (2009), *Causality*, New York: Cambridge University Press. [569,575]
- R Core Team. (2019), *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. URL <https://www.R-project.org/>. [568]
- Ramsahai, R. R. (2012), "Causal Bounds and Observable Constraints for Non-deterministic Models," *Journal of Machine Learning Research*, 13, 829–848. [568]
- Richardson, T. S., and Robins, J. M. (2013), *Single World Intervention Graphs (SWIGs): A Unification of the Counterfactual and Graphical Approaches to Causality*. Center for the Statistics and the Social Sciences, University of Washington Series. Working Paper 128. [575]
- Shpitser, I., and Pearl, J. (2007), "What Counterfactuals can be Tested," in *Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence*, pp. 352–359. [570,575]
- Sjölander, A. (2009), "Bounds on Natural Direct Effects in the Presence of Confounded Intermediate Variables," *Statistics in Medicine*, 28, 558–571. [568]
- Sjölander, A., Lee, W., Källberg, H., and Pawitan, Y. (2014a), "Bounds on Sufficient-Cause Interaction," *European Journal of Epidemiology*, 29, 813–820. [568]
- (2014b), "Bounds on Causal Interactions for Binary Outcomes," *Biometrics*, 70, 500–505. [568]
- Tian, J., and Pearl, J. (2000), "Probabilities of Causation: Bounds and Identification," *Annals of Mathematics and Artificial Intelligence*, 28, 287–313. [568]