

---

# Causal Path Computation

---

**Frederick Eberhardt\***

Institute of Cognitive and Brain Sciences  
University of California, Berkeley  
Berkeley, CA 94720

**Richard Scheines**

Department of Philosophy  
Carnegie Mellon University  
Pittsburgh, PA 15213

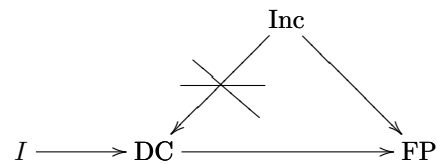
## Abstract

We present an algorithm for causal discovery of linear models using sequences of experiments. The algorithm discovers the true causal structure between the observed variables uniquely, even when there are latent common causes. It detects the latent variables, determines their causal influence, and the output can be analyzed by procedures developed by Silva et al. (2006) for discovery of structure among latent variables. The algorithm works with so-called edge breaking interventions as well as weaker forms of interventions that do not take full control of the intervened variable. Since similar discovery procedures are known to be impossible in discrete models, the results show the power of the linearity assumption.

## 1 INTRODUCTION

A recent BBC article claimed that drug consumption caused family problems.<sup>1</sup> There are two general ways in which one can establish such a causal claim. The first is active, using data collected from an experiment, while the second is passive, using data collected from an observational study. The inferences one can draw about the true causal structure depend on how the data is collected, what tools are available for analysis, and what assumptions one is willing to make about the search space. Concretely, if the analysis is based on (discrete) observational data, and no other information (such as time-order) is available, then the direction of the causal effect is often underdetermined. If there is a correlation between drug consumption and

family problems, and – let us suppose – no other variables are relevant, one can only conclude that drug consumption has an influence on family problems *or* that family problems lead to a change in drug consumption. If, however, the data was obtained from a randomized trial, in which drug doses were randomly assigned to participants (ethical issues aside), and reports of family problems (FP) were found to correlate with drug consumption (DC), then we would have good evidence that drug consumption has a causal effect on family problems. The reasoning is simple: the intervention (I) is “edge-breaking” (Spirtes et al. 2000, Pearl 2000), so randomizing drug consumption makes drug consumption independent of its causes. In particular, if there is a concern that the variables drug consumption and family problems are confounded, by a common cause, such as income (Inc), which may influence both variables, then the randomization would break the influence from income on drug consumption:



It does not matter whether income is a measured or latent variable. In both cases the randomization of drug consumption would break the causal influence of income on drug consumption.

Of course, income may be related to drug consumption and family problems in quite different ways, so we can ask the more general question of whether we can uniquely discover the causal structure between income, family problems and drug consumption, no matter what the causal structure is? Without assumptions about the functional form of the causal relation and with passive observational data alone the answer is No, even when there are no latent variables (i.e. when the set of variables is causally sufficient): There are many causal structures that remain underdetermined

---

Contact: fde@berkeley.edu.

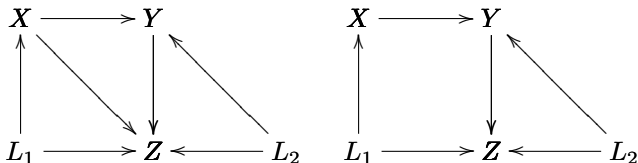
<sup>1</sup>BBC online, 8/28/2007: Drug abuse causes family splits,  
[http://news.bbc.co.uk/2/hi/uk\\_news/6966086.stm](http://news.bbc.co.uk/2/hi/uk_news/6966086.stm).

by independence relations among the measured variables (under the given assumptions) even in the large sample limit. These structures form Markov equivalence classes, in which all structures in the same class imply the same conditional independence relations between the variables.

But suppose we can run a sequence of experiments, each consisting of a randomized trial that varies one variable at a time. Can we then discover the causal structure? If we assume causal sufficiency, the answer is Yes. In Eberhardt et. al. (2006) we proved:

**Theorem 1.1** (Single Interventions, Causally Sufficient).  *$N - 1$  experiments are sufficient and in the worst case necessary to discover the causal structure among  $N$  causally sufficient variables if only a single variable can be randomized per experiment.*

The result provides a worst case bound on the number of experiments for any active search strategy. The bound is determined by the combinatorics of inferences that can be drawn from different experiments, and applies to the large sample limit. Unfortunately, the result is no longer true if there are latent variables: The following causal structures with two latent variables  $L_1$  and  $L_2$  cannot be distinguished by any sequence of experiments randomizing only one variable per experiment (Eberhardt 2007):



So despite the fact that a randomization is an “edge-breaking” intervention, it is impossible to perform a sequence of single intervention experiments that distinguishes the two causal structures. Intuitively, one can see the problem that a randomization of  $X$  will break the incoming  $L_1 \rightarrow X$  edge in both graphs, but that one cannot detect the absence of the  $XZ$ -edge in the second graph because  $X$  and  $Z$  remain dependent even conditional on  $Y$ . The problem, of course, generalizes to many cases beyond this simple counterexample.

One solution is to intervene on more variables simultaneously. In the above graphs, if  $X$  and  $Y$  are simultaneously and independently subject to a randomization in the same experiment, then the two structures can be distinguished, since  $X$  and  $Z$  will appear dependent in the first, but independent in the second. This result can be generalized to arbitrary numbers of variables, but there is a downside: In the worst case any search procedure that reliably discovers the causal structure uniquely (in the large sample limit) requires several

experiments in which a large number of variables are subject to a randomization simultaneously: in at least one experiment all but one variable must be subject to an intervention; if the search procedure is efficient in the number of experiments performed, many more experiments will have intervention set sizes greater than  $N/2$  for  $N$  variables. So the question we explore and answer positively here, is whether we can add other assumptions to the causal search procedure that would make causal discovery in the presence of latent variables feasible without resorting to experiments with large intervention sets.

## 2 TESTING FOR DIFFERENCES IN CORRELATION

The results of the previous section apply to all search procedures based on the assumptions given so far. Traditionally, causal discovery algorithms split into two categories: constraint based and score based algorithms. Score based algorithms, such as the GES-algorithm (Chickering 2002) compute a score for each model given the measured data. The model with the highest score is then deemed to be the most likely or most plausible model given the data. This approach has been generalized to experimental data in Tong & Koller (2001), Murphy (2001) and Cooper and Yoo (1999). In contrast, constraint based algorithms select models according to whether they entail the constraints judged to hold in the data (e.g. the PC- & FCI algorithm: Spirtes et al. 2000). The most general qualitative constraint for causal structure is (conditional) independence. Independence entailments do not depend upon any assumption about the functional form (discrete or continuous) of the causal relation. But of course, other constraints besides independence might be useful and other constraints might well depend on the parameterization of the model space. In particular, we focus here on constraints involving *differences in correlations*.

In the case of a causally sufficient set of variables, the use of correlation differences provides no improvement in the number of experiments needed for causal discovery *in the worst case* – independence tests are all one needs. Of course, there may be particular causal structures that imply certain constraints on correlations (e.g. tetrad constraints) that can be used to identify the causal structure, where independence tests provide no grounds for distinction. But in the case of causal sufficiency, correlation differences do not provide any discovery benefits. However, in the case of causal insufficiency, the impossibility results described in the previous section can be avoided – at least if the true causal model is linear. A causal model is linear if

each variable is a linear combination of its immediate causes and independent noise.

**Theorem 2.1** (Single Intervention, Correlation-Test, Causally Insufficient). *Given a set of  $N$  causally insufficient variables,  $N$  experiments are sufficient and in the worst case necessary to determine the causal graph among the  $N$  observed variables uniquely when only a single intervention is allowed in each experiment and the model is linear.*

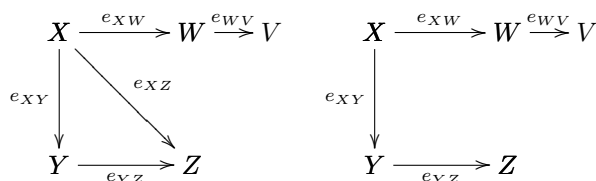
In addition, further analysis allows us to determine the presence and location of latent common causes:

**Theorem 2.2** (Search for Latent Common Causes: Single-Interventions). *Given a set of  $N$  causally insufficient variables and assuming the model is linear,  $N$  single-intervention experiments are sufficient and in the worst case necessary to determine, for each pair of observed variables, whether the pair is confounded by a latent common cause.*

For the worst case, these theorems depend on a sequence of  $N$  experiments in which each experiment  $\mathcal{E}_i$ ,  $0 < i \leq N$  is a randomization of a different variable  $X_i$ . The sequence is – again only for the worst case – order-independent. Proofs of these theorems can be found in Eberhardt (2007)<sup>2</sup>; here we only provide the algorithms that do the job. The theorems signal a significant increase in the information that can be obtained from a sequence of single intervention experiments. Not only do the theorems imply that it is possible to discover the causal structure among the observed variables with single interventions despite the fact that the set of variables is causally insufficient, but one can even discover the location of latent variables. To do so, requires a substantial amount of analysis of the data, as will become evident from the following example and the presentation of the algorithms.

## 2.1 EXAMPLE

We illustrate the result with one very simple example over five variables.



Suppose the true graph over  $\mathbf{V} = \{V, W, X, Y, Z\}$ , with linear edge-coefficients, is given on the left, and assume that each pair of variables is confounded by a latent common cause not shown in the diagram (i.e. there are

<sup>2</sup>See sections 3.3.3 and A.1.3.

$\binom{5}{2} = 10$  latent common causes that are not shown). After performing  $N = 5$  experiments, each involving an intervention on a different variable, we can define a partial ordering ( $\succ$ ) over the variables, such that  $X_i \succ X_j$  if and only if  $X_i \not\perp\!\!\!\perp X_j$  in  $\mathcal{E}_i$ , the experiment in which  $X_i$  is subject to an intervention. In our example we find  $X \succ W \succ V$  and  $X \succ Y \succ Z$ . From this partial order we can construct a partial order graph (POG), which contains a direct edge  $X_i \rightarrow X_j$  if and only if  $X_i \succ X_j$  and there does not exist  $X_k \in \mathbf{V}$  with  $X_i \succ X_k \succ X_j$ . The POG for the partial order here is shown as the graph on the right. It can be shown that the POG is a subgraph of the true graph, connecting any pair of variables that are connected in the true graph. Further, for any two variables connected by a direct edge in the POG, there is no longer directed path between the two variables in the true graph, i.e. if  $X \rightarrow Y$  is in the POG, then there is no path  $X \rightarrow \dots \rightarrow Y$  in the true graph. This knowledge in addition to the fact that a randomization of  $X$  breaks any causal connection between  $X$  and  $Y$  due to a latent or measured common cause, e.g.  $X \leftarrow C \rightarrow Y$ , allows us to infer that the correlation between  $X$  and  $Y$  in the experiment in which  $X$ , but not  $Y$ , is randomized, is entirely due to the direct  $XY$ -edge. Consequently, we can determine the correlation due to the direct edges between any two variables  $X$  and  $Y$  in the POG, by estimating the unconditional correlation between these two variables in the experiment in which the starting point of the edge is subject to intervention. The correlation is equal to the true edge coefficient if it is appropriately normalized.

Now, by considering connections in the POG in an order over the paths that ensures that connections of variables linked by shorter paths in the POG are determined before variables are considered that are further apart<sup>3</sup>, we can test for additional direct edges between any pair of variables  $X_i, X_j$  with  $X_i \succ X_j$ . We compare the total correlation between the two variables in the experiment in which  $X_i$  is subject to an intervention, with the correlation due to the paths between the variables that are already known. Since the intervention on  $X_i$  breaks any influence of latent or measured causes on  $X_i$ , any measured correlation must be due to directed paths from  $X_i$  to  $X_j$  among the observable variables only. The basic idea is to test whether we can account for the correlation between  $X_i$  and  $X_j$  in the experiment in terms of the directed paths from  $X_i$  to  $X_j$  that we already know, and if not, to add appropriate edges. If we start with “closer” connections in the POG, we can ensure that all other direct connections are known before we consider direct connections from vertices high in the graph to ones close to the sink.

<sup>3</sup>See algorithm for details.

Concretely, in our example, we start with paths of length 1 in the POG to determine the edge coefficient of the direct edges. There are four, and we can consider them in any order. Suppose we start from the leaves. We first determine the correlation between  $W$  and  $V$  in the experiment in which  $W$  was subject to an intervention. Due to the way the POG was constructed and the reasons explained above, the estimated unconditional correlation is equal (after normalization) to  $e_{WV}$  on the  $WV$ -edge. Then we estimate the correlation for the  $YZ$ -edge in the experiment in which  $Y$  was subject to an intervention, normalize and have thereby determined  $e_{YZ}$ . Similarly for the  $XY$ - and  $XW$ -edges. We now know the edge coefficients for all the edges in the POG. Next we consider paths of length 2 in the POG:  $X \rightarrow W \rightarrow V$  and  $X \rightarrow Y \rightarrow Z$ . In this case, it does not matter which path we consider first, but this is not generally true.<sup>4</sup> Assume we start with the  $XWV$ -path. We test whether the active correlation  $\rho_{XV}$  between  $X$  and  $V$  in the experiment  $\mathcal{E}_X$ , in which  $X$  was subject to an intervention, is equal to the correlation due to the known path, i.e. whether  $\rho_{XV} = e_{XW}e_{WV}$  (again, assume appropriate normalization). Since this is the case, no direct edge is added from  $X$  to  $V$ . However, in the case of the  $XYZ$ -path,  $\rho_{XZ} \neq e_{XY}e_{YZ}$  and hence an edge  $X \rightarrow Z$  is added and the residual correlation  $e_{XZ} = \rho_{XZ} - e_{XY}e_{YZ}$  is associated with the direct edge. If there were other paths in the POG, they would be considered next, but for this example we are done, we have discovered the structure (and the edge coefficients on that structure) over the observed variables.

We can now search for latent variables by comparing the correlation due to the structure over the observed variables with the passively observed correlation between variables. If they are not equal, there must be latent variables. Let  $L_{AB}$  refer to a latent common cause between variables  $A$  and  $B$  and let  $\delta_{AB}$  be the correlation between  $A$  and  $B$  due to the common cause  $L_{AB}$  (again, like the edge coefficients, appropriately normalized). In this case we start from the root of the graph. If there are several roots, we must check whether any pair of roots is confounded before we consider other pairs of variables. Any pair of roots is confounded just in case they are correlated in the passive observational study.

Next, we first consider whether any direct edge from a root is confounded by a latent variable. Concretely, we test whether the active correlation  $\rho_{XW} = e_{XW}$  is equal to the passively observed correlation  $\tau_{XW}$ . If it

<sup>4</sup>For example if there is a path  $p$  of length 2 and a path  $q$  of length 3 between  $X$  and  $Y$  in a POG, then all subpaths of  $q$  with length 2 need to be considered before the  $p$  is considered. See algorithm.

is, there is no latent common cause between  $X$  and  $W$ . If it is not equal, then we have discovered a latent common cause  $L_{XW}$  and can associate the residual correlation with the latent common cause, i.e. we can determine  $\delta_{XW}$ . Next we consider any paths of length two from the root, and so on. At each step we must consider the correlations due to all (active<sup>5</sup>) paths between variables, including those that involve latent variables that were discovered in previous steps. So when we consider whether there is a latent common cause of  $X$  and  $V$  we have to take into account that there is a causal connection  $X \leftarrow L_{XW} \rightarrow V$  that contributes to the passive observational correlation between  $X$  and  $V$ . But since we know  $\delta_{XW}$  we can take that connection into account.

Once all confounders of the root and any other variable have been considered, pairs of variables in the tier below the root are considered, and then paths, in order of their length, from variables in the tier below the root to variables in lower tiers.

In our example, once we have tested for all latent common causes of the root  $X$  and any other variable, we test in exactly the same way whether there are latent common causes between  $W$  and  $Y$ , since they are both direct descendents of  $X$ , and therefore in the next hierarchical tier. We test whether the passive observational correlation between  $W$  and  $Y$  equals the correlation resulting from all the known paths between the two variables. However, we must take into account correlations due to paths that may include latent variables. Assuming that there are in fact latent common causes between every pair of variables in our example, then we would already have discovered latent variables  $L_{XW}$  and  $L_{XY}$  (among others) in the first round. So for  $Y$  and  $W$  we test whether the passive observational correlation  $\tau_{YW}$  is equal to the sum of  $e_{XW}e_{XY}$  (the correlation due to the  $YXW$ -connection),  $\delta_{XY}e_{XW}$  (the correlation due to the  $YL_{XY}W$  connection) and  $\delta_{XW}e_{XY}$  (the correlation due to the  $YL_{XW}W$  connection). If not, we have discovered  $L_{YW}$ , a latent common cause of  $Y$  and  $W$ , and can establish the correlation that is due to its causal connection. Next we consider latent common causes between the  $YW$ -tier and variables in the tier directly below it, i.e.  $Z$  and  $V$ . In a larger graph we would continue this procedure of considering latent common causes between the  $WY$ -tier and iteratively lower tiers. Once that is completed, we move down one tier, i.e. here to the  $VZ$ -tier, and repeat.

By considering potential latent common causes in a particular top-down order (described precisely in the algorithm), we can ensure that we discover all con-

<sup>5</sup>See algorithm.

founders among variables higher in the graph before we consider confounders of variables closer to the sink. By subtracting the correlations due to all known pathways from the passive observational correlations between two variables, we can identify all latent common causes between two variables, whenever they are present.

### 3 ALGORITHMS

Presentation of the algorithms is simpler if we define some concepts in advance. A post-manipulation graph describes the causal structure among a set of variables given a particular experimental set-up.

**Definition 3.1** (Post Manipulation Graph). *Given a graph  $G = (\mathbf{V}, \mathbf{E})$  and an experiment  $\mathcal{E}$  in which a set  $\mathbf{S} \subseteq \mathbf{V}$  of variables is subject to an intervention, the post-manipulation graph of  $G$  is the graph where all the edges incident on any intervened variable ( $X \in \mathbf{S}$ ) are removed. We will refer to the post-manipulation graph of  $G$  as  $PM(G|\mathcal{E})$ .*

To identify causal connections between two variables we use the notion of an *active path*. An active path is a causal connection that generates a correlation between two variables given a particular conditioning set.

**Definition 3.2** (Active Path). *A vertex  $X$  is active on a path relative to a (conditioning) set of vertices  $\mathbf{C}$  just in case either (i)  $X$  is a collider (common effect), and  $X$  or any of its descendents is in  $\mathbf{C}$ , or (ii)  $X$  is a non-collider and is not in  $\mathbf{C}$ . A path  $p$  is active relative to a set of vertices  $\mathbf{C}$  just in case every vertex on  $p$  is active relative to  $\mathbf{C}$ .*

To start the algorithm, we determine a partial order over the variables using the sequence of experiments:

**Definition 3.3** (Partial Order over Variables:  $O(\mathbf{V}, \succ)$ ). *Given a set of  $N$  experiments on a set of  $N$  (causally insufficient) variables  $\mathbf{V}$  such that each experiment  $\mathcal{E}_i$  is a randomization of a different single variable  $X_i$ , define a partial order ( $\succ$ ) over the set of variables  $\mathbf{V}$ , such that  $X_i \succ X_j$  if and only if  $X_i \not\perp\!\!\!\perp X_j$  in  $\mathcal{E}_i$ .*

We then use a particular order over paths such that shorter paths are considered before longer ones:

**Definition 3.4** (Partial Order over Paths:  $O(\mathcal{P}, \prec)$ ). *Given a set  $\mathcal{P}$  of directed paths, define a partial order over paths in  $\mathcal{P}$  such that for any two paths  $p_1 = X_1 \rightarrow \dots \rightarrow X_r$  and  $p_2 = Y_1 \rightarrow \dots \rightarrow Y_s$ , with  $p_1, p_2 \in \mathcal{P}$  and  $X_1, \dots, X_r, Y_1, \dots, Y_s \in \mathbf{V}$*

1. if  $length(p_1) < length(p_2)$  then  $p_1 \prec p_2$
2. if  $length(p_1) = length(p_2)$  then  $p_1 \prec p_2$  if there exists a path  $p_3 \in \mathcal{P}$ , such that  $p_1 \subset p_3$  and  $p_3 =$

$Y_1 \rightarrow \dots \rightarrow Y_s$  (i.e.  $p_1$  is contained in  $p_3$  and the endpoints of  $p_3$  are the same as those of  $p_2$ ).

We can now specify the algorithm:

**Algorithm 3.5** (Causal Path Computation: Observed Structure). *The algorithm assumes that  $N$  experiments, each involving a randomization of a single variable have been performed and that the true causal model is a linear structural equation model. All correlations are assumed to be appropriately normalized so as to be interchangeable with edge coefficients.*

1. Initialize a graph  $G$  over the variables in  $\mathbf{V}$  that contains no edges.
2. Given the  $N$  experiments, sort the variables in  $\mathbf{V}$  according to the partial order  $O(\mathbf{V}, \succ)$  over variables.
3. For each pair of variables  $X, Y$  such that  $X \succ Y$  and for which there is no other variable  $Z \in \mathbf{V}$  such that  $X \succ Z \succ Y$ , substitute an edge  $X \rightarrow Y$  in  $G$  and determine the correlation  $\rho_{XY}$  from the experiment where  $X$  is subject to an intervention. Let  $e_{XY} = \rho_{XY}$  and associate  $e_{XY}$  with the direct edge  $X \rightarrow Y$ .
4. Sort all directed paths in  $G$  of length greater than two into a partial order over paths  $O(\mathcal{P}, \prec)$ .
5. For each path  $p$  with endpoints  $X$  and  $Y$  in  $O(\mathcal{P}, \prec)$ , starting from the smallest ones in the order, compute the total correlation  $\rho_{XY}$  between  $X$  and  $Y$  from the experiment where  $X$  was subject to an intervention.
6. Let  $\mathcal{P}_{XY}^*$  be the set of all (unconditionally) active paths between  $X$  and  $Y$  in  $PM(G|\mathcal{E}_X)$ , where  $\mathcal{E}_X$  is the experiment in which  $X$  was randomized.
7. Test whether the total correlation  $\rho_{XY}$  between the endpoints  $X$  and  $Y$  in  $\mathcal{E}_X$  can be accounted for in terms of the correlation due to the paths in  $\mathcal{P}_{XY}^*$  alone, i.e. if  $\rho_{XY} = \sum_{p \in \mathcal{P}_{XY}^*} \prod_{e_i \in p} e_i$ , where  $e_i$  is an edge coefficient of an edge on one such path. If so,  $X$  and  $Y$  are determined to be non-adjacent in  $G$ . If not, substitute an edge  $X \rightarrow Y$  in  $G$  and associate the difference in correlation  $e_{XY} = \rho_{XY} - \sum_{p \in \mathcal{P}_{XY}^*} \prod_{e_i \in p} e_i$  with that edge.
8. Paths created by the addition of the new edge are NOT included in  $O(\mathcal{P}, \prec)$ , i.e. the partial ordering over paths is not recomputed.

After all paths in  $O(\mathcal{P}, \prec)$  are considered and the appropriate edges added, the causal structure (and the edge coefficients) among the observed variables are established. We can now proceed to search for latent

common causes. To do so, we define a tier-ordering over the observed variables  $\mathbf{V}$ :

**Definition 3.6** (Tier ordering over observed variables:  $T(\mathbf{V}, \succ_T)$ ). *Given a graph  $G$  a tier ordering  $(T_1 \succ_T \dots \succ_T T_w)$  over  $\mathbf{V}$  is an ordering such that for all  $X \in \mathbf{V}$ ,  $X$  is in exactly one  $T_i$  for  $0 < i \leq w$  and if  $X \in T_i$  for some  $i$ , then either  $i = 1$  or there exists a  $Y \in \mathbf{V}$  with  $Y \in T_{i-1}$  and  $Y \rightarrow X$  in  $G$ .*

The definition implies that all the roots (and unconnected variables) of a graph are in the first tier, and that any variable connected to the root by a directed path of length  $k$  is in the  $(k + 1)^{th}$  tier.

**Algorithm 3.7** (Causal Path Computation: Latent Variables). *Given the graph  $G$  with edge coefficients over the observed variables, determined by the previous algorithm, let  $T(\mathbf{V}, \succ_T)$  be a tier ordering over the variables with  $w$  tiers.*

*For  $i$  from 1 to  $w$  (in that order),*

1. *For all pairs of variables  $X, Y \in T_i$ , with  $X \neq Y$  choose an experiment  $\mathcal{E}_{XY}$  in which  $X$  and  $Y$  are passively observed.*
2. *Let  $\mathcal{P}_{XY}^*$  be the set of all (unconditionally) active paths between  $X$  and  $Y$  in  $PM(G|\mathcal{E}_{XY})$ .*
3. *Compare the correlation  $\tau_{XY}^p$  due to the paths in  $\mathcal{P}_{XY}^*$  with the total passive correlation  $\tau_{XY}$  measured in  $\mathcal{E}_{XY}$ . If  $\tau_{XY}^p = \tau_{XY}$ , then there is no latent common cause of  $X$  and  $Y$ . If not, then add a latent common cause  $X \leftarrow L \rightarrow Y$  to graph  $G$  and associate the difference between the correlations with the path  $X \leftarrow L \rightarrow Y$ .*
4. *For  $j$  in  $i + 1$  to  $w$ , while  $i < j$ ,*
  - (a) *For all pairs of variables  $X, Y$  such that  $X \in T_i$  and  $Y \in T_j$  choose an experiment  $\mathcal{E}_{XY}$  in which  $X$  and  $Y$  are passively observed.*
  - (b) *Let  $\mathcal{P}_{XY}^*$  be the set of all (unconditionally) active paths between  $X$  and  $Y$  in  $PM(G|\mathcal{E}_{XY})$ .*
  - (c) *Compare the correlation  $\tau_{XY}^p$  due to the paths in  $\mathcal{P}_{XY}^*$  with the total passive correlation  $\tau_{XY}$  measured in  $\mathcal{E}_{XY}$ . If  $\tau_{XY}^p = \tau_{XY}$ , then there is no latent common cause of  $X$  and  $Y$ . If not, then add a latent common cause  $X \leftarrow L \rightarrow Y$  to the graph  $G$  and associate the difference between the correlations with the path  $X \leftarrow L \rightarrow Y$ .*

*Return  $G$  with all the latent common causes and all its edge coefficients (or correlations due to the latent common cause).*

Care must be taken when determining the active paths in a particular post-manipulation graph. The experimental intervention might break paths between  $X$  and  $Y$  that would be active if all variables were passively observed. Of course, if an “experiment” is available in which all variables are passively observed, then the determination of active paths between two variables is simple, but there is no necessity for such an experiment, and breaking additional paths by interventions can make the tests of differences in correlation simpler and more reliable.

One of the major advantages of these two algorithms is that correlation tests are unconditional and therefore can be computed on the basis of the entire data set of an experiment. Consequently, the algorithm does not run the risk of any conditional independence tests that – if the conditioning set is large – are impossible, because there is insufficient data. The real work is done in accounting for and testing differences between correlations due to different sets of pathways. These can be intricate tests, since the difference in correlation that a long pathway adds, might be very small. This is less of a problem in the first algorithm – the search for structure among the observed variables – since here the long paths are built up piece by piece from the different experiments and the only tests of differences in correlations occur for additional short direct paths, which would presumably make a significant difference to the total correlation between two variables. The problem of weak correlations due to long paths only bites in the search for the presence of latent variables.

The second algorithm adds a latent common cause only for pairs of variables. Consequently, if there is (in fact) one latent common cause of three variables, the algorithm will render this as three pair-wise common causes. Similar considerations apply if there are more intricate causal relations between latent variables. However, we conjecture, that in certain circumstances we can recover such structure and structure among latent variables by applying the Build-Pure-Clusters-algorithm (Silva et al. 2006). The BPC-algorithm is designed to search for structure among latent variables, and the output our algorithms provide makes that search substantially easier. We have not implemented such a move, but there are two approaches one could take:

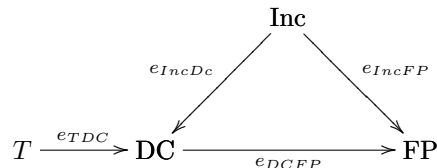
1. Initialize the BPC-algorithm with the known structure and edge coefficients over the observed variables.
2. Remove the correlation due to the observed structure from the overall correlations between variables and run the BPC algorithm on the residual correlations.

If multiple simultaneous interventions per experiment are possible the analysis of the data is largely similar to the above algorithms. However, there are two advantages: First, one can establish the partial ordering over the variables in fewer experiments —  $O(\log_2(N))$  experiments instead of  $N$  experiments — and second, more direct connections between variables can be read off immediately since any indirect path in one experiment can be at most as long as the number of variables not simultaneously subject to an intervention. However, care must be taken that the appropriate passive observational information is available for the second algorithm.<sup>6</sup>

## 4 DISCUSSION

It might appear that the randomization aspect of the interventions is doing the work for these results, i.e. that the results would not be possible if the incoming edges from the latent common causes were not broken by an intervention. We do not always have the luxury of being able to perform randomized interventions. In particular, in the example discussed in the introduction, ethical reasons will prevent a randomized controlled trial on drug consumption. So are our results irrelevant for these cases? The answer is No. As society we do not stand helplessly in light of the problem of drug consumption. We use therapy to treat addicts and while we do not believe that therapy makes drug consumption independent of its normal causes (i.e. it is not an edge-breaking intervention), we do think it can have some influence on the intervened variable. This form of “soft” intervention can be formally characterized as an intervention that influences the conditional distribution of the intervened variable on its causal parents, but does not make the intervened variable independent of its causes: If  $Y$  is a cause of  $X$  then a soft intervention<sup>7</sup>  $I_s$  on  $X$  is such that:  $P(X|Y, I_s = 1) \neq P(X|Y, I_s = 0)$ , but it is not the case (as it would be if  $I_s$  were edge-breaking) that  $P(X|Y, I_s = 1) = P(X|I_s = 1)$ . Of course, a soft intervention of this kind does not break confounding. Can we still use the above algorithms to discover causal structure? – Yes: The above results depend crucially on access to unconfounded variables, not on edge-breaking interventions. Edge-breaking interventions are one way of producing unconfounded variables, but for soft interventions, the *intervention* variables are also unconfounded. Therefore, given that the causal structure is linear there exist sequences of experiments only consisting of soft interventions that

permit the unique discovery of the causal structure for a causally insufficient set of variables. If multiple simultaneous weak interventions are possible, a *single* experiment is in principle sufficient (Eberhardt & Scheines 2006). The analysis proceeds largely analogously to the algorithms presented above, only that differences in correlation need to be tested with regard to the *intervention* variable, not with regard to the *intervened* variable (since the latter remains confounded despite the intervention). That is, if therapy (T) is a soft intervention on drug consumption (DC), as shown in the diagram below,



then the test for the causal effect of DC on FP will be a test of whether there is a correlation between T and FP, i.e. whether  $\rho_{TFP} = 0$ . If it is non-zero, then the causal influence of drug consumption on family problems,  $e_{DCFP}$ , is given by division: i.e.  $e_{DCFP} = \rho_{TFP}/e_{TDC}$ . The technique is very similar to the theory of instrumental variables in econometrics. Sequences of experiments with soft interventions can also be used to detect latent variables. The search algorithm for latent variables is essentially the same as the second algorithm above, but care needs to be taken in how the soft interventions influence the intervened variables.<sup>8</sup>

Without the assumption of linearity, soft interventions are in the worst case of no use to causal discovery in causally insufficient sets of variables, no matter whether single or multiple soft interventions are permitted in each experiment: There always remain causal structures that cannot be distinguished by any sequence of soft-intervention experiments. The ability to use tests of differences in correlations reverses these results despite the softness of the intervention.

We do not believe that these results and algorithms can be extended easily to the discrete case. As our example showed, the results rely on the ability to compute the effect of causal pathways individually, to compute the correlation due to particular subsets of pathways and to be able to compare correlations according to different pathways. This is not always possible for discrete models. Discrete models may contain interactive effects<sup>9</sup>, which prevent an account of a causal effect that can be associated with individual paths.

<sup>6</sup>For more detail, see Eberhardt (2007), sections 3.3.3, 5.3 and 5.4.

<sup>7</sup>For more detail on soft interventions see Eberhardt & Scheines (2006), Korb et al. (2004).

<sup>8</sup>More detail is given in Eberhardt (2007)

<sup>9</sup>Two variables  $X$  and  $Y$  have an interactive effect on variables  $Z$  if the causal effect of  $X$  on  $Z$  is dependent on the state of  $Y$ .

However, we conjecture that these results do hold for discrete models that do not allow interactive causes, such as, for example, noisy-or models. That is, linearity is a sufficient condition, but we conjecture that only some weaker condition such as additivity in the functional form of the model is a necessary assumption.

Lastly, our algorithms consider the worst case. If it is known that not all variables are confounded by latent variables or if the structure among the observed variables is sparse, a variety of adjustments can be made and the number of experiments necessary and sufficient for unique causal discovery can be reduced. The characterization and implementation of the possible adaptive procedures goes beyond this paper. Some results are available in Eberhardt (2007), but many more questions remain to be answered.

## 5 CONCLUSION

We have described an algorithm that uses experimental data to uniquely determine the true causal structure among a set of causally insufficient variables in the large sample limit. It assumes linearity of the causal model, but allows for very weak forms of experimental intervention. We have indicated that linearity is a sufficient condition, but that only much weaker conditions are necessary. However, it is known that for general discrete models similar search procedures are impossible.

Our Causal Pathway Computation algorithms discover the causal structure among the observed variables, they detect the presence and location of latent common causes, and by integrating these results with algorithms that search for structure among latent variables, we conjecture that one is able to discover (within limits) the structure *among* latent variables as well.

## References

- D. M. Chickering (2002). Optimal structure identification with greedy search. *Journal of Machine Learning Research*, 3:507-554.
- G. Cooper and C. Yoo (1999). Causal discovery from a mixture of data. In *Proceedings of the 15th Annual Conference on Uncertainty in Artificial Intelligence*, 116-125, San Francisco, CA.: Morgan Kaufmann.
- F. Eberhardt, C. Glymour, and R. Scheines (2006). N-1 experiments suffice to determine the causal relations among n variables. In D. E. Holmes and L. C. Jain (eds.), *Innovations in Machine Learning*, volume 194 of Theory and Applications Series: Studies in Fuzziness and Soft Computing. Springer-Verlag.
- F. Eberhardt and R. Scheines (2006). Interventions and causal inference. In *Proceedings of the 20th biennial meeting of the Philosophy of Science Association*.
- F. Eberhardt (2007). *Causation and Intervention*, PhD thesis, Carnegie Mellon University, Pittsburgh, PA.
- K. B. Korb, L. R. Hope, A. E. Nicholson, and K. Axnick (2004). Varieties of causal intervention. In C. Zhang, H. W. Guesgen, and W. K. Yeap (eds), *Proceedings of the 8th Pacific Rim International Conferences on Artificial Intelligence*. Springer.
- K. P. Murphy (2001). Active learning of causal bayes net structure. Technical report, Department of Computer Science, U.C. Berkeley.
- J. Pearl (2000). *Causality*. Oxford University Press.
- S. Shimizu, P. O. Hoyer, A. Hyvärinen, and A. J. Kerminen (2006). A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research* 7:2003-2030.
- R. Silva, Glymour, R. Scheines, and P. Spirtes (2006), Learning the Structure of Latent Linear Structure Models, *Journal of Machine Learning Research*, 7, 191-246.
- P. Spirtes, C. Glymour, and R. Scheines (2000). *Causation, Prediction and Search*. MIT Press, 2 edition.
- S. Tong and D. Koller (2001). Active learning for structure in bayesian networks. In *Proceedings of the International Joint Conference on Artificial Intelligence*.