Causal inference with imperfect instrumental variables

Abstract: Instrumental variables allow for quantification of cause and effect relationships even in the absence of interventions. To achieve this, a number of causal assumptions must be met, the most important of which is the independence assumption, which states that the instrument and any confounding factor must be independent. However, if this independence condition is not met, can we still work with imperfect instrumental variables? Imperfect instruments can manifest themselves by violations of the instrumental inequalities that constrain the set of correlations in the scenario. In this article, we establish a quantitative relationship between such violations of instrumental inequalities and the minimal amount of measurement dependence required to explain them for the case of discrete observed variables. As a result, we provide adapted inequalities that are valid in the presence of a relaxed measurement dependence assumption in the instrumental scenario. This allows for the adaptation of existing and new lower bounds on the average causal effect for instrumental scenarios with binary outcomes. Finally, we discuss our findings in the context of quantum mechanics.

Keywords: causal inference with latent variables, instrumental variables, average causal effect, quantum causal models

MSC 2020: 62D20

1 Introduction

Inferring causal relations from data is a central goal in any empirical science. Yet, in spite of its importance, causality has remained a thorny issue. Misled by the commonplace sentence stating that “correlation does not imply causation,” causal inference persists in the view of many as a noble but practically impossible task. Contrary to that, however, the surge and development of the causality theory [1,2] has proven formal conditions under which cause and effect relations can be extracted.

Consider the simplest and fundamental question of deciding whether observed correlations between two variables, $A$ (also known as treatment) and $B$ (also known as effect), are due to some direct causal influence of the first over the second, or due to a common cause, a third, potentially latent (nonobservable)
variable $\Lambda$. Both causal models are observationally equivalent, meaning that both models can generate the same set of possible correlations observed between the target variables. Notwithstanding, causal conclusions can be reached if, instead of passively observing the events, we perform interventions [1,3,4]. In particular, interventions on $A$ put this variable under the experimenter’s control, turning it independent of any latent common cause. If after the intervention, one still observes correlations between $A$ and $B$, then it is possible to unambiguously conclude that $A$ is a cause of $B$. Interventions, however, are often unavailable for various practical, fundamental, or ethical issues.

An elegant way to circumvent such issues is the instrumental variables [1,5–11]. If a proper instrument $X$, correlated with $A$ but statistically independent of $\Lambda$, can be found, then the causal effect of $A$ over $B$ can be estimated even in the absence of interventions or structural equations. Nevertheless, since the instrumental conditions depend on an unobservable variable, identifying an instrument seems to be a matter of judgment that cannot be supported solely by the data. To cope with that, instrumental inequalities have been introduced [12–14], constraints that should be respected by any experiment in compliance with the instrumental assumptions. Thus, the violation of instrumental inequality is an explicit proof that one does not have a proper instrument. Does that mean, however, that no causal inference at all can be made if an instrumental inequality is violated? Or can we still rely on that instrument, even though imperfect, to infer causal relations?

Motivated by these questions, we analyze in detail a generalization of the instrumental causal structure, where we drop the assumption that one has a perfect instrument. More specifically, we relax the assumption that the instrumental variable $X$ should be independent of the latent factor $\Lambda$. Considering the case where $A$ and $B$ are dichotomic and $X$ is also discrete, we derive new instrumental inequalities that take explicitly into account the correlation between $X$ and $\Lambda$. We also generalize the bounds on the average causal effect (ACE) [1,3,4], maximized over all possible realizations of $A$ and $B$, for more general instruments. Although non-standard in the literature, this last maximization comes naturally from the idea of taking the highest possible causal influence from $A$ into $B$.

Finally, we make a connection with the field of quantum foundations, where violation of instrumental inequalities can appear without relaxing the measurement independence assumption [15–18]. Using our results, we establish the minimal measurement dependence needed in the classical instrumental scenario to explain such violations and, as a result, we analyze the robustness of instrumental tests as witnesses of nonclassical behavior.

The article is organized as follows. In Section 2, we discuss how instrumental variables can be employed to put lower bounds on the cause and effect relations between two variables. In Section 3, we discuss the violations of independence assumption and how modified instrumental inequalities and causal bounds on ACE can be derived to take that into account. In Section 4, we discuss quantum violations of the modified inequalities. In Section 5, we discuss our findings and point out interesting questions for future research.

Notations: Throughout the article, we denote random variables by capital letters $A$, $B$, $X$, and $\Lambda$, as well as $\Lambda_x$, $\Lambda_A$, $\Lambda_B$. Without loss of generality, we consider these random variables taking values in the set of non-negative integers $\mathbb{Z}_{0+}$. Probability of an event $E$ is denoted as $p(E)$. We use a common shorthand notation $p(a) = p(A = a)$ to denote the probability of $A$ taking value $a$. Similar shorthand notation is used for conditional probabilities, e.g., $p(a|x) = p(A = a|X = x)$, and interventions, e.g., $p(b|do(a)) = p(B = b|do(A = a))$. We use one exception to this rule for probabilities of the form $p(i, j|k)$ and $p(i|do(m))$, which should be read as $p(A = i, B = j|X = k)$ and $p(B = l|do(A = m))$, respectively, for any $i, j, k, l, m \in \mathbb{Z}_{0+}$. We also use the common notation $|n| = \{0, 1, ..., n - 1\}.$

2 Instrumental variables, instrumental inequalities, and causal bounds

Before getting into details and illustrating the power of an instrumental variable as a causal inference tool, we discuss a simple linear structural model, $b = \beta a + \Lambda$, where $\beta$ can be understood as the strength of the
causal influence of $A$ over $B$ and $\Lambda$ is a latent factor that might affect both $A$ and $B$. By introducing the instrumental variable $X$ and assuming its statistical independence from $\Lambda$, one can infer the causal strength $\beta$. For that aim, it is enough to multiply both sides of the structural equation by $x$ and compute the observed correlations, defined as $\text{corr}(A, B) = \langle A, B \rangle / \langle A \rangle\langle B \rangle$ where $\langle A, B \rangle = \sum_{a,b}(a \cdot b)p(a, b)$ is the expectation value of $A$ and $B$. By doing that, we obtain that $\beta = \text{corr}(X, B)/\text{corr}(X, A)$. It is worth highlighting that the linear structural model is not a requirement. As will be detailed below, under certain assumptions, instrumental variables provide a general tool for estimating causal influences, even in the absence of structural equations.

More formally, an instrumental variable $X$ has only a direct causal influence over $A$ and should be independent of any latent factors acting as a common cause for variables $A$ and $B$. The latter assumption is known by various names such as the independence assumption [8], ignorable treatment assignment [6], no confounding for the effect of $X$ on $B$ [9], and in the literature of quantum foundations is termed as the measurement independence assumption [15–19], an issue of crucial relevance for the violation of Bell inequalities [20,21]. Furthermore, even though $X$ and $B$ might be correlated, those correlations can only be mediated by $A$, that is, there is no direct causal influence of $X$ on $B$, the so-called exchangeability assumption or exclusion restriction [10,11]. See Figure 1(a) for a directed acyclic graph (DAG) description of the instrumental scenario. Altogether, any observed distribution $p(a, b, x)$ compatible with these instrumental conditions should then be decomposable as

$$p(a, b, x) = \sum_{\lambda} p(a|\lambda, x)p(b|\lambda, a)p(x)p(\lambda).$$

Typically, instead of looking at the joint distribution $p(a, b, x)$, one rather considers the conditional distribution $p(a, b|x)$ that, under the same causal assumptions, can be decomposed as

$$p(a, b|x) = \sum_{\lambda} p(a|\lambda, x)p(b|\lambda, a)p(\lambda).$$

The set of probability distributions of the form in equation (2) is bounded in the space of all possible distributions $p(a, b|x)$. These bounds are given by the so-called instrumental inequalities [12–14]. For the simplest case of dichotomic variables there is only one type of instrumental inequalities, which we will call Pearl’s inequality [12] and can be summarized as follows:

$$p(j, 0|0) + p(j, 1|1) \leq 1, \quad p(j, 0|1) + p(j, 1|0) \leq 1, \quad \text{for } j \in \{0, 1\}. \quad (3)$$

Importantly, the instrumental variable can be used for causal inference even in the absence of structural models, something typical in the context of quantum information and refereed there as the device-
independent framework [22,23]. In particular, simply from the observed data distribution \( p(a, b|x) \) one can infer the effect of interventions on the variable \( A \) and thus obtain a lower bound on the \( \text{ACE}_{A \rightarrow B} \) defined as

\[
\text{ACE}_{A \rightarrow B} = \max_{a,a',b} |p(b|\text{do}(a)) - p(b|\text{do}(a'))|,
\]

in which

\[
p(b|\text{do}(a)) = \sum_{\lambda} p(b|a, \lambda)p(\lambda),
\]

and \( \text{do}(a) \) represents the intervention over the variable \( A \). It is worth highlighting that the maximization over \( a, a', b \), as we used in our definition of ACE, is not common in the literature, where it is typically defined as \( E[|B|\text{do}(A = 1)] - E[|B|\text{do}(A = 0)] \), where \( E[|] \) denotes expectation with respect to the probability distribution \( p(b|\text{do}(a)) \). Our aim is to identify a causal influence of one random variable over the other, hence, it is natural to consider the maximum ACE that could occur over possible values of \( a, a', \) and \( b \). The resulting definition is not linear, but piecewise linear (e.g., absolute value for the case of binary \( A \) and \( B \)). However, by fixing the relation between the do-probabilities that enter the ACE definition, e.g., \( p(0|\text{do}(0)) \geq p(0|\text{do}(1)) \), and treating every such case separately, we can effectively linearize the ACE function. Also, since we consider a theoretical study with no particular meaning given to outcomes “0” or “1,” there is always freedom of relabeling the outcomes.

As shown in ref. [3], for the case of binary random variables \( A, B, \) and \( X \) the value of the ACE in equation (4) can be lower-bounded as

\[
\text{ACE}_{A \rightarrow B} \geq 2p(0, 0|0) + p(1, 1|0) + p(0, 1|1) + p(1, 1|1) - 2.
\]

The bound above is particularly relevant because it shows that the effect of interventions can be inferred simply from the observational data. Thus, instrumental variables offer a central tool for situations where interventions are not possible.

The bound in equation (6) is one of the eight expressions given in ref. [3], which provide nontrivial lower bounds on \( \text{ACE}_{A \rightarrow B} \). The three of these eight bounds can be obtained by relabeling the one in equation (6). The remaining four are

\[
\text{ACE}_{A \rightarrow B} \geq p(00|0) + p(11|1) - 1,
\]

and the ones obtained by relabeling of the values of \( X, A, \) and \( B \). However, these four inequalities are not interesting for us, since they continue to hold for any causal structure (including the ones with imperfect instruments). To say it otherwise, these inequalities cannot be violated even if the latent variables (possibly quantum) and the instrumental variables are perfectly correlated.

For the case of more general random variables (not only binary), one can obtain a system of linear inequalities of the form

\[
\text{ACE}_{A \rightarrow B} \geq \max_i C_i,
\]

where \( C_i = \sum_{a,b} c_{i}^{a,b}|p(a, b|x) \) are linear expressions of the probabilities with \( c_{i}^{a,b,x} \in \mathbb{R} \) and the maximum is taken over all such expressions. These lower bounds \( C_i \) can be found using the tools of linear programming [24]. We refer to this type of bounds on the ACE as causal bounds. The causal bound in equation (6) we denote as \( C_1 \). Other bounds studied in this work are given in Section 3.4.

In this work, we focus on the case where the variables \( A \) and \( B \) are binary, i.e., taking values \( a, b \in \{0, 1\} \) but the instrumental variable can take more values (we resort to an arbitrary set of values when discussing the instrumental inequalities and the following two cases \( x \in \{0, 1\} \) and \( x \in \{0, 1, 2\} \), when referring to the problem of causal bounds). At the same time, the methods developed in this article are applicable to the general case where all the random variables take values in arbitrary finite sets. We remark, however, that there are frameworks different from the linear programming approach, we pursue here. For instance, refs [25–27] employ Artstein’s theorem [28] from random set theory to bound continuous functionals of potential outcomes from the literature on treatment effects and offer a valuable tool for when analytical bounds are hard to obtain, as would be the case for variables with high cardinality or even continuous.
3 Relaxing the independence assumption

For the causal bounds such as in equation (6) to hold, one has to guarantee that the instrumental causal assumptions are fulfilled. If any instrumental inequality such as in equation (3) is violated by the observed data distribution \( p(a, b|x) \), then one can unambiguously conclude that at least one of the instrumental assumptions does not hold. Such a violation can have two distinct roots. As shown in refs [23,29–31], even if one imposes the instrumental causal structure to a quantum experiment, still some instrumental inequalities can be violated. This can be seen as a stronger version of Bell’s theorem [20], showing that correlations mediated via quantum entanglement can fail to have a description in terms of standard causal models. The second kind of mechanism, purely classical, and the one we mainly focus on in this article, is the failure of causal assumptions.

For instance, the violation of an instrumental inequality could be motivated by a direct causal influence of \( X \) over \( B \), a violation of the exchangeability assumption (exclusion restriction) shown in Figure 1(b), a scenario for which particular cases have been analyzed in ref. [23] and for which the LP framework we develop can be readily applied. Here, as shown in Figure 1(c) we focus on the violation of the independence assumption. Differently from the typical scenario, we no longer assume that the instrumental variable \( X \) and the common source \( \Lambda \) are independent, that is, \( p(x, \lambda) \neq p(x)p(\lambda) \).

In order to facilitate our analysis, we focus on the DAG including an additional causal link between a latent variable \( \Lambda \) and instrument \( X \) (Figure 1(c)). In this case, the observed probability distribution decomposes as

\[
p(a, b|x) = \frac{1}{p(x)} \sum_{\lambda} p(a|\lambda, x)p(b|\lambda, a)p(x|\lambda)p(\lambda).
\]

(9)

In the following, we argue that without loss of generality, we can consider the response functions \( p(a|\lambda, x), p(b|\lambda, a) \), and \( p(x|\lambda) \) to be deterministic. To see this, consider rewriting \( p(a|\lambda, x) \) using a uniformly distributed auxiliary variable \( \xi_a \) taking values in \([0,1]\) in the following way:

\[
p(a|\lambda, x) = \int \tilde{p}(a|\lambda, x, \xi_a) d\xi_a,
\]

(10)

where \( \tilde{p}(a|\lambda, x, \xi_a) = 1 \) if \( \xi_a \leq p(a|\lambda, x) \) and 0, otherwise. Replacing the original response function with \( \tilde{p}(a|\lambda, x, \xi_a) \) leads to the same observed distribution. We do the same for the other two response functions introducing auxiliary variables \( \xi_b \) and \( \xi_x \). The collection of the newly introduced auxiliary variables can be considered as a part of the global variable \( \Lambda \), resulting in the decomposition involving only deterministic conditional probability distributions. Finally, even though for this argument we made the enlarged confounding variable \( \Lambda \) continuous, it is clear that for the case of discrete observed variables \( A, B \), and \( X \) there is only a finite number of possible combination of deterministic assignments, which means that considering \( \Lambda \) discrete is not a loss of generality. A formal proof of the above argument can be found in refs [32,33] for a related causal scenario of Bell test [20].

It is convenient to treat a realization of \( \Lambda \) as a vector \((\lambda_a, \lambda_u, \lambda_b)\), where \( \lambda_a \in [m_a] \), \( \lambda_u \) takes its values in \([k_u^{m_u}]\) and \( \lambda_b \in [k_b^{m_b}] \). In this relaxed case, any distribution \( p(a, b|x) \) factorizes as follows:

\[
p(a, b|x) = \frac{1}{p(x)} \sum_{\lambda_u, \lambda_b} p(a|x, \lambda_u)p(b|a, \lambda_b)p(x|\lambda_u)p(\lambda_u, \lambda_u, \lambda_b),
\]

where we use the same notation \( p(\cdot) \) for different response functions in order to avoid cumbersome expressions. As argued above, the conditional probabilities \( p(a|x, \lambda_u) \) and \( p(b|a, \lambda_b) \), and \( p(x|\lambda_u) \) can be taken to be deterministic leading to,

\[
p(a, b|x) = \frac{1}{p(x)} \sum_{\lambda_u, \lambda_b} \delta_{f_i(a|x, \lambda_u)}\delta_{g_i(b|a, \lambda_b)}p(x, \lambda_u, \lambda_b),
\]

(11)

where \( f_i(\cdot) \) and \( g_i(\cdot) \) denote deterministic functions, specified by \( \lambda_u \) and \( \lambda_b \), respectively, and we took \( p(x|\lambda_u) = \delta_{x, \lambda_u} \).
In analogy to ref. [17], we use a common measure of dependence between $X$ and $\Lambda$ for the instrumental scenario given by

$$M_{X;\Lambda} = \sum_{x,a,b} |p(x, \lambda_a, \lambda_b) - p(x)p(\lambda_a, \lambda_b)|.$$  \hfill (12)

Crucially to our subsequent analysis, we cast it as the $l_1$-norm of the following vector:

$$M_{X;\Lambda} = \|M\|_1,$$  \hfill (13)

where $q_{a,b,a,b} = p(\lambda, \lambda_a, \lambda_b)$ and for the canonical basis $\{e_{i,j,k}\}_{i,j,k}$ in $\mathbb{R}^{m_k n_k k}$, we have a matrix $M$,

$$M = \sum_{x} \sum_{a,b,a,b} (\delta_{x,\lambda_a} - p(x))e_{x,a,b} \cdot e_{b,a,a,b}.$$  \hfill (14)

Before stating our results, we note that in the literature there are a number of studies dealing with the estimation of causal effects when some assumptions are relaxed. A standard tool is that of the potential outcomes model. Considering a binary treatment $A$, this model implies that

$$B = AB_1 + (1 - A)B_0,$$  \hfill (15)

where $B_0$ and $B_1$ are the potential (nonobserved) outcomes. It is well known that under the assumption of unconfoundedness, implying the conditional independencies $A \perp B_1|W$ and $A \perp B_0|W$, where $W$ is a vector of observed covariates, it is possible to estimate the ACE. Understanding the sensitivity of causal inference to partial conditional independence is a central problem addressed in a number of studies [34–40]. Note, however, that the independence assumption we consider here is native of the instrumental DAG and different of that of the potential outcomes model. Our work is thus complementary to those relaxing unconfoundedness in the potential outcomes model.

Note that the scenario with of a direct cause from $X$ to $A$ is equivalent to the scenario in which $X$ and $A$ are correlated due to confounding variables. This can be verified by noting that the set of probability distributions compatible with the instrumental scenario with the arrow $X \rightarrow A$, described by the following causal model

$$p(a, b, x) = \sum_{\lambda} p(a|x, \lambda)p(b|a, \lambda)p(\lambda),$$  \hfill (16)

is the same set as that of the model where $X \leftarrow \Delta \rightarrow A$, since

$$p(a, b|x) = \sum_{\lambda, \delta} p(a|\lambda, \delta)p(b|a, \lambda)p(\delta|x)p(\lambda) = \sum_{\lambda} p(a|\lambda, x)p(b|a, \lambda)p(\lambda),$$  \hfill (17)

where in the second step we have simply chosen $p(\delta|x) = \delta_{x,\delta}$. Note, however, that the same reasoning does not necessarily apply in the case where we permit the violation of the independence assumption. In this case, one needs to distinguish two situations (Figure 2). In the case when the confounding factor $\Delta$ is directly influenced by $\Lambda$, the analysis of this article can be applied (with a slight modification). As for the second case, in which the confounding factors $\Delta$ and $\Lambda$ are independent, a different method would need to be developed, since the independence condition $p(\Delta \neq i, \Lambda \neq j) = p(\Delta = i)p(\Lambda = j)$ cannot be directly incorporated into a linear program (LP). At the same time, ignoring the independence of $\Delta$ and $\Lambda$ (which are latent factors) would always lead to a valid, but perhaps nontight linear bounds.

### 3.1 Quantifying violation of the independence assumption

The observed correlations in the instrumental experiment given by the observed probability distribution $p(a, b|x)$, as discussed in previous sections, allow us to evaluate the instrumental inequalities or lower bound the strength of the causal influence from $A$ to $B$. Violation of these inequalities implies that the instrumental assumptions were not met in the experiment. As mentioned before, it is important to note that this claim only works if all the latent variables are classical. Crucially, the theory of causality has recently
been generalized to quantum causal modeling [41–47]. In the latter case, the latent variables are quantum states that may be entangled, and the classical variables are obtained through quantum measurements. Quantum causal modeling differs from classical causal modeling in its predictions and as recently demonstrated in refs [23,48,49], if the hidden common cause is allowed to be a quantum entangled state, the bounds obtained for classical instrumental causal structure can be violated. This is true for both instrumental inequalities and causal bounds.

In this article, taking a purely classical perspective on causality, we aim to quantify how much of the above-mentioned violation translates into a relaxation of the independence assumption. More precisely, we aim to find the minimal amount of dependence necessary to explain the violation of either instrumental inequalities or causal bounds.

Given a linear inequality valid for the instrumental scenario $K_{\text{inst}} \geq 0$ (e.g., $K_{\text{inst}} = \text{ACE}_{A \rightarrow B} - \text{C}_i$), if it is violated by a fixed amount $\alpha$, we want to establish what is the minimal amount of dependence, $M_{X;A}$ that could reproduce this violation. Here, we cast this as an optimization problem,

$$\min_{\mathbf{q}} \|M\mathbf{q}\|_1$$

s.t. $K_{\text{inst}} \leq -\alpha,$
$$\sum_{\Lambda_x\Lambda_y\Lambda_z} \mathbf{q}_{x\Lambda_y\Lambda_z} = p(x), \quad \forall x \in [m_x],$$
$$\mathbf{q} \geq 0.$$ (18)

Note that the normalization of $\mathbf{q}$ is implied by the normalization of $p(x)$. We are ready to state our first result.

**Observation 1.** The minimal dependence needed to explain a fixed violation $\alpha$ of a linear inequality valid for the instrumental scenario is a monotonic convex piecewise linear function in $\alpha$.

To see that this statement holds, first we bring the problem in equation (18) to a standard primal form of an LP [50].

$$\max_{\mathbf{q},\mathbf{t}} - \mathbf{t}^\top \cdot \mathbf{t}$$

s.t. $[M, -1] \begin{bmatrix} \mathbf{q} \\ \mathbf{t} \end{bmatrix} \leq \mathbf{0},$
$$[\mathbf{0}, -1] \begin{bmatrix} \mathbf{q} \\ \mathbf{t} \end{bmatrix} \leq \mathbf{0},$$
$$K \cdot \mathbf{t} \cdot \mathbf{q} \leq \begin{bmatrix} \alpha \\ \mathbf{0} \end{bmatrix},$$
$$\Delta \cdot \mathbf{q} \leq \mathbf{p}_x,$$
$$-\Delta \cdot \mathbf{q} \leq -\mathbf{p}_x,$$
$$-\mathbf{q} \leq \mathbf{0}. $$ (19)
In the aforementioned LP, we used the following notations. $\mathbf{1}$ is the vector of 1s, and similarly, $\mathbf{0}$ is the vector of 0s. The matrix $K$ specifies the coefficients in the inequality $K_{\text{inst}}$ and some additional conditions that need to be specified for a specific problem (e.g., the condition on the do-probabilities, under which $\text{ACE}_{A \rightarrow B}$ becomes a linear function. If $\text{ACE}_{A \rightarrow B} = |p(0|do(0)) - p(0|do(1))|$, then the aforementioned condition is either $p(0|do(0)) - p(0|do(1)) \geq 0$ or $p(0|do(0)) - p(0|do(1)) \leq 0$. A matrix $P$ is a probability matrix such that its columns correspond to the deterministic assignments given by $f_{\alpha}(x)$ and $g_{\lambda}(a)$ in equation (11). Finally, $\Delta$ denotes a matrix with entries equal to 1 if the corresponding value of $\lambda_x$ in $q$ is $x$ and 0 otherwise for all values of $x \in [m_x]$. $\mathbf{p}_x$ is a vector of probabilities $p(x)$.

Below, we give the corresponding dual LP to the one in equation (19),

$$
\begin{align*}
\min_{\mathbf{y}, \mathbf{u}, \mathbf{v}, \mathbf{z}} & \quad -\mathbf{a}^T \mathbf{u} + \mathbf{p}_x^T \mathbf{z} \\
\text{s.t.} & \quad \mathbf{M}^T \cdot \mathbf{y} + \mathbf{P}^T \cdot \mathbf{K} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} + \Delta^T \cdot \mathbf{z} \geq \mathbf{0}, \quad \mathbf{0} \leq \mathbf{y} \leq \mathbf{2}, \quad \mathbf{u} \geq 0, \quad \mathbf{v} \geq 0.
\end{align*}
$$

(20)

In the above, we introduced the notation 2, which is a vector of all 2s and $\mathbf{y}, \mathbf{z}, \mathbf{v}$ and $\mathbf{u}$ are the dual variables.

We can see from the above dual formulation of the LP that the solution must be piecewise linear in $\alpha$. Indeed, since the feasibility region of the above LP is a polytope defined by a finite set of constraints, there is a finite set of possibly optimal assignments to $u$ and $z$. Hence, if we change $\alpha$ slowly from 0 to its maximal value, the solution for $u$ might change in at most a finite number of points for $\alpha$. Moreover, it must be clear that for $\alpha = 0$, i.e., in case the inequality $K_{\text{inst}} \geq 0$ is valid, no dependence is required, and thus the output of the optimization problem should be 0. Thus, it must also hold that $\mathbf{p}_x^T \cdot \mathbf{z} = 0$ in the vicinity of $\alpha = 0$.

Since any solution of the above LP, defining the slope $u$ remains a solution for all valid values of $\alpha$, it follows that even though the whole function can be piecewise linear, i.e., have different slopes, these slopes may only increase. In other words, the resulting dependence of $M_{X:A}$ on $\alpha$ is convex and monotonic.

Finally, we must note that the primal problem is feasible, if the violation $\alpha$ is at most the maximum possible, which can be attained by one of the deterministic assignments given by $f_{\alpha}(x)$ and $g_{\lambda}(a)$ in equation (11) expressed as columns of matrix $P$.

### 3.2 Dependencies in the simplest instrumental scenario

Building on the results of this section, here we investigate the minimal required dependence for a fixed violation of instrumental inequalities and bounds on ACE in the simplest instrumental scenario when all the observed random variables are binary. For the both types of inequalities, namely instrumental inequality in equation (3) and causal bound in equation (6), we give exact solutions to the corresponding LPs in equation (20).

**Lemma 2.** For the instrumental scenario with binary observed random variables $X, A, B$ and a latent variable $\Lambda$, the minimal dependence required to explain a violation of the instrumental inequality by $\alpha$ is $M_{X:A} = 4p(X = 0)p(X = 1)\alpha$.

See Appendix A for the proof. We conclude that for a given violation $\alpha$, the uniformly distributed instrumental variable $X$ requires the highest dependence. The reverse also holds true: if the instrumental variable is uniformly random, a given dependence will permit the lowest amount of violation. Our result implies that even though we do not have direct empirical access to the common source between $A$ and $B$, from observational data $p(a, b|x)$ alone we can lower-bound the amount of dependence $M_{X:A}$ present in a given experiment.

Next we investigate how the violation of the lower bound on ACE as in equation (6) translates to the required measurement dependence.
Lemma 3. For the instrumental scenario with binary observed random variables $X, A, B$ and a latent variable $\Lambda$, the minimal measurement dependence required to explain a violation of the lower bound on ACE as in equation (6) by $\alpha$ is $M_{X,\Lambda} = \frac{4p(X = 0)p(X = 1)}{2 - p(X = 0)}\alpha$.

See Appendix B for the proof. Until now, we asked a question which degree of measurement dependence is required to explain violation of a linear inequality (e.g., instrumental inequalities or causal bounds) and we gave an analytical solution for the simplest scenario with binary observed variables. One can, however, ask the reverse question of how the linear inequalities change in the simplest instrumental scenario, if some level of measurement dependence is present in a given setup. This is the inverse problem to the one considered in this section. Since both of these problems aim at estimating the same dependency, they have the same solution, namely the piecewise linear dependence in Observation 1. As a result, we can derive adapted linear inequalities (e.g., binary instrumental inequalities and causal bounds) that accounts for the dependence between $X$ and $\Lambda$, explicitly.

Corollary 3.1. Given a linear inequality valid for the simplest instrumental scenario, $K_{\text{inst}} \geq 0$, the adapted linear inequality in terms of the measurement dependency is

$$K_{\text{inst}} + \frac{M_{X,\Lambda}}{u} \geq 0,$$

where $u$ is the optimization parameter of the dual LP in equation (20).

The above corollary shows that one can still infer cause and effect relations even with nonperfect instruments. Also note that in case of independence, $M_{X,\Lambda} = 0$, we directly recover the inequalities valid in instrumental scenario (Pearl’s inequality in equation (3) and the causal bound in equation (6)).

For a more general case, when the instrumental variable can take more than two values, adapting a linear inequality valid for the perfect instrumental scenario is also possible. However, it is a more involving task as the minimal measurement dependence does not have to be linear in the observed violation, as pointed out in Observation 1. We give numerical treatment for this problem in Section 3.4 and in Figure 3.

3.3 Informational cost

Above, we used the $l_1$-norm (see equation (12)) to quantify the level of dependence in the instrumental scenario. Another common measure used to quantify the dependence between two random variables is the information cost [16,18], given by the Shannon mutual information, a measure of particular relevance in the entropic approach to causal inference [51–53]. In this case, we are interested in quantifying $I(X; \Lambda) = H(X) - H(X|\Lambda)$, where $H(X) = -\sum_x p(x) \log p(x)$ is the Shannon entropy of $X$ and $H(X|\Lambda)$ is the conditional Shannon entropy of $X$ given $\Lambda$, respectively, and logarithm is taken to be base 2. In particular, we ask a question of the minimal required information cost $I(X; \Lambda)$ that would allow for a violation of instrumental inequality in equation (3). For convenience, let us again use the notation

$$K_{\text{inst}} = -p(0, 0|0) - p(0, 1|1) + 1.$$

If no dependence between $X$ and $\Lambda$ is allowed, then $K_{\text{inst}} \geq 0$. We are now ready to present our next result.

Lemma 4. For the instrumental scenario with binary observed random variables $X, A, B$ and a latent variable $\Lambda$, with $X$ uniformly distributed, the minimal informational cost required to explain a value $K_{\text{inst}} < 0$ of instrumental inequality is $I(X; \Lambda) = 1 - h\left(1 - \frac{K_{\text{inst}}}{2}\right)$, where $h(p) = -p \log p - (1 - p) \log(1 - p)$ is the binary entropy.

See Appendix C for the proof. The same result applies to any of the four instrumental inequalities in equation (3).
So far, we have restricted our attention to the case where all variables are binary. Here, we generalize the results for the instrumental variable, which can take more values.

Concerning instrumental inequalities, if the variables $A$ and $B$ are binary, it is known that the instrumental scenario is completely characterized by three inequalities up to the relabelings of the variables, $I_i \leq 0, i = 1, 2, 3$ [11]. The inequality $I_1 \leq 0$ corresponds to Pearl's inequality and was already discussed in the binary case (see equation (3)), the second one is known as Bonet’s inequality [13],

$$ p(0, 1|0) - p(0, 1|1) - p(1, 1|1) - p(1, 0|2) - p(0, 1|2) \leq 0, $$

and the third one is Kedagni’s inequality [11],

$$ p(0, 0|0) + p(1, 0|0) - p(0, 1|1) - p(1, 0|1) - p(0, 0|2) - p(1, 0|2) - p(0, 0|3) - p(1, 1|3) \leq 0. $$

One can obtain other inequalities from refs [11,13] by relabeling inputs and outputs and by coarse graining values of $X$.

Considering the case where $X$ assumes up to three possible values, we obtained two new classes of causal bounds, for which we give two representatives below. All the other causal bounds for three inputs can be obtained by relabeling inputs or outputs in these two inequalities.

$$ C_2 = p(0, 0|0) + p(0, 0|2) + p(1, 0|0) + p(1, 1|1) + p(1, 1|2) - 2. $$

$$ C_3 = p(0, 0|0) + p(0, 0|1) - p(0, 1|1) + p(0, 1|2) + p(1, 0|0) - p(1, 0|1) + p(1, 1|1) + p(1, 1|2) - 2. $$

**Figure 3:** Measurement dependence $M_{X|A}$ for violations $\alpha$ of instrumental inequality $I_2$ (left), and causal bounds $C_2$ (center) and $C_3$ (right). The maximal violation of each inequality attainable in quantum theory is marked by $Q_{\text{max}}$ (Section 4).
For all the causal bounds and the instrumental inequalities we use the LP in equation (18) to estimate the minimal measurement dependency in order to explain the violation by the amount of $\alpha$. The results are summarized in Figure 3.

Even though we only provide analytical solutions of the LPs in the simplest binary case, in more general scenarios, it is sufficient to solve LPs only in a very few points due to the nature of the functional dependence being convex piecewise linear. For example, for the instrumental inequality $I_2 \leq 0$, the numerical results in Figure 3 suggest that for the chosen fixed distributions of $X$, the minimal measurement dependence $M_{X,A}$ is linear in $a$. We could, however, reach the same conclusion by solving the LP for two different values of $a$ in the interval $a \in (0, 1]$ for the same fixed distributions of $X$. The first value can be arbitrary, but the second one must be equal to $a = 1$. Additionally, we know that for $a = 0$, the measurement dependence $M_{X,A} = 0$. If the values of minimal measurement dependency corresponding to these three points belong to the same straight line, we invoke the convexity property, and conclude that $M_{X,A} = ua$, where $u$ is the slope of the obtained straight line. For example, for the uniformly distributed $X$, $M_{X,A} = \frac{2}{3}a$, where the coefficient of $\frac{2}{3}$ can be obtained up to the numerical precision.

4 Quantum violations of instrumental tests

We saw that the instrumental and causal bounds can be violated if a certain amount of measurement dependence is present between an instrumental variable and a common cause $\Lambda$. However, a violation is also possible if we do not assume any relaxation on the instrumental scenario, but instead we consider the case, when the unobserved common cause can be a quantum state [23,48,49]. Since Bell's theorem [20], it is known that classical causal models are unable to explain the correlations observed by measurements on entangled quantum systems. In other terms, if we impose a given causal structure to a quantum experiment, the probability distribution obtained by employing the quantum rules cannot be explained if we assume a classical causal model where each node, including the confounder, corresponds to a random variable.

More precisely, the postulates of quantum theory impose that (i) the state of a physical system is described by a positive, trace-1 linear operator, known as the density matrix $\rho$, acting on a Hilbert space, (ii) an observed outcome is the result of a measurement, characterized by a positive operator-valued measure (POVM), which is a set of linear operators $M_i$ respecting the completeness relation $\sum_i M_i = 1$ (1 being the identity operator in the Hilbert space), (iii) the probability of a given outcome is given by the Born’s rule $p(i) = \text{Tr}(M_i \rho)$, (iv) a composite quantum system is modeled by a tensor product of Hilbert spaces. This set of rules implies a well-defined description of the theory where the probability of all outcomes of an experiment are positive and sum up to one. Applying the quantum postulates to the instrumental scenario considered here, all three observable variables, $X$, $A$, and $B$, are still classical random variables, but instead of the latent variable $\Lambda$, we have a latent quantum state $\rho_{AB}$ of a composite system acting on a tensor product of Hilbert spaces. This type of quantum causal model produces observable correlations given by

$$p_Q(a, b|x) = \text{Tr}((M_a^x \otimes N_b^x)\rho_{AB})$$

where $\rho_{AB}$ is a quantum state of two subsystems, represented by the density matrix acting on the tensor product of two Hilbert spaces $\mathcal{H}_A \otimes \mathcal{H}_B$, $\{M_a^x\}$ is a POVM acting on the first subsystem (Hilbert space $\mathcal{H}_A$) and describes a measurement depending on the choice $x$ with the outcome $a$. Similarly, $\{N_b^x\}$ is a POVM acting on the second subsystem (Hilbert space $\mathcal{H}_B$) and describes a measurement depending on the choice $a$ (which is the measurement outcome obtained on the first subsystem) with the outcome $b$.

Remember that the bounds such as instrumental inequalities and causal bounds derived in refs [3,11,13] assume that the latent node is a classical random variable. Consequently, they can all be systematically obtained using linear programming (for the case of discrete observed random variables). On the other hand,
once we allow quantum common causes and use Born’s rule for probabilities, the description of the set of all possible observed distributions \( p(a, b|x) \) becomes a much more tedious problem [23,48]. In particular, it is not, in general, possible to give a description of the observed probabilities in the form of equations (2) and (5). Since quantum theory encapsulates the classical probability theory as a special case, the former set is known to be strictly larger. Consequently, we cannot expect that the derived inequalities, which are the results of the linear programming, also hold for quantum common causes. Valid bounds for quantum common causes can be obtained using more general convex optimizations tools, such as infinite hierarchies of semidefinite programs [54] or trace inequalities [48]. Finding such valid bounds for quantum common causes is, however, not the aim of this section. Instead, we compare here the violations of classical bounds due to quantum common causes and due to relaxations of the independence assumption. More precisely, we establish the amount of the minimal measurement dependence needed to explain quantum effects in classical causal models.

Interestingly, in the case of the simplest instrumental scenario, the statistics obtained from a latent quantum state cannot violate the instrumental inequalities in equation (3) and no further analysis on minimal measurement dependency is required. Thus, in the simplest instrumental scenario, inequalities can be violated only if causal assumptions are relaxed.

This is no longer the case for the causal bounds on ACE in the simplest instrumental scenario. Refs [23,48] show that in the case of binary observed variables, the causal bound in equation (6) can be violated if the common cause is a quantum state (without any measurement dependence). In order to demonstrate such a violation, in a full analogy with the classical case, one can define (see ref. [47]) interventions on a classical random variable \( A \), which is an outcome of a quantum measurement, as

\[
p_q(\{b\mid do(a)\}) = \text{tr}[(1 \otimes N^a_B)\rho_{AB}],
\]

The identity operator \( I \) in the above formula represents the fact that the outcome of the measurement on the first subsystem, generating \( A \) is discarded, and instead the value \( a \) is supplied to the measurement that produces \( B \). Alternatively, it can be understood as simply “doing nothing” to the first subsystem. The observed quantum average causal effect (qACE) is then given by the difference,

\[
q_{\text{ACE}_{A-B}} = \max_{a,a',b} |\text{tr}[(1 \otimes (N^a_B - N^{a'}_B))\rho_{AB}]|.
\]

Ref. [48] considered the following quantum causal model. The common cause is given by a pure two-qubit entangled state (rank-1 density operator with both Hilbert spaces being two-dimensional complex vector spaces \( \mathbb{C}^2 \)). Such states can be represented as \( \rho_{AB} = |\psi\rangle\langle\psi| \), with \( |\psi\rangle = \sin \alpha |0\rangle \otimes |0\rangle + \cos \alpha |1\rangle \otimes |1\rangle \), where \( |0\rangle, |1\rangle \in \mathbb{C}^2 \) are basis vectors in \( \mathbb{C}^2 \) in the so-called Dirac notation. Note that a quantum state is entangled if it cannot be expressed as a tensor product of the states of the individual subsystems, which is the case for \( 0 < \alpha < \pi/4 \) in the considered example. We say that the state is maximally entangled if \( \sin \alpha = \cos \alpha = 1/\sqrt{2} \). Entanglement is the central phenomenon in quantum mechanics and its applications, and it has only very recently been analyzed in the context of quantum causal discovery [23,48,55,56].

In ref. [48] it was shown that for all \( 0 < \alpha < \pi/4 \), and for appropriately chosen quantum measurements \( \{M^a_{X,A}|N^a_{X,B}\} \), the causal bound in equation (6) can be violated. The degree of violation depends on the parameter \( \alpha \), which governs the amount of entanglement present in the common cause. The maximal possible violation was numerically obtained (and was verified by the hierarchy of semidefinite programs [54]) to be \( 3 - 2\sqrt{2} \approx 0.1716 \).

We use the results of the previous sections to explain the maximal known quantum violation in the classical instrumental scenario with the relaxed measurement independence assumption: the amount of minimum measurement dependency in the classical instrumental causal structure must at least be \( M_{X;A|X} = \frac{q_{\text{p}(X = 0|X = 1)}(3 - 2\sqrt{2})}{2 - p(X = 0)} \). This quantity is maximal for \( P(X = 0) = (2 - \sqrt{2}) \), and equals \( M_{X;A} = (68 - 48\sqrt{2}) \approx 0.1177 \). Note that we are indeed imposing the instrumental causal structure to the quantum experiment. What the violation of instrumental inequality shows is that the only way to find a classical explanation (a causal model described by random variables) for the observed probability
distribution is if some of the assumptions in the instrumental scenario (in this the independence assumption) are relaxed. We see that given a DAG, the quantum description of it will always lead to a larger (or at least equal) set of compatible correlations with the DAG.

In case of more general instrumental scenario, where $X$ can take more than two values, $I_2 \leq 0$ and $I_1 \leq 0$ can be violated by quantum states and measurements [23], both by the maximally entangled state (i.e., when $\sin \alpha = \frac{1}{\sqrt{2}}$) with the amount of \( \left( \frac{1}{\sqrt{2}} - \frac{1}{2} \right) \approx 0.207 \) and $\sqrt{2} - 1 \approx 0.4142$, respectively. See Figure 3 (left) for the relation between the minimal required measurement dependence in classical instrumental scenario and the amount of violation of $I_2 \leq 0$ for various probability distributions of the instrumental variable. In particular, for uniformly distributed instrumental variables the minimal measurement dependence required to explain the quantum violation is $\mathcal{M}_{X,A} = \frac{1}{2} (\sqrt{2} - 1) \approx 0.1381$. The minimal measurement dependence needed to explain the maximal quantum violation of $I_2 \leq 0$ for the uniformly distributed instrumental variables is $\mathcal{M}_{X,A} = \left( \frac{1}{\sqrt{2}} - \frac{1}{2} \right) \approx 0.2071$.

Finally, we consider the quantum violation of the causal bounds for the instrument that takes three values. The inequality $\text{ACE}_{4-B} \geq C_3$ can be violated by the maximally entangled state with the amount of $\left( \frac{1}{\sqrt{2}} - \frac{1}{2} \right) \approx 0.2071$ and the inequality $\text{ACE}_{4-B} \geq C_3$ with the amount of $\sqrt{2} - 1 \approx 0.4142$. In order to explain these violations, the amount of minimum measurement dependency in the classical instrumental causal structure depends on a probability distribution of the instrumental random variable. See Figure 3 (center) and (right) for the particular examples of such distributions. We highlight that even though the violations of causal bounds match with the violations of instrumental inequalities, these quantities are of a very different nature. In particular, the violation of causal bounds required both interventional and observational probability distributions, while the violation of instrumental inequalities relies solely on observational data.

5 Discussion

Instrumental variables offer ways to estimate causal influence even under confounding effects and without the need for interventions. Strikingly, as discovered in ref. [3], one can infer the effect of interventions, without resorting to any structural equations, simply from observational data obtained with the help of an instrument. As already recognized long ago [57], however, "the real difficulty in practice of course is actually finding variables to play the role of instruments." Since the potential correlation of the instrument with any latent variables is in principle unobservable, it might seem that the exogeneity of a given instrument is a matter of trust and intuition rather than a fact supported by the data.

Motivated by this fundamental problem, the data from an instrumental test [11–14] can be employed to benchmark the amount of dependence the instrument can have with a confounding variable. More precisely, we quantify such correlations via a $l_1$-norm, measuring by how much the instrumental variable fails to be exogenous. The violation of an instrumental inequality allows us then to put lower bounds on this dependence. In turn, we derive bounds for the ACE [1] taking into account that some level of dependence, lower bounded by the violation of instrumental inequality, is present. That is, we turn the causal bounds in a reliable tool even if the instrument is not really exogenous.

Relying on an LP description, we obtain fully analytical results for the simplest instrumental scenario where all variables are binary. We study a more general case of trinary instrumental variable numerically using our linear programming technique. In parallel, we also derived new bounds for the ACE (equation (25)), which to the best of our knowledge are new to the literature. We also consider applications of our generalized instrumental inequalities and causal bounds to consider the problem of measurement independence (also known as “free-will”) in the foundations of quantum physics.

It is interesting to compare the lower bounds on the dependence between $\Lambda$ and $X$ obtained either from the instrumental inequality or the causal bound violations, when the observational data include both the distribution $p(a, b|x)$ and the estimated value of the ACE. The required amount of dependence between
instrumental variables and confounders that would explain such violations would then be at least the maximum among the tested inequalities. However, more generally, one would need to solve the LP described in Section 3 with all the tested inequalities included as constraints. The required amount of the relaxation of the independence assumption also provides a way to compare different inequalities in their bounding power in regard to the considered causal structure.

It is worth noting that the effect of imperfect instruments has previously been considered [6,58–61]. Those works, however, relied on a number of additional assumptions, for instance, the study in [59] was limited to regression bivariate models, while here our results are free of any structural equations and valid for any causal mechanisms between the variables. Even though we have focused on the case where treatment and effect variables are binary, the LP framework we propose can also be extended to variables assuming any discrete number of values (limited, of course, by the computational complexity of the problem). Another interesting question for future research is to understand whether similar results may hold for the case of continuous variables or consider statistical tests as in ref. [62], a direction that we hope might be triggered by our results.

Acknowledgments: The authors thank anonymous Referees for providing useful comments and numerous interesting references. We acknowledge partial support by the Foundation for Polish Science (IRAP project, ICTQT, contract no. MAB/2018/5, co-financed by EU within Smart Growth Operational Programme). N.M. is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 441423094. This work was supported by the John Templeton Foundation via the grant Q-CAUSAL No 61084 (the opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the John Templeton Foundation) Grant Agreement No. 61466, by the Serrapilheira Institute (grant number Serra – 1708–15763), by the Simons Foundation (Grant Number 884966, AF), the Brazilian National Council for Scientific and Technological Development (CNPq) via the National Institute for Science and Technology on Quantum Information (INCT-IQ) and Grant Nos. 406574/2018-9 and 307295/2020-6, the Brazilian agencies MCTIC and MEC. M.G. is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – Cluster of Excellence Matter and Light for Quantum Computing (ML4Q) EXC 2004/1 – 390534769.

Conflict of interest: Authors state no conflict of interest.

References

[1] Pearl J. Causality. Cambridge University Press: Cambridge, UK; 2009.
[2] Spirtes P, Glymour CN, Scheines R, Heckerman D. Causation, prediction, and search. MIT Press: Cambridge, MA, USA; 2000.
[3] Balke A, Pearl J. Bounds on treatment effects from studies with imperfect compliance. J Amer Statist Assoc. 1997;92(439):1171–6.
[4] Janzing D, Balduzzi D, Grosse-Wentrup M, Schölkopf B. Quantifying causal influences. Annal Statist. 2013;41(5):2324–58.
[5] Wright PG. Tariff on animal and vegetable oils. New York: Macmillan Company; 1928.
[6] Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. J Amer Statist Assoc. 1996;91(436):444–55.
[7] Greenland S. An introduction to instrumental variables for epidemiologists. Int J Epidemiol. 2000;29(4):722–9.
[8] Rassen JA, Brookhart MA, Glynn RJ, Mittelman MA, Schneeweiss S. Instrumental variables i: instrumental variables exploit natural variation in nonexperimental data to estimate causal relationships. J Clin Epidemiol. 2009;62(12):1226–32.
[9] Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist’s dream? Epidemiology. 2006;17(4):360–72.
[10] Lousdal ML. An introduction to instrumental variable assumptions, validation and estimation. Emerging Themes Epidemiol. 2018;15(1):1–7.
[11] Kédagni D, Mourifié I. Generalized instrumental inequalities: testing the instrumental variable independence assumption. Biometrika. Feb 2020;107:661–75.
Agresti I, Poderini D, Guerini L, Mancusi M, Carvacho G, Aolita L, et al. Experimental device
Bonet B. Instrumentality tests revisited. 2013. arXiv: http://arXiv.org/abs/arXiv:1301.2258.

Abellán C, Acín A, Alarcón A, Alibart O, Andersen CK, Andreoli F, et al. Challenging local realism with human choices.
Pearl J. On the testability of causal models with latent and instrumental variables. In: Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence; 1995. p. 435–43.

Hall MJW. The significance of measurement independence for bell inequalities and locality. In: At the frontier of spacetime. Springer: Cham, Germany; 2016. p. 189–204.

Chaves R, Brask JB, Gross D. Unifying framework for relaxations of the causal assumptions in bell’s theorem. Phys Rev Lett. 2015;114(14):140403.

Chaves R, Moreno G, Polino E, Poderini D, Agresti I, Suprano A, et al. Causal networks and freedom of choice in bell’s theorem. 2021. arXiv: http://arXiv.org/abs/arXiv:2105.05721.

Wood CJ, Spekkens RW. The lesson of causal discovery algorithms for quantum correlations: Causal explanations of bell-inequality violations require fine-tuning. New J Phys. 2015;17(3):033002.

Bell JS. On the Einstein Podolsky Rosen paradox. Phys Physique Fizika. 1964;1(3):195.

Abellán C, Acín A, Alarcón A, Alibart O, Andersen CK, Andreoli F, et al. Challenging local realism with human choices. Nature. 2018;557(7704):212–6.

Pironio S, Scarani V, Vidick T. Focus on device independent quantum information. New J Phys. 2016;18(10):100202.

Chaves R, Carvacho G, Agresti I, DiGiulio V, Aolita L, Giacomini S, et al. Quantum violation of an instrumental test. Nature Phys. 2018;14(3):291–6.

Boyd S, Boyd SP, Vandenberghe L. Convex optimization. Cambridge University Press: Cambridge, UK; 2004.

Molchanov I, Molinari F. Applications of random set theory in econometrics. Annu Rev Econ. 2014;6(1):229–51.

Cheshire A, Rosen AM. Generalized instrumental variable models. Econometrica. 2017;85(3):959–89.

Russell TM. Sharp bounds on functionals of the joint distribution in the analysis of treatment effects. J Business Econ Statist. 2021;39(2):532–46.

Artstein Z. Distributions of random sets and random selections. Israel J Math. 1983;46(4):313–24.

Nery RV, Taddel MM, Chaves R, Aolita L. Quantum steering beyond instrumental causal networks. Phys Rev Lett. 2018;120(14):140408.

Van Himbeeck T, Brask JB, Pironio S, Ramanathan R, Sainz AB, Wolfe E. Quantum violations in the instrumental scenario and their relations to the bell scenario. Quantum. 2019;3:186.

Agresti I, Poderini D, Guerini L, Mancusi M, Carvacho G, Aolita L, et al. Experimental device-independent certified randomness generation with an instrumental causal structure. Commun. Phys. 2020;3(1):1–7.

Peres A. All the bell inequalities. Foundations Phys. 1999;29(4):589–614.

Brunner N, Cavalcanti D, Pironio S, Scarani V, Wehner S. Bell nonlocality. Rev Modern Phys. 2014;86(2):419.

Masten MA, Poirier A. Identification of treatment effects under conditional partial independence. Econometrica. 2018;86(1):317–51.

Masten MA, Poirier A, Zhang L. Assessing sensitivity to unconfoundedness: Estimation and inference, 2020. arXiv: http://arXiv.org/abs/arXiv:2012.15716.

Rosenbaum PR, Rubin DB. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. J R Stat Soc B (Methodological). 1983;45(2):212–8.

Robins JM, Rotnitzky A, Scharfstein DO. Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In: Statistical models in epidemiology, the environment, and clinical trials. Springer: New York, NY, USA; 2000. p. 1–94.

Rosenbaum PR. Overt bias in observational studies. In: Observational studies. Springer: New York, NY, USA; 2002. p. 71–104.

Ichino A, Mealli F, Nannicini T. From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity? J Appl Econom. 2008;23(3):305–27.

de Luna X, Johansson P. Testing for the unconfoundedness assumption using an instrumental assumption. J Causal Infer. 2014;2(2):187–99.

Leifer MS, Spekkens RW. Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference. Phys Rev A. 2013;88(5):052130.

Fritz T. Beyond bell’s theorem ii: Scenarios with arbitrary causal structure. Commun Math Phys. 2016;341(2):391–434.

Henson J, Lal R, Pusey MF. Theory-independent limits on correlations from generalized Bayesian networks. New J Phys. 2014;16(11):113043.

Chaves R, Majenz C, Gross D. Information-theoretic implications of quantum causal structures. Nature Commun. 2015;6(1):1–8.

Pienaar J, Brukner C. A graph-separation theorem for quantum causal models. New J Phys. 2015;17(7):073020.

Costa F, Shrapnel S. Quantum causal modelling. New J Phys. 2016;18(6):063032.
[47] Allen J-MA, Barrett J, Horsman DC, Lee CM, Spekkens RW. Quantum common causes and quantum causal models. Phys Rev X. 2017;7(3):031021.

[48] Gachechiladze M, Miklin N, Chaves R. Quantifying causal influences in the presence of a quantum common cause. Phys Rev Lett. Dec 2020;125:230401.

[49] Agresti I, Poderini D, Polacchi B, Miklin N, Gachechiladze M, Suprano A, et al. Experimental test of quantum causal influences. 2021. arXiv: http://arXiv.org/abs/arXiv:2108.08926.

[50] Chaves R, Kueng R, Brask JB, Gross D. Unifying framework for relaxations of the causal assumptions in bell’s theorem. Phys Rev Lett. Apr 2015;114:140403.

[51] Fritz T, Chaves R. Entropic inequalities and marginal problems. IEEE Trans Inform Theory. 2012;59(2):803–17.

[52] Chaves R, Luft L, Maciel TO, Gross D, Janzing D, Schölkopf B. Inferring latent structures via information inequalities. 2014. arXiv: http://arXiv.org/abs/arXiv:1407.2256.

[53] Budroni C, Miklin N, Chaves R. In distinguishability of causal relations from limited marginals. Phys Rev A. 2016;94(4):042127.

[54] Navascués M, Pironio S, Acín A. Bounding the set of quantum correlations. Phys Rev Lett. 2007;98(1):010401.

[55] Wolfe E, Pozas-Kerstjens A, Grinberg M, Acín A, Navascués M. Quantum inflation: A general approach to quantum causal compatibility. Phys Rev X. 2021;11(2):021043.

[56] Ligthart LT, Gachechiladze M, Gross D. A convergent inflation hierarchy for quantum causal structures. 2021. arXiv: http://arXiv.org/abs/arXiv:2110.14659.

[57] Johnston J, DiNardo J. Econometric methods. Econom Theory. 1963;16:139–42.

[58] Manski CF. Identification problems in the social sciences. Harvard University Press: Cambridge, MA, USA; 1995.

[59] Bartels LM. Instrumental and "quasi-instrumental" variables. Am J Polit Sci. 1991;777–800.

[60] Rosenbaum PR. Using quantile averages in matched observational studies. J R Statist Soc C (Applied Statistics). 1999;48(1):63–78.

[61] Small DS. Sensitivity analysis for instrumental variables regression with overidentifying restrictions. J Amer Statist Assoc. 2007;102(479):1049–58.

[62] Wang L, Robins JM, Richardson TS. On falsification of the binary instrumental variable model. Biometrika. 2017;104(1):229–36.
Appendix A
Proof of Lemma 2

Proof. All the binary instrumental inequalities are given in equation (3). We choose one of them (the results work for any other choice too, due to symmetry present in the problem) and insert it into the primal problem,

\[ K_{\text{inst}} = -p(0|0) - p(0|1) + 1. \]  

(A.1)

In the dual LP in equation (20), the matrix \( K \) is \( 1 \times 8 \), which is a matrix representation of the expression \( K_{\text{inst}} \) above. The matrix \( P \) is \( 8 \times 32 \) with each column corresponding to a deterministic assignment of \( X, A, \) and \( B \) given \( \lambda \). The vector \( z \) has two components, which we call \( z_0 \) and \( z_1 \) and the vector \( y = [y_0, y_1, \ldots, y_3] \) is 32-dimensional. Finally, there is no vector \( v \) in our LP, as there are no additional linear constraints in \( K \).

From the definition of \( M \), we derive that

\[ M_{pX}^{\text{pX}} y \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} \tilde{y}_0 & \cdots & \tilde{y}_{15} \end{bmatrix}, \text{ where } \tilde{y}_i = y_i - y_{i+16}, \text{ i } \in \{0, \ldots, 15\}. \]  

Moreover, note that \( \Delta^T z = \begin{bmatrix} z_0 \\ z_1 \end{bmatrix} \). Taking all the above into account, the LP takes the following form:

\[
\begin{align*}
\min_{\tilde{y}, u, z_0, z_1} & \quad -au + p(X = 0)z_0 + p(X = 1)z_1 \\
\text{s.t.} & \quad p(X = 1)\tilde{y}_i - p(X = 0)\tilde{y}_i + \begin{bmatrix} p(X = 1) \\ p(X = 0) \end{bmatrix} [P^T \cdot K_i] u + z_0 \geq 0, \quad \forall i \in [16], \\
& \quad -p(X = 0)\tilde{y}_i + \begin{bmatrix} p(X = 1) \\ p(X = 0) \end{bmatrix} [P^T \cdot K_i] u + z_1 \geq 0, \quad \forall i \in [16], \\
& \quad -2 \leq \tilde{y}_i \leq 2, \quad u \geq 0.
\end{align*}
\]  

(A.2)

Here \( [P^T \cdot K_i] \) is the \( i \)th term of the vector \( P^T \cdot K \). For \( i \in \{0, \ldots, 15\} \), the expression \( [P^T \cdot K_i] \) can take one of the two possible values, either \( \left( 1 - \frac{1}{p(X = 0)} \right) \) or \( 1 \), and for \( i \in \{16, \ldots, 31\} \), it can take one of the two possible values \( \left( 1 - \frac{1}{p(X = 0)} \right) \) or 1. This simplifies the problem and by erasing redundant constraints we arrive at the final form of the LP which we solve explicitly.

\[
\begin{align*}
\min_{\tilde{y}, u, z_0, z_1} & \quad -au + p(X = 0)z_0 + p(X = 1)z_1 \\
\text{s.t.} & \quad p(X = 1)\tilde{y}_i - p(X = 0)\tilde{y}_i + \begin{bmatrix} p(X = 1) \\ p(X = 0) \end{bmatrix} u + z_0 \geq 0, \quad \forall i \in [1, 2], \\
& \quad -p(X = 0)\tilde{y}_i + \begin{bmatrix} p(X = 1) \\ p(X = 0) \end{bmatrix} u + \frac{p(X = 1)}{p(X = 0)} z_1 \geq 0, \\
& \quad i \in [0, 2], \\
& \quad -2 \leq \tilde{y}_i \leq 2, \quad i \in \{0, 1, 2\}, \quad u \geq 0.
\end{align*}
\]  

(A.3)

By summing the first and the last inequalities for \( i = 0 \), we directly obtain, \( p(X = 0)z_0 + p(X = 1)z_1 \geq 0 \). Finally, summing up the two inequalities, where the variable \( u \) has a negative coefficient, we obtain an upper-bound on \( u \),

\[
\begin{align*}
    u & \geq (\tilde{y}_0 - \tilde{y}_2)p(X = 0)p(X = 1) + p(X = 0)z_0 + p(X = 1)z_1 \\
    & \geq 4p(X = 0)p(X = 1) + p(X = 0)z_0 + p(X = 1)z_1.
\end{align*}
\]  

(A.4)

(A.5)

Using the upper-bound on \( u \) we obtain that the objective function can be lower-bounded by the expression

\[
-4ap(X = 0)p(X = 1) + (p(X = 0)z_0 + p(X = 1)z_1)(1 - a) \geq -4ap(X = 0)p(X = 1).
\]  

(A.6)

As the final step, we note that the assignment: \( u = 4p(X = 0)p(X = 1), \tilde{y}_0 = 2, \tilde{y}_1 = -2, \tilde{y}_2 = 0, \) \( z_0 = 2p(X = 1)(1 - 2p(X = 0)) \), and \( z_1 = -\frac{p(X = 0)}{p(X = 1)}z_0 \) is a feasible point of the LP. Thus, \( M_{X,A} = 4p(X = 0)p(X = 1)a \). \( \square \)
B Proof of Lemma 3

**Proof.** The proof has a similar structure as the one of Lemma 2; however, it is more involving. The main reason for this is that the expression $K_{\text{inst}} = \text{ACE}_{a-b} - C_1$ is written not only in terms of probabilities $p(a, b|x)$, but also in terms of do-probabilities. Additionally, by definition ACE is not linear in do-probabilities, but we can linearize it without loss of generality by requesting that $p(0|do(0)) - p(0|do(1)) \geq 0$. In the dual LP in equation (20), the matrix $K$ is then $2 \times 12$, which is a matrix representation of the expression $K_{\text{inst}} = \text{ACE}_{a-b} - C_1$. The matrix $P$ is $12 \times 32$ with each column corresponding to a deterministic assignment of $X$, $A$ and $B$ given $A$ (which also gives deterministic assignments to the do-probabilities). The vector $z$ has two components, which we call $z_0$ and $z_1$ and the vector $y = [y_0, y_1, \ldots, y_{15}]$ is 32-dimensional. Finally, there is only a single element in vector $v$ in our LP, which corresponds to the positivity of $p(0|do(0)) - p(0|do(1)) \geq 0$.

The matrix $M$ is the same as in Lemma 2,

$$M^T y = \begin{bmatrix} p(X=1) & \cdots & p(X=i) & \cdots & p(X=16) \end{bmatrix} y,$$

where $-2 \leq y \leq 2$ is a column vector,

$$y^T = [\tilde{y}_0 \ldots \tilde{y}_{15}],$$

where $\tilde{y}_i = y_i - y_{i+16}$, $i \in \{0, \ldots, 15\}$, and $\Delta^T z = \begin{bmatrix} z_0 \\ z_1 \end{bmatrix}$. We need to solve the following LP,

$$\begin{align*}
\min & \quad -au + p(X=0)z_0 + p(X=1)z_1 \\
\text{s.t.} & \quad p(X=1)\tilde{y}_i + [P^T \cdot K_{i,0}] u + [P^T \cdot K_{i,1}] v + z_0 \geq 0, \quad \forall i \in [16], \\
& \quad -p(X=1)\tilde{y}_i + \frac{p(X=1)}{p(X=0)}([P^T \cdot K_{i,1}]_{16,0} u + [P^T \cdot K_{i,1}]_{16,1} v) + \frac{p(X=1)}{p(X=0)}z_1 \geq 0, \quad \forall i \in [16], \\
& \quad -2 \leq y \leq 2, \quad u \geq 0, \quad v \geq 0,
\end{align*}$$

(B.1)

where we denoted by $[P^T \cdot K_{i,j}]$ the element of the matrix $P^T \cdot K^T$ on $i$th row and $j$th column (counting from 0).

We give rows of $[K \cdot P]$ here for completeness: $[K \cdot P]_0 = \begin{bmatrix} 2p(X=0) & 3p(X=0) & \cdots & 3p(X=1) \end{bmatrix}$, $[K \cdot P]_1 = \begin{bmatrix} 3p(X=0) & 3p(X=1) & \cdots & 3p(X=1) \end{bmatrix}$, $[K \cdot P]_2 = \begin{bmatrix} 3p(X=0) & 3p(X=1) & \cdots & 3p(X=0) \end{bmatrix}$, $[K \cdot P]_3 = \begin{bmatrix} 3p(X=0) & 3p(X=1) & \cdots & 3p(X=0) \end{bmatrix}$, $[K \cdot P]_4 = \begin{bmatrix} 3p(X=0) & 3p(X=1) & \cdots & 3p(X=0) \end{bmatrix}$, $[K \cdot P]_5 = \begin{bmatrix} 3p(X=0) & 3p(X=1) & \cdots & 3p(X=0) \end{bmatrix}$, $[K \cdot P]_6 = \begin{bmatrix} 3p(X=0) & 3p(X=1) & \cdots & 3p(X=0) \end{bmatrix}$, $[K \cdot P]_7 = \begin{bmatrix} 3p(X=0) & 3p(X=1) & \cdots & 3p(X=0) \end{bmatrix}$.

First, we derive an upper-bound on $u$. For the feasibility region the following must hold true for any $i, j \in [16]$ (which one gets simply by summing the two types of aforementioned constraints),

$$p(X=1)(\tilde{y}_{i,j} - \tilde{y}_i) + u([P^T \cdot K^T]_{i,0} + \frac{p(X=1)}{p(X=0)}([P^T \cdot K^T]_{i,1} + \frac{p(X=1)}{p(X=0)}[P^T \cdot K^T]_{i,16,1} + \frac{p(X=1)}{p(X=0)}z_1$$

$$+ v([P^T \cdot K^T]_{i,1} + \frac{p(X=1)}{p(X=0)}[P^T \cdot K^T]_{i,16,1}) \geq 0.$$  

(B.2)

For $i = 5$ and $j = 5$, the values $[P^T \cdot K^T]_{5,0} = 3p(X=0) - 3p(X=0) / p(X=0)$, $[P^T \cdot K^T]_{5,1} = 3p(X=1) - 3p(X=1) / p(X=0)$, $[P^T \cdot K^T]_{5,16} = -1$, and $[P^T \cdot K^T]_{5,16} = -1$ lead to the condition $p(X=0)z_0 + p(X=1)z_1 \geq v$. For $i = 0$ and $j = 2$, for which $[P^T \cdot K^T]_{0,0} = 2 - \frac{2}{p(X=0)}$ and $[P^T \cdot K^T]_{0,1} = 1 - \frac{1}{p(X=0)}$, we obtain

$$\begin{bmatrix} 1 \应注意 2 - \frac{2}{p(X=0)} \end{bmatrix} (y_0 - 2y_2) + \frac{p(X=0) - 2}{p(X=0)} + \frac{p(X=1)}{p(X=0)} v + z_0 + \frac{p(X=1)}{p(X=0)}z_1 \geq 0,$$

which means that $u \leq \frac{1}{2 - p(X=0)}(4p(X=0)p(X=1) + p(X=1)v + p(X=0)z_0 + p(X=1)z_1)$, since $\tilde{y}_0 - \tilde{y}_2 \leq 4$.

Inserting this value in the objective function, we obtain,

$$\begin{align*}
-\frac{\alpha}{2 - p(X=0)}(4p(X=0)p(X=1) + p(X=1)v + p(X=0)z_0 + p(X=1)z_1) & \geq \frac{-\alpha \cdot 4p(X=0)p(X=1)}{2 - p(X=0)} + (p(X=0)z_0 + p(X=1)z_1)
\end{align*}$$

(B.4)

$$\begin{align*}
\geq \frac{-\alpha \cdot 4p(X=0)p(X=1)}{2 - p(X=0)} + (p(X=0)z_0 + p(X=1)z_1)
\end{align*}$$

(B.5)
\[ \geq -\alpha 4p(X = 0)p(X = 1) + \nu(1 - \alpha) \geq -\alpha 4p(X = 0)p(X = 1) \quad \text{(B.6)} \]

The last step follows as \( \alpha \leq 1 \). As the final step, we note that the assignment: \( u = \frac{4p(X = 0)p(X = 1)}{2 - p(X = 0)}, \nu = 0, \)
\( \hat{y}_0 = \hat{y}_1 = \hat{y}_4 = \hat{y}_5 = \hat{y}_7 = \hat{y}_{10} = \hat{y}_{16} = 2, \quad \hat{y}_2 = \hat{y}_3 = \hat{y}_8 = \hat{y}_9 = \hat{y}_{11} = -2, \quad \hat{y}_6 = \hat{y}_{13} = 2 - \frac{4p(X = 0)}{2 - p(X = 0)}, \)
\( \hat{y}_{11} = \hat{y}_{15} = -\frac{2p(X = 0)}{2 - p(X = 0)}, \) and \( z_0 = -2p(X = 1)(1 - \frac{4p(X = 1)}{2 - p(X = 0)}), \quad z_1 = -\frac{p(X = 0)}{p(X = 0)}, \quad \) is a feasible point of the LP, which means that the lower bound of \( -\alpha \frac{4}{2 - p(X = 0)} \) on the objective function is achievable. \( \square \)

C Proof of Lemma 4

Proof. We rewrite the conditional join probabilities occurring in the expression \( K_{\text{inst}} \) using decomposition in equation (11) and the following notations for the deterministic assignments \( \Lambda_{\beta}^{(0)} = \{\lambda_B | p(B = 0 | A = 0, \lambda_B) = 1\} \) and \( \Lambda_{\beta}^{(1)} = \{\lambda_B | p(B = 1 | A = 0, \lambda_B) = 1\} \).
\[
\frac{1 - K_{\text{inst}}}{2} = \sum_{\lambda_a, \lambda_b \in \Lambda_{\beta}^{(0)}} \sum_{p(X = 0, \lambda_a, \lambda_b)} p(X = 0, \lambda_a, \lambda_b) + \sum_{\lambda_a, \lambda_b \in \Lambda_{\beta}^{(1)}} p(X = 1, \lambda_a, \lambda_b)
\]
\[\leq \sum_{\lambda_a, \lambda_b \in \Lambda_{\beta}^{(0)}} p(X = 0, \lambda_a, \lambda_b) + \sum_{\lambda_a, \lambda_b \in \Lambda_{\beta}^{(1)}} p(X = 1, \lambda_a, \lambda_b)
\]
\[= \sum_{\lambda_b \in \Lambda_{\beta}^{(0)}} p(X = 0, \lambda_b) + \sum_{\lambda_a, \lambda_b \in \Lambda_{\beta}^{(1)}} p(X = 1, \lambda_b)
\]
\[= p(X = 0, \lambda_b \in \Lambda_{\beta}^{(0)}) + p(X = 1, \lambda_b \in \Lambda_{\beta}^{(1)}) = p(X = E),
\]
where \( E \) is a random variable such that \( E = 0 \) if \( \lambda_b \in \Lambda_{\beta}^{(0)} \), and \( E = 1 \) if \( \lambda_b \in \Lambda_{\beta}^{(1)} \). Since \( E \) concerns a particular grouping of latent variable \( \Lambda \), we can first use the data processing inequality and then Fano's inequality to obtain,
\[
I(X; \Lambda) \geq I(X; E) = H(X) - H(X|E) \geq 1 - h(X = E) \geq 1 - h\left(\frac{1 - K_{\text{inst}}}{2}\right).
\]
The last inequality follows since we are only interested in the cases when \( K_{\text{inst}} < 0 \). The above lower bound is tight for all \( K_{\text{inst}} < 0 \), since we can make the following assignments: \( p(A = 0 | X = 0, \lambda_a) = p(A = 0 | X = 1, \lambda_a) = 1, \)
\( p(X = 0 | \lambda_a, \lambda_b \in \Lambda_{\beta}^{(0)} = p(X = 1 | \lambda_a, \lambda_b \in \Lambda_{\beta}^{(0)}) = \frac{1 - K_{\text{inst}}}{2}, \forall \lambda_a, \) and \( p(\lambda_B \in \Lambda_{\beta}^{(0)}) = p(\lambda_B \in \Lambda_{\beta}^{(1)}) = \frac{1}{2}. \) \( \square \)