Elements of Partial Identification

Jakob Zeitler

October 8, 2021

Overview

Causal inference provides the fundamental causal reasoning that machine learning is missing to effectively tackle decision making problems. So far, full identification of causal effects has been the focus of the majority of research: Strong and mostly untestable assumptions, such as no unmeasured confounding, yield point estimates of how a sprint will increase my endurance by 2% or how $10k more in savings will get my loan application accepted. Ideally, we would want to make fewer strong assumptions, but still provide informative suggestions. Partial identification enables this by calculating lower and upper bounds on the true causal effect which are more trustworthy due to more realistic assumptions. Unfortunately, current partial identification methods practically do not scale due to super-exponential parameter growth in the number of variables. Hence, I am developing scalable methods that trade-off computational cost with tightness of bounds. Exact bounding approaches will be crucial to high-stake decision making problems such as AI fairness, which require provable guarantees. Approximate methods will find use in environments valuing execution cost over guarantees, such as personal exercise recommendations or prioritisation of user experience experiments. Both approaches will become fundamental building blocks for trustworthy, causal machine learning.

Contents

1 Introduction

Causal inference provides the fundamental causal reasoning that machine learning is missing to effectively tackle decision making problems. So far, full identification of causal effects has been the focus of the majority of research: Strong and mostly untestable assumptions, such as no unmeasured confounding, yield point estimates of how a sprint will increase my endurance by 2% or how $10k more in savings will get my loan application accepted. Ideally, we would want to make fewer strong assumptions, but still provide informative suggestions. Partial identification enables this by calculating lower and upper bounds on the true causal effect which are more trustworthy due to more realistic assumptions. Unfortunately, current partial identification methods practically do not scale due to super-exponential parameter growth in the number of variables. Hence, I am developing scalable methods that trade-off computational cost with tightness of bounds. Exact bounding approaches will be crucial to high-stake decision making problems such as AI fairness, which require provable guarantees. Approximate methods will find use in environments valuing execution cost over guarantees, such as personal exercise recommendations or prioritisation of user experience experiments. Both approaches will become fundamental building blocks for trustworthy, causal machine learning.

2 Motivation

How many milliliters of intravenous fluid should I inject in my patient who is dying due to multiple organ failure from sepsis?

Despite centuries of medical research, there still is no consensus on the optimal treatment (Byrne and Van Haren 2017). The dominant scientific paradigm that has emerged for these kind of questions suggests to conduct experimental falsification of a hypothesis about the world — also known as the randomized controlled trial (RCT). For example, we can test the hypothesis of treating sepsis patients with 100 milliliter of IV fluid and observe disease progression, recording each recovery and death, comparing it with a group of patients who received no IV fluid.

Assigning a group of patients to a treatment and control group at random is still often misunderstood as the only way to quantify a causal effect of a drug, e.g. aspirin. Indeed, this random assignment ensures that a central assumption of causal inference is given as an empirical fact, i.e. that the calculated effect by randomization is justified to be the true causal effect. In the language of potential outcomes (PO) (Hernán and Robins 2020), randomization ensures *exchangeability* or *ignorability*, i.e. that patients can be exchanged between the treatment and control group without any change in the calculated causal effect. In the language of structural causal models (SCM), randomization breaks the arrows from the confounding factors, preventing the assignment of the patient to either group to be based on attributes such as age or previous diseases.

But even if we do not have data from an RCT, i.e. randomization is not given such that the assignment to treatment or control was based on characteristics of the patients, we can still try to calculate the true causal effect. This is also referred to as an *observational study* as we do not randomize assignment, but only observe data. As long as we have measured all the attributes of the patient, often referred to as *confounders*, that were used during the assignment process, we can adjust for those attributes. The calculated effect is then corrected by the amount of *confounding bias* these variables introduce, hence, they are more accurately described as *deconfounders*, as they de-confound the effect. (Dawid 2015) Having access to deconfounders means that we have knowledge of the treatment/assignment mechanism, also known as fulfilling SCM's *backdoor-criterion* or PO's *conditional exchangeability/ignorability*. With this knowledge, we can then apply a variety of methods to adjust for confounding, e.g. the backdoor-adjustment, inverse probability weighting, propensity score matching and doubly robust methods. (Pearl 2009,Hernán and Robins 2020)

Unfortunately, we can never guarantee that we have measured all confounders, such that our set of deconfounders should never be assumed to be complete. In the best case, this will introduce neglegible bias, and in the worst case reverses the causal direction. Therefore, observational studies require assumptions that are hard to justify, namely *no unmeasured confounding*. If we are able to justify this assumption, it allows *identification* (Pearl 2012) of the causal effect, which in simple terms is a single number with a concrete interpretation. Procedures for identification of effects of interests in different causal models have been widely discussed, most notably via the *do-calculus* which provides a procedure to systematically transform a query for an effect of interest operational on the given data and causal model. (Pearl 2009)

But identification procedures often require strong assumptions, as explained before. As an alternative to identification, one can focus on achieving *partial identification* by making fewer assumptions. Partial identification does not yield a single number, but two numbers: a lower bound and an upper bound that provably include the true causal effect. Because bounds require fewer assumptions, they are easier to justify. But most bounds are not informative, i.e. they include the number zero, such that the we cannot even make a statement about the direction of the causal effect.

This report introduces a method for partial identification that estimates bounds on joint interventions, while only having access to single intervention data, e.g. single treatment RCTs, and observational data. Specifically, the method aims to provide an approach that is feasible with more than 5 variables in an SCM, as an alternative to current bounding methods which have prohibitive computational costs due to super-exponential parameter growth in the number of variables.

3 Related Work

(Saengkyongam and Silva 2020) inspires the causal problem setting of this report. The authors prove full identifiability in the presence of hidden confounding of joint interventions from single and observational regime data. They assume a non-linear continuous structural causal models as well as additive Gaussian noise. This report assumes the same causal problem of estimating unobserved joint interventions, but does not assume non-linear continous SCMs and additive Gaussian noise. As such, it faces the technical problem of partial identification, i.e. calculating bounds on joint interventions.

Partial identification was first extensively discussed in (Manski 1990), who also provides a textbook (Manski 2003). (Balke and Pearl 1994) discuss bounding in SCMs, focusing on the IV model and providing a method to reformulate SCMs with response function variables (RVFs). (Ramsahai 2012) provides a bounding formulation without any PO assumptions, see Appendix A.

Recently, the continuous case has been discussed more widely, albeit reduced to the instrumental variable model. (Kilbertus, Kusner, and Silva 2020) provide an algorithm that allows continuous treatments, while (Zhang and Bareinboim 2021a) present an approch for bounding with continuous outcomes, both in the IV model.

Retrieving constraints to tighten bounds from general graphs is increasing in popularity with varying success. A common starting point is the IV model, as for example in (Sachs, Gabriel, and Sjölander 2020). Focusing on measurement error, (Finkelstein 2020) try to extract linear constraints from a general class of SCMs, while (Zhang and Bareinboim 2021b) characterize a class of canonical graphs using reformulations similar to RFVs, but are not able to provide a procedure to calculate tight bounds from this canonical class. Detailed reviews on partial identification, formulation and implementation are available in (Richardson et al. 2014), (Wu et al. 2019) and (Swanson et al. 2018)

4 Fundamentals of Bounding of Causal Effects

The following section explains the fundamentals of how to bound causal effects.

4.1 Motivation for Bounding of Causal Effects

In the case of unmeasured confounding, when identification of causal effects is not possible without further assumptions, an easily overlooked question is:

Is there still some information in the observational and experimental measurement data a practitioner can act on?

The answer is yes. It is possible to calculate bounds on the effect of interest. For example, to understand how effective aspirin is at removing headaches, facing unmeasured confounding in the data, once can still calculate bounds on the effect, that are possible to act on, which might look like this:

$$(\text{Lower Bound}, \text{Upper Bound}) = (+0.04, +0.42)$$

With these bounds, it is not possible to tell what the *exact* effect of aspirin on headaches will be, but the effect will definitely be positive, which is a piece of information a doctor can still make decisions with. (Pearl 2020) provides an interactive diagram on causal bounds with a recent application to need-based patient treatment.

Most importantly, these bounds are not to be confused with confidence intervals, which make a statement about the certainty of a measured variable of interest to fall between two values. Causal bounds, or partial identification, provide a range where the causal effect will be found, i.e the effect will never fall outside these bounds (ignoring from random variability or causal insufficiency).

4.2 Intuition for Bounding via Response Functions

Bounding of causal effects via response function variables was introduced by (Balke and Pearl 1994). The approach is most easily understood via the example of data with imperfect compliance evaluated with an instrumental variable model.

The problem of *imperfect compliance* When doctors prescribe aspirin, but patients do not stick to their prescription, i.e. what their doctor told them to do, doctors face the problem of *imperfect compliance*. For example, if a doctor sees 100 patients in a week, and tells 50 of them to take aspirin, but only 40 patients take the aspirin as prescribed, then 10 patients did not comply, i.e. they are defiers of the prescription (50-40=10).

Imperfect compliance here leads to unreliable estimates of the causal effect of aspirin on relieving headaches. If patients are treated with a prescription Z and the outcome Y of this aspirin is measured, the effect will be biased as 10 patients will be accounted for as having received the treatment, when in fact they did not as they did not comply to the prescription.

One possible remedy is to keep track of compliance itself, i.e. who ends up taking the aspirin. For example, it is possible to calculate the effect of taking aspirin, i.e. the compliance level X, onto the outcome Y. But in reality, there will still be reasons why people did not comply to the doctor's prescription. If these reasons also influence the effectiveness of the drug for a patient, then there is confounding of X and Y, i.e. there is a (usually unmeasured) confounding factor U, that influences both the compliance to the prescription X and the outcome Y of taking the drug. If there is complete knowledge of these reasons U, they can be adjusted for via the backdoor criterion. If not all the reasons are known, the problem is a case of unmeasured confounding, which is the more realistic case. The instrumental variable model, as described next, will help with this unmeasured confounding.

Causal effects with imperfect compliance via the instrumental variable model While the instrumental variable model (IV model) can be applied to all kinds of settings with unmeasured confounding, it is most easily understood from the perspective of imperfect compliance. The instrumental variable models helps adjust for the bias introduced by imperfect compliance via unmeasured confounding. Instead of looking at the compliance X and the outcome Y alone, it also takes into account the prescription of the patients, denoted by Z. The IV model follows the following structure:

$$X = g(Z, U)$$

i.e. compliance to aspirin prescriptions is influenced by the prescription Z and other factors U.

$$Y = g(X, U)$$

i.e. outcome of aspirin treatment is influenced by taking aspirin.

There are three conditions that an IV model needs to meet for unbiased inference. All conditions concern the actual instrumental variable (IV) Z. A variable is considered an IV if it fulfills the following conditions:

a. $Z \perp\!\!\!\perp U$

- **Condition: Unconfounded Instrument Z**
- Z is independent from confounders (measured and unmeasured) U

b. $Z \not\perp\!\!\!\perp X$

- **Condition: Relevance of Z for X**
- Z has information on X

c. $Z \perp\!\!\!\perp Y | \{X, U\}$

- **Condition: Condition: Exclusion (of the IV Z as part of all variables influencing Z)**
- Given X and U, Y is independent of Z

These conditions have intuitive explanations: The treatment X is confounded with the outcome Y, so calculating the effect from X on Y is susceptible to bias from confounders. If there was an additional variable, call it Z, that has information on X, i.e. a variable that moves together with it, but is not confounded with Y, then it can 'jump in' for X and 'substitute' for it during the effect calculation. If Z and X move together, then the effect of Z on Y should get quite close to the effect of interest of X on Y. It is important to ensure that Z does influence Y directly, but only indirectly through X. Overall, the IV Z needs to have information on X (Relevance), not be confounded with Y via confounders U (Unconfoundedness) and not influence Y directly (Exclusion Criterion)

Causal bounds with imperfect compliance via the instrumental variable model At the core of the imperfect compliance problem are the doctor's prescription and the compliance by the patients. The doctors prescribes aspirin (Z=1) or not (Z=0). The patients ends up taking the aspiring (X=1) or not (X=0). These event split the possible type of responses of patients into **4 unique response groups**:

Group	$Z =$ (prescription)	$X =$ (response to the prescription)
Always Denier	0	$X = 0$ (always deny, so always 0)
	1	$X = 0$ (always deny, so always 0)
Compliers	0	$X = Z = 0$ (comply, so Z)
	1	$X = Z = 1$ (comply, so Z)
Defiers	0	$X = 1 - Z = 1$ (defy, so 1-Z)
	1	$X = 1 - Z = 0$ (defy, so 1-Z)
Always Taker	0	$X = 1$ (always take, so always 1)
	1	$X = 1$ (always take, so always 1)

The 4 groups are defined as follows:

- *Always Deniers* never take aspirin, under any circumstances, with or without prescription, i.e. $X = 0$ ('hardcoded')
- *Compliers* always follow the prescription of the doctor, i.e. $X = Z$
- *Defiers* always defy the prescription of the doctor, always doing the opposite i.e. $X =!Z(= 1 - Z)$
- *Always Takers* always take aspirin, regardless of the prescription, i.e. $X = 1$

Hence, the problem of patients not abiding to prescription is called the problem of *imperfect* compliance, as some of the patients fall out of the group of compliers into any of the other three. Always Deniers would not be a problem if they were not given a prescription, as they behave as the doctor prescribed, and vice versa for Always Takers. But the calculation of a causal of effect of taking aspirin X onto the outcome of reducing headaches Y will always be biased by patients that are Defiers. Usually, most patients follow their prescription, but because of the Defiers of prescriptions, a method to deal with imperfect compliance is requireed.

Because these four groups of patients give us an *exhaustive* and *complete* list of all possible reactions of patients to their prescription Z, the probabilistic response of X can be replaced with a deterministic function of its four behavioural states, i.e. the compliance of patients. That is, U can be replaced with a variable R, that has 4 states (0,1,2,3 for the four groups) which indicate how the patient should **respond** to the value of Z, i.e. with what kind of **function** the patient shall process the prescription. This variable R, representing the influence of U, is called a **response function variable**.

This example of imperfect compliance can be generalised to an abstract definition of response functions. The IV can be seen as encouraging a patient to

prefer one treatment level over another. Being independent of the unmeasured confounding, the perfect IV will also have no impact on the outcome except through encouraging a certain treatment level, i.e. it has exclusively an indirect effect on Y. Generally, in the example of prescribing aspirin and analysing its effect despite imperfect compliance, the focus is on the prescription given by the doctor. The formulation with response function variables is seen as a tool to handle the imperfect compliance of patients to their prescriptions. The next section gives a generalised definition, abstracting away the imperfect compliance setting, of the necessary technical parts for calculating the bounds with response function variables.

4.3 Defintion of Causal Bounds via Response Functions

An established method for estimating causal bounds utilise so called **"response function variables"** (RFVs). RFV formulations have a simple and intutive explanation via the Instrumental Variable model, in the context of imperfect compliance, as described in the earlier section. Put succinctly, the RFVs are used to parameterize the causal model in a more 'granular' way, taking into account the different possibles response of a variable to its ancestors. This yields constraints which can be added to a linear program, for which an objective function, representing the causal effect of interest, is optimized for its maximum and minimum value, i.e. the upper and lower bound of the causal effect. The linear program traverses the space of all possible causal models, as constructed by the constraints, for which the minimum and maximum value represent the lower and upper causal bound.

First, a SCM is assumed, as for example in (Wu et al. 2019,Pearl 2009). Consider an endogenous variable V, with its

- endogenous parents PA_V,
- exogenous parents U_V and
- its associated structural function $v = f_V(pa_V, u_V)$.

U_V can be any variable of any type, with and any domain size. f_V can be any function. As described in the previous section, for each value of U_V, the functional mapping from PA_V to V is a clearly defined deterministic response function, such as denying a prescription or any of the other three. Therefore, each value of U_V can be associated with a unique functional response.

Despite a potentially infinite domain size of U_V, the number of deterministic response functions is predefined by the domain sizes of PA_V and V itself. That is, there is a limited number of ways the variable V can react to its parents, such that they can be fully enumerated.

The **response function variable** is defined as:

$$R_V = 0, ..., N_V - 1$$

where $N_V = |V|^{|PA_V|}$ is the absolute number of response function states that map all possible values of PA_V to V.

A specific value r_V of R_V then represents a specific response of V to its inputs, i.e. the state of its parents. Bounding is the achieved via relating the causal model with these response functions via expressing $P(V)$ as a linear function of $P(R)$:

$$P(V) = \sum_r P(r) \prod_{V \in V} I(v; pa_V, r_V)$$

This is facilitated by an Indicator function I which excludes response function states that do not correspond to the model state in question during each summation. For example for two endogenous X and Y, each with independent U_x and U_y, their joint distribution can be factorised into a sum

$$P(X, Z) = \sum_r P(r_X, r_Z) I(X; r_X) I(Z; X, r_Y)$$

where each of the two Indicators $I(V; pa_V, r_V)$ indicates whether the RFV contributes to the probability state.

Together with the probability simplex, a linear program can be constructed. The objective function is constructed in the same way as above, with RFVs, to represent the causal effect of interest. Minima and maxima of this objective then constitute the lower and upper causal bounds.

4.4 Improving bounds with Additional Assumptions

The causal bounds introduced before are considered to be the 'natural', 'worst case' or 'no assumptions' bounds. Often, these bounds will include 0, such that the effect of interest could be negative or positive, i.e. the bounds are not too informative. Additional assumptions can improve these bounds. Given an application where they can be justified, these bounds can be more informative and possibly even exclude 0 such that the direction of the effect can be extracted from the data. Additional assumptions constrain the space of valid bounds, i.e. previously calculated minima in space are rendered infeasible by additional constraints, such that a new, often *improved* minima is found, and therefore yielding an *improved* lower bound. The same holds for maxima, i.e. upper bounds.

For example, assuming *Monotonicty*, i.e. that there are no defiers, can signifcantly increase the usefulness of the bounds. Bounds calculated with monotonicity obviously are then based on the assumption of a lack of defiers, which generally is hard to verify, but in certain applications easier to argue for, for example in drug trails with a drug known to only ever improve patient health.

Sometimes, these assumptions are testable on the data. For example, the instrumental variable inequality can help falsify modeling assumptions with data (Pearl 1995). Given discrete data, it can falsify the assumption of a IV model, pointing to violation of at least one of the IV conditions:

$$\max_x \sum_y [\max_z Pr(y, x|z)] \leq 1$$

Intuitively, the IV inequality is violated when Z manages to produce changes in Y, given a set value of X, that surpasses the value of 1. If so, the model is falsified by the data. It is not limited to discrete data, though requires more elaborate methods to be extended to continous data (Kédagni and Mourifié 2020).

Though not considered in this report further, at the higher end of the range of possible assumptions are assumptions, or combinations of assumptions, that will lead to point-identification. The lower and upper bounds then collapse onto a single point, such that a scalar value can be calculated for it.

The following section introduces an alternative problem formulation that does not require RFVs. The RFV formulation has a Achilles' heel in its parameter growth: With each variable added to the SCM, the number of parameters increases in a super exponential manner, which renders the bounding calculation prohibitively expensive with current computational resources for 5 or more variables.

A Alternative formulations of bounding

Exclusively with tools from the Potential Outcomes framework

(Robins 1989) and (Manski 1990) independently derived these bounds first.

(Richardson et al. 2014) presents a review, including the following PO based explanation:

$$E[Y(1) - Y(0)] = E[Y(1)] - E[Y(0)]$$

Decomposes into

$$\sum_z E[Y(1)|Z = z]Pr[Z = z] - \sum_z E[Y(0)|Z = z]Pr[Z = z]$$

$$\begin{aligned}(E[Y(1)|Z = 0]Pr[Z = 0] + E[Y(1)|Z = 1]Pr[Z = 1])- \\ (E[Y(0)|Z = 0]Pr[Z = 0] + E[Y(0)|Z = 1]Pr[Z = 1])\end{aligned}$$

This is fully identified except from

$$E[Y(1)|Z = 1 - z]$$

i.e. $E[Y(1)|Z = 0]$ and $E[Y(0)|Z = 1]$

Then, entertain smallest and largest number for $E[Y(1)|Z = 1 - z]$ i.e. $0 \leq E[Y(1)|Z = 1 - z] \leq 1$

This gives the *lower bound*:

$$(Pr[Z=0] + E[Y(1)|Z=1]Pr[Z=1]) - (E[Y(0)|Z=0]Pr[Z=0] + 1Pr[Z=1])$$

$$(E[Y(1)|Z=1]Pr[Z=1]) - (E[Y(0)|Z=0]Pr[Z=0] + Pr[Z=1])$$

$$E[Y(1)|Z=1]Pr[Z=1] - E[Y(0)|Z=0]Pr[Z=0] - Pr[Z=1]$$

In the same way, the *upper bound*:
$0 \leq E[Y(1)|Z=1-z] \leq 1$

$$(1Pr[Z=0] + E[Y(1)|Z=1]Pr[Z=1]) - (E[Y(0)|Z=0]Pr[Z=0] + 0Pr[Z=1])$$

$$E[Y(1)|Z=1]Pr[Z=1] - E[Y(0)|Z=0]Pr[Z=0] + Pr[Z=0]$$

Without any assumptions of the Potential Outcomes framework

(Ramsahai 2012) presents a bound framework that does not rely on the assumptions commonly made in the potential outcomes framework. Instead, the formulation relies solely on the properties of the probability distributions, which is also briefly discussed in (Manski 1990). Ronald Ramsahai's PhD Thesis also provides a comprehensive review of bounding of causal effects in Instrumental Variable models, see (Ramsahai 2012).

References

[1] James M Robins. "The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies". In: *Health service research methodology: a focus on AIDS* (1989), pp. 113–159.

[2] Charles F Manski. "Nonparametric bounds on treatment effects". In: *The American Economic Review* 80.2 (1990), pp. 319–323.

[3] Alexander Balke and Judea Pearl. "Counterfactual Probabilities: Computational Methods, Bounds and Applications". In: *Proceedings of the Tenth international conference on Uncertainty in artificial intelligence* (1994).

[4] Judea Pearl. "On the testability of causal models with latent and instrumental variables". In: *arXiv preprint arXiv:1302.4976* (1995).

[5] Charles F Manski. *Partial identification of probability distributions*. Springer Science & Business Media, 2003.

[6] Judea Pearl. *Causality*. Cambridge University Press, 2009.

[7] Judea Pearl. "Robustness of causal claims". In: *arXiv preprint arXiv:1207.4173* (2012).

[8] Roland R Ramsahai. "Causal bounds and instruments". In: *arXiv preprint arXiv:1206.5262* (2012).

[9] Amy Richardson et al. "Nonparametric bounds and sensitivity analysis of treatment effects". In: *Statistical science: a review journal of the Institute of Mathematical Statistics* 29.4 (2014), p. 596.

[10] A Philip Dawid. "Statistical causality from a decision-theoretic perspective". In: *Annual Review of Statistics and Its Application* 2 (2015), pp. 273–303.

[11] Liam Byrne and Frank Van Haren. "Fluid resuscitation in human sepsis: Time to rewrite history?" In: *Annals of intensive care* 7.1 (2017), pp. 1–8.

[12] Sonja A Swanson et al. "Partial identification of the average treatment effect using instrumental variables: review of methods for binary instruments, treatments, and outcomes". In: *Journal of the American Statistical Association* 113.522 (2018), pp. 933–947.

[13] Yongkai Wu et al. "Pc-fairness: A unified framework for measuring causality-based fairness". In: *Advances in Neural Information Processing Systems*. 2019, pp. 3404–3414.

[14] Noam Finkelstein. "Partial Identifiability in Discrete Data With Measurement Error". In: *arXiv preprint arXiv:2006.06366* (2020).

[15] Miguel A Hernán and James M Robins. *Causal inference: what if*. 2020.

[16] Désiré Kédagni and Ismael Mourifié. "Generalized instrumental inequalities: testing the instrumental variable independence assumption". In: *Biometrika* (2020).

[17] Niki Kilbertus, Matt J Kusner, and Ricardo Silva. "A Class of Algorithms for General Instrumental Variable Models". In: *arXiv preprint arXiv:2006.06366* (2020).

[18] Judea Pearl. *Which Patients are in Greater Need: A counterfactual analysis with reflections on COVID-19*. 2020. URL: `http://causality.cs.ucla.edu/blog/index.php/2020/04/02/which-patients-are-in-greater-need-a-counterfactual-analysis-with-reflections-on-covid-19/`. (accessed: 13.09.2020).

[19] Michael C Sachs, Erin E Gabriel, and Arvid Sjölander. "Symbolic computation of tight causal bounds". In: *arXiv preprint arXiv:2003.10702* (2020).

[20] Sorawit Saengkyongam and Ricardo Silva. "Learning Joint Nonlinear Effects from Single-variable Interventions in the Presence of Hidden Confounders". In: *Conference on Uncertainty in Artificial Intelligence*. PMLR. 2020, pp. 300–309.

[21] Junzhe Zhang and Elias Bareinboim. "Bounding Causal Effects on Continuous Outcome". In: *Proceedings of the 35nd AAAI Conference on Artificial Intelligence*. 2021.

[22] Junzhe Zhang and Elias Bareinboim. "Non-Parametric Methods for Partial Identification of Causal Effects". In: *https://causalai.net/r72.pdf* (2021).