GATEWAYS TO TRACTABILITY FOR SATISFIABILITY IN PEARL'S CAUSAL HIERARCHY

Anonymous authors
Paper under double-blind review

ABSTRACT

Pearl's Causal Hierarchy (PCH) is a central framework for reasoning about probabilistic, interventional, and counterfactual statements, yet the satisfiability problem for PCH formulas is computationally intractable in almost all classical settings. We revisit this challenge through the lens of parameterized complexity and identify the first gateways to tractability. Our results include fixed-parameter and XP-algorithms for satisfiability in key probabilistic and counterfactual fragments, using parameters such as primal treewidth and the number of variables, together with matching hardness results that map the limits of tractability. Technically, we depart from the dynamic programming paradigm typically employed for treewidth-based algorithms and instead exploit structural characterizations of well-formed causal models, providing a new algorithmic toolkit for causal reasoning.

1 INTRODUCTION

Pearl's Causal Hierarchy (PCH) (Shpitser & Pearl, 2008; Pearl, 2009) is a central pillar of the modern theory of causality that is employed in artificial intelligence and other reasoning tasks—see, e.g., the recent survey (Bareinboim et al., 2022) or book (Fenton et al., 2020) on the topic. The PCH is a framework that has three basic layers of depth which capture three fundamental degrees of sophistication for analyzing causal effects and relationships. All of these layers provide a means of formalizing statements via formulas capturing the behavior of a set of probabilistic variables in a *Structural Causal Model* (SCM) (Glymour et al., 1987; Pearl, 2009; Koller & Friedman, 2009; Elwert, 2013), which is a well-established representation of systems with observed as well as hidden variables over a specified domain and their mutual dependencies. As a basic illustrative example, the statement "the likelihood of having both diabetes ($D =$ yes) and blood type B+ ($T =$ B+) is at most 1%" can be expressed by the formula $\psi = \Pr(D = \text{yes} \wedge T = \text{B+}) \leq 0.01$.

The formula ψ above belongs to the first, basic layer of the PCH—that is, the layer $\mathcal{L}_{\mathsf{prob}}$ of probabilistic reasoning that captures direct statements one can make about the probabilities of certain outcomes. The second layer, $\mathcal{L}_{\mathsf{causal}}$, expands on the basic probability terms in $\mathcal{L}_{\mathsf{prob}}$ via the introduction of Pearl's do-operator (Pearl, 2009) which captures interventional causal reasoning. A basic example of an event that can be captured on this layer of the PCH is contracting a disease after being vaccinated against that disease; the probability of this event can be expressed using the term $\Pr([Y = \text{vaccinated}]\ X = \text{contracted})$[1], where the square brackets denote an intervention that is applied before observing the outcome.[2] Hence, the second layer of the PCH allows us to make statements such as $\Pr([Y = \text{vaccinated}]\ X = \text{contracted}) < \Pr(X = \text{contracted})$. The third layer $\mathcal{L}_{\mathsf{counterfact}}$ of the PCH expands on $\mathcal{L}_{\mathsf{causal}}$ by allowing interventions to be chained, and enables complex statements related to counterfactual situations. For instance, a third-layer term such as $\Pr\big(([M = \text{yes}]\ H = \text{no}) \mid (M = \text{no} \wedge H = \text{yes})\big)$ can express the probability that a patient who did not take medication (M) and was hospitalized (H) would have avoided hospitalization if he had taken the medication. Formal definitions of these as well as related notions are available in Section 2.

While the three layers of depth of the PCH focus on the expressivity inside the probability term $\Pr(\cdot)$, there is a second dimension to the PCH—specifically, the *breadth* of operations that can be applied to the probability terms themselves. For $\circledast \in \{\mathsf{prob}, \mathsf{causal}, \mathsf{counterfact}\}$, we distinguish the following fragments of the PCH:

- $\mathcal{L}^{\mathsf{base}}_{\circledast}$: only simple probability terms are allowed, such as $\Pr(\cdot) \leq \Pr(\circ)$ or $\Pr(\cdot) \geq 1$;

[1] Equivalently, $\Pr(X = \text{contracted} \mid do(Y = \text{vaccinated}))$. We follow recent publications in the area (van der Zander et al., 2023; Dörfler et al., 2025) and primarily employ the square-bracket notation.

[2] Interventions are distinct from conditional probability statements such as $\Pr(X = \text{contracted} \mid Y = \text{vaccinated})$. To see this, consider a hypothetical world where the vaccine is ineffective, the disease only exists in a laboratory and an oracle randomly determines whether a person will be infected without vaccination, or receive the vaccine and not come in contact with the disease. In this world, $\Pr(X = \text{contracted} \mid Y = \text{vaccinated}) = 0$ but $\Pr([Y = \text{vaccinated}]\ X = \text{contracted}) > 0$.

- $\mathcal{L}^{\text{lin}}_{\circledast}$: linear combinations of probability terms are allowed, such as $\Pr(\cdot) - \Pr(\circ) \geq 3\Pr(\bullet)$;
- $\mathcal{L}^{\text{poly}}_{\circledast}$: polynomials over probability terms are allowed, such as $\Pr(\cdot)^2 \leq 2\Pr(\circ) \cdot \Pr(\bullet) + 0.1$.

Crucially, the various combinations of depth and breadth give rise to a 3×3 *expressivity matrix* M for PCH (Dörfler et al., 2025, Table 1), (van der Zander et al., 2023, Table 1).

A crucial and well-studied problem in the setting of causal reasoning is SATISFIABILITY—that is, determining whether a given formula (consisting of a set of probability constraints) admits an SCM (Fagin et al., 1990; Ibeling & Icard, 2020; van der Zander et al., 2023; Mossé et al., 2024; Dörfler et al., 2025). We note that there is a high-level parallel between this SATISFIABILITY problem in the causal setting and the well-known BOOLEAN SATISFIABILITY (SAT) and CONSTRAINT SATISFACTION (CSP) problems; the distinction lies in the types of constraints on the input and the nature of the sought-after model. However, solving SATISFIABILITY in our causal reasoning setting is, in general, a much more daunting task. If we let $\text{SAT}^{*}_{\circledast}$ denote the SATISFIABILITY problem for formulas from the fragment $\mathcal{L}^{*}_{\circledast}$ of the PCH, then depending on the choice of $\circledast \in \{\text{prob}, \text{causal}, \text{counterfact}\}$ and $* \in \{\text{base}, \text{lin}, \text{poly}\}$ the problem under consideration will be complete for the complexity classes NP or $\exists\mathbb{R}$—see also the detailed discussion of related work at the end of this section.

Crucially, while previous works have made significant strides towards mapping the classical complexity landscape of the SATISFIABILITY problem, even the "easiest" fragments of the expressivity matrix remain NP-hard. The central aim of this article is to provide a counterweight to this pessimistic perspective and identify fundamental gateways to tractability for SATISFIABILITY, specifically by employing the more refined *parameterized complexity* paradigm (Downey & Fellows, 2013; Cygan et al., 2015). There, one analyzes the running time of algorithms not only in terms of the input size $|I|$, but also with respect to a specified numerical parameter k. The standard notion of tractability used in this setting is then tied to algorithms which run in time $f(k) \cdot |I|^{\mathcal{O}(1)}$ for some computable function f; problems admitting such *fixed-parameter* algorithms are called *fixed-parameter tractable* (FPT). A weaker—but nevertheless still useful—notion of tractability stems from the existence of a so-called XP-*algorithm*, i.e., an algorithm running in time $|I|^{f(k)}$ (this gives rise to the complexity class XP). Our main results include not only the first fixed-parameter and XP-algorithms for the problem, but also matching lower bounds which allow us to identify the limits of parameterized tractability in the expressivity matrix.

Contributions. A loose inspiration for this work stems from the success stories in the aforementioned domains of BOOLEAN SATISFIABILITY and CONSTRAINT SATISFACTION. The parameterized complexity of these two problems is by now very well understood, and perhaps the most classical parameterized algorithms use the *primal treewidth* as the parameter of choice. Essentially, this measures how "tree-like" the interactions between the variables in the instance are—more precisely, this is captured by measuring the *treewidth*, a fundamental graph parameter (Robertson & Seymour, 1984), of the graph obtained by representing variables as vertices and using edges to capture the property of lying in the same "term" (i.e., clause or constraint). In particular, it is known that BOOLEAN SATISFIABILITY is fixed-parameter tractable w.r.t. the primal treewidth (Biere et al., 2009, Chapter 13), while CONSTRAINT SATISFACTION admits an XP-algorithm under the same parameterization (Samer & Szeider, 2010); the latter then becomes fixed-parameter tractable when the parameterization also includes the domain size for the variables (Samer & Szeider, 2010).

Given the above, it is natural to ask whether one can use the primal treewidth to establish tractability for SATISFIABILITY in the PCH setting. As our first set of contributions, we provide a complete answer to this question:

> $\text{SAT}^{\text{lin}}_{\text{prob}}$ is **(1)** in XP w.r.t. the primal treewidth alone, and
> **(2)** fixed-parameter tractable w.r.t. the primal treewidth plus the domain size d.
>
> Moreover, under well-established complexity assumptions one can neither
> **(3)** improve the XP-tractability to FPT (not even for $\text{SAT}^{\text{base}}_{\text{prob}}$), nor
> **(4)** lift **any** of these tractability results to $\text{SAT}^{\text{poly}}_{\text{prob}}$ or $\text{SAT}^{\text{lin}}_{\text{causal}}$.

Furthermore, we remark that parameterizing by the domain size alone does not yield tractability under well-established complexity assumptions (see Theorem 6).

While the above results are comprehensive, they only provide a gateway to tractability for the "shallowest" probabilistic fragment of the PCH. We hence ask whether one can achieve tractability for deeper fragments of the PCH (that is, $\mathcal{L}^{*}_{\text{causal}}$ and $\mathcal{L}^{*}_{\text{counterfact}}$) if the primal treewidth is replaced with a more restrictive parameterization—specifically the number n of variables in the formula. We note that the analogous question in the BOOLEAN SATISFIABILITY and CONSTRAINT SATISFACTION setting is trivial: there, asymptotically optimal (under well-established complexity assumptions) algorithms parameterized by n can merely enumerate all possible models (Impagliazzo et al., 2001; Karthik C. S. et al., 2024). Such an approach is doomed to fail for the causal SATISFIABILITY problem: not only will an SCM contain (potentially many) auxiliary random variables, but also variable dependencies and random distributions that cannot be exhaustively enumerated.

As our second set of contributions, we map the complexity landscape for deeper fragments of the PCH as well:

> $\mathsf{SAT}^{\mathsf{lin}}_{\mathsf{counterfact}}$ is **(5)** in XP w.r.t. n alone, and
> **(6)** fixed-parameter tractable w.r.t. n plus the domain size d.
>
> Moreover, under well-established complexity assumptions one can neither
> **(7)** improve the XP-tractability to FPT (not even for $\mathsf{SAT}^{\mathsf{base}}_{\mathsf{prob}}$), nor
> **(8)** lift **any** of these tractability results to $\mathsf{SAT}^{\mathsf{poly}}_{\mathsf{counterfact}}$.

A schematic overview of our contributions is provided in the mind map below (Figure 1).

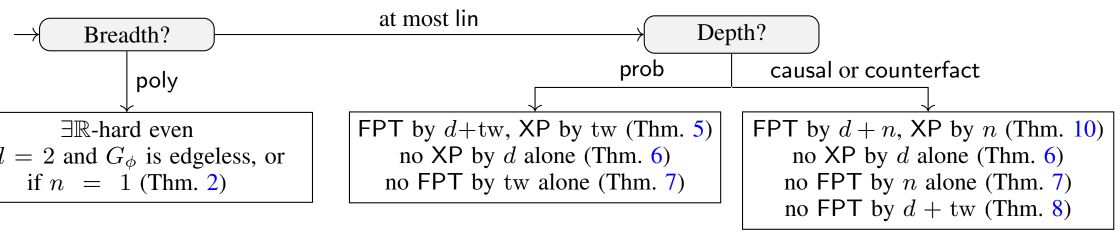


Figure 1: Parameterized complexity of SATISFIABILITY in the PCH based on the position (breadth/depth) in the expressivity matrix M. All results hold under well-established complexity assumptions and refer to an instance with n observed variables over a domain of size d such that the treewidth of the primal graph G_ϕ is tw. We note that the $\exists\mathbb{R}$-hardness of the poly fragment was already established by Mossé et al. (2024), but not under the stated restrictions which rule out tractability in the parameterized setting.

Extensions to Marginalization. In order to efficiently express marginalization, recent works (van der Zander et al., 2023; Dörfler et al., 2025) have extended the classical fragments in M to $\mathcal{L}_{\circledast}^{\mathsf{base}\langle\Sigma\rangle}$, $\mathcal{L}_{\circledast}^{\mathsf{lin}\langle\Sigma\rangle}$, and $\mathcal{L}_{\circledast}^{\mathsf{poly}\langle\Sigma\rangle}$, respectively; the only difference is that these classes additionally include the unary summation operator $\sum$. Depending on the specific fragment considered, including these marginalization operators in the SATISFIABILITY problem yields completeness for the complexity classes $\mathsf{NP}^{\mathsf{PP}}$, PSPACE, NEXP or succ-$\exists\mathbb{R}$—see Dörfler et al. (2025, Table 1). Since the algorithm(s) underlying Results **(5)** and **(6)** can also be used to establish inclusion in the complexity class EXPTIME while $\mathsf{SAT}^{\mathsf{lin}\langle\Sigma\rangle}_{\mathsf{counterfact}}$ is NEXP-complete (Dörfler et al., 2025), under well-established complexity assumptions it is not possible to lift our results towards full marginalization operators as considered in the aforementioned works. Nevertheless, if one were to bound the nesting depth of the unary summation operator $\sum$ to any (arbitrary but fixed) constant, all of our results could be directly translated to the marginalization setting by simply expanding on the respective sums.

Proof Techniques. The standard approach to establishing tractability for problems parameterized by treewidth is to employ dynamic programming—this is the approach used not only for the aforementioned treewidth-based algorithms solving BOOLEAN SATISFIABILITY and CONSTRAINT SATISFACTION, but also for almost every algorithm parameterized by treewidth. From a technical standpoint, it is hence surprising that our results **do not** employ dynamic programming at all; in fact, the SATISFIABILITY problem seems entirely incompatible with the basic tenets of the usual "leaf-to-root" dynamic programming paradigm used for treewidth.

Instead, our proof of Results **(1)** and **(2)** relies on an entirely novel approach. We first prove that every YES-instance of $\mathsf{SAT}^{\mathsf{lin}}_{\mathsf{prob}}$ with primal treewidth k admits a "well-structured" SCM whose hidden variables and dependencies can be neatly mapped onto the tree-like structure of the primal graph and determined in advance. However, this step on its own cannot determine whether an SCM actually exists, as for that we need to compute and verify the probability distributions for the hidden variables. In the second step, we use the tree-likeness of the instance once again to construct a "fixed-parameter sized" linear program which either computes a viable set of probability distributions, or determines that none exists. It is well-known that linear programs can be solved in polynomial time (Papadimitriou & Steiglitz, 1998)—the difficult part lies in building a program that provably verifies the existence of an SCM while avoiding an exponential dependency on the input size.

In order to apply the reduction technique to Results **(5)** and **(6)**, we need to be able to deal with the presence of interventions in the formula. Towards this, we argue that every YES-instance of $\mathsf{SAT}^{\mathsf{lin}}_{\mathsf{counterfact}}$ admits an

SCM with different structural properties than those used for Results **(5)** and **(6)**: in particular, the value of a single hidden variable U determines not just the value of each observed variable but the *function* of how it is determined from the other observed variables. We then define a suitable linear programming formulation that targets the computation of such well-structured SCMs.

For establishing the lower-bound results, we develop three distinct reductions: one from k-MULTICOLORED-CLIQUE which handles **(3)** and **(7)**, one from a restricted variant of the Existential Theory of the Reals problem for Results **(4)** and **(8)**, and a separate reduction from 3-SAT for the remaining lin-causal case of **(4)**.

Related Work. $\mathrm{SAT}^{\mathsf{lin}}_{\mathsf{prob}}$ and $\mathrm{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ were shown to be NP-complete by Fagin et al. (1990), and analogous results for the fragments $\mathrm{SAT}^{\mathsf{lin}}_{\mathsf{causal}}$, $\mathrm{SAT}^{\mathsf{base}}_{\mathsf{causal}}$, $\mathrm{SAT}^{\mathsf{lin}}_{\mathsf{counterfact}}$, and $\mathrm{SAT}^{\mathsf{base}}_{\mathsf{counterfact}}$ were obtained by Mossé et al. (2024). The $\exists\mathbb{R}$-completeness of $\mathrm{SAT}^{\mathsf{poly}}_{\mathsf{prob}}$, $\mathrm{SAT}^{\mathsf{poly}}_{\mathsf{causal}}$ and $\mathrm{SAT}^{\mathsf{poly}}_{\mathsf{counterfact}}$ was also established in the latter work (Mossé et al., 2024). As mentioned above, these complexity-theoretic studies were recently extended to languages containing the summation operator $\sum$ (van der Zander et al., 2023; Dörfler et al., 2025). Other related languages designed to express probabilistic reasoning were developed in the works of, e.g., Nilsson (1986); Georgakopoulos et al. (1988); Ibeling & Icard (2020). Moreover, the existence of solutions to the SATISFIABILITY problem with specific properties has very recently been studied by Bläser et al. (2025).

Beyond SATISFIABILITY, the parameterized complexity paradigm has been employed in several works studying another central problem in the area of causality: BAYESIAN NETWORK STRUCTURE LEARNING. This line of research was initiated by Ordyniak & Szeider (2013), with recent contributions considering a broad range of parameterizations as well as variations of the problem (Ganian & Korchemna, 2021; Grüttemeier et al., 2021a;b; Grüttemeier & Komusiewicz, 2022). The complexity of the related CAUSAL DISCOVERY problem was recently studied by Ganian et al. (2024).

Beyond the aforementioned prominent applications in BOOLEAN SATISFIABILITY and CONSTRAINT SATISFACTION, primal treewidth has been used as a natural means of capturing structural properties of inputs in a variety of other settings as well. Examples of this in the broad AI area include its applications in INTEGER LINEAR PROGRAMMING (Ganian et al., 2017; Ganian & Ordyniak, 2018), HEDONIC GAMES (Peters, 2016; Hanaka & Lampis, 2022), MATRIX COMPLETION (Ganian et al., 2022), ANSWER SET PROGRAMMING (Fichte & Hecher, 2018), and RESOURCE ALLOCATION (Eiben et al., 2023). We note that the treewidth-based algorithms in all of the aforementioned works rely on dynamic programming, which is fundamentally different from the technique employed to achieve our Results **(1)** and **(2)**.

Full proofs and details deferred to the Appendix are marked with (♣).

2 PRELIMINARIES

For $n \in \mathbb{N}$, let $[n] = \{1, \ldots, n\}$. For $i_1, i_2 \in \mathbb{R}$, let $[i_1, i_2] = \{j \in \mathbb{R} \mid i_1 \leq j \leq i_2\}$. We follow established notation as used in (Mossé et al., 2024; van der Zander et al., 2023). By $\mathbf{V}$ we refer to a contingent of random variables and, without loss of generality, assume that each of these share a given domain D of size d.

Syntax of the languages of PCH. For $V \in \mathbf{V}$ and $v \in D$, a statement of the form $V = v$ is called an *atom*. We can combine multiple atoms to obtain *events* over $\mathbf{V}$ by applying the following grammatical rules.

$$\begin{aligned}
\mathcal{E}_{\mathsf{prop}} &::= V = v \mid \neg\mathcal{E}_{\mathsf{prop}} \mid \mathcal{E}_{\mathsf{prop}} \wedge \mathcal{E}_{\mathsf{prop}}, \\
\mathcal{E}_{\mathsf{int}} &::= \top \mid V = v \mid \mathcal{E}_{\mathsf{int}} \wedge \mathcal{E}_{\mathsf{int}}, \\
\mathcal{E}_{\mathsf{post\text{-}int}} &::= [\mathcal{E}_{\mathsf{int}}]\, \mathcal{E}_{\mathsf{prop}}, \\
\mathcal{E}_{\mathsf{counterfact}} &::= \mathcal{E}_{\mathsf{post\text{-}int}} \mid \neg\mathcal{E}_{\mathsf{counterfact}} \mid \mathcal{E}_{\mathsf{counterfact}} \wedge \mathcal{E}_{\mathsf{counterfact}}.
\end{aligned}$$

We call the events in $\mathcal{E}_{\mathsf{prop}}$ *propositions* and the events in $\mathcal{E}_{\mathsf{int}}$ *interventions*. Each event ε can only occur within a probabilistic statement $\Pr(\varepsilon)$, which we call a *term*. The *size of a term* is the number of atoms it contains. For $\mathcal{E} \in \{\mathcal{E}_{\mathsf{prop}}, \mathcal{E}_{\mathsf{post\text{-}int}}, \mathcal{E}_{\mathsf{counterfact}}\}$ and $\varepsilon \in \mathcal{E}$, we define the following valid ways of combining terms.

$$\begin{aligned}
T_{\mathsf{base}}(\mathcal{E}) &::= \Pr(\varepsilon), \\
T_{\mathsf{lin}}(\mathcal{E}) &::= \Pr(\varepsilon) \mid T_{\mathsf{lin}}(\mathcal{E}) + T_{\mathsf{lin}}(\mathcal{E}), \\
T_{\mathsf{poly}}(\mathcal{E}) &::= \Pr(\varepsilon) \mid T_{\mathsf{poly}}(\mathcal{E}) + T_{\mathsf{poly}}(\mathcal{E}) \mid T_{\mathsf{poly}}(\mathcal{E}) \cdot T_{\mathsf{poly}}(\mathcal{E}).
\end{aligned}$$

Lastly, for $* \in \{\mathsf{base}, \mathsf{lin}, \mathsf{poly}\}$ we define $\mathcal{L}^*_{\mathsf{prob}}$ ($\mathcal{L}^*_{\mathsf{causal}}$ and $\mathcal{L}^*_{\mathsf{counterfact}}$, respectively) to be the *language* that contains all sets of inequalities over elements in $T_*(\mathcal{E}_{\mathsf{prop}})$ (in $T_*(\mathcal{E}_{\mathsf{post\text{-}int}})$ and in $T_*(\mathcal{E}_{\mathsf{counterfact}})$, resp.). The elements inside $\mathcal{L}^*_{\mathsf{prob}}$, $\mathcal{L}^*_{\mathsf{causal}}$, and $\mathcal{L}^*_{\mathsf{counterfact}}$ are called *formulas*. Note that tautological and contradictory events can be used to encode comparisons against 1 and 0, such as $\Pr(\varepsilon) \leq 0$. Moreover, the grammars of $\mathcal{L}^{\text{lin}}_{\circledast}$ and $\mathcal{L}^{\text{poly}}_{\circledast}$ support integer coefficients, which can be effectively constructed by summing up multiple probabilities of the same type. Any inequality with rational coefficients can be encoded by multiplying both sides with the smallest common multiple of all non-integer coefficients. At the beginning of the next section, we will compare our syntax to the one used in related work.

Semantics of the Languages of PCH. We define the semantics of the aforementioned languages using the notion of Structural Causal Models as popularized by Glymour et al. (1987) and Pearl (2009, Section 3.2). A *recursive Structural Causal Model* (*SCM*, or simply *model*) over domain D is a tuple $\mathcal{M} = (\mathbf{V}, \mathbf{U}, \mathcal{F}, \mathbb{P})$ with

- a set $\mathbf{V}$ of endogenous (observed) variables, implicitly well-ordered by $\prec$, that range over D,
- a set $\mathbf{U}$ of exogenous (hidden) variables,
- a set $\mathcal{F} = \{f_V\}_{V \in \mathbf{V}}$ of functions where f_V specifies how the value of V can be computed given the values of $\mathbf{U}$ and $\mathbf{V}_{\prec V}$, that is, the subset of $\mathbf{V}$ that precedes $V \in \mathbf{V}$ in $\prec$,
- a probability distribution $\mathbb{P}$ on $\mathbf{U}$.

Note that any model $\mathcal{M}$ whose exogenous variables $\mathbf{U}$ have an infinite or continuous domain is (w.r.t. its evaluation) equivalent to a model $\mathcal{M}'$ where all exogenous variables have discrete and finite domains (Zhang et al., 2022). Consequently, we assume throughout that each variable $U \in \mathbf{U}$ has a discrete and finite domain $\text{Val}(U)$, and let $\text{Val}(\mathbf{U}) = \text{Val}(U_1) \times \ldots \times \text{Val}(U_{|\mathbf{U}|})$ refer to their combined range.

Let $V = v$ be an atom in $\mathcal{E}_{\text{int}}$. We denote by $\mathcal{F}_{V=v}$ the set of functions obtained from $\mathcal{F}$ by replacing f_V with the constant function v. We generalize this definition to arbitrary conjunctions of atoms $\gamma \in \mathcal{E}_{\text{int}}$ in the natural way and denote the set of resulting functions as $\mathcal{F}_\gamma$. Let $\varepsilon \in \mathcal{E}_{\text{prop}}$ and $\overline{u} \in \text{Val}(\mathbf{U})$. We write $\mathcal{F}, \overline{u} \models \varepsilon$ if evaluating $\mathcal{F}$ on input $\overline{u}$ yields an assignment to $\mathbf{V}$ under which ε is satisfied. For $[\gamma]\ \varepsilon \in \mathcal{E}_{\text{post-int}}$, we write $\mathcal{F}, \overline{u} \models [\gamma]\ \varepsilon$ if $\mathcal{F}_\gamma, \overline{u} \models \varepsilon$. Moreover, for all $\varepsilon, \varepsilon_1, \varepsilon_2 \in \mathcal{E}_{\text{counterfact}}$, we write (i) $\mathcal{F}, \overline{u} \models \neg\varepsilon$ if $\mathcal{F}, \overline{u} \not\models \varepsilon$, and (ii) $\mathcal{F}, \overline{u} \models \varepsilon_1 \wedge \varepsilon_2$ if both $\mathcal{F}, \overline{u} \models \varepsilon_1$ and $\mathcal{F}, \overline{u} \models \varepsilon_2$. For a given model $\mathcal{M} = (\mathbf{V}, \mathbf{U}, \mathcal{F}, \mathbb{P})$, we denote $S_{\mathcal{M}} := \{\overline{u} \in \text{Val}(\mathbf{U}) \mid \mathcal{F}, \overline{u} \models \varepsilon\}$. The way $\mathcal{M}$ interprets an expression $\mathbf{t} \in T_{\text{poly}}(\mathcal{E})$ is denoted by $[\![\mathbf{t}]\!]_{\mathcal{M}}$ and recursively defined as follows: $[\![\Pr(\varepsilon)]\!]_{\mathcal{M}} = \sum_{\overline{u} \in S_{\mathcal{M}}(\varepsilon)} \mathbb{P}(\overline{u})$. For two expressions $\mathbf{t}_1, \mathbf{t}_2 \in T_{\text{poly}}(\mathcal{E})$, we define $\mathcal{M} \models \mathbf{t}_1 \leq \mathbf{t}_2$ if and only if $[\![\mathbf{t}_1]\!]_{\mathcal{M}} \leq [\![\mathbf{t}_2]\!]_{\mathcal{M}}$. The semantics for negation and conjunction are defined in the usual way, yielding the semantics for $\mathcal{M} \models \phi$ for any formula $\phi \in \mathcal{L}^{\text{poly}}_{\text{counterfact}}$.

Primal Treewidth. Let $\phi \in \mathcal{L}^{\text{poly}}_{\text{counterfact}}$ be a formula over variables $\mathbf{V}$. By $G_\phi = (\mathbf{V}, E)$, we denote the *primal graph* of ϕ, where $\{V, V'\} \in E$ if and only if $V \neq V'$ and there is a term in ϕ that contains both V and V'. The *treewidth* $\text{tw}(G)$ of a graph G is a well-established measure of how "tree-like" it is; for instance, trees have a treewidth of 1, while an s-vertex complete graph has treewidth $s - 1$. For a definition of treewidth, including the notions of *nice tree decompositions* and *bags* which are used in the formal proof of Theorem 5, we refer to the literature (Cygan et al., 2015, Subsection 7.2). (♣) We let the (*primal*) *treewidth of a formula* ϕ denote the treewidth of its primal graph, that is, $\text{tw}(\phi) = \text{tw}(G_\phi)$.

3 Satisfiability for Languages of PCH and Structural Insights

In this paper, we examine several analogues of the well-known problem Boolean Satisfiability that capture various probabilistic, causal, and counterfactual statements. We denote these problems as $\text{SAT}^*_{\circledast}$, where $\circledast \in \{\text{prob}, \text{causal}, \text{counterfact}\}$ and $* \in \{\text{base}, \text{lin}, \text{poly}\}$, and define them as follows.

$\text{SAT}^*_{\circledast}$	**Input:**	A set D of d domain values and a formula $\phi \in \mathcal{L}^*_{\circledast}$ over variables $\mathbf{V} = \{V_1, \ldots, V_n\}$.
	Task:	Decide if there exists a recursive Structural Causal Model $\mathcal{M}$ over D such that $\mathcal{M} \models \phi$.

The classical computational complexity of $\text{SAT}^*_{\circledast}$ has by now been studied extensively (Fagin et al., 1990; Ibeling & Icard, 2020; van der Zander et al., 2023; Mossé et al., 2024; Dörfler et al., 2025). We remark that our definition of $\text{SAT}^*_{\circledast}$ slightly deviates from the one established in previous works, in the sense that we restrict our attention to input formulas ϕ that are *sets of inequalities* (that is, each inequality forms a constraint) rather than allowing arbitrary Boolean combinations of inequalities. However, this restriction does not affect any of the known complexity-theoretic results, since previous lower-bound proofs did not employ any Boolean combinations beyond sets. However, the situation changes drastically when studying $\text{SAT}^*_{\circledast}$ from the viewpoint of parameterized complexity, as we show next.

Let $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ denote the version of $\mathrm{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ in which ϕ is an arbitrary Boolean combination of inequalities over elements in $T_{\mathsf{base}}(\mathcal{E}_{\mathsf{prob}})$. We justify our restriction to $\mathrm{SAT}^{*}_{\circledast}$ by showing that $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ remains NP-complete in a very restricted setting, thus dashing any hope to exploit structural properties of ϕ.

Theorem 1. $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ *is* NP-*complete even if* G_ϕ *is edgeless and* $d = 2$.

Proof. Fagin et al. (1990) proved that $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ is NP-complete. In order to prove the NP-hardness of $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ even for instances in which each constraint consists of only one variable and all variables have domain $D = \{0, 1\}$, we perform a reduction from 3-SAT. Let $\Phi := \bigwedge_i C_i$ with $C_i := \bigvee_{j\in[3]} \ell_{i_j}$ be a 3-SAT formula over variables $\mathcal{V}$. Define an instance ϕ of $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ with $\mathbf{V} = \{V_v \mid v \in \mathcal{V}\}$ over domain $\{0, 1\}$ as

$$\phi := \bigwedge_i \Big(\bigvee_{j\in[3]} \Pr(g(\ell_{i_j})) = 1 \Big) \wedge \bigwedge_{v\in\mathcal{V}} (\Pr(V_v = 0) = 1 \vee \Pr(V_v = 1) = 1),$$

where $g(\ell_{i_j})$ is replaced by $V_v = 1$ if $\ell_{i_j} = v$, and by $V_v = 0$ if $\ell_{i_j} = \neg v$.

We now argue that Φ is satisfied if and only if ϕ admits an SCM. For the first direction, suppose there exists an assignment $\alpha : \mathcal{V} \to \{0, 1\}^{|\mathcal{V}|}$ under which Φ is satisfied. We construct a model $\mathcal{M} = (\mathbf{V}, \emptyset, \mathcal{F}, \emptyset)$ for $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ as follows. For each $v \in \mathcal{V}$, define $f_{V_v} := \alpha(v)$ as a constant function. To see that $\mathcal{M}$ satisfies all constraints in ϕ, recall that α satisfies at least one literal ℓ_{i_j} in each clause $C_i \in \Phi$, that is, $\alpha(v) = 1$ if $\ell_{i_j} = v$, and $\alpha(v) = 0$ if $\ell_{i_j} = \neg v$. The reduction ensures that the i^{th} conjunct in ϕ contains the disjunct $\Pr(V_v = 1) = 1$ in the first, and $\Pr(V_v = 0) = 1$ in the latter case. Since $[\![\Pr(V_v = \alpha(v))]\!]_{\mathcal{M}} = 1$, this satisfies ϕ.

For the other direction, suppose there exists a model $\mathcal{M}$ satisfying ϕ. Therefore, either $[\![\Pr(V_v = 0)]\!]_{\mathcal{M}} = 1$ or $[\![\Pr(V_v = 1)]\!]_{\mathcal{M}} = 1$ for each $v \in \mathcal{V}$. We obtain an assignment $\alpha : \mathcal{V} \to \{0, 1\}^{|\mathcal{V}|}$ by defining $\alpha(v) = 0$ if and only if $[\![\Pr(V_v = 0)]\!]_{\mathcal{M}} = 1$ for each variable $v \in \mathcal{V}$. Since all conjuncts of ϕ are satisfied by $\mathcal{M}$, it holds by construction that α satisfies all clauses of Φ. $\square$

Intractability of $\mathrm{SAT}^{\mathsf{poly}}_{\mathsf{prob}}$. Our main contributions target the lin and base fragments of the expressivity matrix and are provided in Sections 4 and 5. Here, we show that the tractability results obtained there cannot be lifted to polynomial inequalities. The proof is based on a reduction from the $\exists\mathbb{R}$-complete problem $\mathrm{ETR}^{1/8,+,\times}_{[-1/8,1/8]}$ (Abrahamsen et al., 2017), where the probabilities of individual events are used to encode the values of the variables in a way that conceptually resembles the construction used in van der Zander et al. (2023, Proposition 6.5).

Theorem 2 (♣). $\mathrm{SAT}^{\mathsf{poly}}_{\mathsf{prob}}$ *is* $\exists\mathbb{R}$*-complete even if* $n = 1$*, or if* $d = 2$ *and* G_ϕ *is edgeless.*

Further Structural Insights in $\mathrm{SAT}^{*}_{\circledast}$. In order to facilitate our complexity-theoretic analysis, we emphasize that a Structural Causal Model can be efficiently evaluated, that is, given the values of its hidden variables, it can be decided in polynomial time, whether a certain event happens.

Observation 3 (♣). Given a model $\mathcal{M} = (\mathbf{U}, \mathbf{V}, \mathcal{F}, \mathbb{P})$, an event $\varepsilon \in \mathcal{E}_{\mathsf{counterfact}}$, and some $\overline{u} \in \mathrm{Val}(\mathbf{U})$, let $|\varepsilon|$ denote the number of atoms in ε. Assuming that each function in $\mathcal{F}$ can be evaluated in time $\mathcal{O}(n)$, one can decide whether $\mathcal{F}, \overline{u} \models \varepsilon$ in time in $\mathcal{O}(n^2 + |\varepsilon|)$.

The runtime of our algorithmic results often depends on the size d of the domain D. We remark that assuming that d is not much larger than the size of ϕ does not reduce the generality of our results, as we can always reduce to an equivalent instance where d is bounded from above by $|\phi| + 1$.

Observation 4 (♣). Consider an instance of $\mathrm{SAT}^{*}_{\circledast}$ consisting of a domain D and a formula $\phi \in \mathcal{L}^{*}_{\circledast}$. Let D_ϕ be the set of values in D that are explicitly mentioned in at least one atom in ϕ and choose some $\gamma \notin D_\phi$. Then, it holds that ϕ over domain $D_\phi \cup \{\gamma\}$ is a YES-instance of $\mathrm{SAT}^{*}_{\circledast}$ if and only if so is ϕ over D.

4 Linear Inequalities over Probabilistic Expressions

This section is dedicated to the complexity-theoretic analysis of $\mathrm{SAT}^{*}_{\mathsf{prob}}$, that is, the satisfiability problem for the layer of the PCH that does not allow any interventions. First, we establish the main tractability result of this section, and then proceed by showing that it is tight as outlined in Figure 1.

Theorem 5. $\mathrm{SAT}^{\mathsf{lin}}_{\mathsf{prob}}$ *is in* FPT *w.r.t. the combined parameter* $d + \mathrm{tw}(\phi)$*, and in* XP *w.r.t.* $\mathrm{tw}(\phi)$.

Proof. Consider an instance of $\mathrm{SAT}^{\mathsf{lin}}_{\mathsf{prob}}$ with formula ϕ and domain D. We prove both statements simultaneously by describing an algorithm that runs in time $d^{f(\mathrm{tw}(\phi))}|\phi|^{\mathcal{O}(1)}$, for a computable function f. Consider a nice tree decomposition of G_ϕ consisting of $\mathcal{O}(n)$ nodes with maximum size $w \coloneqq \text{tw}(\phi) + 1$ computed by, e.g., the algorithm of Bodlaender (1996). Without loss of generality, assume that only the bags of leaf nodes are empty and ignore them in the following procedure. For the remaining tree decomposition $\mathbf{T}$, let $D^{|B|}$ be the combined domain of the variables of bag B in $\mathbf{T}$. We construct the following Linear Program (LP). For each bag B and $\overline{v} \in D^{|B|}$, construct an LP-variable $p_{B=\overline{v}}$; this will capture the probability of the event $B = \overline{v}$, that is, each variable in B takes the respective value in $\overline{v}$. To ensure a valid probability distribution over the LP-variables in each bag B, add the LP-constraints

$$p_{B=\overline{v}} \geq 0 \text{ for each LP-variable } p_{B=\overline{v}}, \quad \text{and} \quad \sum\nolimits_{\overline{v} \in D^{|B|}} p_{B=\overline{v}} = 1 \text{ for each bag } B.$$

For every pair of bags B, B' whose nodes are adjacent in $\mathbf{T}$ and $B \neq B'$, note that there is some $V \in \mathbf{V}$ such that, without loss of generality, $B' = B \cup \{V\}$. To guarantee consistency between the probability distributions of B and B', we add for each such pair and each $\overline{v} \in D^{|B|}$ the LP-constraint

$$p_{B=\overline{v}} = \text{sum}\{p_{B'=\overline{v}'} \mid \overline{v}' \in D^{|B'|} \text{ and } \overline{v}' \text{ sets } B \text{ to } \overline{v}\}.$$

Next, for each constraint C in ϕ, consider each of its terms $\Pr(\varepsilon)$ and define $\mathcal{V}_\varepsilon \subseteq \mathbf{V}$ to be the set of variables that occur in $\Pr(\varepsilon)$. By construction, for each ε, all variables in $\mathcal{V}_\varepsilon$ form a clique in G_ϕ. Consequently, there is at least one bag B_ε in $\mathbf{T}$ such that $\mathcal{V}_\varepsilon \subseteq B_\varepsilon$. Consider an arbitrary choice of such B_ε and obtain an LP-constraint from C by replacing each occurrence of term $\Pr(\varepsilon)$ by a sum over all LP-variables $p_{B_\varepsilon=\overline{v}}$ such that $B_\varepsilon = \overline{v}$ satisfies the event ε. Then the LP consists of $\mathcal{O}(n \cdot d^w)$ LP-variables and $\mathcal{O}(|\phi| + n \cdot d^w)$ LP-constraints.

We can find a solution of an LP (or decide that there is none) in polynomial time with respect to its size, that is, the number of its variables plus constraints. Crucially, if ϕ is a YES-instance witnessed by a model $\mathcal{M}$ which induces a probability distribution over $\mathbf{V}$, then the LP admits a solution; indeed, we can satisfy all constraints by setting each LP-variable $p_{B=\overline{v}}$ to the probability of the event $B = \overline{v}$ within that distribution.

For the converse, assume the LP has a solution. We construct a model for ϕ by passing through $\mathbf{T}$ in a breadth-first-search manner, starting from an arbitrary leaf node with some bag $B = \{V\}$. Let U_V be a hidden variable with domain D such that $\mathbb{P}(U_V = v) = p_{B=(v)}$ for all $v \in D$. Whenever we transition from a bag B to a bag B' containing a variable V which we have not yet described in our model, we have $B' = B \cup \{V\}$. For each $\overline{v} \in D^{|B|}$ such that $p_{B=\overline{v}} > 0$, create a hidden variable $U_{V|B=\overline{v}}$ with domain $\text{Val}(U_{V|B=\overline{v}}) = D$ and let

$$\mathbb{P}(U_{V|B=\overline{v}} = x) \coloneqq \frac{p_{B'=\overline{v}+x}}{p_{B=\overline{v}}} \quad \text{for each } x \in D,$$

where $(\overline{v} + x) \in D^{|B'|}$ is such that it sets B to $\overline{v}$ and V to x. This describes a valid probability distribution of $U_{V|B=\overline{v}}$: As $\mathbb{P}(U_{V|B=\overline{v}} = x) \geq 0$ for all x, it remains to show that $\sum_{x \in D} \mathbb{P}(U_{V|B=\overline{v}} = x) = 1$. We have

$$\sum_{x \in D} \mathbb{P}(U_{V|B=\overline{v}} = x) = \sum_{x \in D} \frac{p_{B'=\overline{v}+x}}{p_{B=\overline{v}}} = \frac{1}{p_{B=\overline{v}}} \sum_{x \in D} p_{B'=\overline{v}+x} = 1,$$

as we ensured $p_{B=\overline{v}} = \sum_{x \in D} p_{B'=\overline{v}+x}$ by an LP-constraint. We now define the function f_V such that, for each $\overline{v} \in D^{|B|}$, if $B = \overline{v}$ then $V = U_{B'|B=\overline{v}}$.

It remains to argue that the model $\mathcal{M}$ obtained after visiting every node witnesses ϕ to be a YES-instance. To this end, employ induction over the breadth-first search described above to prove that after visiting a node with bag B, for each $\overline{v} \in D^{|B|}$ the value of $p_{B=\overline{v}}$ describes the probability of $B = \overline{v}$ in the current model $\mathcal{M}'$, that is, $[\![\Pr(B = \overline{v})]\!]_{\mathcal{M}'} = p_{B=\overline{v}}$. As the bag of the first node contains just one variable, the base case is trivial. Now assume that the claim holds for bag B of some node N and model $\mathcal{M}$, and from N we are visiting a node N' with bag B'. If $B' \subseteq B$, then $\mathcal{M}$ is not changed and the claim follows immediately. Otherwise $B' = B \cup \{V\}$ for a variable V and $\mathcal{M}$ is extended to a model $\mathcal{M}'$ as described above. Consider any $\overline{v}' \in D^{|B'|}$. Let $\overline{v}$ equal $\overline{v}'$ when restricted to the variables in B and let x be the value of variable V in $\overline{v}'$. By the construction of $\mathcal{M}'$ and the induction hypothesis, we have that $[\![\Pr(B = \overline{v})]\!]_{\mathcal{M}'} = [\![\Pr(B = \overline{v})]\!]_{\mathcal{M}} = p_{B=\overline{v}}$. If $[\![\Pr(B = \overline{v})]\!]_{\mathcal{M}'} = 0$, then $[\![\Pr(B' = \overline{v}')]\!]_{\mathcal{M}'} = 0$, which is correct by the consistency constraints $p_{B'=\overline{v}'} \leq p_{B=\overline{v}} = 0$. Otherwise,

$$[\![\Pr(B' = \overline{v}')]\!]_{\mathcal{M}'} = [\![\Pr(B = \overline{v})]\!]_{\mathcal{M}'} \cdot \mathbb{P}(U_{V|B=\overline{v}} = x) = p_{B=\overline{v}} \cdot \frac{p_{B'=\overline{v}'}}{p_{B=\overline{v}}} = p_{B'=\overline{v}'}.$$

Given a nice tree decomposition, the LP can be constructed and solved in time in $(|\phi| + n \cdot d^w)^{\mathcal{O}(1)}$. In case of a YES-instance, this time also suffices to construct a suitable model.

An illustrative example of the employed construction is provided in the appendix (♣). □

Next, we show that under well-established complexity assumptions, parameterization by d alone cannot yield tractability, even when the primal graph G_ϕ has bounded degree.

Theorem 6 (♣). $\mathsf{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ *is* NP*-complete even if* $d = 2$ *and the maximum degree of* G_ϕ *is* 8.

Proof Sketch. The containment in NP follows from $\mathsf{arbSAT}^{\mathsf{base}}_{\mathsf{prob}} \in \mathsf{NP}$ (Fagin et al., 1990). To show hardness in our restricted setting, we perform a reduction from 3-SAT. Note that 3-SAT remains NP-hard when restricted to formulas in which each variable occurs exactly twice negated and twice non-negated (Darmann & Döcker, 2021). Thus, w.l.o.g., we assume that our formula $\Phi := \bigwedge_i C_i$ with $C_i := \bigvee_{j\in[3]} \ell_{i_j}$ has this property. We construct an instance ϕ of $\mathsf{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ with $\mathbf{V} = \{V_v \mid v \in \mathcal{V}\}$ and $D = \{0, 1\}$ such that for each $C_i \in \Phi$, the constraint $\Pr(\bigvee_{j\in[3]} g(\ell_{i_j})) = 1$ is added to ϕ, where g is defined as in Theorem 1. Since each variable occurs in at most 4 clauses, there are at most 8 other variables it co-occurs with; consequently, G_ϕ has a maximum degree of 8. It remains to argue that Φ is satisfied if and only if ϕ admits an SCM. For this, the crucial insight is that the values of the endogenous variables in any SCM that satisfies ϕ correspond to a satisfying assignment of the variables in Φ. □

The following result complements Theorem 6 by ruling out fixed-parameter tractable algorithms for $\mathsf{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ under a different parameterization, namely the number of variables n. Note that since $\mathrm{tw}(\phi) \leq n$, this implies that we should not expect the primal treewidth of a graph to yield fixed-parameter tractability for $\mathsf{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ alone.

Theorem 7 (♣). $\mathsf{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ *is* $\mathsf{W}[1]$*-hard parameterized by* n.

Proof Sketch. We perform a reduction from k-MULTICOLORED-CLIQUE, which asks, given a properly vertex-colored graph G with colors $1, \ldots, k$ and vertices $v_1, \ldots, v_r$, whether G contains a k-clique. Given G, we construct an instance ϕ of $\mathsf{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ as follows. Let $\mathbf{V} = \{V_1, \ldots, V_k\}$ and $D = \{v_1, \ldots, v_r\}$. For each $i \in [k]$ and $a \in [r]$, add the constraint $\Pr(V_i = v_a) \leq 0$, unless v_a has color i. For each non-adjacent v_a, v_b with $a < b$ and colors i, j, add the constraint $\Pr(V_i = v_a \wedge V_j = v_b) \leq 0$. The construction takes polynomial time and sets $n = k$. It remains to show that G contains a k-clique if and only if ϕ admits an SCM. □

5 LINEAR INEQUALITIES OVER CAUSAL OR COUNTERFACTUAL EXPRESSIONS

In this section, we turn our attention to interventional causal reasoning. We initiate our study by showing that the FPT-tractability that was established in Theorem 5 does not carry over.

Theorem 8 (♣). $\mathsf{SAT}^{\mathsf{lin}}_{\mathsf{causal}}$ *is* NP*-complete even if* $d = 2$ *and* G_ϕ *consists of vertex-disjoint paths of length* 2.

Proof Sketch. Fagin et al. (1990) showed that $\mathsf{arbSAT}^{\mathsf{lin}}_{\mathsf{causal}}$ is NP-complete. To show NP-hardness even for instances ϕ with $d = 2$ and for which G_ϕ consists of vertex-disjoint paths of length 2, we reduce from 3-SAT. Consider a 3-SAT formula Φ with r variables. We define an instance ϕ of $\mathsf{SAT}^{\mathsf{lin}}_{\mathsf{causal}}$ with domain $D = \{0, 1\}$ as follows. For each variable $x_i \in \Phi$, we introduce endogenous variables V_i and $\overline{V}_i$, as well as the constraints

$$\Pr([V_i = 1]\, \overline{V}_i = 1) = 0, \text{ and } \Pr([\overline{V}_i = 1]\, V_i = 1) = 0.$$

Furthermore, for each clause $\ell_1 \vee \ell_2 \vee \ell_3$ in Φ, we add $\Pr(L_1 = 1) + \Pr(L_2 = 1) + \Pr(L_3 = 1) \geq 1$ to ϕ, where, $L_j = V_i$ if $\ell_j = x_i$, and $L_j = \overline{V}_i$ if $\ell_j = \overline{x}_i$ for $j \in [3]$. Note that G_ϕ consists only of edges between V_i and $\overline{V}_i$, for $i \in [r]$. For the proof of correctness, note that for every model $\mathcal{M}$ of ϕ, either $[\![\Pr(V_i = 1)]\!]_\mathcal{M} = 0$ or $[\![\Pr(\overline{V}_i = 1)]\!]_\mathcal{M} = 0$ holds, depending on the relative order of V_i and $\overline{V}_i$ in the well-order $\prec$ in $\mathcal{M}$. □

We contrast the hardness obtained in Theorem 8 by considering the number of variables in $\mathbf{V}$ as a new parameter. Towards this goal, Lemma 9 establishes the existence of a well-structured model for every YES-instance.

Lemma 9. *Let* ϕ *over domain* D *be a YES-instance of* $\mathsf{SAT}^{\mathsf{poly}}_{\mathsf{counterfact}}$ *over variables* $\mathbf{V}$*. There is an ordering* $V_1, \ldots, V_n$ *of* $\mathbf{V}$ *such that the* ϕ *is satisfied by a model* $\mathcal{M} = (\mathbf{V}, \mathbf{U}, \mathcal{F}, \mathbb{P})$ *with the following properties: for each* $i \in [n]$*, let* Q_i *be the set of all possible functions mapping the values of* $V_1, \ldots, V_{i-1}$ *to a value of* V_i*, that is, the set of all functions from* D^{i-1} *to* D *(with* Q_1 *simply being a set of constant functions). Then*

- $\mathbf{U} = \{U\}$ *with* $\mathrm{Val}(U) = Q_1 \times \ldots \times Q_n$*, where* $U[i] \in Q_i$ *denotes the* i^{th} *entry in* U*; and*
- $\mathcal{F} = \{f_{V_i} \mid i \in [n]\}$ *with* $f_{V_i}(U, V_1, \ldots, V_{i-1}) := U[i](V_1, \ldots, V_{i-1})$.

Proof. Let $\mathcal{M}' = (\mathbf{V}, \mathbf{U}', \mathcal{F}', \mathbb{P}')$ be any model witnessing that ϕ is a YES-instance. Without loss of generality, we assume $\mathbf{U}'$ to consist of a single variable U': If there were multiple variables $U_1, U_2, \dots, U_\ell$ in $\mathbf{U}'$, we could replace them by some U' with $\mathrm{Val}(U') = \mathrm{Val}(U_1) \times \mathrm{Val}(U_2) \times \dots \times \mathrm{Val}(U_\ell)$, and update $\mathbb{P}'$ and $\mathcal{F}'$ accordingly.

Consider an ordering $V_1, \dots, V_n$ of $\mathbf{V}$ that respects the implicit well-order $\prec$ of $\mathcal{M}'$. Then, for each $i \in [n]$, $f'_{V_i} \in \mathcal{F}$ describes the value of V_i as a function of U' and $V_1, \dots, V_{i-1}$. Partition $\mathrm{Val}(U')$ such that there is a class C_q for each $q = (q_1, \dots, q_n) \in (Q_1 \times \dots \times Q_n)$ and let it contain $u \in \mathrm{Val}(U')$ if and only if for all $i \in [n]$ and $\overline{s} \in D^{i-1}$, we have that $f'_{V_i}(u, \overline{s}) = q_i(\overline{s})$. Now each $u \in \mathrm{Val}(U')$ belongs to precisely one class C_q.

We construct the model $\mathcal{M} = (\mathbf{V}, \mathbf{U}, \mathcal{F}, \mathbb{P})$ where $\mathbf{U}$ and $\mathcal{F}$ are defined as specified above and $\mathbb{P}$ is such that for each $q \in \mathrm{Val}(U)$ we have $\mathbb{P}(U = q) = \sum_{u' \in C_q} \mathbb{P}'(U' = u')$ (with $\mathbb{P}(U = q) = 0$ if $C_q = \emptyset$). Note that this yields a well-defined probability distribution over U. The model $\mathcal{M}$ satisfies ϕ since every term $\Pr(\varepsilon)$ over $\mathbf{V}$ has the same probability in $\mathcal{M}$ and $\mathcal{M}'$. Indeed, for every class C_q and each event ε, we have that $\mathcal{F}', u \models \varepsilon$ either for all $u \in C_q$ or no $u \in C_q$, as by definition all these u result in the exact same values for the variables in $\mathbf{V}$, even under interventions. Furthermore, $\mathcal{F}$ is such that $\mathcal{F}, q \models \varepsilon$ if and only if $\mathcal{F}', u \models \varepsilon$ for all $u \in C_q$. For any event $\varepsilon \in \mathcal{E}_{\mathsf{counterfact}}$, recall that $S_{\mathcal{M}} \subseteq \mathrm{Val}(U)$ and $S_{\mathcal{M}'} \subseteq \mathrm{Val}(U')$ denote the sets of values of hidden variables such that ε happens in the respective model. We proved that $S_{\mathcal{M}'} = \bigcup_{q \in S_{\mathcal{M}}} C_q$ and thus, by definition of $\mathbb{P}$, event ε happens in both models with the same probability, that is, $[\![\Pr(\varepsilon)]\!]_{\mathcal{M}} = [\![\Pr(\varepsilon)]\!]_{\mathcal{M}'}$. Hence, $\mathcal{M}$ witnesses ϕ to be a YES-instance as well. □

Theorem 10. $\mathsf{SAT}^{\mathsf{lin}}_{\mathsf{counterfact}}$ *is in* FPT *w.r.t. the combined parameter* $d + n$*, and in* XP *w.r.t.* n*.*

Proof. Given an instance ϕ over domain D, we perform the following for each of the $n!$ orders of variables. If we do not find a model for any of these orders, we reject the instance. Consider a total order $V_1, \dots, V_n$ of $\mathbf{V}$ and let U, $\mathcal{F}$, and the Q_i be defined as in Lemma 9. We test whether ϕ admits a model as described in Lemma 9 with respect to that order using the following LP. Create an LP-variable p_q for each $q \in Q_1 \times \dots \times Q_n$. These represent the probability distribution over U, so we add an LP-constraint $\sum_{q \in Q_1 \times \dots \times Q_n} p_q = 1$ and, for each such q, we add an LP-constraint $p_q \geq 0$. Furthermore, using Observation 3, we transform each constraint in ϕ into an LP-constraint by replacing each term $\Pr(\varepsilon)$ by a sum over all variables p_q for which $\mathcal{F}, q \models \varepsilon$. We accept the instance ϕ if and only if there is at least one ordering of $\mathbf{V}$ for which the constructed LP has a solution. In each constructed LP there are $\mathcal{O}(|\phi| + |Q_1 \times \dots \times Q_n|)$ LP-constraints and the number of variables is at most

$$|Q_1 \times \dots \times Q_n| = \prod_{i=1}^{n} |Q_i| \leq \prod_{i=1}^{n} d^{\,i-1} \cdot d \leq d^{\,n^2}.$$

As we can find a solution to an LP (or decide that there is none) in polynomial time with respect to its size (that is, the number of variables plus constraints), the total runtime of the algorithm is $n!(|\phi|^{\mathcal{O}(1)} + d^{\mathcal{O}(n^2)})$.

It remains to prove correctness. If one of the constructed LPs has a solution, let $\mathbb{P}$ be the probability distribution over U as described by the variables p_q in that solution. It is straight-forward to verify that $(\mathbf{V}, \{U\}, \mathcal{F}, \mathbb{P})$ is a model for ϕ, so we correctly accept the instance. For the other direction, suppose ϕ is a YES-instance. Then there is a well-structured model $\mathcal{M} = (\mathbf{V}, \{U\}, \mathcal{F}, \mathbb{P})$ for some ordering of $\mathbf{V}$ by Lemma 9. Observe that the LP constructed for that ordering of the variables has a solution by setting $p_q := \mathbb{P}(U = q)$ for each $q \in \mathrm{Val}(U)$, so the algorithm correctly accepts the instance. □

6 Concluding Remarks

While previous works have focused on mapping the complexity lower bounds for Satisfiability in Pearl's Causal Hierarchy, the parameterized paradigm allows us to identify islands of tractability for the problem. Our contributions include not only these positive results, but also lower bounds which show that the obtained complexity classifications are tight. The presented findings open up several avenues for future work, such as:

- **The Impact of Marginalization.** As mentioned in Section 1, recent works (van der Zander et al., 2023; Dörfler et al., 2025) have proposed an enrichment of the classical fragments in the expressivity matrix M via the summation operator $\sum$ in order to capture marginalization. It would be interesting to explore possible extensions of our approaches to this setting—in particular, can we obtain tractability for these enriched fragments of the PCH without bounding the nesting depth of $\sum$?
- **Treewidth-Guided Linear Programming.** To the best of our knowledge, the proofs of Results **(1)** and **(2)** rely on an entirely novel approach to establishing parameterized tractability with respect to treewidth. Since this approach is particularly tailored to problems that combine discrete structures (graphs) with non-discrete elements (e.g., probabilities), we would not be surprised to see further applications of the technique in the domains of artificial intelligence and machine learning.

References

Mikkel Abrahamsen, Anna Adamaszek, and Tillmann Miltzow. The art gallery problem is $\exists\mathbb{R}$-complete. *CoRR*, abs/1704.06969, 2017.

Elias Bareinboim, Juan D. Correa, Duligur Ibeling, and Thomas Icard. On pearl's hierarchy and the foundations of causal inference. In Hector Geffner, Rina Dechter, and Joseph Y. Halpern (eds.), *Probabilistic and Causal Inference: The Works of Judea Pearl*, volume 36 of *ACM Books*, pp. 507–556. ACM, 2022.

Armin Biere, Marijn Heule, and Hans van Maaren. *Handbook of satisfiability*, volume 185. IOS press, 2009.

Markus Bläser, Julian Dörfler, Maciej Liskiewicz, and Benito van der Zander. Probabilistic and causal satisfiability: Constraining the model. In Keren Censor-Hillel, Fabrizio Grandoni, Joël Ouaknine, and Gabriele Puppis (eds.), *52nd International Colloquium on Automata, Languages, and Programming, ICALP 2025*, volume 334 of *LIPIcs*, pp. 144:1–144:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2025.

Hans L. Bodlaender. A linear-time algorithm for finding tree-decompositions of small treewidth. *SIAM J. Comput.*, 25(6):1305–1317, 1996.

Marek Cygan, Fedor V. Fomin, Łukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michał Pilipczuk, and Saket Saurabh. *Parameterized algorithms*. Springer, Cham, 2015.

Andreas Darmann and Janosch Döcker. On simplified np-complete variants of monotone3-sat. *Discrete Applied Mathematics*, 292:45–58, 2021.

Julian Dörfler, Benito van der Zander, Markus Bläser, and Maciej Liskiewicz. From probability to counterfactuals: the increasing complexity of satisfiability in pearl's causal hierarchy. In *The Thirteenth International Conference on Learning Representations, ICLR 2025*, 2025.

Rodney G. Downey and Michael R. Fellows. *Fundamentals of Parameterized Complexity*. Texts in Computer Science. Springer, 2013.

Eduard Eiben, Robert Ganian, Thekla Hamm, and Sebastian Ordyniak. Parameterized complexity of envy-free resource allocation in social networks. *Artificial Intelligence*, 315:103826, 2023.

Felix Elwert. *Graphical Causal Models*, pp. 245–273. Springer Netherlands, Dordrecht, 2013.

Ronald Fagin, Joseph Y. Halpern, and Nimrod Megiddo. A logic for reasoning about probabilities. *Information and Computation*, 87(1/2):78–128, 1990.

Norman E. Fenton, Martin Neil, and Anthony C. Constantinou. The book of why: The new science of cause and effect, judea pearl, dana mackenzie. basic books (2018). *Artificial Intelligence*, 284:103286, 2020.

Johannes Klaus Fichte and Markus Hecher. Exploiting treewidth for counting projected answer sets. In Michael Thielscher, Francesca Toni, and Frank Wolter (eds.), *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference, KR 2018*, pp. 639–640. AAAI Press, 2018.

Robert Ganian and Viktoriia Korchemna. The complexity of bayesian network learning: Revisiting the superstructure. In Marc'Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021*, pp. 430–442, 2021.

Robert Ganian and Sebastian Ordyniak. The complexity landscape of decompositional parameters for ILP. *Artificial Intelligence*, 257:61–71, 2018.

Robert Ganian, Sebastian Ordyniak, and M. S. Ramanujan. Going beyond primal treewidth for (M)ILP. In Satinder Singh and Shaul Markovitch (eds.), *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp. 815–821. AAAI Press, 2017.

Robert Ganian, Thekla Hamm, Viktoriia Korchemna, Karolina Okrasa, and Kirill Simonov. The complexity of k-means clustering when little is known. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato (eds.), *International Conference on Machine Learning, ICML 2022*, volume 162 of *Proceedings of Machine Learning Research*, pp. 6960–6987. PMLR, 2022.

Robert Ganian, Viktoriia Korchemna, and Stefan Szeider. Revisiting causal discovery from a complexity-theoretic perspective. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI 2024*, pp. 3377–3385. ijcai.org, 2024.

George F. Georgakopoulos, Dimitris J. Kavvadias, and Christos H. Papadimitriou. Probabilistic satisfiability. *Journal of Complexity*, 4(1):1–11, 1988.

Clark Glymour, Richard Scheines, Peter Spirtes, and Kevin Kelly. *Discovering causal structure: Artificial intelligence, philosophy of science, and statistical modeling*. Academic Press, 1987.

Niels Grüttemeier and Christian Komusiewicz. Learning bayesian networks under sparsity constraints: A parameterized complexity analysis. *Journal of Artificial Intelligence Research*, 74:1225–1267, 2022.

Niels Grüttemeier, Christian Komusiewicz, and Nils Morawietz. Efficient bayesian network structure learning via parameterized local search on topological orderings. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021*, pp. 12328–12335. AAAI Press, 2021a.

Niels Grüttemeier, Christian Komusiewicz, and Nils Morawietz. On the parameterized complexity of polytree learning. In Zhi-Hua Zhou (ed.), *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI 2021*, pp. 4221–4227. ijcai.org, 2021b.

Tesshu Hanaka and Michael Lampis. Hedonic games and treewidth revisited. In Shiri Chechik, Gonzalo Navarro, Eva Rotenberg, and Grzegorz Herman (eds.), *30th Annual European Symposium on Algorithms, ESA 2022*, volume 244 of *LIPIcs*, pp. 64:1–64:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

Duligur Ibeling and Thomas Icard. Probabilistic reasoning across the causal hierarchy. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020*, pp. 10170–10177. AAAI Press, 2020.

Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *Journal of Computer and System Science*, 63(4):512–530, 2001.

Karthik C. S., Dániel Marx, Marcin Pilipczuk, and Uéverton S. Souza. Conditional lower bounds for sparse parameterized 2-csp: A streamlined proof. In Merav Parter and Seth Pettie (eds.), *2024 Symposium on Simplicity in Algorithms, SOSA 2024*, pp. 383–395. SIAM, 2024.

Daphne Koller and Nir Friedman. *Probabilistic Graphical Models - Principles and Techniques*. MIT Press, 2009.

Milan Mossé, Duligur Ibeling, and Thomas Icard. Is causal reasoning harder than probabilistic reasoning? *The Review of Symbolic Logic*, 17(1):106–131, 2024.

Nils J. Nilsson. Probabilistic logic. *Artificial Intelligence*, 28(1):71–87, 1986.

Sebastian Ordyniak and Stefan Szeider. Parameterized complexity results for exact bayesian network structure learning. *Journal of Artificial Intelligence Research*, 46:263–302, 2013.

Christos H Papadimitriou and Kenneth Steiglitz. *Combinatorial optimization: algorithms and complexity*. Courier Corporation, 1998.

Judea Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, USA, 2nd edition, 2009.

Dominik Peters. Graphical hedonic games of bounded treewidth. In Dale Schuurmans and Michael P. Wellman (eds.), *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, pp. 586–593. AAAI Press, 2016.

Neil Robertson and Paul D. Seymour. Graph minors. III. planar tree-width. *Journal of Combinatorial Theory B*, 36(1):49–64, 1984.

Marko Samer and Stefan Szeider. Constraint satisfaction with bounded treewidth revisited. *Journal of Computer and System Science*, 76(2):103–114, 2010.

Ilya Shpitser and Judea Pearl. Complete identification methods for the causal hierarchy. *Jornal of Machine Learning Research*, 9:1941–1979, 2008.

Benito van der Zander, Markus Bläser, and Maciej Liskiewicz. The hardness of reasoning about probabilities and causality. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI 2023*, pp. 5730–5738, 2023.

Junzhe Zhang, Jin Tian, and Elias Bareinboim. Partial counterfactual identification from observational and experimental data. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato (eds.), *International Conference on Machine Learning, ICML 2022*, volume 162 of *Proceedings of Machine Learning Research*, pp. 26548–26558. PMLR, 2022.

GATEWAYS TO TRACTABILITY FOR SATISFIABILITY IN PEARL'S CAUSAL HIERARCHY (APPENDIX: FULL VERSION)

Anonymous authors
Paper under double-blind review

ABSTRACT

Pearl's Causal Hierarchy (PCH) is a central framework for reasoning about probabilistic, interventional, and counterfactual statements, yet the satisfiability problem for PCH formulas is computationally intractable in almost all classical settings. We revisit this challenge through the lens of parameterized complexity and identify the first gateways to tractability. Our results include fixed-parameter and XP-algorithms for satisfiability in key probabilistic and counterfactual fragments, using parameters such as primal treewidth and the number of variables, together with matching hardness results that map the limits of tractability. Technically, we depart from the dynamic programming paradigm typically employed for treewidth-based algorithms and instead exploit structural characterizations of well-formed causal models, providing a new algorithmic toolkit for causal reasoning.

1 INTRODUCTION

Pearl's Causal Hierarchy (PCH) (Shpitser & Pearl, 2008; Pearl, 2009) is a central pillar of the modern theory of causality that is employed in artificial intelligence and other reasoning tasks—see, e.g., the recent survey (Bareinboim et al., 2022) or book (Fenton et al., 2020) on the topic. The PCH is a framework that has three basic layers of depth which capture three fundamental degrees of sophistication for analyzing causal effects and relationships. All of these layers provide a means of formalizing statements via formulas capturing the behavior of a set of probabilistic variables in a *Structural Causal Model* (SCM) (Glymour et al., 1987; Pearl, 2009; Koller & Friedman, 2009; Elwert, 2013), which is a well-established representation of systems with observed as well as hidden variables over a specified domain and their mutual dependencies. As a basic illustrative example, the statement "the likelihood of having both diabetes (D = yes) and blood type B+ (T = B+) is at most 1%" can be expressed by the formula $\psi = \Pr(D = \text{yes} \wedge T = \text{B+}) \leq 0.01$.

The formula ψ above belongs to the first, basic layer of the PCH—that is, the layer $\mathcal{L}_{\mathsf{prob}}$ of probabilistic reasoning that captures direct statements one can make about the probabilities of certain outcomes. The second layer, $\mathcal{L}_{\mathsf{causal}}$, expands on the basic probability terms in $\mathcal{L}_{\mathsf{prob}}$ via the introduction of Pearl's do-operator (Pearl, 2009) which captures interventional causal reasoning. A basic example of an event that can be captured on this layer of the PCH is contracting a disease after being vaccinated against that disease; the probability of this event can be expressed using the term $\Pr([Y = \text{vaccinated}]\ X = \text{contracted})$[1], where the square brackets denote an intervention that is applied before observing the outcome.[2] Hence, the second layer of the PCH allows us to make statements such as $\Pr([Y = \text{vaccinated}]\ X = \text{contracted}) < \Pr(X = \text{contracted})$. The third layer $\mathcal{L}_{\mathsf{counterfact}}$ of the PCH expands on $\mathcal{L}_{\mathsf{causal}}$ by allowing interventions to be chained, and enables complex statements related to counterfactual situations. For instance, a third-layer term such as $\Pr\big(([M = \text{yes}]\ H = \text{no}) \mid (M = \text{no} \wedge H = \text{yes})\big)$ can express the probability that a patient who did not take medication (M) and was hospitalized (H) would have avoided hospitalization if he had taken the medication. Formal definitions of these as well as related notions are available in Section 2.

While the three layers of depth of the PCH focus on the expressivity inside the probability term $\Pr(\cdot)$, there is a second dimension to the PCH—specifically, the *breadth* of operations that can be applied to the probability terms themselves. For $\circledast \in \{\mathsf{prob}, \mathsf{causal}, \mathsf{counterfact}\}$, we distinguish the following fragments of the PCH:

[1]Equivalently, $\Pr(X = \text{contracted} \mid do(Y = \text{vaccinated}))$. We follow recent publications in the area (van der Zander et al., 2023; Dörfler et al., 2025) and primarily employ the square-bracket notation.

[2]Interventions are distinct from conditional probability statements such as $\Pr(X = \text{contracted} \mid Y = \text{vaccinated})$. To see this, consider a hypothetical world where the vaccine is ineffective, the disease only exists in a laboratory and an oracle randomly determines whether a person will be infected without vaccination, or receive the vaccine and not come in contact with the disease. In this world, $\Pr(X = \text{contracted} \mid Y = \text{vaccinated}) = 0$ but $\Pr([Y = \text{vaccinated}]\ X = \text{contracted}) > 0$.

- $\mathcal{L}^{\text{base}}_{\circledast}$: only simple probability terms are allowed, such as $\Pr(\cdot) \leq \Pr(\circ)$ or $\Pr(\cdot) \geq 1$;
- $\mathcal{L}^{\text{lin}}_{\circledast}$: linear combinations of probability terms are allowed, such as $\Pr(\cdot) - \Pr(\circ) \geq 3\Pr(\bullet)$;
- $\mathcal{L}^{\text{poly}}_{\circledast}$: polynomials over probability terms are allowed, such as $\Pr(\cdot)^2 \leq 2\Pr(\circ) \cdot \Pr(\bullet) + 0.1$.

Crucially, the various combinations of depth and breadth give rise to a 3×3 *expressivity matrix* M for PCH (Dörfler et al., 2025, Table 1), (van der Zander et al., 2023, Table 1).

A crucial and well-studied problem in the setting of causal reasoning is SATISFIABILITY—that is, determining whether a given formula (consisting of a set of probability constraints) admits an SCM (Fagin et al., 1990; Ibeling & Icard, 2020; van der Zander et al., 2023; Mossé et al., 2024; Dörfler et al., 2025). We note that there is a high-level parallel between this SATISFIABILITY problem in the causal setting and the well-known BOOLEAN SATISFIABILITY (SAT) and CONSTRAINT SATISFACTION (CSP) problems; the distinction lies in the types of constraints on the input and the nature of the sought-after model. However, solving SATISFIABILITY in our causal reasoning setting is, in general, a much more daunting task. If we let $\text{SAT}^{*}_{\circledast}$ denote the SATISFIABILITY problem for formulas from the fragment $\mathcal{L}^{*}_{\circledast}$ of the PCH, then depending on the choice of $\circledast \in \{\text{prob}, \text{causal}, \text{counterfact}\}$ and $* \in \{\text{base}, \text{lin}, \text{poly}\}$ the problem under consideration will be complete for the complexity classes NP or $\exists\mathbb{R}$—see also the detailed discussion of related work at the end of this section.

Crucially, while previous works have made significant strides towards mapping the classical complexity landscape of the SATISFIABILITY problem, even the "easiest" fragments of the expressivity matrix remain NP-hard. The central aim of this article is to provide a counterweight to this pessimistic perspective and identify fundamental gateways to tractability for SATISFIABILITY, specifically by employing the more refined *parameterized complexity* paradigm (Downey & Fellows, 2013; Cygan et al., 2015). There, one analyzes the running time of algorithms not only in terms of the input size $|I|$, but also with respect to a specified numerical parameter k. The standard notion of tractability used in this setting is then tied to algorithms which run in time $f(k) \cdot |I|^{\mathcal{O}(1)}$ for some computable function f; problems admitting such *fixed-parameter* algorithms are called *fixed-parameter tractable* (FPT). A weaker—but nevertheless still useful—notion of tractability stems from the existence of a so-called XP-*algorithm*, i.e., an algorithm running in time $|I|^{f(k)}$ (this gives rise to the complexity class XP). Our main results include not only the first fixed-parameter and XP-algorithms for the problem, but also matching lower bounds which allow us to identify the limits of parameterized tractability in the expressivity matrix.

Contributions. A loose inspiration for this work stems from the success stories in the aforementioned domains of BOOLEAN SATISFIABILITY and CONSTRAINT SATISFACTION. The parameterized complexity of these two problems is by now very well understood, and perhaps the most classical parameterized algorithms use the *primal treewidth* as the parameter of choice. Essentially, this measures how "tree-like" the interactions between the variables in the instance are—more precisely, this is captured by measuring the *treewidth*, a fundamental graph parameter (Robertson & Seymour, 1984), of the graph obtained by representing variables as vertices and using edges to capture the property of lying in the same "term" (i.e., clause or constraint). In particular, it is known that BOOLEAN SATISFIABILITY is fixed-parameter tractable w.r.t. the primal treewidth (Biere et al., 2009, Chapter 13), while CONSTRAINT SATISFACTION admits an XP-algorithm under the same parameterization (Samer & Szeider, 2010); the latter then becomes fixed-parameter tractable when the parameterization also includes the domain size for the variables (Samer & Szeider, 2010).

Given the above, it is natural to ask whether one can use the primal treewidth to establish tractability for SATISFIABILITY in the PCH setting. As our first set of contributions, we provide a complete answer to this question:

> $\text{SAT}^{\text{lin}}_{\text{prob}}$ is **(1)** in XP w.r.t. the primal treewidth alone, and
> **(2)** fixed-parameter tractable w.r.t. the primal treewidth plus the domain size d.
>
> Moreover, under well-established complexity assumptions one can neither
> **(3)** improve the XP-tractability to FPT (not even for $\text{SAT}^{\text{base}}_{\text{prob}}$), nor
> **(4)** lift **any** of these tractability results to $\text{SAT}^{\text{poly}}_{\text{prob}}$ or $\text{SAT}^{\text{lin}}_{\text{causal}}$.

Furthermore, we remark that parameterizing by the domain size alone does not yield tractability under well-established complexity assumptions (see Theorem 6).

While the above results are comprehensive, they only provide a gateway to tractability for the "shallowest" probabilistic fragment of the PCH. We hence ask whether one can achieve tractability for deeper fragments of the PCH (that is, $\mathcal{L}^{*}_{\text{causal}}$ and $\mathcal{L}^{*}_{\text{counterfact}}$) if the primal treewidth is replaced with a more restrictive parameterization—specifically the number n of variables in the formula. We note that the analogous question in the BOOLEAN SATISFIABILITY and CONSTRAINT SATISFACTION setting is trivial: there, asymptotically optimal (under well-established complexity assumptions) algorithms parameterized by n can merely enumerate all possible models (Impagliazzo et al., 2001; Karthik C. S. et al., 2024). Such an approach is doomed to fail for the causal SATISFIABILITY problem: not only will an SCM contain (potentially many) auxiliary random variables, but also variable dependencies and random distributions that cannot be exhaustively enumerated.

As our second set of contributions, we map the complexity landscape for deeper fragments of the PCH as well:

> $\mathsf{SAT}^{\mathsf{lin}}_{\mathsf{counterfact}}$ is **(5)** in XP w.r.t. n alone, and
> **(6)** fixed-parameter tractable w.r.t. n plus the domain size d.
>
> Moreover, under well-established complexity assumptions one can neither
> **(7)** improve the XP-tractability to FPT (not even for $\mathsf{SAT}^{\mathsf{base}}_{\mathsf{prob}}$), nor
> **(8)** lift **any** of these tractability results to $\mathsf{SAT}^{\mathsf{poly}}_{\mathsf{counterfact}}$.

A schematic overview of our contributions is provided in the mind map below (Figure 1).

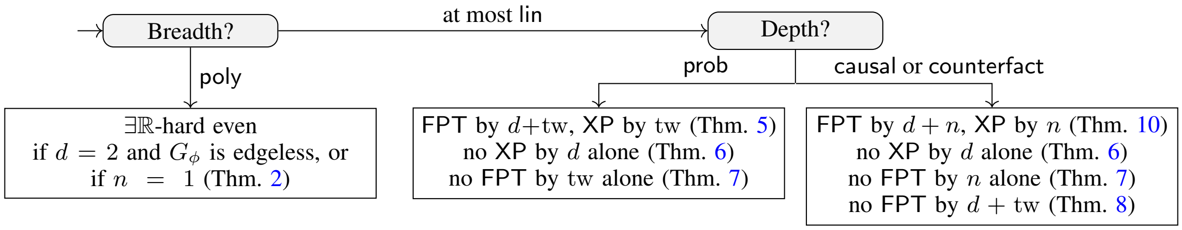


Figure 1: Parameterized complexity of SATISFIABILITY in the PCH based on the position (breadth/depth) in the expressivity matrix M. All results hold under well-established complexity assumptions and refer to an instance with n observed variables over a domain of size d such that the treewidth of the primal graph G_ϕ is tw. We note that the $\exists\mathbb{R}$-hardness of the poly fragment was already established by Mossé et al. (2024), but not under the stated restrictions which rule out tractability in the parameterized setting.

Extensions to Marginalization. In order to efficiently express marginalization, recent works (van der Zander et al., 2023; Dörfler et al., 2025) have extended the classical fragments in M to $\mathcal{L}^{\mathsf{base}\langle\Sigma\rangle}_{\circledast}$, $\mathcal{L}^{\mathsf{lin}\langle\Sigma\rangle}_{\circledast}$, and $\mathcal{L}^{\mathsf{poly}\langle\Sigma\rangle}_{\circledast}$, respectively; the only difference is that these classes additionally include the unary summation operator $\sum$. Depending on the specific fragment considered, including these marginalization operators in the SATISFIABILITY problem yields completeness for the complexity classes $\mathsf{NP}^{\mathsf{PP}}$, PSPACE, NEXP or succ-$\exists\mathbb{R}$—see Dörfler et al. (2025, Table 1). Since the algorithm(s) underlying Results **(5)** and **(6)** can also be used to establish inclusion in the complexity class $\mathsf{EXPTIME}$ while $\mathsf{SAT}^{\mathsf{lin}\langle\Sigma\rangle}_{\mathsf{counterfact}}$ is NEXP-complete (Dörfler et al., 2025), under well-established complexity assumptions it is not possible to lift our results towards full marginalization operators as considered in the aforementioned works. Nevertheless, if one were to bound the nesting depth of the unary summation operator $\sum$ to any (arbitrary but fixed) constant, all of our results could be directly translated to the marginalization setting by simply expanding on the respective sums.

Proof Techniques. The standard approach to establishing tractability for problems parameterized by treewidth is to employ dynamic programming—this is the approach used not only for the aforementioned treewidth-based algorithms solving BOOLEAN SATISFIABILITY and CONSTRAINT SATISFACTION, but also for almost every algorithm parameterized by treewidth. From a technical standpoint, it is hence surprising that our results **do not** employ dynamic programming at all; in fact, the SATISFIABILITY problem seems entirely incompatible with the basic tenets of the usual "leaf-to-root" dynamic programming paradigm used for treewidth.

Instead, our proof of Results **(1)** and **(2)** relies on an entirely novel approach. We first prove that every YES-instance of $\mathsf{SAT}^{\mathsf{lin}}_{\mathsf{prob}}$ with primal treewidth k admits a "well-structured" SCM whose hidden variables and dependencies can be neatly mapped onto the tree-like structure of the primal graph and determined in advance. However, this step on its own cannot determine whether an SCM actually exists, as for that we need to compute and verify the probability distributions for the hidden variables. In the second step, we use the tree-likeness of the instance once again to construct a "fixed-parameter sized" linear program which either computes a viable set of probability distributions, or determines that none exists. It is well-known that linear programs can be solved in polynomial time (Papadimitriou & Steiglitz, 1998)—the difficult part lies in building a program that provably verifies the existence of an SCM while avoiding an exponential dependency on the input size.

In order to apply the reduction technique to Results **(5)** and **(6)**, we need to be able to deal with the presence of interventions in the formula. Towards this, we argue that every YES-instance of $\text{SAT}^{\text{lin}}_{\text{counterfact}}$ admits an SCM with different structural properties than those used for Results **(5)** and **(6)**: in particular, the value of a single hidden variable U determines not just the value of each observed variable but the *function* of how it is determined from the other observed variables. We then define a suitable linear programming formulation that targets the computation of such well-structured SCMs.

For establishing the lower-bound results, we develop three distinct reductions: one from k-MULTICOLORED-CLIQUE which handles **(3)** and **(7)**, one from a restricted variant of the Existential Theory of the Reals problem for Results **(4)** and **(8)**, and a separate reduction from 3-SAT for the remaining lin-causal case of **(4)**.

Related Work. $\text{SAT}^{\text{lin}}_{\text{prob}}$ and $\text{SAT}^{\text{base}}_{\text{prob}}$ were shown to be NP-complete by Fagin et al. (1990), and analogous results for the fragments $\text{SAT}^{\text{lin}}_{\text{causal}}$, $\text{SAT}^{\text{base}}_{\text{causal}}$, $\text{SAT}^{\text{lin}}_{\text{counterfact}}$, and $\text{SAT}^{\text{base}}_{\text{counterfact}}$ were obtained by Mossé et al. (2024). The $\exists\mathbb{R}$-completeness of $\text{SAT}^{\text{poly}}_{\text{prob}}$, $\text{SAT}^{\text{poly}}_{\text{causal}}$ and $\text{SAT}^{\text{poly}}_{\text{counterfact}}$ was also established in the latter work (Mossé et al., 2024). As mentioned above, these complexity-theoretic studies were recently extended to languages containing the summation operator $\sum$ (van der Zander et al., 2023; Dörfler et al., 2025). Other related languages designed to express probabilistic reasoning were developed in the works of, e.g., Nilsson (1986); Georgakopoulos et al. (1988); Ibeling & Icard (2020). Moreover, the existence of solutions to the SATISFIABILITY problem with specific properties has very recently been studied by Bläser et al. (2025).

Beyond SATISFIABILITY, the parameterized complexity paradigm has been employed in several works studying another central problem in the area of causality: BAYESIAN NETWORK STRUCTURE LEARNING. This line of research was initiated by Ordyniak & Szeider (2013), with recent contributions considering a broad range of parameterizations as well as variations of the problem (Ganian & Korchemna, 2021; Grüttemeier et al., 2021a;b; Grüttemeier & Komusiewicz, 2022). The complexity of the related CAUSAL DISCOVERY problem was recently studied by Ganian et al. (2024).

Beyond the aforementioned prominent applications in BOOLEAN SATISFIABILITY and CONSTRAINT SATISFACTION, primal treewidth has been used as a natural means of capturing structural properties of inputs in a variety of other settings as well. Examples of this in the broad AI area include its applications in INTEGER LINEAR PROGRAMMING (Ganian et al., 2017; Ganian & Ordyniak, 2018), HEDONIC GAMES (Peters, 2016; Hanaka & Lampis, 2022), MATRIX COMPLETION (Ganian et al., 2022), ANSWER SET PROGRAMMING (Fichte & Hecher, 2018), and RESOURCE ALLOCATION (Eiben et al., 2023). We note that the treewidth-based algorithms in all of the aforementioned works rely on dynamic programming, which is fundamentally different from the technique employed to achieve our Results **(1)** and **(2)**.

2 PRELIMINARIES

For $n \in \mathbb{N}$, let $[n] = \{1, \dots, n\}$. For $i_1, i_2 \in \mathbb{R}$, let $[i_1, i_2] = \{j \in \mathbb{R} \mid i_1 \le j \le i_2\}$. We follow established notation as used in (Mossé et al., 2024; van der Zander et al., 2023). By $\mathbf{V}$ we refer to a contingent of random variables and, without loss of generality, assume that each of these share a given domain D of size d.

Syntax of the languages of PCH. For $V \in \mathbf{V}$ and $v \in D$, a statement of the form $V = v$ is called an *atom*. We can combine multiple atoms to obtain *events* over $\mathbf{V}$ by applying the following grammatical rules.

$$\begin{aligned}
\mathcal{E}_{\text{prop}} &::= V = v \mid \neg\mathcal{E}_{\text{prop}} \mid \mathcal{E}_{\text{prop}} \wedge \mathcal{E}_{\text{prop}},\\
\mathcal{E}_{\text{int}} &::= \top \mid V = v \mid \mathcal{E}_{\text{int}} \wedge \mathcal{E}_{\text{int}},\\
\mathcal{E}_{\text{post-int}} &::= [\mathcal{E}_{\text{int}}]\, \mathcal{E}_{\text{prop}},\\
\mathcal{E}_{\text{counterfact}} &::= \mathcal{E}_{\text{post-int}} \mid \neg\mathcal{E}_{\text{counterfact}} \mid \mathcal{E}_{\text{counterfact}} \wedge \mathcal{E}_{\text{counterfact}}.
\end{aligned}$$

We call the events in $\mathcal{E}_{\text{prop}}$ *propositions* and the events in $\mathcal{E}_{\text{int}}$ *interventions*. Each event ε can only occur within a probabilistic statement $\Pr(\varepsilon)$, which we call a *term*. The *size of a term* is the number of atoms it contains. For $\mathcal{E} \in \{\mathcal{E}_{\text{prop}}, \mathcal{E}_{\text{post-int}}, \mathcal{E}_{\text{counterfact}}\}$ and $\varepsilon \in \mathcal{E}$, we define the following valid ways of combining terms.

$$\begin{aligned}
T_{\text{base}}(\mathcal{E}) &::= \Pr(\varepsilon),\\
T_{\text{lin}}(\mathcal{E}) &::= \Pr(\varepsilon) \mid T_{\text{lin}}(\mathcal{E}) + T_{\text{lin}}(\mathcal{E}),\\
T_{\text{poly}}(\mathcal{E}) &::= \Pr(\varepsilon) \mid T_{\text{poly}}(\mathcal{E}) + T_{\text{poly}}(\mathcal{E}) \mid T_{\text{poly}}(\mathcal{E}) \cdot T_{\text{poly}}(\mathcal{E}).
\end{aligned}$$

Lastly, for $* \in \{\text{base}, \text{lin}, \text{poly}\}$ we define $\mathcal{L}^*_{\text{prob}}$ ($\mathcal{L}^*_{\text{causal}}$ and $\mathcal{L}^*_{\text{counterfact}}$, respectively) to be the *language* that contains all sets of inequalities over elements in $T_*(\mathcal{E}_{\text{prop}})$ (in $T_*(\mathcal{E}_{\text{post-int}})$ and in $T_*(\mathcal{E}_{\text{counterfact}})$, resp.). The elements inside $\mathcal{L}^*_{\text{prob}}$, $\mathcal{L}^*_{\text{causal}}$, and $\mathcal{L}^*_{\text{counterfact}}$ are called *formulas*. Note that tautological and contradictory events can be used to encode comparisons against 1 and 0, such as $\Pr(\varepsilon) \leq 0$. Moreover, the grammars of $\mathcal{L}^{\text{lin}}_{\circledast}$ and $\mathcal{L}^{\text{poly}}_{\circledast}$ support integer coefficients, which can be effectively constructed by summing up multiple probabilities of the same type. Any inequality with rational coefficients can be encoded by multiplying both sides with the smallest common multiple of all non-integer coefficients. At the beginning of the next section, we will compare our syntax to the one used in related work.

Semantics of the Languages of PCH. We define the semantics of the aforementioned languages using the notion of Structural Causal Models as popularized by Glymour et al. (1987) and Pearl (2009, Section 3.2). A *recursive Structural Causal Model* (*SCM*, or simply *model*) over domain D is a tuple $\mathcal{M} = (\mathbf{V}, \mathbf{U}, \mathcal{F}, \mathbb{P})$ with

- a set $\mathbf{V}$ of endogenous (observed) variables, implicitly well-ordered by $\prec$, that range over D,
- a set $\mathbf{U}$ of exogenous (hidden) variables,
- a set $\mathcal{F} = \{f_V\}_{V \in \mathbf{V}}$ of functions where f_V specifies how the value of V can be computed given the values of $\mathbf{U}$ and $\mathbf{V}_{\prec V}$, that is, the subset of $\mathbf{V}$ that precedes $V \in \mathbf{V}$ in $\prec$,
- a probability distribution $\mathbb{P}$ on $\mathbf{U}$.

Note that any model $\mathcal{M}$ whose exogenous variables $\mathbf{U}$ have an infinite or continuous domain is (w.r.t. its evaluation) equivalent to a model $\mathcal{M}'$ where all exogenous variables have discrete and finite domains (Zhang et al., 2022). Consequently, we assume throughout that each variable $U \in \mathbf{U}$ has a discrete and finite domain $\mathrm{Val}(U)$, and let $\mathrm{Val}(\mathbf{U}) = \mathrm{Val}(U_1) \times \ldots \times \mathrm{Val}(U_{|\mathbf{U}|})$ refer to their combined range.

Let $V = v$ be an atom in $\mathcal{E}_{\text{int}}$. We denote by $\mathcal{F}_{V=v}$ the set of functions obtained from $\mathcal{F}$ by replacing f_V with the constant function v. We generalize this definition to arbitrary conjunctions of atoms $\gamma \in \mathcal{E}_{\text{int}}$ in the natural way and denote the set of resulting functions as $\mathcal{F}_\gamma$. Let $\varepsilon \in \mathcal{E}_{\text{prop}}$ and $\overline{u} \in \mathrm{Val}(\mathbf{U})$. We write $\mathcal{F}, \overline{u} \models \varepsilon$ if evaluating $\mathcal{F}$ on input $\overline{u}$ yields an assignment to $\mathbf{V}$ under which ε is satisfied. For $[\gamma]\, \varepsilon \in \mathcal{E}_{\text{post-int}}$, we write $\mathcal{F}, \overline{u} \models [\gamma]\, \varepsilon$ if $\mathcal{F}_\gamma, \overline{u} \models \varepsilon$. Moreover, for all $\varepsilon, \varepsilon_1, \varepsilon_2 \in \mathcal{E}_{\text{counterfact}}$, we write (i) $\mathcal{F}, \overline{u} \models \neg\varepsilon$ if $\mathcal{F}, \overline{u} \not\models \varepsilon$, and (ii) $\mathcal{F}, \overline{u} \models \varepsilon_1 \wedge \varepsilon_2$ if both $\mathcal{F}, \overline{u} \models \varepsilon_1$ and $\mathcal{F}, \overline{u} \models \varepsilon_2$. For a given model $\mathcal{M} = (\mathbf{V}, \mathbf{U}, \mathcal{F}, \mathbb{P})$, we denote $S_{\mathcal{M}} := \{\overline{u} \in \mathrm{Val}(\mathbf{U}) \mid \mathcal{F}, \overline{u} \models \varepsilon\}$. The way $\mathcal{M}$ interprets an expression $\mathbf{t} \in T_{\text{poly}}(\mathcal{E})$ is denoted by $[\![\mathbf{t}]\!]_{\mathcal{M}}$ and recursively defined as follows: $[\![\Pr(\varepsilon)]\!]_{\mathcal{M}} = \sum_{\overline{u} \in S_{\mathcal{M}}(\varepsilon)} \mathbb{P}(\overline{u})$. For two expressions $\mathbf{t}_1, \mathbf{t}_2 \in T_{\text{poly}}(\mathcal{E})$, we define $\mathcal{M} \models \mathbf{t}_1 \leq \mathbf{t}_2$ if and only if $[\![\mathbf{t}_1]\!]_{\mathcal{M}} \leq [\![\mathbf{t}_2]\!]_{\mathcal{M}}$. The semantics for negation and conjunction are defined in the usual way, yielding the semantics for $\mathcal{M} \models \phi$ for any formula $\phi \in \mathcal{L}^{\text{poly}}_{\text{counterfact}}$.

Primal Treewidth. Let $\phi \in \mathcal{L}^{\text{poly}}_{\text{counterfact}}$ be a formula over variables $\mathbf{V}$. By $G_\phi = (\mathbf{V}, E)$, we denote the *primal graph* of ϕ, where $\{V, V'\} \in E$ if and only if $V \neq V'$ and there is a term in ϕ that contains both V and V'. A *nice tree decomposition* of G_ϕ is a pair $(\mathbf{T}, \chi)$, where $\mathbf{T}$ is a tree (whose vertices are called *nodes*) rooted at a node N_0 and χ is a function that assigns to each node N a set $\chi(N) \subseteq \mathbf{V}$ such that:

- For every $\{V, V'\} \in E$, there is a node N such that $\{V, V'\} \subseteq \chi(N)$.
- For every vertex $V \in \mathbf{V}$, the set of nodes N satisfying $V \in \chi(N)$ forms a subtree of $\mathbf{T}$.
- $|\chi(N)| = 0$ if N is a leaf of $\mathbf{T}$ or $N = N_0$.
- There are only three kinds of non-leaf nodes in $\mathbf{T}$:
  - *introduce*: a node N with exactly one child N' such that $\chi(N) = \chi(N') \cup \{V\}$ for a $V \notin \chi(N')$.
  - *forget*: a node N with exactly one child N' such that $\chi(N) = \chi(N') \setminus \{V\}$ for a $V \in \chi(N')$.
  - *join*: a node N with two children N_1, N_2 such that $\chi(N) = \chi(N_1) = \chi(N_2)$.

We call each set $\chi(N)$ a *bag*. The width of a nice tree decomposition $(\mathbf{T}, \chi)$ is the size of the largest bag $\chi(N)$ minus 1, and the *treewidth of* G_ϕ, denoted by $\mathrm{tw}(G_\phi)$, is the minimum width of a nice tree decomposition of G_ϕ. We let the (*primal*) *treewidth of a formula* ϕ denote the treewidth of its primal graph, that is, $\mathrm{tw}(\phi) = \mathrm{tw}(G_\phi)$.

The Class $\exists\mathbb{R}$. The Existential Theory of the Reals (ETR) contains all true sentences of the form

$$\exists x_1 \ldots \exists x_k \phi(x_1, \ldots, x_k),$$

where ϕ is a quantifier-free Boolean formula over the basis $\{\wedge, \vee, \neg\}$ and a signature consisting of constant symbols (0 and 1), function symbols ($+$ and $\cdot$), and predicates ($<, \leq$, and $=$). The sentence is interpreted over the real numbers in the standard way. The closure of ETR under polynomial time many-one reductions yields the complexity class $\exists\mathbb{R}$ (Grigoriev & Jr., 1988; Canny, 1988). For a comprehensive compendium on $\exists\mathbb{R}$, see Schaefer et al. (2024); here, we only require the class for the lower bound established in Theorem 2.

3 Satisfiability for Languages of PCH and Structural Insights

In this paper, we examine several analogues of the well-known problem Boolean Satisfiability that capture various probabilistic, causal, and counterfactual statements. We denote these problems as $\mathrm{SAT}^{*}_{\circledast}$, where $\circledast \in \{\mathsf{prob}, \mathsf{causal}, \mathsf{counterfact}\}$ and $* \in \{\mathsf{base}, \mathsf{lin}, \mathsf{poly}\}$, and define them as follows.

$\mathrm{SAT}^{*}_{\circledast}$	**Input:**	A set D of d domain values and a formula $\phi \in \mathcal{L}^{*}_{\circledast}$ over variables $\mathbf{V} = \{V_1, \ldots, V_n\}$.
	Task:	Decide if there exists a recursive Structural Causal Model $\mathcal{M}$ over D such that $\mathcal{M} \models \phi$.

The classical computational complexity of $\mathrm{SAT}^{*}_{\circledast}$ has by now been studied extensively (Fagin et al., 1990; Ibeling & Icard, 2020; van der Zander et al., 2023; Mossé et al., 2024; Dörfler et al., 2025). We remark that our definition of $\mathrm{SAT}^{*}_{\circledast}$ slightly deviates from the one established in previous works, in the sense that we restrict our attention to input formulas ϕ that are *sets of inequalities* (that is, each inequality forms a constraint) rather than allowing arbitrary Boolean combinations of inequalities. However, this restriction does not affect any of the known complexity-theoretic results, since previous lower-bound proofs did not employ any Boolean combinations beyond sets. However, the situation changes drastically when studying $\mathrm{SAT}^{*}_{\circledast}$ from the viewpoint of parameterized complexity, as we show next.

Let $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ denote the version of $\mathrm{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ in which ϕ is an arbitrary Boolean combination of inequalities over elements in $T_{\mathsf{base}}(\mathcal{E}_{\mathsf{prob}})$. We justify our restriction to $\mathrm{SAT}^{*}_{\circledast}$ by showing that $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ remains NP-complete in a very restricted setting, thus dashing any hope to exploit structural properties of ϕ.

Theorem 1. $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ *is* NP*-complete even if G_ϕ is edgeless and $d = 2$.*

Proof. Fagin et al. (1990) proved that $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ is NP-complete. In order to prove the NP-hardness of $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ even for instances in which each constraint consists of only one variable and all variables have domain $D = \{0, 1\}$, we perform a reduction from 3-SAT. Let $\Phi := \bigwedge_i C_i$ with $C_i := \bigvee_{j \in [3]} \ell_{i_j}$ be a 3-SAT formula over variables $\mathcal{V}$. Define an instance ϕ of $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ with $\mathbf{V} = \{V_v \mid v \in \mathcal{V}\}$ over domain $\{0, 1\}$ as

$$\phi := \bigwedge_i \Big(\bigvee_{j \in [3]} \Pr(g(\ell_{i_j})) = 1 \Big) \wedge \bigwedge_{v \in \mathcal{V}} (\Pr(V_v = 0) = 1 \vee \Pr(V_v = 1) = 1),$$

where $g(\ell_{i_j})$ is replaced by $V_v = 1$ if $\ell_{i_j} = v$, and by $V_v = 0$ if $\ell_{i_j} = \neg v$.

We now argue that Φ is satisfied if and only if ϕ admits an SCM. For the first direction, suppose there exists an assignment $\alpha : \mathcal{V} \to \{0, 1\}^{|\mathcal{V}|}$ under which Φ is satisfied. We construct a model $\mathcal{M} = (\mathbf{V}, \emptyset, \mathcal{F}, \emptyset)$ for $\mathrm{arbSAT}^{\mathsf{base}}_{\mathsf{prob}}$ as follows. For each $v \in \mathcal{V}$, define $f_{V_v} := \alpha(v)$ as a constant function. To see that $\mathcal{M}$ satisfies all constraints in ϕ, recall that α satisfies at least one literal ℓ_{i_j} in each clause $C_i \in \Phi$, that is, $\alpha(v) = 1$ if $\ell_{i_j} = v$, and $\alpha(v) = 0$ if $\ell_{i_j} = \neg v$. The reduction ensures that the i^{th} conjunct in ϕ contains the disjunct $\Pr(V_v = 1) = 1$ in the first, and $\Pr(V_v = 0) = 1$ in the latter case. Since $[\![\Pr(V_v = \alpha(v))]\!]_{\mathcal{M}} = 1$, this satisfies ϕ.

For the other direction, suppose there exists a model $\mathcal{M}$ satisfying ϕ. Therefore, either $[\![\Pr(V_v = 0)]\!]_{\mathcal{M}} = 1$ or $[\![\Pr(V_v = 1)]\!]_{\mathcal{M}} = 1$ for each $v \in \mathcal{V}$. We obtain an assignment $\alpha : \mathcal{V} \to \{0, 1\}^{|\mathcal{V}|}$ by defining $\alpha(v) = 0$ if and only if $[\![\Pr(V_v = 0)]\!]_{\mathcal{M}} = 1$ for each variable $v \in \mathcal{V}$. Since all conjuncts of ϕ are satisfied by $\mathcal{M}$, it holds by construction that α satisfies all clauses of Φ. $\square$

Intractability of $\mathrm{SAT}^{\mathsf{poly}}_{\mathsf{prob}}$. Our main contributions target the lin and base fragments of the expressivity matrix and are provided in Sections 4 and 5. Here, we show that the tractability results obtained there cannot be lifted to polynomial inequalities.

Theorem 2. $\mathrm{SAT}^{\mathsf{poly}}_{\mathsf{prob}}$ *is $\exists\mathbb{R}$-complete even if $n = 1$, or if $d = 2$ and G_ϕ is edgeless.*

Proof. Mossé et al. (2024) proved that $\mathrm{SAT}^{\mathsf{poly}}_{\mathsf{prob}}$ is $\exists\mathbb{R}$-complete. Using a different reduction, we first prove our second statement by showing hardness on instances in which no two variables co-occur in a term and all variables have a binary domain. Our proof conceptually resembles the construction used in van der Zander et al. (2023, Proposition 6.5). Consider the following problem.

$\text{ETR}^{1/8,+,\times}_{[-1/8,1/8]}$	**Input:**	A set S of equations over real variables $x_1, \ldots, x_n \in [-\frac{1}{8}, \frac{1}{8}]$ in which each equation is of the form $x_i = \frac{1}{8}$ or $x_{i_1} + x_{i_2} = x_{i_3}$ or $x_{i_1} x_{i_2} = x_{i_3}$, and $i, i_1, i_2, i_3 \in [n]$.
	Task:	Decide if S has a solution.

Note that $\text{ETR}^{1/8,+,\times}_{[-1/8,1/8]}$ is known to be $\exists\mathbb{R}$-complete (Abrahamsen et al., 2017). We show that every instance of $\text{ETR}^{1/8,+,\times}_{[-1/8,1/8]}$ can be reduced to an instance of $\text{SAT}^{\text{poly}}_{\text{prob}}$, in which all variables have a binary domain and each term speaks about just one variable.

Let S be an instance of $\text{ETR}^{1/8,+,\times}_{[-1/8,1/8]}$ that contains variables $x_1, \ldots, x_n$ and let $V_1, \ldots, V_n$ be binary random variables. We obtain an instance ϕ over domain $D = \{0, 1\}$ of $\text{SAT}^{\text{poly}}_{\text{prob}}$ in time $O(|\phi|)$ by replacing for all $i \in [n]$ each occurrence of x_i in S by the expression $e_i := \frac{1}{4}\Pr(V_i = 0) - \frac{1}{8}$. Note that in ϕ, no two random variables co-occur in the same term and $d = 2$. We now argue that S is satisfiable if and only if ϕ has a model. First, suppose there exists a solution for S, i.e., a function f that maps each variable $x_i, i \in [n]$ to a value in $[-\frac{1}{8}, \frac{1}{8}]$ such that all equations are satisfied under f. We obtain a model $\mathcal{M}$ for ϕ by introducing a hidden variable U_i for each binary random variable V_i in ϕ and define f_{V_i} such that $V_i = U_i$. Then, letting $\mathbb{P}(U_i = 0) := 4f(x_i) + \frac{1}{2}$ for all $i \in [n]$ will satisfy ϕ, since it enforces $0 \leq [\![\Pr(V_i = 0)]\!]_{\mathcal{M}} \leq 1$ and ensures that $e_i = f(x_i)$. Likewise, if there exists a model $\mathcal{M}$ for ϕ, then we can determine $[\![\Pr(V_i = 0)]\!]_{\mathcal{M}}$ and derive e_i. Setting x_i to the value e_i for all $i \in [n]$ thus solves all equations in S and ensures $x_i \in [-\frac{1}{8}, \frac{1}{8}]$.

The above construction is easily adapted to show hardness even if $n = 1$. To construct ϕ, let $\mathbf{V} = \{V\}$ with $d = n+1$. For each $i \in [n]$, add a constraint $\Pr(V = i) \leq \frac{1}{n}$. Furthermore, add a constraint for each constraint in S that, for $i \in [n]$, replaces each occurrence of x_i by $e_i := \frac{n}{4}\Pr(V = i) - \frac{1}{8}$, which scales $\Pr(V = i)$ to be in $[-\frac{1}{8}, \frac{1}{8}]$. This construction holds by the same arguments as employed above, where the value 0 in the range of V serves as a buffer so that for $i \in [n]$ the probability $\Pr(V = i)$ can take arbitrary values in $[0, \frac{1}{n}]$. $\square$

Despite the hardness of sets of polynomial inequalities even in the absence of interventions, we remark that one can still obtain exponential time algorithms by employing the constructions in Theorems 5 and 10. Using the same approaches now requires solving systems of polynomial inequalities instead of LPs. These $\exists\mathbb{R}$ instances can be solved, for example, by invoking Renegar's Theorem (Renegar, 1992a;b;c).

Further Structural Insights in $\text{SAT}^*_{\circledast}$. In order to facilitate our complexity-theoretic analysis, we emphasize that a Structural Causal Model can be efficiently evaluated, that is, given the values of its hidden variables, it can be decided in polynomial time, whether a certain event happens.

Observation 3. Given a model $\mathcal{M} = (\mathbf{U}, \mathbf{V}, \mathcal{F}, \mathbb{P})$, an event $\varepsilon \in \mathcal{E}_{\text{counterfact}}$, and some $\overline{u} \in \text{Val}(\mathbf{U})$, let $|\varepsilon|$ denote the number of atoms in ε. Assuming that each function in $\mathcal{F}$ can be evaluated in time $\mathcal{O}(n)$, one can decide whether $\mathcal{F}, \overline{u} \models \varepsilon$ in time in $\mathcal{O}(n^2 + |\varepsilon|)$.

Proof. The only randomness in a model stems from the hidden variables $\mathbf{U}$. Fixing their values thus deterministically settles whether ε holds true or not. To decide which one is the case, it suffices to show how to evaluate events from $\mathcal{E}_{\text{post-int}}$, as $\mathcal{E}_{\text{counterfact}}$ is simply a Boolean formula over these events that—having evaluated each event of type $\mathcal{E}_{\text{post-int}}$—can be evaluated in time in $\mathcal{O}(|\varepsilon|)$. To evaluate an event $[\gamma]\,\varepsilon' \in \mathcal{E}_{\text{post-int}}$, we compute the value of each variable $V \in \mathbf{V}$ following the implicit well-order $\prec$ of the model. Note that the value of V is either fixed by the intervention γ or can be computed from $\overline{u}$ and the values of $\mathbf{V}_{\prec V}$. Once the values of all variables in $\mathbf{V}$ are determined, the probabilistic event ε' can be evaluated in time in $\mathcal{O}(|\varepsilon'|)$. $\square$

The runtime of our algorithmic results often depends on the size d of the domain D. We remark that assuming that d is not much larger than the size of ϕ does not reduce the generality of our results, as we can always reduce to an equivalent instance where d is bounded from above by $|\phi| + 1$.

Observation 4. Consider an instance of $\text{SAT}^*_{\circledast}$ consisting of a domain D and a formula $\phi \in \mathcal{L}^*_{\circledast}$. Let D_ϕ be the set of values in D that are explicitly mentioned in at least one atom in ϕ and choose some $\gamma \notin D_\phi$. Then, it holds that ϕ over domain $D_\phi \cup \{\gamma\}$ is a YES-instance of $\text{SAT}^*_{\circledast}$ if and only if so is ϕ over D.

Proof. To show this equivalence, intuitively, we contract all values in $D \setminus D_\phi$ into γ.

Suppose there is a model $\mathcal{M}$ solving $(\phi, D_\phi \cup \{\gamma\})$. As ϕ does not differentiate between values in $\{\gamma\} \cup D \setminus D_\phi$, the model $\mathcal{M}$ also witnesses (ϕ, d) to be a YES-instance.

For the other direction, suppose there is a model $\mathcal{M} = (\mathbf{V}, \mathbf{U}, \mathcal{F}, \mathbb{P})$ solving (ϕ, D) with an implicit well-order $\prec$, and let $V_1, V_2, \ldots$ denote $\mathbf{V}$ as ordered by $\prec$. Note that, without loss of generality, for each $i \in [n]$ we can assume $f_{V_i} \in \mathcal{F}$ to be represented by a case distinction over the values of $\mathbf{U}$ and $\mathbf{V}_{\prec V}$, where the result of each case is stated as an element in D. Exhaustively repeat the following. Let i be minimal such that there is some function $f_V \in \mathcal{F}$ which has a condition $V_i = v$ in the case distinction with $v \notin D_\phi$. Replace the condition by $\underline{f_{V_i}} = v$, where $\underline{f_{V_i}}$ denotes the term used to compute f_{V_i}. Rewrite the updated f_V to once more be a case distinction over the values of $\mathbf{U}$ and $\mathbf{V}_{\prec V}$. Note that this preserves the probability distribution over f_V, even under interventions: For every fixed $\overline{u} \in \mathbf{U}$, the variables $V_1, \ldots, V_{i-1}$ are computed the same way as before. Further, by assumption there is no intervention with an atom setting V_i to v. Hence, the only way that $V_i = v$ happens is if $\underline{f_{V_i}} = v$.

If at some point there is no i satisfying the condition, we have an alternative but equivalent representation of the functions in $\mathcal{F}$ which do not compare any variable $V \in \mathbf{V}$ to any value outside D_ϕ. At this point, update the functions once more such that whenever a function f_V would output some value in $D \setminus D_\phi$, it now outputs γ. As the functions computing the other variables do not compare such a variable V to a value outside of D_ϕ, all variables are computed in the exact same way as before except that for each variable V all outputs in $D \setminus D_\phi$ are contracted into γ. This yields an updated model $\mathcal{M}'$ where the domain is restricted to $D_\phi \cup \{\gamma\}$. This model satisfies ϕ if so does $\mathcal{M}$. □

4 Linear Inequalities over Probabilistic Expressions

This section is dedicated to the complexity-theoretic analysis of $\mathrm{SAT}^*_{\mathsf{prob}}$, that is, the satisfiability problem for the layer of the PCH that does not allow any interventions. First, we establish the main tractability result of this section, and then proceed by showing that it is tight as outlined in Figure 1.

Theorem 5. *$\mathrm{SAT}^{\mathsf{lin}}_{\mathsf{prob}}$ is in* FPT *w.r.t. the combined parameter $d + \mathrm{tw}(\phi)$, and in* XP *w.r.t.* $\mathrm{tw}(\phi)$.

Proof. Consider an instance of $\mathrm{SAT}^{\mathsf{lin}}_{\mathsf{prob}}$ with formula ϕ and domain D. We prove both statements simultaneously by describing an algorithm that runs in time $d^{f(\mathrm{tw}(\phi))}|\phi|^{\mathcal{O}(1)}$, for a computable function f. Consider a nice tree decomposition of G_ϕ consisting of $\mathcal{O}(n)$ nodes with maximum size $w := \mathrm{tw}(\phi) + 1$ computed by, e.g., the algorithm of Bodlaender (1996). Without loss of generality, assume that only the bags of leaf nodes are empty and ignore them in the following procedure. For the remaining tree decomposition $\mathbf{T}$, let $D^{|B|}$ be the combined domain of the variables of bag B in $\mathbf{T}$. We construct the following Linear Program (LP). For each bag B and $\overline{v} \in D^{|B|}$, construct an LP-variable $p_{B=\overline{v}}$; this will capture the probability of the event $B = \overline{v}$, that is, each variable in B takes the respective value in $\overline{v}$. To ensure a valid probability distribution over the LP-variables in each bag B, add the LP-constraints

$$p_{B=\overline{v}} \geq 0 \text{ for each LP-variable } p_{B=\overline{v}}, \quad \text{and} \quad \sum\nolimits_{\overline{v} \in D^{|B|}} p_{B=\overline{v}} = 1 \text{ for each bag } B.$$

For every pair of bags B, B' whose nodes are adjacent in $\mathbf{T}$ and $B \neq B'$, note that there is some $V \in \mathbf{V}$ such that, without loss of generality, $B' = B \cup \{V\}$. To guarantee consistency between the probability distributions of B and B', we add for each such pair and each $\overline{v} \in D^{|B|}$ the LP-constraint

$$p_{B=\overline{v}} = \mathrm{sum}\{p_{B'=\overline{v}'} \mid \overline{v}' \in D^{|B'|} \text{ and } \overline{v}' \text{ sets } B \text{ to } \overline{v}\}.$$

Next, for each constraint C in ϕ, consider each of its terms $\Pr(\varepsilon)$ and define $\mathcal{V}_\varepsilon \subseteq \mathbf{V}$ to be the set of variables that occur in $\Pr(\varepsilon)$. By construction, for each ε, all variables in $\mathcal{V}_\varepsilon$ form a clique in G_ϕ. Consequently, there is at least one bag B_ε in $\mathbf{T}$ such that $\mathcal{V}_\varepsilon \subseteq B_\varepsilon$. Consider an arbitrary choice of such B_ε and obtain an LP-constraint from C by replacing each occurrence of term $\Pr(\varepsilon)$ by a sum over all LP-variables $p_{B_\varepsilon=\overline{v}}$ such that $B_\varepsilon = \overline{v}$ satisfies the event ε. Then the LP consists of $\mathcal{O}(n \cdot d^w)$ LP-variables and $\mathcal{O}(|\phi| + n \cdot d^w)$ LP-constraints. An exemplary instance is constructed in Example 1.

We can find a solution of an LP (or decide that there is none) in polynomial time with respect to its size, that is, the number of its variables plus constraints. Crucially, if ϕ is a YES-instance witnessed by a model $\mathcal{M}$ which induces a probability distribution over $\mathbf{V}$, then the LP admits a solution; indeed, we can satisfy all constraints by setting each LP-variable $p_{B=\overline{v}}$ to the probability of the event $B = \overline{v}$ within that distribution.

For the converse, assume the LP has a solution. We construct a model for ϕ by passing through $\mathbf{T}$ in a breadth-first-search manner, starting from an arbitrary leaf node with some bag $B = \{V\}$. Let U_V be a hidden variable with domain D such that $\mathbb{P}(U_V = v) = p_{B=(v)}$ for all $v \in D$. Whenever we transition from a bag B to a bag B' containing a variable V which we have not yet described in our model, we have $B' = B \cup \{V\}$. For each $\overline{v} \in D^{|B|}$ such that $p_{B=\overline{v}} > 0$, create a hidden variable $U_{V|B=\overline{v}}$ with domain $\mathrm{Val}(U_{V|B=\overline{v}}) = D$ and let

$$\mathbb{P}(U_{V|B=\overline{v}} = x) := \frac{p_{B'=\overline{v}+x}}{p_{B=\overline{v}}} \quad \text{for each } x \in D,$$

where $(\overline{v}+x) \in D^{|B'|}$ is such that it sets B to $\overline{v}$ and V to x. This describes a valid probability distribution of $U_{V|B=\overline{v}}$: As $\mathbb{P}(U_{V|B=\overline{v}} = x) \geq 0$ for all x, it remains to show that $\sum_{x\in D} \mathbb{P}(U_{V|B=\overline{v}} = x) = 1$. We have

$$\sum_{x\in D} \mathbb{P}(U_{V|B=\overline{v}} = x) = \sum_{x\in D} \frac{p_{B'=\overline{v}+x}}{p_{B=\overline{v}}} = \frac{1}{p_{B=\overline{v}}} \sum_{x\in D} p_{B'=\overline{v}+x} = 1,$$

as we ensured $p_{B=\overline{v}} = \sum_{x\in D} p_{B'=\overline{v}+x}$ by an LP-constraint. We now define the function f_V such that, for each $\overline{v} \in D^{|B|}$, if $B=\overline{v}$ then $V = U_{B'|B=\overline{v}}$.

It remains to argue that the model $\mathcal{M}$ obtained after visiting every node witnesses ϕ to be a YES-instance. To this end, employ induction over the breadth-first search described above to prove that after visiting a node with bag B, for each $\overline{v} \in D^{|B|}$ the value of $p_{B=\overline{v}}$ describes the probability of $B = \overline{v}$ in the current model $\mathcal{M}'$, that is, $[\![\Pr(B=\overline{v})]\!]_{\mathcal{M}'} = p_{B=\overline{v}}$. As the bag of the first node contains just one variable, the base case is trivial. Now assume that the claim holds for bag B of some node N and model $\mathcal{M}$, and from N we are visiting a node N' with bag B'. If $B' \subseteq B$, then $\mathcal{M}$ is not changed and the claim follows immediately. Otherwise $B' = B \cup \{V\}$ for a variable V and $\mathcal{M}$ is extended to a model $\mathcal{M}'$ as described above. Consider any $\overline{v}' \in D^{|B'|}$. Let $\overline{v}$ equal $\overline{v}'$ when restricted to the variables in B and let x be the value of variable V in $\overline{v}'$. By the construction of $\mathcal{M}'$ and the induction hypothesis, we have that $[\![\Pr(B=\overline{v})]\!]_{\mathcal{M}'} = [\![\Pr(B=\overline{v})]\!]_{\mathcal{M}} = p_{B=\overline{v}}$. If $[\![\Pr(B=\overline{v})]\!]_{\mathcal{M}'} = 0$, then $[\![\Pr(B'=\overline{v}')]\!]_{\mathcal{M}'} = 0$, which is correct by the consistency constraints $p_{B'=\overline{v}'} \leq p_{B=\overline{v}} = 0$. Otherwise,

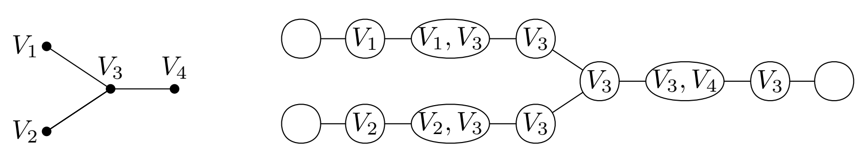
$$[\![\Pr(B'=\overline{v}')]\!]_{\mathcal{M}'} = [\![\Pr(B=\overline{v})]\!]_{\mathcal{M}'} \cdot \mathbb{P}(U_{V|B=\overline{v}} = x) = p_{B=\overline{v}} \cdot \frac{p_{B'=\overline{v}'}}{p_{B=\overline{v}}} = p_{B'=\overline{v}'}.$$

Given a nice tree decomposition, the LP can be constructed and solved in time in $(|\phi| + n \cdot d^w)^{\mathcal{O}(1)}$. In case of a YES-instance, this time also suffices to construct a suitable model. $\square$

Example 1 (Construction in Theorem 5)**.** *Consider the following instance ϕ of $\mathrm{SAT}^{\mathsf{lin}}_{\mathsf{base}}$ with endogenous variables $\mathbf{V} = \{V_1, V_2, V_3, V_4\}$ over domain $D = \{0,1\}$.*

$$\begin{aligned} \Pr(V_1 = 1 \wedge V_3 = 1) &\geq \tfrac{1}{2} \\ \Pr(V_2 = 1 \vee V_3 = 1) - 2\Pr(V_3 = 1 \vee V_4 = 1) &\geq 0 \\ \Pr(V_4 = 1) &\geq \tfrac{1}{3} \end{aligned}$$

The corresponding primal graph G_ϕ and a nice tree decomposition of G_ϕ are as follows.



We construct a linear program as follows. For each non-empty bag, we create one LP-variable for each possible assignment of the variables in the bag. This yields variables $p_{V_i=0}$ and $p_{V_i=1}$ for $i \in [3]$ as well as $p_{V_1=x,V_3=y}$ and $p_{V_2=x,V_3=y}$ and $p_{V_3=x,V_4=y}$ for all pairs $x, y \in \{0,1\}$. Note that for clarity we here write $p_{V_i=x,V_j=y}$ instead of $p_{B=(x,y)}$ with $B = \{V_i, V_j\}$, $i < j$. We introduce the following LP-constraints

$$\begin{aligned} p_{V_i=0} \geq 0, \quad p_{V_i=1} \geq 0, \quad p_{V_i=0} + p_{V_i=1} &= 1 && \textit{for } i \in [3], \\ p_{V_a=x,V_b=y} &\geq 0 && \textit{for } (a,b) \in \{(1,3),(2,3),(3,4)\}, x,y \in \{0,1\}, \\ p_{V_a=0,V_b=0} + p_{V_a=0,V_b=1} + p_{V_a=1,V_b=0} + p_{V_a=1,V_b=1} &= 1 && \textit{for } (a,b) \in \{(1,3),(2,3),(3,4)\}, \end{aligned}$$

which we extend by the following LP-constraints that ensure consistency between bags

$$\begin{aligned} p_{V_a=x} &= p_{V_a=x,V_b=0} + p_{V_a=x,V_b=1} && \textit{for } (a,b) \in \{(1,3),(2,3),(3,4)\}, x \in \{0,1\}, \\ p_{V_b=x} &= p_{V_a=0,V_b=x} + p_{V_a=1,V_b=x} && \textit{for } (a,b) \in \{(1,3),(2,3)\}, x \in \{0,1\}. \end{aligned}$$

Last, the following three LP-constraints encode the constraints in ϕ:

$$\begin{aligned} p_{V_1=1,V_3=1} &\geq \tfrac{1}{2} \\ p_{V_2=1,V_3=0} + p_{V_2=1,V_3=1} + p_{V_2=0,V_3=1} - 2(p_{V_3=1,V_4=0} + p_{V_3=1,V_4=1} + p_{V_3=0,V_4=1}) &\geq 0 \\ p_{V_3=0,V_4=1} + p_{V_3=1,V_4=1} &\geq \tfrac{1}{3}. \end{aligned}$$

Here, ϕ is a YES-instance of $\mathrm{SAT}^{\mathsf{lin}}_{\mathsf{base}}$ and satisfied by an SCM $\mathcal{M} = (\mathbf{V}, \{U\}, \mathcal{F}, \mathbb{P})$ such that $\mathbb{P}(U=1) = \frac{1}{2}$

holds [3] *as well as*

$$f_{V_1}(U) := U, \qquad f_{V_2} := 1, \qquad f_{V_3}(V_1) := V_1, \quad \textit{and} \quad f_{V_4}(V_1) := V_1,$$

which corresponds to the LP-solution where all LP-variables have value 0 *except for*

$$p_{V_2=1} = 1, \quad p_{V_i=0} = p_{V_i=1} = \tfrac{1}{2} \quad \textit{for } i \in \{1,3\},$$
$$p_{V_1=0,V_3=0} = p_{V_1=1,V_3=1} = p_{V_2=1,V_3=0} = p_{V_2=1,V_3=1} = p_{V_3=0,V_4=0} = p_{V_3=1,V_4=1} = \tfrac{1}{2}.$$

Likewise, this solution to the LP yields a model $\mathcal{M}' = (\mathbf{V}, \mathbf{U}'\mathcal{F}', \mathbb{P}')$ *witnessing* ϕ *to be a YES-instance, as constructed by passing through the tree decomposition in order* $\{V_1\}, \{V_1, V_3\}, \{V_3\}, \{V_3, V_4\}, \{V_2, V_3\}, \{V_2\}$. *We illustrate the first steps towards constructing* $\mathcal{M}'$. *First, we introduce a hidden variable* U_{V_1} *with* $\mathbb{P}'(U_{V_1} = 1) = \frac{1}{2}$ *and let* $f'_{V_1}(U_{V_1}) := U_{V_1}$. *Next we construct hidden variables* $U_{V_3|V_1=0}$ *and* $U_{V_3|V_1=1}$, *where* $\mathbb{P}'(U_{V_3|V_1=0} = 1) = 0$ *and* $\mathbb{P}'(U_{V_3|V_1=1} = 1) = 1$, *and define*

$$f_{V_3}(V_1, U_{V_3|V_1=0}, U_{V_3|V_1=1}) := \begin{cases} U_{V_3|V_1=0}, & \textit{if } V_1 = 0; \\ U_{V_3|V_1=1}, & \textit{if } V_1 = 1. \end{cases}$$

Defining the remaining observed variables analogously yields an SCM $\mathcal{M}'$ *which satisfies* ϕ.

Next, we show that under well-established complexity assumptions, parameterization by d alone cannot yield tractability, even when the primal graph G_ϕ has bounded degree.

Theorem 6. $\text{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ *is* NP*-complete even if* $d = 2$ *and the maximum degree of* G_ϕ *is* 8.

Proof. The containment in NP follows from $\text{arbSAT}^{\mathsf{base}}_{\mathsf{prob}} \in \mathsf{NP}$ (Fagin et al., 1990). To show hardness in our restricted setting, we perform a reduction from 3-SAT. Note that 3-SAT remains NP-hard when restricted to formulas in which each variable occurs exactly twice negated and twice non-negated (Darmann & Döcker, 2021). Thus, w.l.o.g., we assume that our formula $\Phi := \bigwedge_i C_i$ with $C_i := \bigvee_{j\in[3]} \ell_{i_j}$ over variables $\mathcal{V}$ has this property. We construct an instance ϕ of $\text{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ with $\mathbf{V} = \{V_v \mid v \in \mathcal{V}\}$ and $D = \{0,1\}$ such that for each $C_i \in \Phi$, the constraint $\Pr(\bigvee_{j\in[3]} g(\ell_{i_j})) = 1$ is added to ϕ, where g is defined as in Theorem 1. Since each variable occurs in at most 4 clauses, there are at most 8 other variables it co-occurs with; consequently, G_ϕ has a maximum degree of 8. We now argue that Φ is satisfied if and only if ϕ admits an SCM.

Suppose there exists an assignment $\alpha : \mathcal{V} \to \{0,1\}^{|\mathcal{V}|}$ that satisfies Φ. We construct a model $\mathcal{M}$ satisfying ϕ as in the proof of Theorem 1. The proof that $\mathcal{M}$ satisfies all constraints in ϕ is analogous.

For the other direction, suppose there exists a model $\mathcal{M} = (\mathbf{V}, \mathbf{U}, \mathcal{F}, \mathbb{P})$ satisfying ϕ. From the existence of $\mathbb{P}$, we conclude that there exists an assignment $\alpha_{\mathbf{U}}$ to $\mathbf{U}$ that has a non-zero probability. Moreover, by construction, fixing such an $\alpha_{\mathbf{U}}$ yields a full assignment $\alpha_{\mathbf{V}}$ to $\mathbf{V}$ of non-zero probability. We obtain an assignment $\alpha_\Phi : \mathcal{V} \to \{0,1\}^{|\mathcal{V}|}$ by enforcing $\alpha_\Phi(v) = 0 \Leftrightarrow \alpha_{\mathbf{V}}(V_v) = 0$ for each variable $v \in \mathcal{V}$. We claim that α_Φ satisfies all clauses in Φ. Suppose the contrary, that is, there exists a clause $C_i = \bigvee_{j\in[3]} \ell_{i_j}$ in Φ that is not satisfied by α_Φ. Then $\bigvee_{j\in[3]} g(\ell_{i_j})$ is not satisfied by $\alpha_{\mathbf{V}}$. However, since $\alpha_{\mathbf{V}}$ has non-zero probability, it follows that $[\![\Pr(\bigvee_{j\in[3]} g(\ell_{i_j}))]\!]_{\mathcal{M}} < 1$, thus, $\mathcal{M}$ fails to satisfy all constraints in ϕ which contradicts the fact that it is a model. We can therefore conclude that α_Φ is a satisfying assignment for Φ. □

The following result complements Theorem 6 by ruling out fixed-parameter tractable algorithms for $\text{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ under a different parameterization, namely the number of variables n. Note that since $\mathrm{tw}(\phi) \le n$, this implies that we should not expect the primal treewidth of a graph to yield fixed-parameter tractability for $\text{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ alone.

Theorem 7. $\text{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ *is* $\mathsf{W}[1]$*-hard parameterized by* n.

Proof. We perform a reduction from k-MULTICOLORED-CLIQUE, which asks, given a properly vertex-colored graph G with colors $1, \dots, k$ and vertices $v_1, \dots, v_r$, whether G contains a k-clique. Given G, we construct an instance ϕ of $\text{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ as follows. Let $\mathbf{V} = \{V_1, \dots, V_k\}$ and $D = \{v_1, \dots, v_r\}$. For each $i \in [k]$ and $a \in [r]$, add the constraint $\Pr(V_i = v_a) \le 0$, unless v_a has color i. For each non-adjacent v_a, v_b with $a < b$ and colors i, j, add the constraint $\Pr(V_i = v_a \wedge V_j = v_b) \le 0$. The construction takes polynomial time and sets $n = k$.

Suppose G contains a clique of size k. Let $\alpha \colon [k] \to \{v_1, \dots, v_r\}$ be such that for every $i \in [k]$ the clique contains vertex $\alpha(i)$ of color i. Consider the model $(\mathbf{V}, \emptyset, \mathcal{F}, \emptyset)$, where $\mathcal{F}$ is such that $f_{V_i} := \alpha(i)$ is a constant function for each $V_i \in \mathbf{V}$. Clearly, this satisfies each constraint of the form $\Pr(V_i = v_a) \le 0$. Assume this

[3] Here, for binary variables we just state the probability of one case; the probability for the other immediately follows.

would not satisfy a constraint of the form $\Pr(V_i = v_a \wedge V_j = v_b) \leq 0$. Then $\alpha(i) = v_a$ and $\alpha(j) = v_b$, so both v_a and v_b are in the clique and thereby adjacent, which contradicts the existence of the constraint.

For the other direction, assume there is a model $\mathcal{M}$ satisfying the $\mathrm{SAT}^{\mathsf{base}}_{\mathsf{prob}}$ instance. Then there is at least one assignment $\overline{v} \in D^k$ such that $[\![\Pr(\mathbf{V} = \overline{v})]\!]_{\mathcal{M}} > 0$. Let $\alpha\colon [k] \to \{v_1, \ldots, v_r\}$ be such that for each $i \in [k]$ we have $V_i = \alpha(i)$ in this assignment. We argue that the vertices $\alpha(1), \ldots, \alpha(k)$ form a clique. Towards a contradiction, assume there are $i, j \in [k], i \neq j$ such that $\alpha(i)$ and $\alpha(j)$ are non-adjacent. Then there is a constraint $\Pr(V_i = \alpha(i) \wedge V_j = \alpha(j)) \leq 0$, which contradicts the model being a solution to the instance as $[\![\Pr(V_i = \alpha(i) \wedge V_j = \alpha(j))]\!]_{\mathcal{M}} \geq [\![\Pr(\mathbf{V} = \overline{v})]\!]_{\mathcal{M}} > 0$. $\square$

5 Linear Inequalities over Causal or Counterfactual Expressions

In this section, we turn our attention to interventional causal reasoning. We initiate our study by showing that the FPT-tractability that was established in Theorem 5 does not carry over.

Theorem 8. $\mathrm{SAT}^{\mathsf{lin}}_{\mathsf{causal}}$ *is* NP*-complete even if* $d = 2$ *and* G_ϕ *consists of vertex-disjoint paths of length* 2*.*

Proof. Fagin et al. (1990) showed that $\mathrm{arbSAT}^{\mathsf{lin}}_{\mathsf{causal}}$ is NP-complete. To show NP-hardness even for instances ϕ with $d = 2$ and for which G_ϕ consists of vertex-disjoint paths of length 2, we reduce from 3-SAT. Consider a 3-SAT formula Φ with r variables. We define an instance ϕ of $\mathrm{SAT}^{\mathsf{lin}}_{\mathsf{causal}}$ with domain $D = \{0, 1\}$ as follows. For each variable $x_i \in \Phi$, we introduce endogenous variables V_i and $\overline{V}_i$, as well as the constraints

$$\Pr([V_i = 1]\, \overline{V}_i = 1) = 0, \text{ and } \Pr([\overline{V}_i = 1]\, V_i = 1) = 0.$$

Furthermore, for each clause $\ell_1 \vee \ell_2 \vee \ell_3$ in Φ, we add $\Pr(L_1 = 1) + \Pr(L_2 = 1) + \Pr(L_3 = 1) \geq 1$ to ϕ, where, $L_j = V_i$ if $\ell_j = x_i$, and $L_j = \overline{V}_i$ if $\ell_j = \overline{x}_i$ for $j \in [3]$. Note that G_ϕ consists only of edges between V_i and $\overline{V}_i$, for $i \in [r]$. We now argue that Φ is satisfiable if and only if there is a model for ϕ. Suppose there exists an assignment $\alpha : \mathrm{Var}(\Phi) \to \{0, 1\}^r$ under which Φ is satisfied. We construct a model $\mathcal{M} = (\mathbf{V}, \emptyset, \mathcal{F}, \emptyset)$ for ϕ as follows: First, note that by construction, $\mathbf{V} = \{V_i, \overline{V}_i \mid i \in [r]\}$. For each $i \in [r]$, if $\alpha(x_i) = 1$ then let $\mathcal{F}$ be such that $f_{\overline{V}_i} := 0$, which satisfies $\Pr([V_i = 1]\, \overline{V}_i = 1) = 0$, and $f_{V_i} := 1 - \overline{V}_i$, which satisfies $\Pr([\overline{V}_i = 1]\, V_i = 1) = 0$. If instead $\alpha(x_i) = 0$, let $f_{V_i} := 0$ and $f_{\overline{V}_i} := 1 - V_i$, which analogously satisfies both constraints. This yields an SCM $\mathcal{M}$. It remains to show that each clause-constraint is satisfied. Consider a clause $\ell_1 \vee \ell_2 \vee \ell_3$ and let, without loss of generality, ℓ_1 be TRUE in α. Assume that $\ell_1 = x_i$ for some $i \in [r]$ (the case of $\ell_1 = \overline{x}_i$ holds analogously). Then, in the model without interventions, it holds that $\overline{V}_i = 0$ and $V_i = 1 - 0 = 1$ with probability 1, in other words, $[\![\Pr(V_i = 1)]\!]_{\mathcal{M}} = 1$. As $\Pr(V_i = 1)$ is one of the summands in the constraint for clause i and the other summands are non-negative, this satisfies the clause constraint.

For the other direction, suppose there exists a model $\mathcal{M} = (\mathbf{V}, \mathbf{U}, \mathcal{F}, \mathbb{P})$ for ϕ with an associated well-order $\prec$ over $\mathbf{V}$. Consider the assignment α obtained by letting $x_i = 1$ if $\overline{V}_i \prec V_i$ and $x_i = 0$, otherwise. Note that if $\overline{V}_i \prec V_i$ then $f_{\overline{V}_i}$ does not depend on V_i and thus the constraint $\Pr([V_i = 1]\, \overline{V}_i = 1) = 0$ implies $[\![\Pr(\overline{V}_i = 1)]\!]_{\mathcal{M}} = 0$. Vice versa, if $V_i \prec \overline{V}_i$, we have $[\![\Pr(V_i = 1)]\!]_{\mathcal{M}} = 0$. As for each clause $\ell_1 \vee \ell_2 \vee \ell_3$ we have that $\Pr(L_1 = 1) + \Pr(L_2 = 1) + \Pr(L_3 = 1) \geq 1$, there is $j \in [3]$ such that L_j does *not* precede its counterpart and thus ℓ_i is set to TRUE by α. $\square$

We contrast the hardness obtained in Theorem 8 by considering the number of variables in $\mathbf{V}$ as a new parameter. Towards this goal, Lemma 9 establishes the existence of a well-structured model for every YES-instance.

Lemma 9. *Let* ϕ *over domain* D *be a YES-instance of* $\mathrm{SAT}^{\mathsf{poly}}_{\mathsf{counterfact}}$ *over variables* $\mathbf{V}$*. There is an ordering* $V_1, \ldots, V_n$ *of* $\mathbf{V}$ *such that the* ϕ *is satisfied by a model* $\mathcal{M} = (\mathbf{V}, \mathbf{U}, \mathcal{F}, \mathbb{P})$ *with the following properties: for each* $i \in [n]$*, let* Q_i *be the set of all possible functions mapping the values of* $V_1, \ldots, V_{i-1}$ *to a value of* V_i*, that is, the set of all functions from* D^{i-1} *to* D *(with* Q_1 *simply being a set of constant functions). Then*

- $\mathbf{U} = \{U\}$ *with* $\mathrm{Val}(U) = Q_1 \times \ldots \times Q_n$*, where* $U[i] \in Q_i$ *denotes the* i^{th} *entry in* U*; and*
- $\mathcal{F} = \{f_{V_i} \mid i \in [n]\}$ *with* $f_{V_i}(U, V_1, \ldots, V_{i-1}) := U[i](V_1, \ldots, V_{i-1})$*.*

Proof. Let $\mathcal{M}' = (\mathbf{V}, \mathbf{U}', \mathcal{F}', \mathbb{P}')$ be any model witnessing that ϕ is a YES-instance. Without loss of generality, we assume $\mathbf{U}'$ to consist of a single variable U': If there were multiple variables $U_1, U_2, \ldots, U_\ell$ in $\mathbf{U}'$, we could replace them by some U' with $\mathrm{Val}(U') = \mathrm{Val}(U_1) \times \mathrm{Val}(U_2) \times \ldots \times \mathrm{Val}(U_\ell)$, and update $\mathbb{P}'$ and $\mathcal{F}'$ accordingly.

Consider an ordering $V_1, \ldots, V_n$ of $\mathbf{V}$ that respects the implicit well-order $\prec$ of $\mathcal{M}'$. Then, for each $i \in [n]$, $f'_{V_i} \in \mathcal{F}$ describes the value of V_i as a function of U' and $V_1, \ldots, V_{i-1}$. Partition $\mathrm{Val}(U')$ such that there is a class C_q for each $q = (q_1, \ldots, q_n) \in (Q_1 \times \ldots \times Q_n)$ and let it contain $u \in \mathrm{Val}(U')$ if and only if for all $i \in [n]$ and $\bar{s} \in D^{i-1}$, we have that $f'_{V_i}(u, \bar{s}) = q_i(\bar{s})$. Now each $u \in \mathrm{Val}(U')$ belongs to precisely one class C_q.

We construct the model $\mathcal{M} = (\mathbf{V}, \mathbf{U}, \mathcal{F}, \mathbb{P})$ where $\mathbf{U}$ and $\mathcal{F}$ are defined as specified above and $\mathbb{P}$ is such that for each $q \in \mathrm{Val}(U)$ we have $\mathbb{P}(U = q) = \sum_{u' \in C_q} \mathbb{P}'(U' = u')$ (with $\mathbb{P}(U = q) = 0$ if $C_q = \emptyset$). Note that this yields a well-defined probability distribution over U. The model $\mathcal{M}$ satisfies ϕ since every term $\Pr(\varepsilon)$ over $\mathbf{V}$ has the same probability in $\mathcal{M}$ and $\mathcal{M}'$. Indeed, for every class C_q and each event ε, we have that $\mathcal{F}', u \models \varepsilon$ either for all $u \in C_q$ or no $u \in C_q$, as by definition all these u result in the exact same values for the variables in $\mathbf{V}$, even under interventions. Furthermore, $\mathcal{F}$ is such that $\mathcal{F}, q \models \varepsilon$ if and only if $\mathcal{F}', u \models \varepsilon$ for all $u \in C_q$. For any event $\varepsilon \in \mathcal{E}_{\mathsf{counterfact}}$, recall that $S_{\mathcal{M}} \subseteq \mathrm{Val}(U)$ and $S_{\mathcal{M}'} \subseteq \mathrm{Val}(U')$ denote the sets of values of hidden variables such that ε happens in the respective model. We proved that $S_{\mathcal{M}'} = \bigcup_{q \in S_{\mathcal{M}}} C_q$ and thus, by definition of $\mathbb{P}$, event ε happens in both models with the same probability, that is, $[\![\Pr(\varepsilon)]\!]_{\mathcal{M}} = [\![\Pr(\varepsilon)]\!]_{\mathcal{M}'}$. Hence, $\mathcal{M}$ witnesses ϕ to be a YES-instance as well. □

Theorem 10. $\mathrm{SAT}^{\mathsf{lin}}_{\mathsf{counterfact}}$ *is in* FPT *w.r.t. the combined parameter $d + n$, and in* XP *w.r.t. n.*

Proof. Given an instance ϕ over domain D, we perform the following for each of the $n!$ orders of variables. If we do not find a model for any of these orders, we reject the instance. Consider a total order $V_1, \ldots, V_n$ of $\mathbf{V}$ and let U, $\mathcal{F}$, and the Q_i be defined as in Lemma 9. We test whether ϕ admits a model as described in Lemma 9 with respect to that order using the following LP. Create an LP-variable p_q for each $q \in Q_1 \times \ldots \times Q_n$. These represent the probability distribution over U, so we add an LP-constraint $\sum_{q \in Q_1 \times \ldots \times Q_n} p_q = 1$ and, for each such q, we add an LP-constraint $p_q \geq 0$. Furthermore, using Observation 3, we transform each constraint in ϕ into an LP-constraint by replacing each term $\Pr(\varepsilon)$ by a sum over all variables p_q for which $\mathcal{F}, q \models \varepsilon$. We accept the instance ϕ if and only if there is at least one ordering of $\mathbf{V}$ for which the constructed LP has a solution. In each constructed LP there are $\mathcal{O}(|\phi| + |Q_1 \times \ldots \times Q_n|)$ LP-constraints and the number of variables is at most

$$|Q_1 \times \ldots \times Q_n| = \prod_{i=1}^{n} |Q_i| \leq \prod_{i=1}^{n} d^{i-1} \cdot d \leq d^{n^2}.$$

As we can find a solution to an LP (or decide that there is none) in polynomial time with respect to its size (that is, the number of variables plus constraints), the total runtime of the algorithm is $n!(|\phi|^{\mathcal{O}(1)} + d^{\mathcal{O}(n^2)})$.

It remains to prove correctness. If one of the constructed LPs has a solution, let $\mathbb{P}$ be the probability distribution over U as described by the variables p_q in that solution. It is straight-forward to verify that $(\mathbf{V}, \{U\}, \mathcal{F}, \mathbb{P})$ is a model for ϕ, so we correctly accept the instance. For the other direction, suppose ϕ is a YES-instance. Then there is a well-structured model $\mathcal{M} = (\mathbf{V}, \{U\}, \mathcal{F}, \mathbb{P})$ for some ordering of $\mathbf{V}$ by Lemma 9. Observe that the LP constructed for that ordering of the variables has a solution by setting $p_q := \mathbb{P}(U = q)$ for each $q \in \mathrm{Val}(U)$, so the algorithm correctly accepts the instance. □

6 Concluding Remarks

While previous works have focused on mapping the complexity lower bounds for Satisfiability in Pearl's Causal Hierarchy, the parameterized paradigm allows us to identify islands of tractability for the problem. Our contributions include not only these positive results, but also lower bounds which show that the obtained complexity classifications are tight. The presented findings open up several avenues for future work, such as:

- **The Impact of Marginalization.** As mentioned in Section 1, recent works (van der Zander et al., 2023; Dörfler et al., 2025) have proposed an enrichment of the classical fragments in the expressivity matrix M via the summation operator $\sum$ in order to capture marginalization. It would be interesting to explore possible extensions of our approaches to this setting—in particular, can we obtain tractability for these enriched fragments of the PCH without bounding the nesting depth of $\sum$?
- **Treewidth-Guided Linear Programming.** To the best of our knowledge, the proofs of Results **(1)** and **(2)** rely on an entirely novel approach to establishing parameterized tractability with respect to treewidth. Since this approach is particularly tailored to problems that combine discrete structures (graphs) with non-discrete elements (e.g., probabilities), we would not be surprised to see further applications of the technique in the domains of artificial intelligence and machine learning.

REFERENCES

Mikkel Abrahamsen, Anna Adamaszek, and Tillmann Miltzow. The art gallery problem is $\exists\mathbb{R}$-complete. *CoRR*, abs/1704.06969, 2017.

Elias Bareinboim, Juan D. Correa, Duligur Ibeling, and Thomas Icard. On pearl's hierarchy and the foundations of causal inference. In Hector Geffner, Rina Dechter, and Joseph Y. Halpern (eds.), *Probabilistic and Causal Inference: The Works of Judea Pearl*, volume 36 of *ACM Books*, pp. 507–556. ACM, 2022.

Armin Biere, Marijn Heule, and Hans van Maaren. *Handbook of satisfiability*, volume 185. IOS press, 2009.

Markus Bläser, Julian Dörfler, Maciej Liskiewicz, and Benito van der Zander. Probabilistic and causal satisfiability: Constraining the model. In Keren Censor-Hillel, Fabrizio Grandoni, Joël Ouaknine, and Gabriele Puppis (eds.), *52nd International Colloquium on Automata, Languages, and Programming, ICALP 2025*, volume 334 of *LIPIcs*, pp. 144:1–144:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2025.

Hans L. Bodlaender. A linear-time algorithm for finding tree-decompositions of small treewidth. *SIAM J. Comput.*, 25(6):1305–1317, 1996.

John F. Canny. Some algebraic and geometric computations in PSPACE. In Janos Simon (ed.), *Proceedings of the 20th Annual ACM Symposium on Theory of Computing*, pp. 460–467. ACM, 1988.

Marek Cygan, Fedor V. Fomin, Łukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michał Pilipczuk, and Saket Saurabh. *Parameterized algorithms*. Springer, Cham, 2015.

Andreas Darmann and Janosch Döcker. On simplified np-complete variants of monotone3-sat. *Discrete Applied Mathematics*, 292:45–58, 2021.

Julian Dörfler, Benito van der Zander, Markus Bläser, and Maciej Liskiewicz. From probability to counterfactuals: the increasing complexity of satisfiability in pearl's causal hierarchy. In *The Thirteenth International Conference on Learning Representations, ICLR 2025*, 2025.

Rodney G. Downey and Michael R. Fellows. *Fundamentals of Parameterized Complexity*. Texts in Computer Science. Springer, 2013.

Eduard Eiben, Robert Ganian, Thekla Hamm, and Sebastian Ordyniak. Parameterized complexity of envy-free resource allocation in social networks. *Artificial Intelligence*, 315:103826, 2023.

Felix Elwert. *Graphical Causal Models*, pp. 245–273. Springer Netherlands, Dordrecht, 2013.

Ronald Fagin, Joseph Y. Halpern, and Nimrod Megiddo. A logic for reasoning about probabilities. *Information and Computation*, 87(1/2):78–128, 1990.

Norman E. Fenton, Martin Neil, and Anthony C. Constantinou. The book of why: The new science of cause and effect, judea pearl, dana mackenzie. basic books (2018). *Artificial Intelligence*, 284:103286, 2020.

Johannes Klaus Fichte and Markus Hecher. Exploiting treewidth for counting projected answer sets. In Michael Thielscher, Francesca Toni, and Frank Wolter (eds.), *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference, KR 2018*, pp. 639–640. AAAI Press, 2018.

Robert Ganian and Viktoriia Korchemna. The complexity of bayesian network learning: Revisiting the superstructure. In Marc'Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021*, pp. 430–442, 2021.

Robert Ganian and Sebastian Ordyniak. The complexity landscape of decompositional parameters for ILP. *Artificial Intelligence*, 257:61–71, 2018.

Robert Ganian, Sebastian Ordyniak, and M. S. Ramanujan. Going beyond primal treewidth for (M)ILP. In Satinder Singh and Shaul Markovitch (eds.), *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp. 815–821. AAAI Press, 2017.

Robert Ganian, Thekla Hamm, Viktoriia Korchemna, Karolina Okrasa, and Kirill Simonov. The complexity of k-means clustering when little is known. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato (eds.), *International Conference on Machine Learning, ICML 2022*, volume 162 of *Proceedings of Machine Learning Research*, pp. 6960–6987. PMLR, 2022.

Robert Ganian, Viktoriia Korchemna, and Stefan Szeider. Revisiting causal discovery from a complexity-theoretic perspective. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI 2024*, pp. 3377–3385. ijcai.org, 2024.

George F. Georgakopoulos, Dimitris J. Kavvadias, and Christos H. Papadimitriou. Probabilistic satisfiability. *Journal of Complexity*, 4(1):1–11, 1988.

Clark Glymour, Richard Scheines, Peter Spirtes, and Kevin Kelly. *Discovering causal structure: Artificial intelligence, philosophy of science, and statistical modeling*. Academic Press, 1987.

Dima Grigoriev and Nicolai N. Vorobjov Jr. Solving systems of polynomial inequalities in subexponential time. *Journal of Symbolic Computation*, 5(1/2):37–64, 1988.

Niels Grüttemeier and Christian Komusiewicz. Learning bayesian networks under sparsity constraints: A parameterized complexity analysis. *Journal of Artificial Intelligence Research*, 74:1225–1267, 2022.

Niels Grüttemeier, Christian Komusiewicz, and Nils Morawietz. Efficient bayesian network structure learning via parameterized local search on topological orderings. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021*, pp. 12328–12335. AAAI Press, 2021a.

Niels Grüttemeier, Christian Komusiewicz, and Nils Morawietz. On the parameterized complexity of polytree learning. In Zhi-Hua Zhou (ed.), *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI 2021*, pp. 4221–4227. ijcai.org, 2021b.

Tesshu Hanaka and Michael Lampis. Hedonic games and treewidth revisited. In Shiri Chechik, Gonzalo Navarro, Eva Rotenberg, and Grzegorz Herman (eds.), *30th Annual European Symposium on Algorithms, ESA 2022*, volume 244 of *LIPIcs*, pp. 64:1–64:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

Duligur Ibeling and Thomas Icard. Probabilistic reasoning across the causal hierarchy. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020*, pp. 10170–10177. AAAI Press, 2020.

Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *Journal of Computer and System Science*, 63(4):512–530, 2001.

Karthik C. S., Dániel Marx, Marcin Pilipczuk, and Uéverton S. Souza. Conditional lower bounds for sparse parameterized 2-csp: A streamlined proof. In Merav Parter and Seth Pettie (eds.), *2024 Symposium on Simplicity in Algorithms, SOSA 2024*, pp. 383–395. SIAM, 2024.

Daphne Koller and Nir Friedman. *Probabilistic Graphical Models - Principles and Techniques*. MIT Press, 2009.

Milan Mossé, Duligur Ibeling, and Thomas Icard. Is causal reasoning harder than probabilistic reasoning? *The Review of Symbolic Logic*, 17(1):106–131, 2024.

Nils J. Nilsson. Probabilistic logic. *Artificial Intelligence*, 28(1):71–87, 1986.

Sebastian Ordyniak and Stefan Szeider. Parameterized complexity results for exact bayesian network structure learning. *Journal of Artificial Intelligence Research*, 46:263–302, 2013.

Christos H Papadimitriou and Kenneth Steiglitz. *Combinatorial optimization: algorithms and complexity*. Courier Corporation, 1998.

Judea Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, USA, 2nd edition, 2009.

Dominik Peters. Graphical hedonic games of bounded treewidth. In Dale Schuurmans and Michael P. Wellman (eds.), *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, pp. 586–593. AAAI Press, 2016.

James Renegar. On the computational complexity and geometry of the first-order theory of the reals, part I: introduction. preliminaries. the geometry of semi-algebraic sets. the decision problem for the existential theory of the reals. *Journal of Symbolic Computation*, 13(3):255–300, 1992a.

James Renegar. On the computational complexity and geometry of the first-order theory of the reals, part II: the general decision problem. preliminaries for quantifier elimination. *Journal of Symbolic Computation*, 13(3): 301–328, 1992b.

James Renegar. On the computational complexity and geometry of the first-order theory of the reals, part III: quantifier elimination. *Journal of Symbolic Computation*, 13(3):329–352, 1992c.

Neil Robertson and Paul D. Seymour. Graph minors. III. planar tree-width. *Journal of Combinatorial Theory B*, 36(1):49–64, 1984.

Marko Samer and Stefan Szeider. Constraint satisfaction with bounded treewidth revisited. *Journal of Computer and System Science*, 76(2):103–114, 2010.

Marcus Schaefer, Jean Cardinal, and Tillmann Miltzow. The existential theory of the reals as a complexity class: A compendium. *CoRR*, abs/2407.18006, 2024.

Ilya Shpitser and Judea Pearl. Complete identification methods for the causal hierarchy. *Jornal of Machine Learning Research*, 9:1941–1979, 2008.

Benito van der Zander, Markus Bläser, and Maciej Liskiewicz. The hardness of reasoning about probabilities and causality. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI 2023*, pp. 5730–5738, 2023.

Junzhe Zhang, Jin Tian, and Elias Bareinboim. Partial counterfactual identification from observational and experimental data. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato (eds.), *International Conference on Machine Learning, ICML 2022*, volume 162 of *Proceedings of Machine Learning Research*, pp. 26548–26558. PMLR, 2022.