



Lifted algorithms for symmetric weighted first-order model sampling

Yuanhong Wang^{a,b,*}, Juhua Pu^{a,b}, Yuyi Wang^{c,d}, Ondřej Kuželka^{e,**}

^a State Key Laboratory of Complex and Critical Software Environment, Beihang University, Beijing, China

^b International Innovation Institute, Beihang University, Hangzhou, China

^c CRRC Zhuzhou Institute, Zhuzhou, China

^d ETH Zurich, Zurich, Switzerland

^e Czech Technical University in Prague, Prague, Czech Republic

ARTICLE INFO

Keywords:

Model sampling
First-order logic
Domain-liftability
Counting quantifier

ABSTRACT

Weighted model counting (WMC) is the task of computing the weighted sum of all satisfying assignments (i.e., models) of a propositional formula. Similarly, weighted model sampling (WMS) aims to randomly generate models with probability proportional to their respective weights. Both WMC and WMS are hard to solve exactly, falling under the #P-hard complexity class. However, it is known that the counting problem may sometimes be tractable, if the propositional formula can be compactly represented and expressed in first-order logic. In such cases, model counting problems can be solved in time polynomial in the domain size, and are known as *domain-liftable*. The following question then arises: Is it also the case for WMS? This paper addresses this question and answers it affirmatively. Specifically, we prove the *domain-liftability under sampling* for the two-variables fragment of first-order logic with counting quantifiers in this paper, by devising an efficient sampling algorithm for this fragment that runs in time polynomial in the domain size. We then further show that this result continues to hold even in the presence of cardinality constraints. To empirically validate our approach, we conduct experiments over various first-order formulas designed for the uniform generation of combinatorial structures and sampling in statistical-relational models. The results demonstrate that our algorithm outperforms a state-of-the-art WMS sampler by a substantial margin, confirming the theoretical results.

1. Introduction

Given a propositional formula and a weight for each truth assignment, *weighted model counting* (WMC) aims to compute the cumulative weight of all satisfying assignments (i.e., models) of the input formula. A closely related problem to WMC is *weighted model sampling* (WMS), which samples models of the propositional formula, where the probability of choosing a model is proportional to its weight. These problems find applications in various domains, including machine learning [35,37], probabilistic inference [6], statistics, planning and combinatorics [9,47], and constrained random verification [3,4,40]. Unfortunately, both WMC and WMS are computationally challenging, falling within the #P-hard complexity class [43,10]. Nevertheless, a glimmer of hope emerges when

* Corresponding author at: State Key Laboratory of Complex and Critical Software Environment, Beihang University, Beijing, China.

** Corresponding author.

E-mail addresses: lucienwang@buaa.edu.cn (Y. Wang), ondrej.kuzelka@fel.cvut.cz (O. Kuželka).

the input propositional formula can be naturally and compactly represented using first-order logic. In such cases, the WMC problem may become tractable by exploiting the symmetries present in the problem.

When the input formula Γ is expressed in first-order logic, the problem of calculating the Weighted Model Count (WMC) is known as Weighted First-Order Model Counting (WFOMC). In WFOMC, a domain Δ containing a finite number of objects is provided as part of the input. The models of Γ over Δ are usually defined as the models of the grounding of Γ over Δ . The weights are assigned to atomic facts and their negations in the models, and the overall weight of a model is determined by the product of the weights of its constituent facts. The objective of a WFOMC problem is to compute the weighted summation over all models of the formula Γ over Δ . Many real-world problems, including probabilistic inference and weight learning in various *statistical-relational learning* (SRL) models, can be directly reduced to WFOMC [53,56].

Compared with WMC, an important advantage of WFOMC is the existence of *domain-lifted* (or simply *lifted*) algorithms for certain fragments of first-order logic, which are algorithms that have a polynomial runtime with respect to the domain size [49,1,27,28]. The intuition behind lifted algorithms is to lift the computation from the propositional (ground) level to the higher level of first-order logic, avoiding the need to ground the input formula [52]. For instance, the model count of $\Phi = \forall x \exists y : R(x, y)$ over a domain of size n can be easily computed as $(2^n - 1)^n$ using the fact that all objects in the domain are indistinguishable. However, if we ground Φ to a propositional logic formula, the indistinguishability of the objects will be lost.

The existence of lifted algorithms for WFOMC raises an intriguing question: Are there likewise lifted algorithms for WMS variations in first-order logic? In this work, we answer this question by studying the problem of *weighted first-order model sampling* (WFOMS). This problem, which is a sampling counterpart of WFOMC, aims to generate a model of the input first-order sentence Γ based on a probability that is proportional to its weight. The WFOMS problem offers a natural reduction for a wide range of sampling problems without the necessity of converting them into WMS by grounding the input first-order sentence. Many problems relating to the generation of combinatorial structures can be readily formulated as WFOMS. For instance, suppose we are interested in uniformly sampling labeled 2-colored graphs with n nodes, it is equivalent to solving the WFOMS problem on the sentence:

$$\begin{aligned} \forall x : \neg E(x, x) \wedge \\ \forall x \forall y : E(x, y) \Rightarrow E(y, x) \wedge \\ \forall x : \text{Red}(x) \vee \text{Black}(x) \wedge \\ \forall x : \neg \text{Red}(x) \vee \neg \text{Black}(x) \wedge \\ \forall x \forall y : E(x, y) \Rightarrow \neg(\text{Red}(x) \wedge \text{Red}(y)) \wedge \neg(\text{Black}(x) \wedge \text{Black}(y)) \end{aligned} \quad (1)$$

over a domain of size n , where *Red* and *Black* are two unary predicates representing the two colors of vertices. Moreover, sampling problems in SRL can be also reduced to WFOMS as well. An illustrative example is the sampling of possible worlds from a Markov logic network (MLN) [36]. This can be easily transformed into a WFOMS problem using the same reduction technique employed for converting probabilistic inference to WFOMC (please see Appendix E.1).

The answer to the question of whether there exist lifted algorithms for WFOMS depends on the input sentence of the problem. Indeed, by the direct reduction from the spectrum membership problem [21], one can easily show that it is unlikely for every first-order sentence to have a lifted sampling algorithm.¹ In the context of model counting, a first-order sentence or a class of sentences that allows for a lifted counting algorithm is referred to as *domain-liftable* (or simply *liftable*). Similarly, for WFOMS, we use the terminology of *domain-liftable under sampling* (or simply *sampling liftable*) to describe the fragments that allow for a lifted sampling algorithm. The objective of this paper is to investigate the sampling liftable for certain fragments of first-order logic.

The main contribution of this work is establishing the sampling liftable for the two-variables fragment of first-order logic with counting quantifiers $\exists_{=k}$, $\exists_{\leq k}$ and $\exists_{\geq k}$ (see, e.g., [17]), which stand for *exist exactly k*, *exist at most k*, and *exist at least k* respectively. We note that this fragment, denoted by C^2 , is expressive enough to encode various interesting sampling problems. For instance, the sentence for sampling 2-colored graphs mentioned above contains only two variables and is in C^2 . The uniform generation of k -regular graphs, a problem that has been widely studied in the combinatorics community [7,13], can be also solved by a lifted sampling algorithm for C^2 . This problem can be formulated as a WFOMS on the following C^2 sentence:

$$\forall x \forall y : (E(x, y) \Rightarrow E(y, x)) \wedge \forall x : \neg E(x, x) \wedge \forall x \exists_{=k} y : E(x, y), \quad (2)$$

where $\forall x \exists_{=k} y : E(x, y)$ expresses that every vertex x has exactly k incident edges.

Our proof for the sampling liftable of C^2 proceeds by progressively demonstrating the sampling liftable for the following fragments:

- **UFO²**: The first-order fragment comprising of universally quantified sentences of the form $\forall x \forall y : \psi(x, y)$ with some quantifier-free formula $\psi(x, y)$;

¹ Recall that the spectrum, $\text{Spec}(\Gamma)$, of a formula Γ is the set of numbers n for which Γ has a model over a domain of size n . The spectrum membership problem, “is $n \in \text{Spec}(\Gamma)$?”, can be reduced to WFOMS by checking whether the sampling algorithm fails to sample a model. Even if the sampling problem is guaranteed to have models, the spectrum decision problem can still be reduced to WFOMS. Consider a WFOMS on $\Gamma \vee A()$ over a domain of size n , where A is a nullary predicate not in Γ . If the sampler can generate models with $\text{False } A()$, we can confirm that $n \in \text{Spec}(\Gamma)$, or if the sampler always generates models with $\text{True } A()$, we can conclude that $n \notin \text{Spec}(\Gamma)$ with high probability.

- \mathbf{FO}^2 : The two-variable fragment of first-order logic obtained by restricting the variable vocabulary to $\{x, y\}$;
- \mathbf{C}^2 : The two-variable fragment of first-order logic with counting quantifiers $\exists_{=k}$, $\exists_{\leq k}$ and $\exists_{\geq k}$.

Note that \mathbf{UFO}^2 is a sub-fragment of \mathbf{FO}^2 , which in turn is a sub-fragment of \mathbf{C}^2 . The analysis of sampling liftability for smaller fragments can serve as a basis for that of the larger ones. This is analogous to the case of WFOMC, where the liftability for \mathbf{C}^2 is established by first proving the liftability for \mathbf{UFO}^2 [49] and \mathbf{FO}^2 [54], and then extending the result to \mathbf{C}^2 [28].

This manuscript is an expanded version of a conference paper [58] that appeared in LICS 2023. It includes all the technical preliminaries and provides comprehensive details of the proofs. Additionally, this journal version extends the positive result therein to the \mathbf{C}^2 fragment. The rest of the paper is organized as follows. Section 2 presents the essential concepts used throughout this paper. In Section 3, we formally define the problem of weighted first-order model sampling. The domain-liftability under sampling for \mathbf{UFO}^2 , \mathbf{FO}^2 and \mathbf{C}^2 is then established in Sections 4, 5 and 6, respectively. In Section 7, we extend the result to the case with additional cardinality constraints, which in turn provides a more practically efficient sampling algorithm for a particular subfragment of \mathbf{C}^2 . Section 9 presents the experimental results, which demonstrate the efficiency of our proposed algorithm. Section 10 discusses the relevant literature concerning the WFOMS problems. Finally, Section 11 provides the concluding remarks and outlines potential directions for future research.

2. Preliminaries

In this section, we briefly review the main necessary technical concepts that we will use in the paper.

2.1. First-order logic

We consider the function-free fragment of first-order logic. An *atom* of arity k takes the form $P(x_1, \dots, x_k)$ where P/k is from a vocabulary of predicates (also called relations), and x_1, \dots, x_k are logical variables from a vocabulary of variables. A *literal* is an atom or its negation. A *formula* is defined inductively as an atom, the negation of a single formula, or the conjunction or disjunction of two formulas. A formula may optionally be surrounded by one or more quantifiers of the form $\forall x$ or $\exists x$, where x is a logical variable. A logical variable in a formula is said to be *free* if it is not bound by any quantifier. A formula with no free variables is called a *sentence*. The vocabulary of a formula α is taken to be \mathcal{P}_α .

Given a predicate vocabulary \mathcal{P} , a \mathcal{P} -structure \mathcal{A} is a tuple (Δ, I) , where Δ , called *domain*, is an arbitrary finite set, and I interprets each predicate in \mathcal{P} over Δ . In the context of this paper, the domain is usually predefined, and thus we can leave out the domain from a structure and instead treat a structure as either a set of ground literals in I or their conjunction. Given a \mathcal{P} -structure \mathcal{A} and $\mathcal{P}' \subseteq \mathcal{P}$, we write $\langle \mathcal{A} \rangle_{\mathcal{P}'}$ for the \mathcal{P}' -reduct of \mathcal{A} that only contains the literals for predicates in \mathcal{P}' . We follow the standard semantics of first-order logic for determining whether a structure is a model of a formula. We denote the set of all models of a formula α over the domain Δ by $\mathcal{M}_{\alpha, \Delta}$. A set L of ground literals is said to be *valid* w.r.t. (α, Δ) , if there exists a model μ in $\mathcal{M}_{\alpha, \Delta}$ such that $L \subseteq \mu$.

2.2. Weighted first-order model counting

The *first-order model counting problem* [53] asks, when given a domain Δ and a sentence Γ , how many models Γ has over Δ . The *weighted first-order model counting (WFOMC) problem* adds a pair of weighting functions (w, \bar{w}) to the input, that both map the set of all predicates in Γ to a set of weights: $\mathcal{P}_\Gamma \rightarrow \mathbb{R}$.

Definition 1 (*Weight of literals*). Given a pair of weighting function (w, \bar{w}) and a set L of literals, the weight of L under (w, \bar{w}) is defined as

$$\langle w, \bar{w} \rangle(L) := \prod_{l \in L_T} w(\text{pred}(l)) \cdot \prod_{l \in L_F} \bar{w}(\text{pred}(l))$$

where L_T (resp. L_F) denotes the set of true ground (resp. false) literals in L , and $\text{pred}(l)$ maps a literal l to its corresponding predicate name.

Definition 2 (*Weighted first-order model counting*). Let (w, \bar{w}) be a weighting on a sentence Γ . The WFOMC problem on Γ over a finite domain Δ under (w, \bar{w}) is to compute

$$\text{WFOMC}(\Gamma, \Delta, w, \bar{w}) := \sum_{\mu \in \mathcal{M}_{\Gamma, \Delta}} \langle w, \bar{w} \rangle(\mu).$$

Note that since these weightings are defined on the predicate level, all groundings of the same predicate get the same weights. For this reason, the notion of WFOMC defined here is also referred to as *symmetric* WFOMC.

Given a sentence, or a class of sentences, prior research has mainly focused on its *data complexity* for WFOMC—the complexity of computing $\text{WFOMC}(\Gamma, \Delta, w, \bar{w})$ when fixing the input sentence Γ and weighting (w, \bar{w}) , and treating the domain size n as a unary

input. A sentence, or class of sentences, that exhibits polynomial-time data complexity is said to be *domain-liftable* (or *liftable*). Various fragments of first-order logic have been proven to be liftable, such as UFO^2 [49], FO^2 [54], S^2FO^2 [25], and S^2RU [25].

For technical purposes, when the domain is fixed, we allow the input sentence to contain some ground literals, e.g., $(\forall x \forall y : fr(x, y) \wedge sm(x) \Rightarrow sm(y)) \wedge sm(e_1) \wedge \neg fr(e_1, e_3)$ over a fixed domain of $\{e_1, e_2, e_3\}$. These ground literals are often called evidence, and WFOMC on such sentences is known as WFOMC with evidence [51,50]. In this paper, we also call this counting problem *conditional* WFOMC. An important result of conditional WFOMC is its maintenance of polynomial complexity when the ground literals are unary. This result was provided in [51] for a specific lifted counting algorithm, called first-order knowledge compilation (refer to **Positive Result** in Section 4). We generalize the result to any lifted counting algorithm in Proposition 1 based on a similar technique. Please find the details in Appendix A.

Proposition 1. *Let Γ be a domain-liftable first-order sentence, and let Δ be a domain. For any set L of unary literals grounding on Δ , and any weighting functions (w, \bar{w}) , $\text{WFOMC}(\Gamma \wedge \bigwedge_{l \in L} l, \Delta, w, \bar{w})$ can be computed in time polynomial in both the domain size and the size of L .*

2.3. Types and tables

We define a *1-literal* as an atomic predicate or its negation using only the variable x , and a *2-literal* as an atomic predicate or its negation using both variables x and y . An atom like $R(x, x)$ or its negation is considered a 1-literal, even though R is a binary relation. A 2-literal is always of the form $R(x, y)$ and $R(y, x)$, or their respective negations.

Let \mathcal{P} be a finite vocabulary. A *1-type* over \mathcal{P} is a maximally consistent set² of 1-literals formed by \mathcal{P} . Denote the set of all 1-types over \mathcal{P} as $U_{\mathcal{P}}$. The size of $U_{\mathcal{P}}$ is finite and only depends on the size of \mathcal{P} . We often view a 1-type τ as a conjunction of its elements, and write $\tau(x)$ when viewing τ as a formula.

A *2-table* over \mathcal{P} is a maximally consistent set of 2-literals formed by \mathcal{P} . We often identify a 2-table π with a conjunction of its elements and write it as a formula $\pi(x, y)$. The total number of 2-tables over \mathcal{P} also only depends on the size of \mathcal{P} .

Consider a structure \mathcal{A} defined over a predicate vocabulary \mathcal{P} . An element e in the domain of \mathcal{A} *realizes* the 1-type τ if $\mathcal{A} \models \tau(e)$. Every domain element in \mathcal{A} realizes exactly one 1-type over \mathcal{P} , which we call the *1-type of the element*. The *2-table of an element tuple* $(a, b) \in \Delta^2$ is the unique 2-table π that (a, b) satisfies in \mathcal{A} : $\mathcal{A} \models \pi(a, b)$. A structure is fully characterized by the 1-types of its elements and the 2-tables of its element tuples.

Example 1. Consider the vocabulary $\mathcal{P} = \{F/2, G/1\}$ and the structure over \mathcal{P}

$$\{F(a, a), G(a), F(b, b), \neg G(b), F(a, b), \neg F(b, a)\}$$

with the domain $\{a, b\}$. The 1-types of the elements a and b are $F(x, x) \wedge G(x)$ and $F(x, x) \wedge \neg G(x)$ respectively. The 2-table of the element tuples (a, b) and (b, a) are $F(x, y) \wedge \neg F(y, x)$ and $\neg F(x, y) \wedge F(y, x)$ respectively.

2.4. Notations

We will use $[n]$ to denote the set of $\{1, 2, \dots, n\}$. The notation $\{x_i\}_{i \in [n]}$ represents the set of terms $\{x_1, x_2, \dots, x_n\}$, and $(x_i)_{i \in [n]}$ the vector of (x_1, x_2, \dots, x_n) , which is also denoted by \mathbf{x} . We use $\mathbf{x}^{\mathbf{y}}$ to denote the product over element-wise power of two vectors \mathbf{x} and \mathbf{y} : $\mathbf{x}^{\mathbf{y}} = \prod_{i \in [n]} x_i^{y_i}$. Using the vector notation, we write the multinomial coefficient $\binom{N}{x_1, x_2, \dots, x_n}$ as $\binom{N}{\mathbf{x}}$.

An important operation in our sampling algorithms is *partitioning*. A partition of a set is defined as a grouping of its elements into disjoint subsets. In this paper, all partitions under consideration are presumed to be *order dependent*, and are represented by a vector of subsets $(S_i)_{i \in [m]}$. Given a partition $S = (S_i)_{i \in [m]}$ of a finite set, we refer to the vector of cardinalities $(|S_i|)_{i \in [m]}$ as the *configuration* of S .³ We adopt the term *configuration space* to refer to the set of all partition configurations with a constant length over a given domain.

Definition 3 (Configuration space). Given a non-negative integer M and a positive integer m , we define the configuration space $\mathcal{T}_{M,m}$ as

$$\mathcal{T}_{M,m} = \left\{ (n_i)_{i \in [m]} \mid \sum_{i \in [m]} n_i = M, n_1, n_2, \dots, n_m \in \mathbb{N} \right\}.$$

We remark that the size of $\mathcal{T}_{M,m}$ is given by $\binom{M+m-1}{m-1}$, which is polynomial in M (while exponential in m).

² A set of literals is maximally consistent if it is consistent (does not contain both a literal and its negation) and cannot be extended to a larger consistent set.

³ The notion of the partition configuration is analogous to many other concepts in the lifted inference literature, e.g., the histogram in lifted variable elimination algorithms [30,41].

3. Weighted first-order model sampling

We are now ready to formally define the problem of *weighted first-order model sampling* (WFOMS).

Definition 4 (*Weighted first-order model sampling*). Let (w, \bar{w}) be a pair of weighting functions: $\mathcal{P}_\Gamma \rightarrow \mathbb{R}_{\geq 0}$.⁴ The symmetric WFOMS problem on Γ over a domain Δ under (w, \bar{w}) is to generate a model $\mu \in \mathcal{M}_{\Gamma, \Delta}$ according to the probability

$$\mathbb{P}[\mu] = \frac{\langle w, \bar{w} \rangle(\mu)}{\text{WFOMC}(\Gamma, \Delta, w, \bar{w})}. \quad (3)$$

We call a probabilistic algorithm that realizes a solution to the WFOMS a *weighted model sampler*. Throughout this paper, when the context is clear, we simply refer to the weighted model sampler as a *sampler*. We use $G(\Gamma, \Delta, w, \bar{w})$ to denote a sampler of the WFOMS problem on Γ over Δ under (w, \bar{w}) . We adapt the notion of data complexity of WFOMC to the sampling problem, and say a sampler is *domain-lifted* (or simply *lifted*) if the model sampling algorithm runs in time polynomial in the domain size n . We call a sentence, or class of sentences, that admits a domain-lifted sampler *domain-liftable under sampling* (or simply *sampling liftable*).

Example 2. A sampler of the sentence $\forall x \forall y : (E(x, y) \rightarrow E(y, x)) \wedge \neg E(x, x)$ over a domain of size n under the weighting $w(E) = 3, \bar{w}(E) = 1$ samples undirected graphs, where the probability of each edge is

$$\frac{w(E) \cdot w(E)}{w(E) \cdot w(E) + w(\neg E) \cdot w(\neg E)} = 0.9.$$

The sampler actually corresponds to an Erdős-Rényi graph $\mathcal{G}_{n, p=0.9}$ [11].

We define the probability of a sentence Φ conditional on another sentence Γ over a domain Δ under (w, \bar{w}) as

$$\mathbb{P}[\Phi \mid \Gamma; \Delta, w, \bar{w}] := \frac{\text{WFOMC}(\Phi \wedge \Gamma, \Delta, w, \bar{w})}{\text{WFOMC}(\Gamma, \Delta, w, \bar{w})}.$$

With a slight abuse of notation, we also write the probability of a set L of ground literals conditional on a sentence Γ over a domain Δ under (w, \bar{w}) in the same form:

$$\mathbb{P}[L \mid \Gamma; \Delta, w, \bar{w}] := \mathbb{P}\left[\bigwedge_{l \in L} l \mid \Gamma; \Delta, w, \bar{w}\right].$$

Then, the required sampling probability of a model μ in the WFOMS problem can be written as $\mathbb{P}[\mu] = \mathbb{P}[\mu \mid \Gamma; \Delta, w, \bar{w}]$. When the context is clear, we omit Δ and (w, \bar{w}) in the conditional probability.

In this paper, we often convert one WFOMS problem into another, which is commonly referred to as a *reduction*. The essential property of such reductions extensively used in this paper is their *soundness*.

Definition 5 (*Sound reduction*). A reduction of the WFOMS problem of $(\Gamma, \Delta, w, \bar{w})$ to $(\Gamma', \Delta', w', \bar{w}')$ is sound iff there exists a polynomial-time deterministic function f , such that f is a mapping from $\mathcal{M}_{\Gamma', \Delta'}$ to $\mathcal{M}_{\Gamma, \Delta}$, and for every model $\mu \in \mathcal{M}_{\Gamma, \Delta}$,

$$\mathbb{P}[\mu \mid \Gamma; \Delta, w, \bar{w}] = \sum_{\substack{\mu' \in \mathcal{M}_{\Gamma', \Delta'} : \\ f(\mu') = \mu}} \mathbb{P}[\mu' \mid \Gamma'; \Delta', w', \bar{w}']. \quad (4)$$

Through a sound reduction, we can easily transform a sampler G' of $(\Gamma', w', \bar{w}', \Delta')$ to a sampler G of $(\Gamma, w, \bar{w}, \Delta)$ by

$$G(\Gamma, \Delta, w, \bar{w}) := f(G'(\Gamma', \Delta', w', \bar{w}')).$$

The soundness is transitive, i.e., if the reductions from a WFOMS problem \mathfrak{S}_1 to \mathfrak{S}_2 and from \mathfrak{S}_2 to \mathfrak{S}_3 are both sound, the reduction from \mathfrak{S}_1 to \mathfrak{S}_3 is also sound. For sound reductions in this paper, the most used mapping function is the \mathcal{P}_Γ -reduct $f(\mu') = \langle \mu' \rangle_{\mathcal{P}_\Gamma}$, where $\mathcal{P}_\Gamma \subseteq \mathcal{P}_{\Gamma'}$.

4. Universally quantified FO^2 is sampling liftable

In this section, we provide our first result of domain-liftability under sampling. We consider the fragment of UFO^2 , a fragment containing all sentences in the form $\forall x \forall y : \psi(x, y)$, where $\psi(x, y)$ is a quantifier-free formula.

Theorem 1. *The fragment of UFO^2 is domain-liftable under sampling.*

⁴ The non-negative weights assumption ensures that the sampling probability of a model is well-defined.

We prove the sampling liftability of UFO^2 by constructing a lifted sampler. Suppose that we wish to sample models from some input UFO^2 sentence $\Gamma = \forall x \forall y : \psi(x, y)$ over a domain $\Delta = \{e_i\}_{i \in [n]}$ under weights (w, \bar{w}) . Given a \mathcal{P}_Γ -structure \mathcal{A} over Δ , we denote τ_i the 1-type of the domain element e_i and $\pi_{i,j}$ the 2-table of the elements tuple (e_i, e_j) . The structure \mathcal{A} is fully characterized by the ground 1-types $\tau_i(e_i)$ and 2-tables $\pi_{i,j}(e_i, e_j)$, and we can write the sampling probability of \mathcal{A} as

$$\mathbb{P}[\mathcal{A} \mid \Gamma] = \mathbb{P} \left[\bigwedge_{i \in [n]} \tau_i(e_i) \wedge \bigwedge_{i,j \in [n]: i < j} \pi_{i,j}(e_i, e_j) \mid \Gamma \right].$$

By the definition of conditional probability, the sampling probability can be further decomposed as

$$\mathbb{P}[\mathcal{A} \mid \Gamma] = \underbrace{\mathbb{P} \left[\bigwedge_{\substack{i,j \in [n]: \\ i < j}} \pi_{i,j}(e_i, e_j) \mid \Gamma \wedge \bigwedge_{i \in [n]} \tau_i(e_i) \right]}_{\mathfrak{P}_2} \cdot \underbrace{\mathbb{P} \left[\bigwedge_{i \in [n]} \tau_i(e_i) \mid \Gamma \right]}_{\mathfrak{P}_1}. \quad (5)$$

This decomposition naturally gives rise to a two-phase sampling algorithm:

1. sample the 1-types τ_i of all elements according to the probability \mathfrak{P}_1 , and
2. sample the 2-tables $\pi_{i,j}$ of all elements tuples according to \mathfrak{P}_2 .

The weighted model sampler is shown in Algorithm 1 with `OneTypesSampler` and `TwoTablesSamplerForUFO2` being the two sampling algorithms for the two phases, respectively.

Algorithm 1 `WeightedModelSamplerForUFO2`($\Gamma, \Delta, w, \bar{w}$).

Input: A UFO^2 sentence Γ , a domain Δ of size n and a weighting (w, \bar{w})

Output: A model μ of Γ sampled according to the probability $\mathbb{P}[\mu \mid \Gamma]$

```

1:  $\{\tau_i(e_i)\}_{i \in [n]} \leftarrow \text{OneTypesSampler}(\Gamma, \Delta, w, \bar{w})$ 
2:  $\{\pi_{i,j}(e_i, e_j)\}_{i,j \in [n]} \leftarrow \text{TwoTablesSamplerForUFO2}(\Gamma, \Delta, w, \bar{w}, \{\tau_i\}_{i \in [n]})$ 
3: return  $\{\tau_i(e_i)\}_{i \in [n]} \cup \{\pi_{i,j}(e_i, e_j)\}_{i,j \in [n]}$ 

```

Example 3. Consider the sampling of the 2-colored graphs mentioned in the Introduction. Let the number of vertices be 4. It corresponds to a WFOMS problem on the sentence (1) over a domain of size 4 with the weights being 1 for all predicates. The 1-types that can be sampled are $\text{Red}(x) \wedge \neg \text{Black}(x)$ and $\neg \text{Red}(x) \wedge \text{Black}(x)$, which correspond to coloring the vertices as red and black, respectively. The 2-tables that can be sampled are $E(x, y) \wedge E(y, x)$ and $\neg E(x, y) \wedge \neg E(y, x)$, indicating the presence or absence of an edge between two vertices. This sampling problem can be decomposed into two steps: coloring of the vertices and sampling of the edges.

















4.1. Sampling 1-types

Recall that the number of all possible 1-types over a predicates vocabulary \mathcal{P} is $|U_{\mathcal{P}}|$. Any assignment of 1-types to elements can be viewed as a $|U_{\mathcal{P}}|$ -length partition of the domain, where each disjoint subset precisely contains the elements realizing the corresponding 1-type. Therefore, sampling 1-types is equivalent to randomly partitioning the domain Δ into subsets of size $|U_{\mathcal{P}}|$. Furthermore, the symmetry property of the weighting function guarantees that any permutation of the elements in the domain does not impact the satisfaction or weight of the models. Thus, partitions with the same configuration are equally likely to be sampled. This allows us to further split the sampling problem of 1-types into two stages: 1) sampling a partition configuration and 2) randomly partitioning the domain according to the sampled configuration. The latter stage of random partitioning is trivial, and we will demonstrate that the first stage of sampling a partition configuration can be accomplished in time polynomial in the domain size.

Example 4. Consider 1-types sampling for Example 3. Let $w(\text{Red}) = 2$ and all other weights be 1. The sampling of 1-types is to color the vertices as red and black, and realized by a binary partition. The probability of a given partition or coloring scheme is proportional to the summed weight of all graphs with the coloring scheme. As shown in Table 1, the coloring schemes with the same partition configuration have the same weight W , and thus have the same sampling probability.

We first observe that the overall number of partition configurations is given by the size of the configuration space $\mathcal{T}_{n, |U_{\mathcal{P}}|}$, which is polynomial in the domain size. This property allows us to use the *enumerative sampling* approach to generate the partition configuration randomly. In enumerative sampling, we enumerate all possible values of the random variable, and then pick one of them according to the probability distribution. To this end, we need to compute the probability of each possible partition configuration. For any partition configuration $\mathbf{m} = (m_1, m_2, \dots, m_{|U_{\mathcal{P}}|})$, there are a total of $\binom{n}{\mathbf{m}}$ partitions. The symmetry of the weighting function

Table 1
1-types sampling for 2-colored graphs.

| Configuration | Coloring scheme | W | $ \mathcal{C} $ |
|---------------|---|-----|-----------------|
| (0, 4) |  | 1 | $2^0 = 1$ |
| (1, 3) |     | 2 | $2^3 = 8$ |
| (2, 2) |       | 4 | $2^4 = 16$ |
| (3, 1) |     | 8 | $2^3 = 8$ |
| (4, 0) |  | 16 | $2^0 = 1$ |

* filled: black vertices; empty: red vertices.

guarantees that all of these partitions have the same sampling probability. The probability $\mathfrak{P}_1^{\mathbf{m}}$ can be written in the form of \mathfrak{P}_1 , shown below, due to the equivalence between partitions and 1-types assignments of domain elements:

$$\mathfrak{P}_1^{\mathbf{m}} = \frac{\text{WFOMC}(\Gamma \wedge \bigwedge_{i=1}^n \tau_i(e_i), \Delta, w, \bar{w})}{\text{WFOMC}(\Gamma, \Delta, w, \bar{w})}, \quad (6)$$

where the 1-types τ_i correspond to a specific partition with the size \mathbf{m} . Thus, the sampling probability of the partition size \mathbf{m} can be calculated as $\mathfrak{P}_1^{\mathbf{m}} \cdot \binom{n}{\mathbf{m}}$. Finally, let us demonstrate that the computation of the probability $\mathfrak{P}_1^{\mathbf{m}}$ is polynomial-time in the domain size. In (6), the computation of the denominator WFOMC on Γ is in polynomial-time in the domain size due to the liftability of Γ for WFOMC [49]. The numerator can be viewed as a WFOMC problem of Γ conditioned on the unary facts in all ground 1-types $\tau_i(e_i)$, whose size is clearly polynomial in the size of the domain. By Proposition 1 and the liftability of Γ , it follows that the complexity of computing $\mathfrak{P}_1^{\mathbf{m}}$ is polynomial in the domain size. The algorithm for sampling 1-types is presented in Algorithm 2.

Algorithm 2 OneTypesSampler($\Gamma, \Delta, w, \bar{w}$).

Input: A UFO² sentence Γ , a domain $\Delta = \{e_i\}_{i \in [n]}$ and a weighting (w, \bar{w})

Output: 1-types $\{\tau_i(e_i)\}_{i \in [n]}$

```

1: for  $\mathbf{m} \in \mathcal{T}_{n, |U_{P_r}|}$  do
2:    $p(\mathbf{m}) \leftarrow \mathfrak{P}_1^{\mathbf{m}} \cdot \binom{n}{\mathbf{m}}$  ▷ Compute the probability of  $\mathbf{m}$ 
3: end for
4: Sample  $\mathbf{m}^*$  from  $\mathcal{T}_{n, |U_{P_r}|}$  according to the probability distribution  $p$ 
5: Randomly partition  $\Delta$  into  $|U_{P_r}|$  subsets of size  $m_i^*$  for  $i \in [|U_{P_r}|]$ 
6: for  $i \in [n]$  do
7:    $\tau_i \leftarrow$  the 1-type of the element  $e_i$  in the partition
8: end for
9: return  $\{\tau_i(e_i)\}_{i \in [n]}$ 

```

4.2. Sampling 2-tables

For sampling $\pi_{i,j}$ according to the probability

$$\mathfrak{P}_2 = \mathbb{P} \left[\bigwedge_{\substack{i,j \in [n]: \\ i < j}} \pi_{i,j}(e_i, e_j) \mid \Gamma \wedge \bigwedge_{i \in [n]} \tau_i(e_i) \right],$$

we first ground out $\Gamma = \forall x \forall y : \psi(x, y)$ over the domain Δ :

$$\bigwedge_{i,j \in [n]: i < j} \psi(e_i, e_j) \wedge \psi(e_j, e_i),$$

where $\psi(x, y)$ is the quantifier-free formula in Γ . Let $\psi_{i,j}(x, y)$ be the simplified formula of $\psi(x, y) \wedge \psi(y, x)$ obtained by replacing the ground 1-literals with their truth value given by the 1-types τ_i and τ_j . The probability \mathfrak{P}_2 can be written as

$$\mathbb{P} \left[\bigwedge_{i,j \in [n]: i < j} \pi_{i,j}(e_i, e_j) \mid \bigwedge_{i,j \in [n]: i < j} \psi_{i,j}(e_i, e_j) \right]. \quad (7)$$

In this conditional probability, all ground 2-tables $\pi_{i,j}(e_i, e_j)$ are independent in the sense that they do not share any ground literals. The independence also holds for the ground formulas $\psi_{i,j}(e_i, e_j)$, because all ground 1-literals were replaced by their truth values. It follows that (7) can be factorized into

$$\prod_{i,j \in [n]: i < j} \mathbb{P}[\pi_{i,j}(e_i, e_j) \mid \psi_{i,j}(e_i, e_j)].$$

Hence, sampling the 2-tables $\pi_{i,j}$ can be solved separately. The outline of the algorithm is presented in Algorithm 3. The overall computational complexity is clearly polynomial in the domain size.

Algorithm 3 TwoTablesSamplerForUFO2($\Gamma, \Delta, w, \bar{w}, (\tau_i)_{i \in [n]}$).

Input: A \mathbf{UFO}^2 sentence Γ , a domain $\Delta = \{e_i\}_{i \in [n]}$, a weighting (w, \bar{w}) , and 1-types $(\tau_i)_{i \in [n]}$

Output: 2-tables $\{\pi_{i,j}(e_i, e_j)\}_{i,j \in [n]}$

```

1:  $\mathcal{A} \leftarrow \emptyset$ 
2:  $\Pi \leftarrow$  all 2-tables over  $\mathcal{P}_\Gamma$  ▷ Note that the size of  $\Pi$  only depends on  $\mathcal{P}_\Gamma$ 
3: for  $i \in [n]$  do
4:   for  $j \in [n]$  do
5:     if  $i < j$  then
6:       Simplify  $\psi(x, y) \wedge \psi(y, x)$  to  $\psi_{i,j}(x, y)$  by replacing the 1-literals with their truth values given by  $\tau_i$  and  $\tau_j$ 
7:       for  $\pi \in \Pi$  do
8:          $p(\pi) \leftarrow \mathbb{P}[\pi(e_i, e_j) \mid \psi_{i,j}(e_i, e_j)]$ 
9:       end for
10:      Sample  $\pi_{i,j}$  from  $\Pi$  according to the probability distribution  $p$ 
11:       $\mathcal{A} \leftarrow \mathcal{A} \cup \{\pi_{i,j}(e_i, e_j)\}$ 
12:    end if
13:  end for
14: end for
15: return  $\mathcal{A}$ 

```

Example 5. Consider the 2-tables sampling in the WFOMS problem presented in Example 3. It is clear that the sampling of an edge is fully determined by the colors of its endpoints. There are only two cases: if the endpoints share the same color, no edge can exist between them, otherwise the edge is sampled with a probability of 1/2 (recall that $w(E) = \bar{w}(E) = 1$).

Proof of Theorem 1. The procedures presented above for sampling τ_i and $\pi_{i,j}$ are both polynomial in the domain size, which forms a lifted sampler for Γ , and thus complete the proof. \square

Remark 1. Directly extending the approach above to the case of \mathbf{FO}^2 necessitates another novel and more sophisticated strategy, specifically for the sampling of 2-tables. This is due to the fact that it is impossible to decouple the grounding of $\forall x \exists y : \varphi(x, y)$ into a conjunction of independent formulas, even when conditioned on the sampled 1-types.

5. \mathbf{FO}^2 is sampling liftable

We now show the domain-liftability under sampling of the \mathbf{FO}^2 fragment. It is common for logical algorithms to operate on normal form representations instead of arbitrary sentences. The normal form of \mathbf{FO}^2 used in our sampling algorithm is the *Scott normal form (SNF)* [38]; an \mathbf{FO}^2 sentence is in SNF, if it is written as:

$$\Gamma = \forall x \forall y : \psi(x, y) \wedge \bigwedge_{k \in [m]} \forall x \exists y : \varphi_k(x, y), \quad (8)$$

where the formulas $\psi(x, y)$ and $\varphi_k(x, y)$ are quantifier-free formulas. It is well-known that one can convert any \mathbf{FO}^2 sentence Γ in polynomial-time into a formula Γ_{SNF} in SNF such that Γ and Γ_{SNF} are equisatisfiable [16]. The principal idea is to substitute, starting from the atomic level and working upwards, any subformula $\varphi(x) = Qy : \phi(x, y)$, where $Q \in \{\forall, \exists\}$ and ϕ is quantifier-free, with an atomic formula A_φ , where A_φ is a fresh predicate symbol. This novel atom $A_\varphi(x)$ is then separately “axiomatized” to be equivalent to $\varphi(x)$. If the weight of A_φ is set to be $w(A_\varphi) = \bar{w}(A_\varphi) = 1$, we have that such reduction is also sound (recall the soundness definition in Definition 5).

Lemma 1. For any WFOMS problem $\mathfrak{S} = (\Gamma, \Delta, w, \bar{w})$ where Γ is an \mathbf{FO}^2 sentence, there exists a WFOMS problem $\mathfrak{S}' = (\Gamma', \Delta, w', \bar{w}')$, where Γ' is in SNF, such that the reduction from \mathfrak{S} to \mathfrak{S}' is sound.

The proof is straightforward, as every novel predicate (e.g., P_φ) introduced in the SNF transformation is axiomatized to be equivalent to the subformula $(\varphi(x))$, and thus fully determined by the subformula in every model of the resulting SNF sentence (see the details in Appendix B.1).

Theorem 2. *The fragment \mathbf{FO}^2 is domain-liftable under sampling.*

We demonstrate the sampling liftability of \mathbf{FO}^2 through the development of a lifted sampler that bears resemblance to the framework presented in Section 4. Specifically, the approach involves a two-stage algorithm derived from the probability decomposition of 1-types and 2-tables in (5), which comprises the sampling of 1-types τ_i in the first stage, followed by the sampling of 2-tables $\pi_{i,j}$ in the second stage. In the first stage, the same technique used for \mathbf{UFO}^2 as discussed in Section 4 can be used. The time complexity of this process remains polynomial in the size of the domain following the same reasoning and the domain-liftability of \mathbf{FO}^2 [54]. The second stage, which is to sample 2-tables conditional on the sampled 1-types, however, is the most challenging aspect of the sampling problem and will be the main focus of the remainder of this section.

5.1. A working example

Before delving into the details of the algorithm, we provide a working example to illustrate the basic idea for sampling 2-tables. The example is to sample an undirected graph of size n without any isolated vertex uniformly at random. Its corresponding sentence can be written in SNF:

$$\Gamma_G := (\forall x \forall y : (E(x, y) \Rightarrow E(y, x)) \wedge \neg E(x, x)) \wedge (\forall x \exists y : E(x, y)),$$

and the sampling problem corresponds to a WFOMS problem on Γ_G under $w(E) = \bar{w}(E) = 1$ over a domain of vertices $V = \{v_i\}_{i \in [n]}$. In this sentence, the only 1-type that can be sampled is $\neg E(x, x)$, which does not require any sampling. The 2-tables that can be sampled are $\pi^1(x, y) = E(x, y) \wedge E(y, x)$ and $\pi^2(x, y) = \neg E(x, y) \wedge \neg E(y, x)$ representing the connectedness of two vertices. In the following, we will focus on the sampling problem of 2-tables (i.e., edges).

We first apply the following transformation on Γ_G resulting in Γ_{GT} :

1. introduce an auxiliary *Tseitin predicate* $Z/1$ with the weight $w(Z) = \bar{w}(Z) = 1$ that indicates the non-isolation of vertices,
2. append $\forall x : Z(x) \Leftrightarrow \exists y : E(x, y)$ to Γ_G , and
3. remove $\forall x \exists y : E(x, y)$.

We then consider a slightly more general WFOMS problem on $\hat{\Gamma}_G := \Gamma_{GT} \wedge \bigwedge_{v \in V_\exists} Z(v)$ over a domain V_\forall , where $V_\exists \subseteq V_\forall \subseteq V$ and V_\exists represents the set of vertices that should be non-isolated in the graph induced by V_\forall . The original WFOMS problem of Γ_G can be clearly reduced to the more general problem by setting $V_\exists = V_\forall = V$, and the reduction is sound with the mapping function $f(\mu') = \langle \mu' \rangle_{\{E\}}$. Given a $\mathcal{P}_{\hat{\Gamma}_G}$ -structure \mathcal{A} , the interpretation of the predicate E fully determines the interpretation of Z . Therefore, in the subsequent discussion, any $\mathcal{P}_{\hat{\Gamma}_G}$ -structure should be understood as a $\mathcal{P}_{\{E\}}$ -structure, where Z is omitted.

Given an $\mathcal{P}_{\hat{\Gamma}_G}$ -structure \mathcal{A} , we denote the substructure of \mathcal{A} concerning a vertex $v_i \in V_\forall$ by \mathcal{A}_i :

$$\mathcal{A}_i := \bigcup_{v_j \in V_\forall : j \neq i} \pi_{i,j}(v_i, v_j),$$

where $\pi_{i,j}$ is the 2-table of (v_i, v_j) . We then proceed to choose a vertex v_i from V_\forall , and decompose the sampling probability of \mathcal{A} as follows:

$$\mathbb{P}[\mathcal{A} | \hat{\Gamma}_G] = \mathbb{P}[\mathcal{A} | \hat{\Gamma}_G \wedge \mathcal{A}_i] \cdot \mathbb{P}[\mathcal{A}_i | \hat{\Gamma}_G]. \quad (9)$$

The decomposition leads to two successive sub-problems of the general WFOMS problem: the first one is to sample a substructure \mathcal{A}_i from $\hat{\Gamma}_G$; and the second can be viewed as a new WFOMS problem on $\hat{\Gamma}_G$ conditioned on the sampled \mathcal{A}_i .

We first show that if \mathcal{A}_i is valid w.r.t. $(\hat{\Gamma}_G, V_\forall)$, the new WFOMS problem can be reduced to the general WFOMS problem but with the smaller domain $V'_\forall = V_\forall \setminus \{v_i\}$ and

$$V'_\exists = V_\exists \setminus (\{v_i\} \cup \{v_i | v_i \in V_\exists : \pi_{i,i} = \pi^1\}).$$

The set V'_\exists removes the vertices that become non-isolated after the sampling of \mathcal{A}_i from V_\exists . It is easy to check that the reduction is sound, because every model of $\Gamma_{GT} \wedge \bigwedge_{v \in V'_\exists} Z(v)$ over V'_\forall can be mapped to a unique model of $\hat{\Gamma}_G \wedge \mathcal{A}_i$ over V_\forall , and vice versa, without affecting the satisfaction and weight of the models.

By the reduction above, the decomposition of (9) can be performed recursively on the WFOMS problem on $\hat{\Gamma}_G$ over V_\forall . Specifically, the recursive algorithm takes V_\forall and V_\exists as input,

1. selects a vertex v_i from V_\forall ,
2. samples its substructure \mathcal{A}_i according to the probability $\mathbb{P}[\mathcal{A}_i | \hat{\Gamma}_G]$, and
3. obtains a new problem with updated V'_\forall and V'_\exists for recursion.

The recursion terminates when all substructures \mathcal{A}_i are sampled (i.e., V_\forall contains a single vertex), or the problem degenerates to a WFOMS problem on \mathbf{UFO}^2 sentence (i.e., V_\exists is empty). The number of recursions is less than $|V|$, the total number of vertices. An

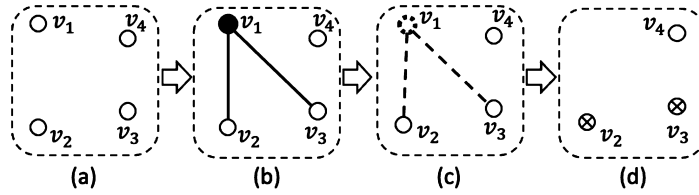


Fig. 1. A sampling step for an undirected graph with no isolated vertices: (a) begins with an initial graph that has no edges, and in the more general sampling problem, $V_\forall = V_\exists = \{v_1, v_2, v_3, v_4\}$; (b) sample edges for the vertex v_1 ; (c) remove the vertex v_1 with its sampled edges; (d) and obtain a graph with some vertices already non-isolated (v_2 and v_3), resulting in a new sampling problem with $V'_\forall = \{v_2, v_3, v_4\}$ and $V'_\exists = \{v_4\}$.

example of a single recursive step is shown in Fig. 1: the vertex v_1 from $V_\forall = V_\exists = \{v_1, v_2, v_3, v_4\}$ is selected with the edges incident to it sampled; the vertex v_1 is then removed from V_\forall , and the vertices v_2 and v_3 that become non-isolated are also removed from V_\exists .

The remaining problem is to sample the substructure \mathcal{A}_t according to $\mathbb{P}[\mathcal{A}_t | \hat{\Gamma}_G]$. Recall that \mathcal{A}_t determines the edges between v_t and vertices in V'_\forall . Let $V_1 = V'_\forall \setminus V_\exists$ and $V_2 = V'_\forall \setminus V_1$. We can effectively generate a sample of \mathcal{A}_t by sampling two binary partitions of V_1 and V_2 , respectively, yielding the sets V_{11}, V_{12} and V_{21}, V_{22} ; the vertices in V_{11} and V_{21} will be connected to v_t , while the vertices in V_{12} and V_{22} will be disconnected to v_t . It can be demonstrated that the sampling probability of a substructure \mathcal{A}_t only depends on the two partition configurations ($|V_{11}|, |V_{12}|$) and ($|V_{21}|, |V_{22}|$). The proof of this claim can be found in Section 5.2.4, where the more general case of \mathbf{FO}^2 sampling is addressed. As a result, the sampling of \mathcal{A}_t can be achieved by first sampling the two partition configurations, followed by two random partitions on V_1 and V_2 with the respective sampled configurations. To sample a tuple of partition configurations, we use the enumerative sampling method. The number of all possible tuples of partition configurations is clearly polynomial in $|V'_\forall|$, and it will be shown in Section 5.2.4 that the sampling probability of each configuration tuple can be computed in time polynomial in $|V'_\forall|$. Therefore, the complexity of the sampling algorithm is polynomial in the number of vertices. This, together with the complexity of the recursion procedure, which is also polynomial in the number of vertices, implies that the whole sampling algorithm is lifted.

5.2. Domain recursive sampling for 2-tables

We now present our algorithm for sampling 2-tables, which uses the technique illustrated in the previous subsection. The core idea, called *domain recursion*, involves considering individual objects from the domain at a time, sampling their corresponding atomic facts, and subsequently obtaining a new sampling problem that has a similar form to the original one but with a smaller domain and potentially fewer existentially-quantified formulas. This process is repeated recursively on the reduced sampling problems until the domain has become a singleton or all the existentially quantified formulas have been eliminated.

Let us consider the WFOMS problem with fixed 1-types ($\Gamma \wedge \bigwedge_{i \in [n]} \tau_i(e_i), \Delta, w, \bar{w}$), where Γ is a sentence in SNF (8), $\Delta = \{e_i\}_{i \in [n]}$ is a domain of size n , and each τ_i is the sampled 1-type of the element e_i . W.l.o.g.,⁵ we suppose that each formula $\varphi_k(x, y)$ in the SNF sentence (8) is an atomic formula $R_k(x, y)$, where each R_k is a binary predicate with weights $w(R_k) = \bar{w}(R_k) = 1$.

5.2.1. A more general sampling problem

We first construct the following sentence from Γ :

$$\Gamma_T := \forall x \forall y : \psi(x, y) \wedge \bigwedge_{k \in [m]} \forall x : Z_k(x) \Leftrightarrow (\exists y : R_k(x, y)), \quad (10)$$

where each $Z_k/1$ is a Tseitin predicate with the weight $w(Z_k) = \bar{w}(Z_k) = 1$. Note that in (10), the interpretation of R_k fully determines the interpretation of Z_k . Once the 2-tables are sampled, the interpretation of Z_k is also fixed, adding no additional cost to the sampling problem. Therefore, for ease of presentation, we will omit the handling of Z_k in the following discussion.

We then consider a more general WFOMS problem of the following sentence

$$\Gamma_T \wedge \bigwedge_{i \in [n]} (\tau_i(e_i) \wedge C_i) \quad (11)$$

over a domain of $\{e_i\}_{i \in [n]}$, where each C_i is a conjunction over a subset of the ground atoms $\{Z_k(e_i)\}_{k \in [m]}$. We call C_i the *existential constraint* on the element e_i and allow $C_i = \top$, which means e_i is not existentially quantified.

The more general sampling problem has the necessary structure for the domain recursion algorithm to be performed. To verify it, the original WFOMS problem on $\Gamma \wedge \bigwedge_{i \in [n]} \tau_i(e_i)$ can be reduced to the more general problem by setting all existential constraints to be

⁵ Any SNF sentence can be transformed into such form by introducing an auxiliary predicate R_k with weights $w(R_k) = \bar{w}(R_k) = 1$ for each $\varphi_k(x, y)$, appending $\forall x \forall y : R_k(x, y) \Leftrightarrow \varphi_k(x, y)$ to the sentence, and replacing $\varphi_k(x, y)$ with $R_k(x, y)$. The transformation is obviously sound when viewing it as a reduction in WFOMS.

$\bigwedge_{k \in [m]} Z_k(x)$. On the other hand, the WFOMS problem on the **UFO**² sentence $\forall x \forall y : \psi(x, y) \wedge \bigwedge_{i \in [n]} \tau_i(e_i)$ is also reducible to the problem with $C_i = \top$ for all $i \in [n]$. It is easy to check that these two reductions are both sound.

5.2.2. Block and cell types

It is worth noting that the Tseitin predicates Z_k introduced in the more general sampling problem are not contained in the given 1-types τ_i . Therefore, to incorporate these predicates into the sampling problem, we introduce the notions of block and cell types.

Consider a set of Tseitin predicates $\{Z_k\}_{k \in [m]}$. A *block type* β is a subset of the atoms $\{Z_k(x)\}_{k \in [m]}$. The number of the block types is 2^m , where m is the number of existentially-quantified formulas. We often represent a block type as $\beta(x)$ and view it as a conjunctive formula over the atoms within the block. It is important to note that the block types only indicate which Tseitin atoms should hold for a given element, while the Tseitin atoms not covered by the block types are left unspecified. In contrast, the 1-types explicitly determine the truth values of all unary and reflexive atoms, excluding the Tseitin atoms.

The block type and 1-type work together to define the cell type. A *cell type* $\eta = (\beta, \tau)$ is a pair of a block type β and a 1-type τ . We also write a cell type as a conjunctive formula of $\eta(x) = \beta(x) \wedge \tau(x)$. The cell type of an element is simply the tuple of its block type (given by the sentence $\hat{\Gamma}$) and 1-type (already sampled in the first stage of the algorithm). The cell types of elements naturally produce a partition of the domain, similar to how a 1-types assignment divides a domain into disjoint subsets. Each disjoint subset of elements is called a *cell*. The *cell configuration* is defined as the configuration of the corresponding partition.

With the notion of block and cell type, we can write the existential constraint C_i in (11) as $\beta_i(e_i)$, where β_i is the block type of the element e_i , and the sentence (11) as $\Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i)$, where $\eta_i = (\beta_i, \tau_i)$.

5.2.3. Domain recursion

We now show how to apply the domain recursion scheme to the WFOMS problem $(\Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i), \Delta, w, \bar{w})$. Let \mathcal{A}_i be the set of ground 2-tables concerning the element e_i :

$$\mathcal{A}_i := \bigcup_{j \in [n]: j \neq i} \pi_{i,j}(e_i, e_j). \quad (12)$$

By the domain recursion, we select an arbitrary element e_t from Δ and decompose the sampling probability of a structure \mathcal{A} into

$$\mathbb{P} \left[\mathcal{A} \mid \Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i) \wedge \mathcal{A}_t \right] \cdot \mathbb{P} \left[\mathcal{A}_t \mid \Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i) \right].$$

We always assume that the substructure \mathcal{A}_t is valid w.r.t. $(\Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i), \Delta)$, i.e., $\Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i) \wedge \mathcal{A}_t$ is satisfiable, or the sampling probability of \mathcal{A}_t is zero. We will demonstrate that the WFOMS problem specified by the first probability in the above equation can be reduced to a new WFOMS problem of the same form as the original sampling problem, but over a smaller domain $\Delta' = \Delta \setminus e_t$.

We first introduce the notion of the *relaxing* on block types and cell types, which is the basic operation of the reduction. Given a 2-table π and a block type β , the relaxed block of β under π is defined as

$$\beta \downarrow \pi := \beta \setminus \{Z_k(x) \mid k \in [m] : R_k(y, x) \in \pi\}.$$

Similarly, we can apply the relaxation under π on a cell type $\eta = (\beta, \tau)$, resulting in a relaxed cell type $\eta \downarrow \pi = (\beta \downarrow \pi, \tau)$.

Example 6. Consider a WFOMS problem on the sentence

$$\forall x : (Z_1(x) \Leftrightarrow (\exists y : R_1(x, y))) \wedge$$

$$\forall x : (Z_2(x) \Leftrightarrow (\exists y : R_2(x, y)))$$

$$\wedge \dots \wedge Z_1(a) \wedge Z_2(a) \wedge \dots,$$

where only the block type $\beta = \{Z_1(x), Z_2(x)\}$ of the element a is shown. Suppose the sampled 2-table of (e, a) is

$$\pi = \{\neg R_1(x, y), R_1(y, x), R_2(x, y), \neg R_2(y, x)\}.$$

Then the relaxed block of the element a is $\beta \downarrow \pi = \{Z_2(x)\}$, as the corresponding quantified formula $\exists y : R_1(a, y)$ of $Z_1(a)$ is satisfied by the fact $R_1(a, e_i)$ in $\pi(e_t, a)$.

Let

$$\tilde{\Gamma} := \Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i) \wedge \mathcal{A}_t, \quad (13)$$

and

$$\tilde{\Gamma}' := \Gamma_T \wedge \bigwedge_{i \in [n] \setminus \{t\}} (\eta_i \downarrow \pi_{t,i})(e_i). \quad (14)$$

We have the following lemma.

Lemma 2. If $\tilde{\Gamma}$ is satisfiable, the reduction from the WFOMS problem $(\tilde{\Gamma}, \Delta, w, \bar{w})$ to $(\tilde{\Gamma}', \Delta', w, \bar{w})$ is sound.

Proof. We prove the soundness of the reduction by constructing a mapping function from models of $\tilde{\Gamma}'$ to models of $\tilde{\Gamma}$. This function is defined as $f(\mu') = \mu' \cup \mathcal{A}_t \cup \tau_t(e_t)$. The function f is clearly deterministic, polynomial-time. To simplify the remaining arguments of the proof, we first show that f is bijective. Write the sentences $\tilde{\Gamma}$ and $\tilde{\Gamma}'$ as

$$\tilde{\Gamma} = \forall x \forall y : \psi(x, y) \wedge \bigwedge_{i \in [n]} (\tau_i(e_i) \wedge \beta_i(e_i)) \wedge \mathcal{A}_t \wedge \Lambda,$$

$$\tilde{\Gamma}' = \forall x \forall y : \psi(x, y) \wedge \bigwedge_{i \in [n] \setminus \{t\}} (\tau_i(e_i) \wedge (\beta_i \downarrow \pi_{t,i})(e_i)) \wedge \Lambda,$$

where $\Lambda = \bigwedge_{k \in [m]} \forall x : Z_k(x) \Leftrightarrow (\exists y : R_k(x, y))$.

(\Rightarrow). For any model μ' in $\mathcal{M}_{\tilde{\Gamma}', \Delta'}$, we prove that $f(\mu') = \mu' \cup \mathcal{A}_t \cup \tau_t(e_t)$ is a model in $\mathcal{M}_{\tilde{\Gamma}, \Delta}$. First, one can easily check that $f(\mu')$ satisfies $\forall x \forall y : \psi(x, y) \wedge \bigwedge_{i \in [n]} \tau_i(e_i)$. Next, we show that $f(\mu')$ also satisfies $\bigwedge_{i \in [n]} \beta_i(e_i) \wedge \Lambda$. For any element $i \in [n] \setminus \{t\}$, we have that $\exists y : R_k(e_i, y)$ is true in μ' (and also $f(\mu')$) for all $k \in [m]$ such that $Z_k(x) \in \beta_i \downarrow \pi_{t,i}$. By the definition of the relaxed block, for any $k \in [m]$ such that $Z_k(x) \in \beta_i \setminus (\beta_i \downarrow \pi_{t,i})$, the ground relation $R_k(e_i, e_t)$ is in \mathcal{A}_t , and thus $\exists y : R_k(e_i, y)$ is also satisfied in $f(\mu')$. It follows that $\exists y : R_k(e_i, y)$ is true in $f(\mu')$ for all $k \in [m]$ such that $Z_k(x) \in \beta_i$. Furthermore, the satisfaction of $\exists y : R_k(e_t, y)$ for all $k \in [m]$ such that $Z_k(x) \in \beta_t$ is guaranteed by the satisfiability of $\tilde{\Gamma}$. Therefore, it is easy to show that $f(\mu')$ satisfies $\bigwedge_{i \in [n]} \beta_i(e_i) \wedge \Lambda$, which together with the satisfaction of $\forall x \forall y : \psi(x, y) \wedge \bigwedge_{i \in [n]} \tau_i(e_i)$ implies that $f(\mu')$ is a model of $\tilde{\Gamma}$.

(\Leftarrow). For any model μ in $\mathcal{M}_{\tilde{\Gamma}, \Delta}$, we show that there exists a unique model μ' in $\mathcal{M}_{\tilde{\Gamma}', \Delta'}$ such that $f(\mu') = \mu$. Let the respective structure be $\mu' = \mu \setminus ((\mathcal{A}_t) \cup \tau_t(e_t))$, and the uniqueness is clear from the definition of f . First, it is easy to check that μ' satisfies $\forall x \forall y : \psi(x, y) \wedge \bigwedge_{i \in [n] \setminus \{t\}} \tau_i(e_i)$. Then we show that μ' also satisfies $\bigwedge_{i \in [n] \setminus \{t\}} (\beta_i \downarrow \pi_{t,i})(e_i) \wedge \Lambda$. For any $i \in [n] \setminus \{t\}$, we have that $\exists y : R_k(e_i, e)$ is true in μ . Grounding $\exists y : R_k(e_i, e)$ over Δ and replacing the atoms $R_k(e_t, e_j)$ with their truth values in \mathcal{A}_t , we have that $\bigvee_{j \in [n] \setminus \{t\}} R_k(e_i, e_j)$ is true in μ' for all $k \in [m]$ such that $Z_k(x) \in \beta_i \downarrow \pi_{t,i}$. Thus, μ' also satisfies $\bigwedge_{i \in [n] \setminus \{t\}} (\beta_i \downarrow \pi_{t,i})(e_i) \wedge \Lambda$. This, along with the satisfaction of $\forall x \forall y : \psi(x, y) \wedge \bigwedge_{i \in [n] \setminus \{t\}} \tau_i(e_i)$, leads to the conclusion.

Now, we are prepared to demonstrate the consistency of sampling probability through the mapping function. Since f is bijective, it is sufficient to show

$$\mathbb{P}[f(\mu') \mid \tilde{\Gamma}; \Delta, w, \bar{w}] = \mathbb{P}[\mu' \mid \tilde{\Gamma}'; \Delta', w, \bar{w}]$$

for any model in $\mathcal{M}_{\tilde{\Gamma}', \Delta'}$. By the definition of the mapping function f , we have

$$\langle w, \bar{w} \rangle(f(\mu')) = \langle w, \bar{w} \rangle(\mu') \cdot \langle w, \bar{w} \rangle(\mathcal{A}_t) \cdot \langle w, \bar{w} \rangle(\tau_t).$$

Moreover, due to the bijectivity of f , we have

$$\begin{aligned} \text{WFOMC}(\tilde{\Gamma}, \Delta, w, \bar{w}) &= \sum_{\mu \in \mathcal{M}_{\tilde{\Gamma}, \Delta}} \langle w, \bar{w} \rangle(\mu) \\ &= \sum_{\mu' \in \mathcal{M}_{\tilde{\Gamma}', \Delta'}} \langle w, \bar{w} \rangle(f(\mu')) \\ &= \langle w, \bar{w} \rangle(\mathcal{A}_t) \cdot \langle w, \bar{w} \rangle(\tau_t) \cdot \sum_{\mu' \in \mathcal{M}_{\tilde{\Gamma}', \Delta'}} \langle w, \bar{w} \rangle(\mu') \\ &= \langle w, \bar{w} \rangle(\mathcal{A}_t) \cdot \langle w, \bar{w} \rangle(\tau_t) \cdot \text{WFOMC}(\tilde{\Gamma}', \Delta', w, \bar{w}). \end{aligned} \tag{15}$$

Finally, by the definition of conditional probability, we can write

$$\begin{aligned} \mathbb{P}[f(\mu') \mid \tilde{\Gamma}; \Delta, w, \bar{w}] &= \frac{\langle w, \bar{w} \rangle(f(\mu'))}{\text{WFOMC}(\tilde{\Gamma}, \Delta, w, \bar{w})} \\ &= \frac{\langle w, \bar{w} \rangle(\mu') \cdot \langle w, \bar{w} \rangle(\mathcal{A}_t) \cdot \langle w, \bar{w} \rangle(\tau_t)}{\langle w, \bar{w} \rangle(\mathcal{A}_t) \cdot \langle w, \bar{w} \rangle(\tau_t) \cdot \text{WFOMC}(\tilde{\Gamma}', \Delta', w, \bar{w})} \\ &= \frac{\langle w, \bar{w} \rangle(\mu')}{\text{WFOMC}(\tilde{\Gamma}', \Delta', w, \bar{w})} \\ &= \mathbb{P}[\mu' \mid \tilde{\Gamma}'; \Delta', w, \bar{w}], \end{aligned} \tag{16}$$

completing the proof. \square

With the sound reduction presented above, one can readily perform the domain recursion on the more general WFOMS problem. The only remaining task is the sampling of the substructure \mathcal{A}_t at each recursive step, which will be discussed in the following subsection.

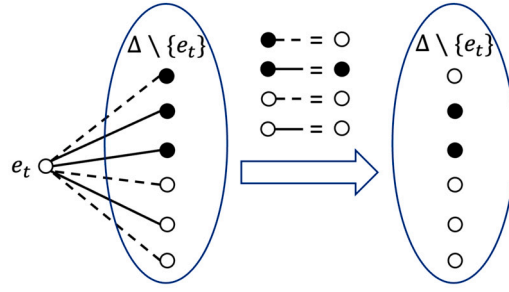


Fig. 2. An example of sampling a substructure. There are two cell types, denoted by \bullet and \circ , and two 2-tables, denoted by $---$ and $---$. The sampled substructure \mathcal{A}_t is represented on the left, whose 2-tables configuration is $(g_{\bullet, ---}^{\mathcal{A}_t}, g_{\bullet, ---}^{\mathcal{A}_t}, g_{\circ, ---}^{\mathcal{A}_t}, g_{\circ, ---}^{\mathcal{A}_t}) = (1, 2, 2, 1)$. With the relaxations of cell types defined above the arrow, where, e.g., $\bullet --- = \circ$ means the relaxed cell type of \bullet under $---$ is \circ , the reduced cell types for each element are presented on the right. One can easily obtain the reduced cell configuration from the 2-table configuration $(g_{\bullet, ---}^{\mathcal{A}_t}, g_{\bullet, ---}^{\mathcal{A}_t}, g_{\circ, ---}^{\mathcal{A}_t}, g_{\circ, ---}^{\mathcal{A}_t})$.

5.2.4. Sampling substructures

For the sake of brevity, we shall limit our focus to the initial recursive step involving sampling \mathcal{A}_t from $\Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i)$. The subsequent steps can be executed by the same process.

We follow a similar approach as in the 1-types sampling, which samples \mathcal{A}_t through random partitions on cells. Let N_c be the total number of cell types, and fix the linear order of cell types as $\eta^1, \eta^2, \dots, \eta^{N_c}$. For all $i \in [N_c]$, denote by $C_{\eta^i} \subseteq \Delta'$ the cells of η^i constructed from the cell types $\eta_1, \eta_2, \dots, \eta_{n-1}$:

$$C_{\eta^i} := \{e_j \mid j \in [n-1] : \eta_j = \eta^i\}.$$

Let N_b be the number of all 2-tables, and fix the linear order of 2-tables as $\pi^1, \pi^2, \dots, \pi^{N_b}$. Recall that \mathcal{A}_t consists of the ground 2-tables of all tuples of e_t and the elements in Δ' . Any substructure \mathcal{A}_t can be represented as a collection of N_c partitions; each partition is applied to a cell C_{η^i} to split it into N_b disjoint subsets; each of these subsets is associated with a specific 2-table π^j , and contains precisely the elements e , whose combination (e_t, e) with e_t has the 2-table π^j .

Consider a substructure \mathcal{A}_t and its corresponding partitions. For any cell C_{η^i} , let $g_{\eta^i, \pi^j}^{\mathcal{A}_t}$ represent the cardinality of the subset in C_{η^i} associated with the 2-table π^j . We can write the configuration of the partition of cell C_{η^i} as $(g_{\eta^i, \pi^j}^{\mathcal{A}_t})_{j \in [N_b]}$, and denote it by the vector $\mathbf{g}_{\eta^i}^{\mathcal{A}_t}$. The 2-tables configuration \mathbf{g} of \mathcal{A}_t is then defined as the concatenation of partition configurations over all cells, i.e., $\mathbf{g}^{\mathcal{A}_t} := \bigoplus_{i \in [N_c]} \mathbf{g}_{\eta^i}^{\mathcal{A}_t}$, where \bigoplus denotes the concatenation operator. Fig. 2 illustrates an example of a substructure and its 2-tables configuration. In the following, we show that the sampling probability of \mathcal{A}_t is entirely determined by its corresponding 2-tables configuration.

To begin, it is clear that the sampling probability of \mathcal{A}_t is proportional to the value of $\text{WFOMC}(\Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i) \wedge \mathcal{A}_t, \Delta, w, \bar{w})$. As stated by (15), we can write it as

$$\text{WFOMC}(\tilde{\Gamma}', \Delta', w, \bar{w}) \cdot \langle w, \bar{w} \rangle(\tau_t) \cdot \langle w, \bar{w} \rangle(\mathcal{A}_t) \quad (17)$$

where $\tilde{\Gamma}' = \Gamma_T \wedge \bigwedge_{i \in [n] \setminus \{t\}} (\eta_i \downarrow \pi_{t,i})(e_i)$ is the reduced sentence given the substructure \mathcal{A}_t . Denote by $\mathbf{w} = (\langle w, \bar{w} \rangle(\pi^i))_{i \in [N_b]}$ the weight vector of 2-tables. We can write (17) as

$$\text{WFOMC}(\tilde{\Gamma}', \Delta', w, \bar{w}) \cdot \langle w, \bar{w} \rangle(\tau_t) \cdot \prod_{i \in [N_c]} \mathbf{w}^{\mathbf{g}_{\eta^i}^{\mathcal{A}_t}}. \quad (18)$$

In the expression above, the 1-types τ_t is fixed, and the last term fully depends on the 2-tables configurations $\mathbf{g}^{\mathcal{A}_t}$. The only part that needs further analysis is the WFOMC problem on $\tilde{\Gamma}'$.

The WFOMC problem on $\tilde{\Gamma}'$ can be viewed as the WFOMC problem on Γ_T conditioned on the cell types $\eta_i \downarrow \pi_{t,i}$. By the symmetry of the weighting function, its value does not depend on the specific cell types $\eta_i \downarrow \pi_{t,i}$ assigned to each element, but instead relies on the overall cell configuration corresponding to the cell types $\eta_i \downarrow \pi_{t,i}$. We denote this cell configuration by a vector $(n_{\eta^i})_{i \in [N_c]}$, where n_{η^i} represents the number of elements whose cell type is η^i under the relaxation of 2-tables in \mathcal{A}_t . By the definition of $g_{\eta^i, \pi^j}^{\mathcal{A}_t}$, which is the number of elements in the cell C_{η^i} that are relaxed by the 2-table π^j , we can write each n_{η^i} as

$$n_{\eta^i} = \sum_{j \in [N_c]} \sum_{k \in [N_b] : \eta^j \downarrow \pi^k = \eta^i} g_{\eta^j, \pi^k}^{\mathcal{A}_t}. \quad (19)$$

For the example in Fig. 2, there are two cell types \bullet and \circ and two block types $---$ and $---$. The relaxed cell type of \circ under both $---$ and $---$ is \circ , while the relaxed cell types of \bullet under $---$ and $---$ are \circ and \bullet , respectively. By (19), we have $n_{\bullet} = g_{\bullet, ---}^{\mathcal{A}_t} + g_{\bullet, ---}^{\mathcal{A}_t}$.

and $n_{\circ} = g_{\bullet, \dots}^{\mathcal{A}_t} + g_{\circ, \dots}^{\mathcal{A}_t} + g_{\circ, \dots}^{\mathcal{A}_t}$. Therefore, we have that the 2-tables configuration fully determines the cell configuration and, consequently, the value of $\text{WFOMC}(\tilde{\Gamma}', \Delta', w, \bar{w})$.

According to the aforementioned reasoning, the sampling probability of a substructure is completely determined by its 2-tables configuration. Thus, we can adopt a similar approach to the one we used to sample 1-types in Section 4.1 and sample \mathcal{A}_t by first sampling a 2-tables configuration, and then randomly partitioning the cells accordingly. The sampling for the 2-tables configuration can be achieved by the enumerative sampling method. For any 2-tables configuration $\mathbf{g}^{\mathcal{A}_t}$, its sampling probability is proportional to the value of (18) multiplied by $\prod_{i \in [N_c]} \binom{|C_{\eta^i}|}{g_{\eta^i}^{\mathcal{A}_t}}$, where $|C_{\eta^i}|$ is the cardinality of C_{η^i} .

Finally, we need to ensure that the sampled substructure \mathcal{A}_t is valid w.r.t. $(\Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i), \Delta)$. It can be easily achieved by imposing some constraints on the 2-tables configuration to be sampled. A 2-table π is called *coherent* with a 1-types tuple (τ, τ') if, for some domain elements a and b , the interpretation of $\tau(a) \cup \pi(a, b) \cup \tau'(b)$ satisfies the formula $\psi(a, b) \wedge \psi(b, a)$. The following two constraints on the 2-tables configuration can make the sampled substructure valid:

- Any 2-table $\pi_{t,i}$ contained in \mathcal{A}_t must be coherent with τ_t and τ_i , the 1-types of e_t and e_i . This translates to a requirement on the 2-tables configuration that when sampling a configuration of partition of a cell C_{η^i} , the cardinality of 2-tables that are not coherent with τ_t and τ_i is restricted to be 0;
- For any Tseitin Z_k in the block type β_t , the substructure \mathcal{A}_t must contain at least one ground atom $R_k(e_t, a)$, where a is a domain element, to make \mathcal{A}_t satisfy the existentially quantified formula $\exists y : R_k(e_t, y)$. This means that there must be at least one 2-table π such that $R_k(x, y) \in \pi$, whose cardinality in some cells is nonzero.

5.2.5. Sampling algorithm

Algorithm 4 TwoTablesSamplerForFO2($\Gamma_T, \Delta, w, \bar{w}, (\eta_i)_{i \in [n]}$).

Input: A sentence Γ_T of the form (10), a domain $\Delta = \{e_i\}_{i \in [n]}$, a weighting (w, \bar{w}) and the cell type $\eta_i = (\tau_i, \beta_i)$ of each domain element e_i

Output: 2-tables $\{\pi_{i,j}\}_{i,j \in [n]}$

```

1: if  $n = 1$  then
2:   return  $\emptyset$ 
3: end if
4: if all block types  $\beta_1, \dots, \beta_n$  are  $\top$  then
5:   return TwoTablesSamplerForUFO2( $\forall x \forall y : \psi(x, y), \Delta, w, \bar{w}, (\tau_i)_{i \in [n]}$ )
6: end if
7: Choose  $t \in [n]$ ;  $\Delta' \leftarrow \Delta \setminus \{e_t\}$ ;  $\mathbf{G} \leftarrow \emptyset$ 
8: Obtain the cells  $C_{\eta^1}, \dots, C_{\eta^{n_t}}$  given by  $\eta_1, \dots, \eta_{n-1}$ 
9: for  $(\mathbf{g}_{\eta^i})_{i \in [N_c]} \leftarrow \bigotimes_{i \in [N_c]} \mathcal{T}_{|C_{\eta^i}|, N_b}$  do
10:   $\mathbf{g} \leftarrow \bigoplus_{i \in [N_c]} \mathbf{g}_{\eta^i}$ 
11:  if  $\text{ExSat}(\mathbf{g}, \eta_t)$  then
12:    Obtain the reduced cell configuration  $(n_{\eta^i})_{i \in [N_c]}$  from  $\mathbf{g}$  by (19)
13:     $p(\mathbf{g}) \leftarrow (18) \cdot \prod_{i \in [N_c]} \binom{n_{\eta^i}}{g_{\eta^i}}$ 
14:     $\mathbf{G} \leftarrow \mathbf{G} \cup \{\mathbf{g}\}$ 
15:  end if
16: end for
17:  $\forall \mathbf{g} \in \mathbf{G}, p(\mathbf{g}) \leftarrow p(\mathbf{g}) / \sum_{\mathbf{g}' \in \mathbf{G}} p(\mathbf{g}')$ 
18: Sample a 2-tables configuration  $\mathbf{g}^*$  from  $\mathbf{G}$  with the probability  $p$ 
19: Remove  $e_t$  from the cell  $C_{\eta_t}$ 
20:  $\mathcal{A} \leftarrow \emptyset$ 
21: for  $i \in [N_c]$  do
22:  Randomly partition the cell  $C_{\eta^i}$  into  $\{G_{\eta^i, \pi^j}\}_{j \in [N_b]}$  according to  $\mathbf{g}_{\eta^i}^*$ 
23:  for  $j \in [N_b]$  do
24:     $\mathcal{A} \leftarrow \mathcal{A} \cup \{\pi^j(e_t, e)\}_{e \in G_{\eta^i, \pi^j}}$ 
25:     $\forall e_s \in G_{\eta^i, \pi^j}, \eta_s' \leftarrow \eta_s \downarrow \pi^j$ 
26:  end for
27: end for
28:  $\mathcal{A} \leftarrow \mathcal{A} \cup \text{TwoTablesSamplerForFO2}(\Gamma_T, \Delta', w, \bar{w}, (\eta_i')_{i \in [n-1]})$ 
29: return  $\mathcal{A}$ 

```

$\triangleright \bigotimes$ is the Cartesian product

By combining all the ingredients discussed above, we now present our sampling algorithm for 2-tables, as shown in Algorithm 4. The overall structure of the algorithm follows a recursive approach, where a recursive call with a smaller domain and relaxed cell types is invoked at Line 31. The algorithm terminates when the input domain contains a single element (at Line 1) or there are no existential constraints on the elements (at Line 4). In the latter case, the algorithm resorts to the 2-tables sampling algorithm for UFO^2 presented in Section 4.2. In Lines 10-22, all possible 2-tables configurations are enumerated. For each configuration, we compute its corresponding weight in Lines 13-14 and decide whether it should be sampled in Lines 15-20, where $\text{Uniform}(0, 1)$ is a random number uniformly distributed in $[0, 1]$. When the 2-tables configuration has been sampled, we randomly partition the cells, and then update the sampled 2-tables and the cell type of each element respectively at Line 27 and 28. The submodule $\text{ExSat}(\mathbf{g}, \eta)$ at

Line 12 is used to check whether the 2-tables configuration \mathbf{g} guarantees the validity of the sampled substructures. Its pseudo-code can be found in Appendix D.2.

Lemma 3. *The complexity of TwoTablesSamplerForFO2 in Algorithm 4 is polynomial in the size of the input domain.*

Proof. The algorithm TwoTablesSamplerForFO2 is invoked at most n times, where n is the size of the domain. The main computation of each recursive call is the for-loop, where we need to iterate over all $\prod_{i \in [N_c]} |\mathcal{T}_{n_{\eta^i}, N_b}|$ possible configurations. Recall that the size of the configuration space $\mathcal{T}_{M,m}$ is polynomial in M . Thus the number of iterations executed by the loop is also polynomial in the domain size. The other complexity arises from the WFOMC problems on W and (18). These problems can be viewed as conditional WFOMC on a set of unary facts, whose size is clearly polynomial in the domain size. Therefore, according to Proposition 1 and the liftability of \mathbf{FO}^2 , the aforementioned counting problems can be solved in polynomial time in the domain size. As a result, the complexity of the entire algorithm is polynomial in the size of the input domain. \square

Remark 2. We note that there are several optimizations to TwoTablesSamplerForFO2 in our implementation, e.g., heuristically choosing the domain element for recursion so that the algorithm can quickly reach the terminal condition. However, the algorithm as described here is easier to understand and efficient enough to prove our main result, so we leave the discussion of some of the optimizations to Appendix D.1.

Algorithm 5 WeightedModelSamplerForFO2($\Gamma, \Delta, w, \tilde{w}$).

Input: A SNF sentence Γ , a domain Δ of size n and a weighting (w, \tilde{w})

Output: A model μ of Γ sampled according to the probability $\mathbb{P}[\mu | \Gamma]$

- 1: $\{\tau_i(e_i)\}_{i \in [n]} \leftarrow \text{OneTypesSampler}(\Gamma, \Delta, w, \tilde{w})$
 - 2: Transform Γ into Γ_T by (10)
 - 3: Set the block and cell types $\beta_i(x)$ and $\eta_i(x)$ for each e_i as $\bigwedge_{k \in [m]} Z_k(x)$ and (β_i, τ_i) respectively
 - 4: $\{\pi_{i,j}(e_i, e_j)\}_{i,j \in [n]} \leftarrow \text{TwoTablesSamplerForFO2}(\Gamma_T, \Delta, w, \tilde{w}, (\eta_i)_{i \in [n]})$
 - 5: **return** $\{\tau_i(e_i)\}_{i \in [n]} \cup \{\pi_{i,j}(e_i, e_j)\}_{i,j \in [n]}$
-

With the proposed TwoTablesSamplerForFO2, we can now prove Theorem 2.

Proof of Theorem 2. By Lemma 1, it is sufficient to demonstrate that all SNF sentences are sampling liftable. To achieve this, we construct a lifted sampler for SNF sentences as shown in Algorithm 5. The sampler consists of two stages, one for sampling 1-types and the other for 2-tables, in a manner similar to the sampler for \mathbf{UFO}^2 . The sampling of 1-types can be realized using the 1-types sampler for \mathbf{UFO}^2 presented in Section 4.1. Due to the domain-liftability of \mathbf{FO}^2 , this approach retains its polynomial complexity w.r.t. the domain size. The sampling of 2-tables is handled by TwoTablesSamplerForFO2, whose complexity has been proved polynomial in the domain size, according to Lemma 3. Therefore, the entire sampling algorithm runs in time polynomial in the domain size, and hence is domain-lifted. \square

6. \mathbf{C}^2 is sampling liftable

In this section, we extend the sampling liftability of \mathbf{FO}^2 to \mathbf{C}^2 , the 2-variables fragment of first-order logic with counting quantifiers $\exists_{=k}$, $\exists_{\leq k}$ and $\exists_{\geq k}$. These counting quantifiers are defined as follows. Let \mathcal{A} be a structure defined on a domain Δ . Then the sentence $\exists_{=k} x : \varphi(x)$ is true in \mathcal{A} if there are exactly k distinct elements $t_1, \dots, t_k \in \Delta$ such that $\mathcal{A} \models \varphi(t_i)$. For example, the sentence

$$\forall x \forall y : (E(x, y) \Rightarrow E(y, x)) \wedge \forall x : \neg E(x, x) \wedge \forall x \exists_{=2} y : E(x, y)$$

encodes 2-regular graphs, i.e., graphs where each vertex has exactly two neighbors. The other two counting quantifiers can be defined: $(\exists_{\leq k} x : \varphi(x)) \Leftrightarrow (\forall x : \neg \varphi(x)) \vee \bigvee_{i \in [k]} (\exists_{=i} x : \varphi(x))$ and $(\exists_{\geq k} x : \varphi(x)) \Leftrightarrow \neg (\exists_{\leq k-1} x : \varphi(x))$. For ease of presentation, we allow the counting parameter $k = 0$, and define the quantifier $\exists_{=0}$ by $(\exists_{=0} x : \varphi(x)) \Leftrightarrow (\forall x : \neg \varphi(x))$. Note that the existential quantifiers \exists can be always written as $\exists_{\geq 1}$, and thus we omit \exists in the following discussion and assume that \mathbf{C}^2 is obtained by adding the counting quantifiers to \mathbf{UFO}^2 . The notation of $[\cdot]$ is extended such that $[i, j]$ denotes the set of integers $\{i, i+1, \dots, j\}$.

The sampling liftability and the lifted sampling algorithm for \mathbf{C}^2 are built upon the framework for \mathbf{FO}^2 as in Section 5, which includes the following components.

Normal form. We first introduce the following sentence as a normal form of \mathbf{C}^2 :

$$\Gamma_{\forall} \wedge \bigwedge_{i \in [m]} (\forall x : A_i(x) \Leftrightarrow (\exists_{=k_i} y : R_i(x, y))), \quad (20)$$

where Γ_{\forall} is a \mathbf{UFO}^2 sentence, each k_i is a non-negative integer, $R_k(x, y)$ is an atomic formula, and A_k is a unary predicate. Any \mathbf{C}^2 sentence can be converted into this normal form and the corresponding reduction is sound. This process involves converting all $\exists_{\leq k}$

and $\exists_{\geq k}$ quantifiers into $\exists_{=k}$ quantifiers according to the definition of counting quantifiers.⁶ Then a similar approach for converting \mathbf{FO}^2 sentence into SNF in Section 5 is used to substitute each $\exists_{=k}y : R_i(x, y)$ with an auxiliary atom $A_i(x)$. The details can be found in Appendix B.2.

Lemma 4. *For any WFOMS problem $\mathfrak{S} = (\Gamma, \Delta, w, \bar{w})$ where Γ is a \mathbf{C}^2 sentence, there exists a WFOMS problem $\mathfrak{S}' = (\Gamma', \Delta, w', \bar{w}')$ where Γ' is of the form (20) and $\max_{i \in [m]} k_i \leq |\Delta|$, such that the reduction from \mathfrak{S} to \mathfrak{S}' is sound.*

Sampling 1-types. The 1-type for each element e_i can be sampled by the same approach as in Section 5. Let $\tau_1, \tau_2, \dots, \tau_n$ be the sampled 1-types. The predicates A_i are contained in these 1-types, and thus are determined after the 1-types sampling.

A more general WFOMS problem for 2-tables sampling. Similar to the case of \mathbf{FO}^2 , we need to transform the 2-tables sampling problem into a more general form to apply the domain recursion scheme. For each formula $\exists_{=k_j}y : R_j(x, y)$, we introduce $2(k_j + 1)$ new unary predicates $Z_{j,0}^\exists, Z_{j,1}^\exists, \dots, Z_{j,k_j}^\exists, Z_{j,0}^\sharp, Z_{j,1}^\sharp, \dots, Z_{j,k_j}^\sharp$, and append the conjunction

$$\forall x : \left(Z_{j,q}^\exists(x) \Leftrightarrow (\exists_{=q}y : R_j(x, y)) \right) \wedge \left(Z_{j,q}^\sharp(x) \Leftrightarrow \neg(\exists_{=q}y : R_j(x, y)) \right) \quad (21)$$

over $q \in [0, k_j]$ to Γ_\forall , resulting in a new sentence Γ_T . Let

$$\mathcal{Z}_j := \{ Z_{j,0}^\exists(x), \dots, Z_{j,k_j}^\exists(x), Z_{j,0}^\sharp(x), \dots, Z_{j,k_j}^\sharp(x), \top \},$$

and

$$\mathcal{Z} := \{ \{ Z_1(x), Z_2(x), \dots, Z_m(x) \} \mid Z_1(x) \in \mathcal{Z}_1, \dots, Z_m(x) \in \mathcal{Z}_m \}. \quad (22)$$

The more general WFOMS problem is then defined on the sentence

$$\Gamma_T \wedge \bigwedge_{i \in [n]} \tau_i(e_i) \wedge \bigwedge_{i \in [n]} v_i(e_i), \quad (23)$$

where each $v_i(x)$ is a conjunction over a set of atomic formulas in \mathcal{Z} :

$$v_i(x) \in \left\{ \bigwedge_{j \in [m]} Z_j(x) \mid \{ Z_1(x), Z_2(x), \dots, Z_m(x) \} \in \mathcal{Z} \right\}.$$

Let $w(Z_{j,q}^*) = \bar{w}(Z_{j,q}^*) = 1$ for all $j \in [m], q \in [0, k_j]$ and $* \in \{\exists, \sharp\}$. It is easy to check that the 2-tables sampling problem is reducible to the more general WFOMS problem, and the reduction is sound.

Example 7. Consider the WFOMS problem on the following sentence:

$$\begin{aligned} & \forall x \forall y : (E(x, y) \Rightarrow E(y, x)) \wedge \forall x : \neg E(x, x) \wedge \\ & \forall x : Red(x) \Leftrightarrow (\exists_{=2}y : E(x, y)). \end{aligned}$$

It encodes the colored graphs where a vertex is colored red if and only if it has exactly two neighbors. Suppose the domain is $\{v_1, v_2, v_3, v_4\}$, and the sampled 1-types for each element are $\tau_1 = \{Red(x)\}$, $\tau_2 = \{\neg Red(x)\}$, $\tau_3 = \{\neg Red(x)\}$, $\tau_4 = \{Red(x)\}$. The transformed sentence for the more general WFOMS problem is

$$\begin{aligned} & \forall x \forall y : E(x, y) \Rightarrow E(y, x) \wedge \forall x : \neg E(x, x) \wedge \\ & Red(v_1) \wedge \neg Red(v_2) \wedge Red(v_3) \wedge Red(v_4) \wedge \\ & \forall x : (Z_{1,0}^\exists(x) \Leftrightarrow (\exists_{=0}y : E(x, y))) \wedge \forall x : (Z_{1,0}^\sharp(x) \Leftrightarrow \neg(\exists_{=0}y : E(x, y))) \wedge \\ & \forall x : (Z_{1,1}^\exists(x) \Leftrightarrow (\exists_{=1}y : E(x, y))) \wedge \forall x : (Z_{1,1}^\sharp(x) \Leftrightarrow \neg(\exists_{=1}y : E(x, y))) \wedge \\ & \forall x : (Z_{1,2}^\exists(x) \Leftrightarrow (\exists_{=2}y : E(x, y))) \wedge \forall x : (Z_{1,2}^\sharp(x) \Leftrightarrow \neg(\exists_{=2}y : E(x, y))) \wedge \\ & Z_{1,2}^\exists(v_1) \wedge Z_{1,2}^\sharp(v_2) \wedge Z_{1,2}^\sharp(v_3) \wedge Z_{1,2}^\exists(v_4). \end{aligned}$$

The block types for each element are $v_1(x) = \{Z_{1,2}^\exists(x)\}$, $v_2(x) = \{Z_{1,2}^\sharp(x)\}$, $v_3(x) = \{Z_{1,2}^\sharp(x)\}$, $v_4(x) = \{Z_{1,2}^\exists(x)\}$.

Domain recursion. The domain recursion scheme is still applicable to the WFOMS problem on the sentence (23), where we view the sets in $\mathcal{Z}(x)$ as “block types”, and $v_i(x)$ as the block type of e_i . When the substructure \mathcal{A}_i of an element e_i has been sampled, the

⁶ We stress that the formula $\exists_{\geq k}x : R(x, y)$ is converted to $\neg(\bigvee_{i \in [0, k-1]} \exists_{=i}x : R(x, y))$ rather than $\bigvee_{i \in [k, n]} \exists_{=i}x : R(x, y)$, since the latter one depends on the domain size n .

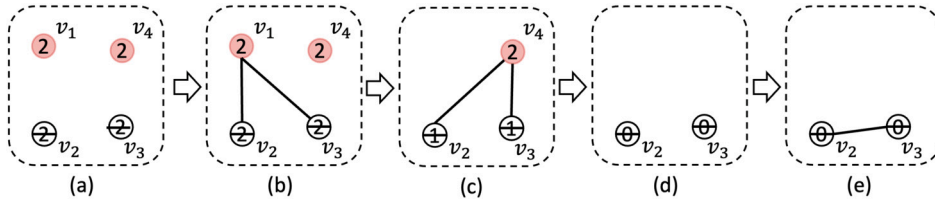


Fig. 3. An illustration of domain recursion for the WFOMS problem in Example 7. The vertices v_1 and v_4 are colored red, while the vertices v_2 and v_3 are not colored, according to the sampled 1-types. The number shown in each vertex corresponds to its block type, denoting the number of neighbors that the vertex should or should not have, e.g., v_1 should have exactly two neighbors, while v_2 should not have two neighbors. In each domain recursion step, the block types of the selected element are relaxed according to the sampled edges. Note that in the final step, the edge between v_1 and v_4 is always sampled, as the block type of v_1 and v_3 requires them to have at least one neighbor.

block types of the remaining elements are relaxed by the 2-tables in \mathcal{A}_t . This leads to a new WFOMS problem on the sentence in the same form of (23), but with a smaller domain. We can show that this reduction to the new WFOMS problem is sound. The argument is similar to what we have done for FO^2 , and is deferred to Appendix F. An example of this reduction is illustrated in Fig. 3.

Lemma 5. Define the relaxed block type $v \downarrow \pi$ of a block type v under a 2-table π as a set in \mathcal{Z} such that for each $Z_{j,q}^* \in v$,

- if $* = \#$, $q = 0$ and $R_j(y, x) \in \pi$, then $\top \in v \downarrow \pi$;
- otherwise, if $R_j(y, x) \in \pi$, then⁷ $Z_{j,q-1}^* \in v \downarrow \pi$, and if $R_j(y, x) \notin \pi$, then $Z_{j,q}^* \in v \downarrow \pi$.

For any substructure \mathcal{A}_t defined in (12), let

$$\tilde{\Gamma} := \Gamma_T \wedge \bigwedge_{i \in [n]} \tau_i(e_i) \wedge \bigwedge_{i \in [n]} v_i(e_i) \wedge \mathcal{A}_t, \quad (24)$$

and

$$\tilde{\Gamma}' := \Gamma_T \wedge \bigwedge_{i \in [n] \setminus \{t\}} \tau_i(e_i) \wedge \bigwedge_{i \in [n] \setminus \{t\}} (v_i \downarrow \pi_{t,i})(e_i). \quad (25)$$

If $\tilde{\Gamma}$ is satisfiable, the reduction from the WFOMS problem of $(\tilde{\Gamma}, \Delta, w, \bar{w})$ to $(\tilde{\Gamma}', \Delta \setminus \{e_t\}, w, \bar{w})$ is sound.

Sampling algorithm. By the domain recursion scheme, it is easy to devise a recursive algorithm for sampling 2-tables in a manner similar to Algorithm 4. In fact, the procedure in Algorithm 4 remains the same except for the following modifications:

- **M1:** All block types in the algorithm, including those in the cell types, are changed with $v(x)$,
- **M2:** The termination condition on block types becomes that all block types are $\{\top\}$.
- **M3:** The subroutine ExSat now includes an additional check for the satisfaction of block types. Specifically, the sampled 2-tables concerning the selected element must satisfy the block type of the element, i.e., for any $Z_{j,q}^\exists(x)$ (resp. $Z_{j,q}^\#(x)$) in $v_t(x)$, there must exist exactly q (resp. must not exist q) elements $e \in \Delta \setminus \{e_t\}$ such that $R_j(e_t, e) \in \pi$.

Remark 3. As mentioned at the beginning of the section, the existential quantifiers \exists can be replaced by $\exists_{\geq 1}$. Then the auxiliary predicates $Z_{j,0}^\#$ are exactly the Tseitin predicates Z_j introduced in (10). The block types $v(x)$, which can only contain \top and $Z_{j,0}^\#(x)$, degenerates to the ones we defined in Section 5.2.2 for FO^2 . Lemma 5 and the sampling algorithm above is equivalent to Lemma 2 and Algorithm 4 respectively.

Theorem 3. The C^2 fragment is domain-liftable under sampling.

Proof. By Lemma 4, it is sufficient to demonstrate that all C^2 sentences in the normal form (20) are sampling liftable. We prove it by showing that the sampling algorithm presented above is lifted. The sampling of 1-types is clearly polynomial-time in the domain size by the same argument as in the proof of Theorem 2 and the liftability of C^2 [28]. For sampling 2-tables, we show that the modifications **M1**, **M2** and **M3** on Algorithm 4 do not change the polynomial-time complexity in the domain size. The complexity is not affected by **M2**, since the number of domain recursion steps is still the domain size. Furthermore, the modification **M3** only introduces a minimal overhead to each recursion step, which is not dependent on the domain size, and thus does not change the polynomial-time complexity. For **M1**, we need some additional arguments. First, by the definition of block types, the number of all possible block types is $\prod_{j \in [m]} 2(k_j + 2)$, which is independent of the domain size. Therefore, the complexity of the main loop

⁷ The corner case where $q = 0$ and $R_j(y, x) \in \pi$, resulting in negative counting parameters, cannot occur during the domain recursion, and thus we ignore it here.

in Algorithm 4 is still polynomial-time in the domain size. With **M1**, the algorithm now needs to solve the WFOMC problems on sentences in the form of (23). These counting problems can be again viewed as conditional WFOMC with unary evidence, whose complexity is clearly polynomial in the domain size, following from Proposition 1 and the liftability of \mathbf{C}^2 . As a result, the complexity of the modified algorithm is still polynomial in the domain size, and hence the \mathbf{C}^2 fragment is domain-liftable under sampling. \square

7. Sampling liftability with cardinality constraints

In this section, we extend our result to the case containing the *cardinality constraints*. A single cardinality constraint is a statement of the form $|P| \star q$, where \star is a comparison operator (e.g., $=, \leq, \geq, <, >$) and q is a natural number. These constraints are imposed on the number of distinct positive ground literals in a structure \mathcal{A} formed by the predicate P . For example, a structure \mathcal{A} satisfies the constraint $|P| \leq q$ if there are at most q literals for P that are true in \mathcal{A} . For illustration, we allow cardinality constraints as atomic formulas in the FO formulas, e.g., $(|E| = 2) \wedge (\forall x \forall y : E(x, y) \Rightarrow E(y, x))$ (its models can be interpreted as undirected graphs with exactly one edge) and the satisfaction relation \models is extended naturally. Note the difference between cardinality constraints and counting quantifiers: the former is a statement about the number of positive ground literals for a given predicate that are true in a structure, while the latter is a statement about the number of elements in a structure that satisfy certain properties. Another important difference is in the data complexity that we will show: the counting parameter k_i of counting quantifiers is regarded as a part of the sentence and is fixed when considering the data complexity, while the cardinality constraints allow the parameter q to be a part of the input instance, while still guaranteeing polynomial runtime.

7.1. \mathbf{C}^2 with cardinality constraints

We first establish the domain-liftability under sampling for the fragment \mathbf{C}^2 augmented with cardinality constraints. Since \mathbf{C}^2 is a superset of \mathbf{FO}^2 and \mathbf{UFO}^2 , this result also implies the domain-liftability under sampling of \mathbf{FO}^2 and \mathbf{UFO}^2 with cardinality constraints. Let Γ be a \mathbf{C}^2 sentence and

$$Y := \varphi(|P_1| \star q_1, \dots, |P_M| \star q_M), \quad (26)$$

where φ is a Boolean formula, $\{P_i\}_{i \in [M]} \subseteq \mathcal{P}_\Gamma$, and $\forall i \in [M], q_i \in \mathbb{N}$. Let us consider the WFOMS problem on $\Gamma \wedge Y$ over the domain Δ under (w, \bar{w}) .

The sampling algorithm for $\Gamma \wedge Y$ keeps the same structure as those for \mathbf{UFO}^2 , \mathbf{FO}^2 and \mathbf{C}^2 , containing two successive sampling routines for 1-types and 2-tables respectively. As usual, we only focus on the sampling of 2-tables in the following, since the process for sampling 1-types is identical to those for \mathbf{UFO}^2 , \mathbf{FO}^2 , and \mathbf{C}^2 .

We first show that the domain recursive property still holds in the sampling problem of 2-tables. Given a set L of ground literals and a predicate P , we define $N(P, L)$ as the number of positive ground literals for P in L . Given a substructure \mathcal{A}_i of the element e_i , denote the 1-type of e_i by τ_i as usual, let $q'_i = q_i - N(P_i, \mathcal{A}_i) - N(P_i, \tau_i(e_i))$ for every $i \in [M]$, and define

$$Y' = \varphi(|P_1| \star q'_1, \dots, |P_M| \star q'_M). \quad (27)$$

Let $\tilde{\Gamma}_C = \tilde{\Gamma} \wedge Y$ and $\tilde{\Gamma}'_C = \tilde{\Gamma}' \wedge Y'$, where $\tilde{\Gamma}$ and $\tilde{\Gamma}'$ are the original and reduced sentences defined as (24) and (25) respectively. Then the reduction from the WFOMS problem on $\tilde{\Gamma}_C$ to $\tilde{\Gamma}'_C$ is sound. The proof follows the same argument for Lemma 2 and Lemma 5, and is deferred to Appendix F.

Lemma 6. *The reduction from the WFOMS problem of $(\tilde{\Gamma}_C, \Delta, w, \bar{w})$ to $(\tilde{\Gamma}'_C, \Delta', w, \bar{w}')$ is sound.*

By Lemma 6, we develop a recursive sampling algorithm for 2-tables in Algorithm 6. This algorithm is derived from Algorithm 4 with the redundant lines not shown in the pseudocode. The differences from the original algorithm are:

- **M1:** All WFOMC problems now contain the cardinality constraints, e.g., Y in Line 4 and Y' in Line 10,
- **M2:** The terminal condition, which previously checked the block types, is removed,⁸ and
- **M3:** The validity check for the sampled 2-tables configuration in ExSat now includes an additional check for the well-definedness of the reduced cardinality constraints Y' , returning False if there is any $q'_i \notin \mathbb{N}$ for $i \in [M]$.

Theorem 4. *Let Γ be a \mathbf{C}^2 sentence and Y be of the form (26). Then $\Gamma \wedge Y$ is domain-liftable under sampling.*

Proof. We prove the sampling liftability of \mathbf{C}^2 with cardinality constraints by showing the sampler presented above is lifted. The complexity of sampling 1-types is polynomial in the domain size by the same argument as in the proof of Theorem 1 and the

⁸ We can keep this terminal condition and invoke a more efficient sampler for \mathbf{UFO}^2 with cardinality constraints, e.g., the one from our previous work [57]. However, removing the condition does not change the polynomial-time complexity of the algorithm.

Algorithm 6 TwoTablesSamplerForCC($\Gamma_T, Y, \Delta, w, \bar{w}, (\eta_i)_{i \in [n]}$).

```

1: ...
2: for  $(\mathbf{g}_{\eta'})_{i \in [N_c]} \leftarrow \text{Prod}(\mathcal{T}_{\eta_{a_1}, N_b}, \dots, \mathcal{T}_{\eta_{a_{N_c}}, N_b})$  do
3:   if ExSat( $\mathbf{g}_{\eta'}, Y$ ) then
4:     Obtain the reduced cell configuration  $(\eta_{\eta'})_{i \in [N_c]}$  from  $\mathbf{g}$  by (19)
5:     Get the new cardinality constraints  $Y'$  w.r.t.  $\mathbf{g}$  by (27)
6:      $p(\mathbf{g}) \leftarrow \text{WFOMC}(\bar{\Gamma}' \wedge Y', w, \bar{w}, \Delta) \cdot \langle w, \bar{w} \rangle(\tau_i) \cdot \prod_{i \in [N_c]} \binom{\eta_{\eta'}}{\mathbf{g}_{\eta'}} \cdot \mathbf{w}^{\mathbf{g}_{\eta'}}$ 
7:      $\mathbf{G} \leftarrow \mathbf{G} \cup \{\mathbf{g}\}$ 
8:   end if
9: end for
10: ...
11: Obtain the reduced cardinality constraints  $Y'$  w.r.t.  $\mathbf{g}^*$  by (27)
12:  $\mathcal{A} \leftarrow \mathcal{A} \cup \text{TwoTablesSamplerForCC}(\Gamma_T, Y', \Delta', w, \bar{w}, (\eta'_i)_{i \in [n-1]})$ 
13: return  $\mathcal{A}$ 

```

liftability of \mathbf{C}^2 with cardinality constraints [28]. Next, we show that the modifications **M1**, **M2** and **M3** to Algorithm 4 do not affect the polynomial-time complexity of the algorithm. First, it can be observed that **M3** has a negligible impact on the algorithm's complexity, and **M2** does not affect the polynomial-time complexity of the algorithm. Then, by the liftability of \mathbf{C}^2 with cardinality constraints and Proposition 1, we have that the new WFOMC problems with additional cardinality constraints are still liftable. Therefore, the entire complexity of Algorithm 6 remains polynomial in the domain size, and thus the sampler combining the 1-types sampling and Algorithm 6 is lifted, which completes the proof. \square

7.2. A more efficient sampler for \mathbf{SC}^2

With the lifted sampler for \mathbf{FO}^2 with cardinality constraints, we can provide a practically more efficient sampler for some subfragment of \mathbf{C}^2 .⁹ Specifically, we focus on the fragment \mathbf{SC}^2 , which consists of sentences of the form

$$\Gamma_{\forall} \wedge (\forall x \exists_{=k_1} y : \phi_1(x, y)) \wedge \dots \wedge (\forall x \exists_{=k_m} y : \phi_{m'}(x, y)) \\ \wedge (\exists_{=k'_1} x \forall y : \phi'_1(x, y)) \wedge \dots \wedge (\exists_{=k'_{m'}} x \forall y : \phi'_{m'}(x, y)),$$

where Γ_{\forall} is a \mathbf{UFO}^2 sentence. We call this fragment *two-variable logic with counting in SNF* (\mathbf{SC}^2), as the conjunction apart from Γ_{\forall} resembles SNFs. The sentence (2) for encoding k-regular graphs is an example of \mathbf{SC}^2 sentences.

The more efficient sampler for \mathbf{SC}^2 draws inspiration from the work conducted by Kuzelka [28]. The primary findings of their study were partially obtained through a reduction from the WFOMC problem on \mathbf{SC}^2 sentences to \mathbf{UFO}^2 sentences with cardinality constraints. We demonstrate that this reduction can be also applied to the sampling problem and it is sound. The proof follows a similar technique used in [28], and the details are deferred to Appendix C.

Lemma 7. *For any WFOMS problem $\mathfrak{S} = (\Phi, \Delta, w, \bar{w})$ where Φ is a \mathbf{SC}^2 sentence, there exists a WFOMS problem $\mathfrak{S}' = (\Gamma' \wedge Y, \Delta, w', \bar{w}')$, where Γ' is a \mathbf{UFO}^2 sentence, Y denotes cardinality constraints of the form (26), such that the reduction from \mathfrak{S} to \mathfrak{S}' is sound.*

Using the Lemma above and the lifted sampler in Algorithm 6 for \mathbf{UFO}^2 with cardinality constraints, it is easy to devise a lifted sampler for \mathbf{SC}^2 without involving the counting quantifiers. It is known that the counting algorithm for \mathbf{C}^2 sentences usually needs more complicated and sophisticated techniques than the one for \mathbf{UFO}^2 sentences with cardinality constraints [28]. Thus, the new sampler for \mathbf{SC}^2 based on Lemma 7 and Algorithm 6 is more efficient and easier to implement than the one based on the sampling algorithm presented in Section 6. We also note that further generalizing this technique to the \mathbf{C}^2 sentences is infeasible, since the reduction from \mathbf{C}^2 to \mathbf{SC}^2 for WFOMC used in [28] introduced some negative weights on predicates, which would make the corresponding sampling problem ill-defined. As a result, for WFOMS problems on general \mathbf{C}^2 sentences, one has to resort to the lifted sampling algorithm presented in Section 6.

8. Complexity discussion

In this section, we delve further into the exact complexity of our sampling algorithms, and provide some insights into the factors that affect the complexity of the algorithms. Given that our sampling algorithms heavily depend on the WFOMC solver, which is the primary bottleneck in the algorithms' execution, we will primarily focus on evaluating complexity in terms of the number of calls made to the WFOMC solver. The domain size is denoted by n .

For the sampling of 1-types, all the algorithms presented in this paper employ the same technique, i.e., enumerating all possible partition configurations over the domain and sampling from them according to their weights. Computing the weight of every partition configuration requires a call to the WFOMC solver (Line 2 in Algorithm 2). There are at most $\mathcal{T}_{n, |U_{\Gamma}|}$ partition configurations, where

⁹ Although the sampler for \mathbf{C}^2 proposed in Section 6 has been proved to be lifted, which means its complexity is polynomial in the domain size, the exponents of the polynomials are usually very large.

Table 2
The complexity of the sampling algorithms for different fragments.

| Fragments (with constraints) | #WFOMC |
|------------------------------|--|
| UFO ² | $O(n^u)$ |
| FO ² | $O(n^{ub2^m+1})$ |
| C ² | $O(n^{ub2^{m'}(k+1)^{m'}+1})$ |
| Cardinality constraints | The same as the corresponding fragment |

U_{P_Γ} is the set of 1-types in the sentence. Observe that not all 1-types have to always be valid in the sentence, and the number of valid 1-types is usually much smaller than $|U_{P_\Gamma}|$ in practice. For instance, in the sentence (1) of 2-colored graphs, the number of valid 1-types is 2, i.e., red and black nodes, while the total number of 1-types is $2^3 = 8$. Instead of enumerating configurations on all 1-types, the algorithm can enumerate configurations just on the valid 1-types. Let u be the number of valid 1-types in the sentence, then the enumeration complexity is bounded by $O(n^u)$.

For the sampling of 2-tables, the complexity of the algorithms varies depending on the fragment to which the sentence belongs.

- For **UFO**², the sampling of 2-tables is decomposed into independent sampling problems on pairs of elements, where no WFOMC call is needed, and the complexity of each problem is constant in the domain size.
- For **FO**², there are at most n recursion steps. In each step, the algorithm needs to solve one WFOMC problem for each reduced cell configuration. The number of reduced cell configurations is bounded by $\prod_{i \in [N_c]} \mathcal{T}_{|C_{\eta^i}|, N_b}$, where N_c and N_b is the number of cell types and 2-tables respectively, and C_{η^i} is the set of elements whose cell types are η^i . Similar to the analysis for 1-types, we consider the number of valid cells and valid 2-tables. Let m be the number of existentially quantified formulas in the SNF sentence. The number of valid cells is bounded by $u2^m$. For 2-tables, we consider the maximum number of 2-tables coherent with any 1-types tuple and denote it by b . With these notations, the size of $C_{\eta^i} \leq n$, $N_c \leq u2^m$, $N_b \leq b$, and thus the number $\prod_{i \in [N_c]} \mathcal{T}_{|C_{\eta^i}|, N_b}$ is in $O((n^b)^{u2^m}) = O(n^{ub2^m})$. The total number of WFOMC calls for sampling 2-tables is in $O(n^{ub2^m+1})$.
- For **C**², the complexity is similar to that of **FO**², except that more block types are considered. Suppose there are m' counting quantifiers in the normal form (20) of **C**², and the counting parameters are bounded by k . The notations of u and b are the same as those in **FO**². Then the number of block types is bounded by $(2(k+1))^{m'}$, and the number of cell types is bounded by $u(2(k+1))^{m'}$. The total number of WFOMC calls for sampling 2-tables is in $O(n \cdot (n^b)^{u(2(k+1))^{m'}}) = O(n^{ub2^{m'}(k+1)^{m'}+1})$.
- For fragments with cardinality constraints, the cardinality constraints do not introduce additional WFOMC calls (see Algorithm 6), and thus the complexity is the same as the complexity for the corresponding fragment without cardinality constraints.

The overall complexity of the sampling algorithms for different fragments is summarized in Table 2. It is important to note that the complexity provided is measured in terms of the number of WFOMC calls, and the actual runtime of the algorithms is influenced by the performance of the WFOMC solver.

9. Experimental results

To assess the efficacy of our sampling algorithms, we conducted a series of experiments aimed at evaluating their performance. We also conducted statistical tests to check the correctness of our implementation. All algorithms were implemented in Python¹⁰ and the experiments were performed on a computer with an 8-core Intel i7 3.60GHz processor and 32 GB of RAM.

Many sampling problems can be expressed as WFOMS problems. Here we consider two typical ones:

- **Sampling combinatorial structures:** The uniform generation of some combinatorial structures can be directly reduced to a WFOMS problem, e.g., the uniform generation of *graphs with no isolated vertices* and *k-regular graphs* in Section 5.1 and the introduction. We added four more combinatorial sampling problems to these two for evaluation: *functions*, *functions w/o fix-points* (i.e., the functions f satisfying $f(x) \neq x$), *permutations* and *permutations w/o fix-points*. The details of these problems are described in Appendix E.1.
- **Sampling from MLNs:** Our algorithms can be also applied to sample possible worlds from MLNs. An MLN defines a distribution over structures (i.e., possible worlds in SRL literature), and its respective sampling problem is to randomly generate possible worlds according to this distribution. There is a standard reduction from the sampling problem of an MLN to a WFOMS problem (see Appendix E.1 and also [57]). We use three MLNs in our experiments:
 - A variant of the classic *friends-smokers* MLN with the constraint that every person has at least one friend:

$$\begin{aligned} & \{(+\infty, \forall x : \neg fr(x, x)), \\ & (+\infty, \forall x \forall y : fr(x, y) \Rightarrow fr(y, x)), \end{aligned}$$

¹⁰ The code can be found in https://github.com/lucienwang1009/lifted_sampling_fo2.

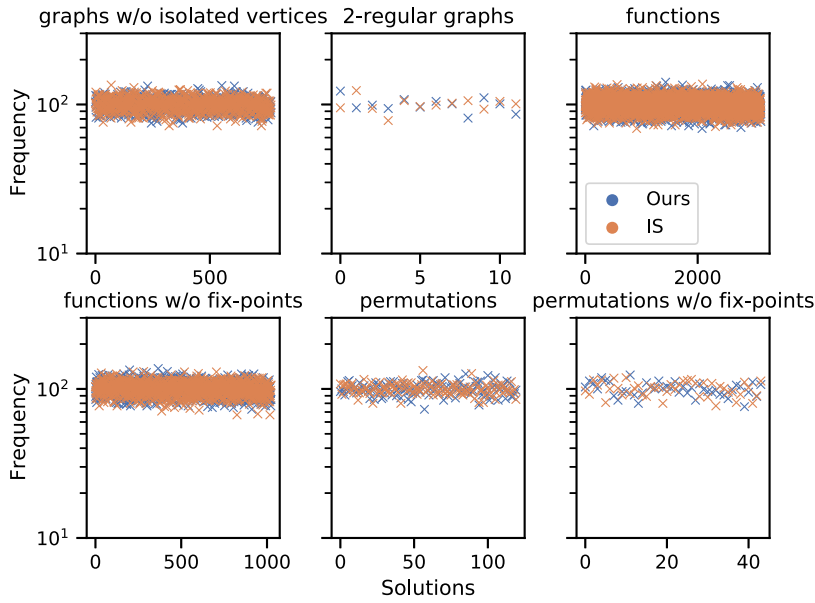


Fig. 4. Uniformity comparison between an ideal sampler (IS) and our weighted model sampler.

$$(+\infty, \forall x \exists y : fr(x, y))$$

$$(0, sm(x)),$$

$$(0.2, fr(x, y) \wedge sm(x) \Rightarrow sm(y))\}.$$

- The *employment* MLN used in [54]:

$$\{(1.3, \exists y : workfor(x, y) \vee boss(x))\},$$

which states that with high probability, every person either is employed by a boss or is a boss.

- The *deskmate* MLN:

$$\{(+\infty, \forall x : \neg mate(x, x) \wedge \neg fr(x, x)),$$

$$(+\infty, \forall x \forall y : mate(x, y) \Rightarrow mate(y, x)),$$

$$(+\infty, \forall x \exists_{=1} y : mate(x, y)),$$

$$(+\infty, \forall y \exists_{=1} x : mate(x, y)),$$

$$(1.0, mate(x, y) \Rightarrow fr(x, y))\},$$

which states that every student has exactly one deskmate, and if two students are deskmates, then they are probably friends. The details about the reduction from sampling from MLNs to WFOMS problems and the resulting WFOMS problems of these two MLNs can be found in Appendix E.1.

9.1. Correctness of the implementation

We first performed a statistical test on our implementation by focusing on the uniform generation of combinatorial structures within small domains, where exact sampling is feasible via enumeration-based techniques; we choose the domain size of 5 for evaluation. To serve as a benchmark, we have implemented a simple ideal uniform sampler, denoted by IS, by enumerating all the models and then drawing samples uniformly from these models; this is also why we use such a small domain consisting only of five elements in this experiment. For each combinatorial structure encoded into an \mathbf{FO}^2 sentence Γ , a total of $100 \times |\mathcal{M}_{\Gamma, \Delta}|$ models were generated from both IS and our weighted model sampler. Fig. 4 depicts the model distribution produced by these two algorithms—the horizontal axis represents models numbered lexicographically, while the vertical axis represents the generated frequencies of models. The figure suggests that the distribution generated by our weighted model sampler is indistinguishable from that of IS. Furthermore, a statistical test on the distributions produced by the weighted model sampler was performed, and no statistically significant difference from the uniform distribution was found. The details of this test can be found in Appendix E.2.

For sampling problems from MLNs, enumerating all the models is infeasible even for a domain of size 5, e.g., there are $2^{5^2+5} = 2^{30}$ models in the employment MLN. That is why we test the *count distribution* of predicates from the problems. Instead of specifying the

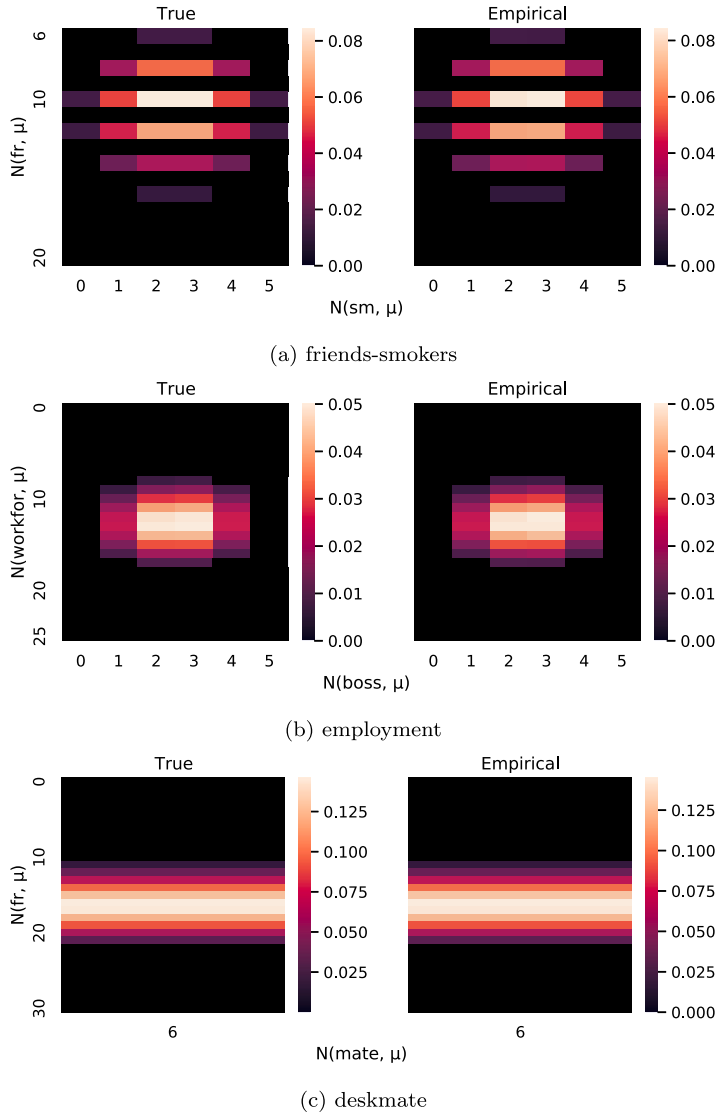


Fig. 5. Conformity testing for the count distribution of MLNs.

probability of each model, the count distribution only tells us the probability that a certain number of predicates are interpreted to be true in the models. An advantage of testing count distributions is that they can be efficiently computed for our MLNs. Please refer to [28] for more details about count distributions. We also note that the conformity of count distribution is a necessary condition for the correctness of algorithms. We kept the domain size to 5 for friends-smokers and employment MLNs and set it to 6 for the deskmate MLN (due to the counting quantifiers in the MLN). We would like to emphasize that the choice of small domains was not made for scalability purposes, but rather to facilitate the statistical tests. In the following section, we will demonstrate that our approach can indeed scale to larger domains. A total number of 10^5 models were generated from the weighted model sampler for each MLN. The empirical distributions of count-statistics, along with the true count distributions, are shown in Fig. 5. It is easy to check the conformity of the empirical distribution to the true one from the figure. The statistical test was also performed on the count distribution, and the results confirm the conclusion drawn from the figure (also see Appendix E.2).

9.2. Performance

To evaluate the performance, we compare our weighted model samplers with Unigen¹¹ [3,40], the state-of-the-art approximate sampler for Boolean formulas. Note that there is no guarantee of polynomial complexity for Unigen, and its runtime highly depends

¹¹ <https://github.com/meelgroup/unigen>.

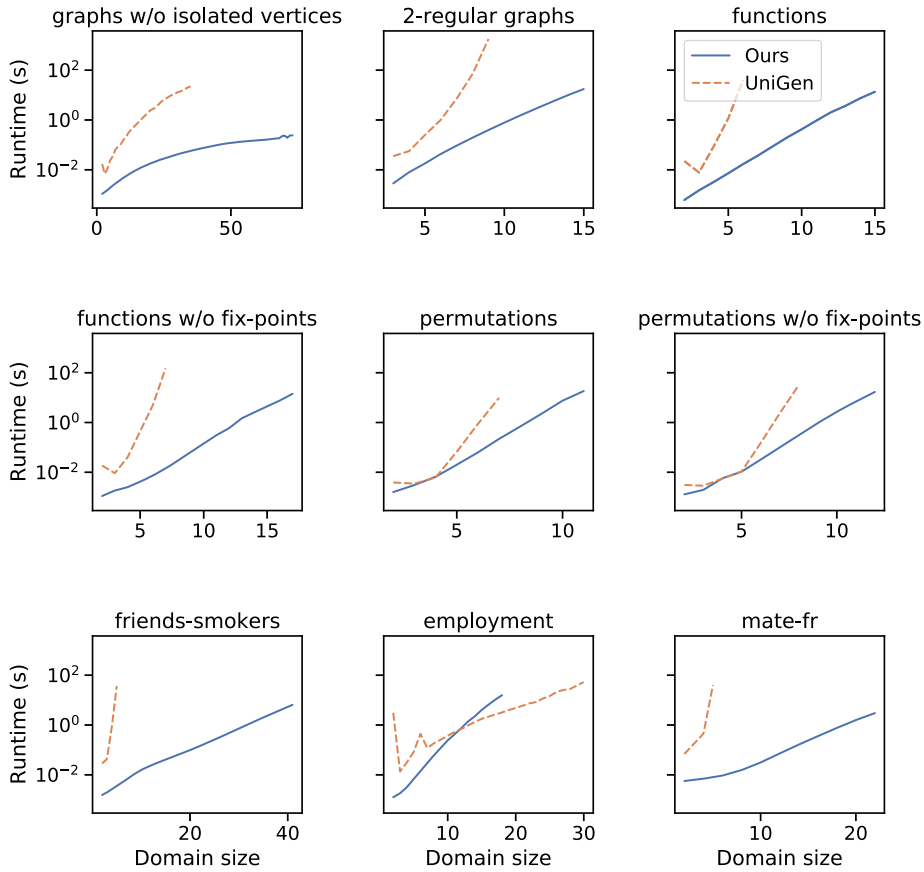


Fig. 6. Performance of weighted model sampler versus UniGen.

on the underlying SAT solver it uses. We reduce the WFOMS problems to the sampling problems of Boolean formulas by grounding the input first-order sentence over the given domain. Since Unigen only works for uniform sampling, we employ the technique in [5] to encode the weighting function in the WFOMS problem into a Boolean formula.

For each sampling problem, we randomly generated 1000 models by our weighted model sampler and Unigen respectively, and computed the average sampling time per one model. The performance comparison is shown in Fig. 6. In most cases, our approach is much faster than UniGen. The exception in the employment MLN, where UniGen performed better than the weighted model sampler, is likely due to the simplicity of this specific instance (almost every structure is a model of this MLN). This coincides with the theoretical result that our weighted model sampler is polynomial-time in the domain size, while UniGen usually needs a significant number of expensive SAT calls on the grounded formulas.

The full C^2 sampler proposed in Section 6 was not implemented in our study due to the extremely slow computational speed of the existing WFOMC solver [28] for full C^2 , despite its polynomial complexity. Developing faster WFOMC algorithms is orthogonal to the present work where we only use the existing solver as a black box.

10. Related work

The studies on model counting and sampling problems date back to the 1970s [43,44,22]. In the seminal paper by Jerrum et al. [22], a significant connection was established between the random generation of combinatorial structures (i.e., model sampling) and the problem of model counting. Specifically, this paper showed that the random generation of combinatorial structures can be reduced in polynomial time to the model counting problem, under the condition of *self-reducibility*. Self-reducibility refers to the property where the solution set for a given instance of a problem can be expressed in terms of the solution sets of smaller instances of the same problem. In the context of propositional logic, self-reducibility naturally holds as the solution set of a Boolean formula Γ can be expressed in terms of the solution sets of Γ_a and $\Gamma_{\neg a}$, where a is an arbitrary atom in Γ , and Γ_a and $\Gamma_{\neg a}$ are the formulas obtained from Γ by setting a to be true and false respectively. This property allows for a polynomial-time reduction from WMS to WMC, where the truth value of each atom in Γ is incrementally sampled based on the ratio of the WMC of Γ_a and $\Gamma_{\neg a}$. However, in first-order logic, self-reducibility is not generally guaranteed. Conditioning on a ground atom in a first-order sentence may make the resulting problem intractable, even if the original problem is tractable. For example, Van den Broeck and Davis [51] have proven that there exists a UFO^2 sentence with *binary* ground atoms, whose model count cannot be computed in time polynomial in the size

of the binary atoms, unless $P = NP$. This result implies that the reduction derived from self-reducibility is not applicable to WFOMS, motivating this paper to develop new techniques for sampling from first-order formulas.

The approach taken in this paper, as well as the formal *liftability* notions considered here, are inspired by the *lifted inference* literature [55]. In lifted inference, the goal is to perform probabilistic inference in statistical-relational models in a way that exploits symmetries in the high-level structure of the model. Models that are amenable to scalable inference as domain size (or population size) increases are dubbed *domain-liftable*, in a similar spirit to the notion of *domain-liftability under sampling* presented here. There exists a very extensive literature on lifted inference, which has been explored from both graphical models and logic perspectives. In the graphical models literature, the focus is on leveraging the symmetries of random variables and factor graphs. This involves grouping symmetric variables and factors into their parameterized counterparts, leading to the construction of a relational graphical model on which lifted inference can be performed [34,36,23]. The objective of these approaches is typically to extend the inference algorithms used in traditional graphical models to relational models. Examples of such algorithms include lifted variable elimination [34,8,30,41], lifted belief propagation [39,26], and lifted junction tree [2]. Another strand of research in the lifted inference models the relations between random variables using propositional or first-order logic, and then performs inference by exploiting the symmetries in the logical structure of the model [53,48,54,45,18]. In these studies, the WFOMC problem serves as an “assembly language” for the probabilistic inference problem, and the goal is to develop efficient algorithms for solving the WFOMC problem. Of particular interest is the work by Beame et al. [1, Appendix C], which studies the data complexity of WFOMC of UFO^2 . The general argument used there—namely, the analysis of a two-variable sentence in terms of its cell types—forms a basis for our sampling approaches discussed in the paper. Finally, though not directly related to our work, we note that the lifting notion, which groups symmetric components in a problem and performs computation on a higher level, has also been explored in other areas, such as probabilistic theorem proving [14] as well as linear and convex quadratic programming [31,32].

The domain recursion scheme, another important approach adopted in this paper, is similar to the *domain recursion rule* used in weighted first-order model counting [49,24,25,42]. The domain recursion rule for WFOMC is a technique that uses a gradual grounding process on the input first-order sentence, where only one element of the domain is grounded at a time. As each element is grounded, the partially grounded sentence is simplified until the element is entirely removed, resulting in a new WFOMC problem with a smaller domain. With the domain recursion rule, one can apply the principle of induction on the domain size, and solve the counting problem by dynamic programming. A closely related work to this paper is the approach presented by Kazemi et al. [24], where they used the domain recursion rule to solve the WFOMC problem on FO^2 sentences without using Skolemization [54], which introduces negative weights. However, it is important to note that their approach can be only applied to some specific first-order formulas, whereas the domain recursion scheme presented in this paper, mainly designed for eliminating the existentially-quantified formulas, supports the entire C^2 fragment with cardinality constraints.

It is also worth mentioning that weighted model sampling is a relatively well-studied area [15,3,4]. However, many real-world problems can be represented more naturally and concisely in first-order logic, and suffer from a significant increase in formula size when grounded out to propositional logic. For example, a formula of the form $\forall x \exists y : \varphi$ is encoded as a Boolean formula of the form $\bigwedge_{i=1}^n \bigvee_{j=1}^n l_{i,j}$, whose length is quadratic in the domain size n . Since even finding a solution to a such large ground formula is challenging, most sampling approaches for propositional logic instead focus on designing approximate samplers. We also note that these approaches are not polynomial-time in the length of the input formula, and rely on access to an efficient SAT solver. An alternative strand of research [19,20,12] on combinatorial sampling, focuses on the development of near-uniform and efficient sampling algorithms. However, these approaches can only be employed for specific Boolean formulas that satisfy a particular technical requirement known as the Lovász Local Lemma. The WFOMS problems studied in this paper do not typically meet the requisite criteria for the application of these techniques.

11. Conclusion and future work

In this paper, we prove the domain-liftability under sampling for the C^2 fragment. The result is further extended to the fragment of C^2 with the presence of cardinality constraints. The widespread applicability of WFOMS renders the proposed approach a promising candidate to serve as a universal paradigm for a plethora of sampling problems.

A potential avenue for further research is to expand the methodology presented in this paper to encompass more expressive first-order languages. Specifically, the domain recursion scheme employed in this paper could be extended to other tractable fragments, as its analogous technique, the domain recursion rule, has been demonstrated to be effective in proving the domain-liftability of the fragments S^2FO^2 and S^2RU for WFOMC [25].

In addition to extending the input logic, other potential directions for future research include incorporating elementary axioms, such as tree axiom [46] and linear order axiom [42], as well as more general weighting functions that involve negative weights. However, it is important to note that these extensions would likely require a more advanced and nuanced approach than the one proposed in this paper, and may present significant challenges.

Finally, the lower complexity bound of WFOMS is also an interesting open problem. We have discussed in the introduction that it is unlikely for an (even approximate) lifted weighted model sampler to exist for full first-order logic. However, the establishment of a tighter lower bound for fragments of FO, such as FO^3 , remains an unexplored and challenging area that merits further investigation.

CRediT authorship contribution statement

Yuanhong Wang: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **Juhua Pu:** Funding acquisition, Resources. **Yuyi Wang:** Conceptualization, Formal analysis, Supervision, Writing – review & editing. **Ondřej Kuželka:** Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Yuanhong Wang, Juhua Pu reports financial support was provided by The National Key R&D Program of China, the National Science Foundation of China. Ondřej Kuželka reports financial support was provided by The Czech Science Foundation.

Data availability

No data was used for the research described in the article.

Acknowledgement

Yuanhong Wang and Juhua Pu are supported by the National Science Foundation of China (62177002). Ondřej Kuželka's work is supported by the Czech Science Foundation project 20-19104Y and partially also 23-07299S (most of the work was done before the start of the latter project).

Appendix A. WFOMC with unary evidence

In this section, we show how to deal with unary evidence in conditional WFOMC.

Van den Broeck and Davis [51] handled the unary evidence by the following transformation of the input sentence Γ . For any unary predicate P appearing in the evidence, they split the domain into Δ_{\top}^P , Δ_{\perp}^P and Δ_{\emptyset}^P , where Δ_{\top}^P and Δ_{\perp}^P contains precisely the elements with evidence $P(a)$ and $\neg P(a)$, respectively, and $\Delta_{\emptyset}^P = \Delta \setminus (\Delta_{\top}^P \cup \Delta_{\perp}^P)$ is the remaining elements. Then the sentence Γ was transformed into

$$(\forall x \in \Delta_{\top}^P : \Gamma_{\top}^P) \wedge (\forall x \in \Delta_{\perp}^P : \Gamma_{\perp}^P) \wedge (\forall x \in \Delta_{\emptyset}^P : \Gamma),$$

where Γ_{\top}^P and Γ_{\perp}^P were obtained from Γ by replacing all occurrences of P with True and False, respectively, and $\forall x \in \Delta'$ was a domain constraint that restricts the quantifier to the domain Δ' . The procedure could be repeated to support multiple unary predicates, and the resulting sentence was then compiled into an FO d-DNNF circuits [53] for model counting. The domain constraints are natively supported by FO d-DNNF circuits, and the compilation and model counting have been shown to be in time polynomial in the domain size.

However, not all WFOMC algorithms can effectively support the domain constraints, and efficiently count the model of the transformed sentence. In the following, we provide a simpler approach to deal with the unary evidence, without the need for domain constraints.

We first introduce the notion of evidence type. An *evidence type* σ over a predicate vocabulary \mathcal{P} is a consistent set of 1-literals formed by \mathcal{P} . For instance, both $\{P(x), \neg Q(x)\}$ and $\{P(x)\}$ are evidence types over $\{P/1, Q/1\}$. The evidence type can be also viewed as a conjunction of its elements, where $\sigma(x)$ denotes a quantifier-free formula. If the evidence type σ is an empty set, then $\sigma(x)$ is defined as \top . The number of evidence types over \mathcal{P} is finite, and independent of the size of the domain. Given a set L of ground 1-literals, the evidence type of an element is defined as the set of all literals in L that are associated with the element. For example, if $L = \{P(a), \neg Q(a), R(b)\}$ and $\Delta = \{a, b, c\}$, then the evidence type of a is $\{P(x), \neg Q(x)\}$, the evidence type of b is $\{R(x)\}$, and the evidence type of c is \emptyset .

Proposition 1. Let Γ be a domain-liftable first-order sentence, and let Δ be a domain. For any set L of unary literals grounding on Δ , and any weighting functions (w, \bar{w}) , $\text{WFOMC}(\Gamma \wedge \bigwedge_{l \in L} l, \Delta, w, \bar{w})$ can be computed in time polynomial in both the domain size and the size of L .

Proof. Let $\sigma^1, \sigma^2, \dots, \sigma^m$ be the distinct evidence types of elements given by L . The number of elements with evidence type σ^i is denoted by n_i . We first transform the input sentence into

$$\Gamma' := \Gamma \wedge \bigwedge_{i \in [m]} (\forall x : \xi^i(x) \Rightarrow \sigma^i(x)) \wedge \Lambda,$$

where each $\xi^i/1$ is a fresh predicate for σ^i with the weight $w(\xi^i) = \bar{w}(\xi^i) = 1$, and

$$\Lambda = \left(\forall x : \bigvee_{i \in [m]} \xi^i(x) \right) \wedge \left(\forall x : \bigwedge_{i, j \in [m]: i \neq j} (\neg \xi^i(x) \vee \neg \xi^j(x)) \right). \quad (\text{A.1})$$

The formula Λ states that each element a has exactly one $\xi^i(a)$ that is True. In other words, the interpretation of predicates ξ^i can be seen as a partition of the domain, where each disjoint subset in the partition contains precisely the elements with evidence type σ^i . Then we have

$$\text{WFOMC}(\Gamma \wedge \bigwedge_{l \in L} l, \Delta, w, \bar{w}) = \frac{1}{\binom{n}{n_1, n_2, \dots, n_m}} \cdot \text{WFOMC}(\Gamma' \wedge \bigwedge_{i \in [m]} (|\xi^i| = n_i), \Delta, w, \bar{w}), \quad (\text{A.2})$$

where $|\xi^i| = n_i$ is a cardinality constraint that restricts the number of True ξ^i in the model to be n_i .

The reasoning is as follows. Denote by k_1, k_2, \dots, k_n the indices of evidence type of each element in the domain, i.e., σ^{k_i} is the evidence type of i -th element in the domain, and ξ^{k_i} is its corresponding predicate. We first show that

$$\text{WFOMC}(\Gamma \wedge \bigwedge_{l \in L} l, \Delta, w, \bar{w}) = \text{WFOMC}(\Gamma' \wedge \bigwedge_{i \in [n]} \xi^{k_i}(e_i), \Delta, w, \bar{w}), \quad (\text{A.3})$$

where e_i is the i -th element in the domain. The proof is built on the fact that the reduct mapping $\langle \cdot \rangle_{\mathcal{P}_\Gamma}$ is a bijective function from the models of $\Psi' = \Gamma' \wedge \bigwedge_{i \in [n]} \xi^{k_i}(e_i)$ to the models of $\Psi = \Gamma \wedge \bigwedge_{l \in L} l$. First, it is easy to show that the mapping $\langle \cdot \rangle_{\mathcal{P}_\Gamma}$ is a function from the models of Ψ' to the models of Ψ , since for any model μ' of Ψ' , $\langle \mu' \rangle_{\mathcal{P}_\Gamma}$ is a model of Ψ . Second, the mapping $\langle \cdot \rangle_{\mathcal{P}_\Gamma}$ is injective. This is because, for any two models μ'_1 and μ'_2 of Ψ' , they must share the same interpretation of the predicates ξ^i due to the constraint Λ and the evidence $\xi^{k_i}(e_i)$; and if $\langle \mu'_1 \rangle_{\mathcal{P}_\Gamma} = \langle \mu'_2 \rangle_{\mathcal{P}_\Gamma}$, it follows that these two models also have the same interpretation of the predicates in \mathcal{P}_Γ , and thus $\mu'_1 = \mu'_2$. Finally, the mapping is also surjective: for any model μ of Ψ , we can always construct a model μ' of Ψ' by expanding μ with the evidence $\xi^{k_i}(e_i)$ for each $i \in [n]$, such that $\langle \mu' \rangle_{\mathcal{P}_\Gamma} = \mu$. Thus, we have (A.3).

Next, let us consider the two WFOMC problems of $\Gamma' \wedge \bigwedge_{i \in [m]} (|\xi^i| = n_i)$ and $\Gamma' \wedge \bigwedge_{i \in [n]} \xi^{k_i}(e_i)$. Observe that any interpretation of ξ^i that satisfies Γ' can be viewed as an order-dependent partition of the domain with the configuration of the partition being $(|\xi^1|, |\xi^2|, \dots, |\xi^m|)$. Thus, $\text{WFOMC}(\Gamma' \wedge \bigwedge_{i \in [m]} (|\xi^i| = n_i), \Delta, w, \bar{w})$ can be written as the summation of $\text{WFOMC}(\Gamma' \wedge \bigwedge_{i \in [n]} \xi^{k_i}(e_i), \Delta, w, \bar{w})$ over all possible interpretations $\xi^{k_i}(e_i)$, whose corresponding partition configuration is (n_1, \dots, n_m) . There are totally $\binom{n}{n_1, n_2, \dots, n_m}$ such interpretations, and the interpretation $\xi^{k_i}(e_i)$ is one of them. Moreover, due to the symmetry of WFOMC, all these interpretations result in the same result of WFOMC. Thus, we can write

$$\text{WFOMC}(\Gamma' \wedge \bigwedge_{i \in [m]} (|\xi^i| = n_i), \Delta, w, \bar{w}) = \binom{n}{n_1, n_2, \dots, n_m} \cdot \text{WFOMC}(\Gamma' \wedge \bigwedge_{i \in [n]} \xi^{k_i}(e_i), \Delta, w, \bar{w}). \quad (\text{A.4})$$

Combining (A.3) and (A.4) yields (A.2). Finally, computing (A.2) has been shown to be in time polynomial in the domain size in [53] for any domain-liftable sentence Γ , which completes the proof. \square

Appendix B. Normal forms

B.1. Scott normal form

We briefly describe the transformation of \mathbf{FO}^2 formulas to SNF and prove the soundness of its corresponding reduction on the WFOMS problems. The process is well-known, so we only sketch the related details. Interested readers can refer to Kuusisto and Lutz [27] for more details.

Let Γ be a sentence of \mathbf{FO}^2 . To put it into SNF, consider a subformula $\varphi(x) = Qy : \phi(x, y)$, where $Q \in \{\forall, \exists\}$ and ϕ is quantifier-free. Let A_φ be a fresh unary predicate¹² and consider the sentence

$$\forall x : (A_\varphi(x) \Leftrightarrow (Qy : \phi(x, y)))$$

which states that $\varphi(x)$ is equivalent to $A_\varphi(x)$. Let Q' denote the dual of Q , i.e., $Q' \in \{\forall, \exists\} \setminus \{Q\}$, this sentence can be seen equivalent to

$$\Gamma' := \forall x Qy : (A_\varphi(x) \Rightarrow \phi(x, y)) \wedge \forall x Q'y : (\phi(x, y) \Rightarrow A_\varphi(x)).$$

Let

$$\Gamma'' := \Gamma' \wedge \Gamma[\varphi(x)/A_\varphi(x)],$$

where $\Gamma[\varphi(x)/A_\varphi(x)]$ is obtained from Γ by replacing $\varphi(x)$ with $A_\varphi(x)$. For any domain Δ , every model of Γ'' over Δ can be mapped to a unique model of Γ over Δ . The bijective mapping function is simply the projection $\langle \cdot \rangle_{\mathcal{P}_\Gamma}$. Let both the positive and negative weights of A_φ be 1 and denote the new weighting functions as w' and \bar{w}' . It is clear that the reduction from $(\Gamma, \Delta, w, \bar{w})$

¹² If $\varphi(x)$ has no free variables, e.g., $\exists x : \phi(x)$, the predicate A_φ is nullary.

to $(\Gamma'', \Delta, w', \bar{w}')$ is sound. Repeat this process from the atomic level and work upwards until the sentence is in SNF. The whole reduction remains sound due to the transitivity of soundness.

B.2. Normal form for C^2

We show that any C^2 sentence can be converted into the normal form:

$$\Gamma_{\forall} \wedge \bigwedge_{i \in [m]} (\forall x : A_i(x) \Leftrightarrow (\exists_{=k_i} y : R_i(x, y))),$$

where Γ_{\forall} is a FO^2 sentence, each k_i is a non-negative integer, $R_i(x, y)$ is an atomic formula, and A_i is an unary predicate. The process is as follows:

1. Convert each counting-quantified formula of the form $\exists_{\geq k} y : \phi(x, y)$ to $\neg(\exists_{\leq k-1} y : \phi(x, y))$.
2. Decompose each $\exists_{\leq k} y : \phi(x, y)$ into $\bigvee_{i \in [0, k]} (\exists_{=i} y : \phi(x, y))$.
3. Replace each subformula $\exists_{=k} y : \phi(x, y)$, where $k > |\Delta|$, with False.
4. Starting from the atomic level and working upwards, replace any subformula $\exists_{=k} \phi(x, y)$, where $\phi(x, y)$ is a formula that does not contain any counting quantifier, with $A(x)$; and append $\forall x \forall y : R(x, y) \Leftrightarrow \phi(x, y)$ and $\forall x : A(x) \Leftrightarrow (\exists_{=k} y : R(x, y))$, where R is an auxiliary binary predicate, to the original sentence.

It is easy to check that the reduction presented above is sound and independent of the domain size if the domain size is greater than the maximum counting parameter k in the input sentence.¹³

Appendix C. A sound reduction from SC^2 to FO^2 with cardinality constraints

In this section, we show the sound reduction from a WFOMS problem on SC^2 sentence to a WFOMS problem on FO^2 sentence with cardinality constraints.

We first need the following two lemmas, which are based on the transformations from [28].

Lemma 8. *Let Γ be a first-order logic sentence, and let Δ be a domain. Let Φ be a first-order sentence with cardinality constraints, defined as follows:*

$$\begin{aligned} \Pi := & (|P| = k \cdot |\Delta|) \\ & \wedge (\forall x \forall y : P(x, y) \Leftrightarrow (R_1^P(x, y) \vee \dots \vee R_k^P(x, y))) \\ & \wedge \bigwedge_{i \in [k]} (\forall x \exists y : R_i^P(x, y)) \\ & \wedge \bigwedge_{i, j \in [k] : i \neq j} (\forall x \forall y : \neg R_i^P(x, y) \vee \neg R_j^P(x, y)), \end{aligned}$$

where R_i^P are auxiliary predicates not in \mathcal{P}_{Γ} with weight $w(R_i^P) = \bar{w}(R_i^P) = 1$. Then the reduction from the WFOMS $(\Gamma \wedge \forall x \exists_{=k} y : P(x, y), \Delta, w, \bar{w})$ to $(\Gamma \wedge \Pi, \Delta, w, \bar{w})$ is sound.

Proof. Let $f(\cdot) = \langle \cdot \rangle_{\mathcal{P}_{\Gamma} \cup \{P\}}$ be a mapping function. We first show that f is from $\mathcal{M}_{\Gamma \wedge \Pi, \Delta}$ to $\mathcal{M}_{\Gamma \wedge \forall x \exists_{=k} y : P(x, y), \Delta}$: if $\mathcal{A} \models \Gamma \wedge \Pi$ then $f(\mathcal{A}) \models \Gamma \wedge \forall x \exists_{=k} y : P(x, y)$.

The sentence Π means that for every $c_1, c_2 \in \Delta$ such that $P(c_1, c_2)$ is true, there is exactly one $i \in [k]$ such that $R_i^P(c_1, c_2)$ is true. Thus we have that $\sum_{i \in [k]} |R_i^P| = |P| = k \cdot |\Delta|$, which together with $\bigwedge_{i \in [k]} \forall x \exists y : R_i^P(x, y)$ implies that $|R_i^P| = k$ for $i \in [k]$. We argue that each R_i^P is a function predicate in the sense that $\forall x \exists_{=1} y : R_i^P(x, y)$ holds in any model of $\Gamma \wedge \Pi$. Let us suppose, for contradiction, that $(\forall x \exists y : R_i^P(x, y)) \wedge (|R_i^P| = k)$ holds but there is some $a \in \Delta$ such that $R_i^P(a, b)$ and $R_i^P(a, b')$ are true for some $b \neq b' \in \Delta$. We have $|\{(x, y) \in \Delta^2 \mid R_i^P(x, y) \wedge x \neq a\}| \geq |\Delta| - 1$ by the fact $\forall x \exists y : R_i^P(x, y)$. It follows that $|R_i^P| \geq |\{(x, y) \in \Delta^2 \mid R_i^P(x, y) \wedge x \neq a\}| + 2 > |\Delta|$, which leads to a contradiction. Since all of R_i^P are function predicates, it is easy to check $\forall x \exists_{=k} y : P(x, y)$ must be true in any model μ of $\Gamma \wedge \Pi$, i.e., $f(\mu) \models \Gamma \wedge \forall x \exists_{=k} y : P(x, y)$.

To finish the proof, one can easily show that, for every model $\mu \in \mathcal{M}_{\Gamma \wedge \forall x \exists_{=k} y : P(x, y), \Delta}$, there are exactly $(k!)^{|\Delta|}$ models $\mu' \in \mathcal{M}_{\Gamma \wedge \Pi, \Delta}$ such that $f(\mu') = \mu$. The reason for this is that 1) if, for any $a \in \Delta$, we permute b_1, b_2, \dots, b_k in $R_1^P(a, b_1), R_2^P(a, b_2), \dots, R_k^P(a, b_k)$ in the model μ' , we get another model of $\Gamma \wedge \Pi$, and 2) up to these permutations, the predicates R_i^P in μ' are determined uniquely by μ . Finally, the weights of all these μ' are the same as those of μ , and we can write

¹³ This condition does not change the data complexity of the problem, as parameters k of the counting quantifiers are considered constants but not the input of the problem.

$$\begin{aligned}
\sum_{\substack{\mu' \in \mathcal{M}_{\Gamma \wedge \Pi, \Delta} : \\ f(\mu') = \mu}} \mathbb{P}[\mu' \mid \Gamma \wedge \Pi] &= \frac{\sum_{\mu' \in \mathcal{M}_{\Gamma \wedge \Pi, \Delta} : \langle w, \bar{w} \rangle(\mu')} f(\mu')}{\text{WFOMC}(\Gamma \wedge \Pi, \Delta, w, \bar{w})} \\
&= \frac{(k!)^{|\Delta|} \cdot \langle w, \bar{w} \rangle(\mu)}{(k!)^{|\Delta|} \cdot \text{WFOMC}(\Gamma \wedge \forall x \exists_{=k} y : P(x, y), \Delta, w, \bar{w})} \\
&= \frac{\langle w, \bar{w} \rangle(\mu)}{\text{WFOMC}(\Gamma \wedge \forall x \exists_{=k} y : P(x, y), \Delta, w, \bar{w})} \\
&= \mathbb{P}[\mu \mid \Gamma \wedge \forall x \exists_{=k} y : P(x, y)],
\end{aligned}$$

which completes the proof. \square

Lemma 9. Let Γ be a first-order logic sentence, Δ be a domain, and P be a predicate. Then the WFOMS $(\Gamma \wedge \forall_{=k} \forall y : P(x, y), \Delta, w, \bar{w})$ can be reduced to $(\Gamma \wedge (|U| = k) \wedge (\forall x : U(x) \Leftrightarrow (\forall y : P(x, y))), \Delta, w, \bar{w})$, where U is an auxiliary unary predicate with weight $w(U) = \bar{w}(U) = 1$, and the reduction is sound.

Proof. The proof is straightforward. \square

Proof of Lemma 7. We can first get rid of all formulas of the form $\exists_{=k} x \forall y : P(x, y)$ by repeatedly using Lemma 9. Then we can use Lemma 8 repeatedly to eliminate the formulas of the form $\forall x \exists_{=k} y : P(x, y)$. The whole reduction is sound due to the transitivity of soundness. \square

Appendix D. Missing details of weighted model sampler

D.1. Optimizations for weighted model sampler

There exist several optimizations to make the sampling algorithm more practical. Here, we present some of them that are used in our implementation.

Heuristic element selection The complexity of TwoTablesSamplerForFO2 heavily depends on the recursion depth. In our implementation, when selecting a domain element e_i for sampling its substructure, we always chose the element with the “strongest” existential constraint that contains the most Tseitin atoms $Z_k(x)$. It would help TwoTablesSamplerForFO2 reach the condition that the existential constraint for all elements is \top . In this case, TwoTablesSamplerForFO2 will invoke the more efficient sampler for UFO^2 to sample the remaining substructures.

Further decomposition of 2-tables sampling Let \mathcal{P}_{\exists} be the union of predicates vocabularies of the existentially quantified formulas

$$\mathcal{P}_{\exists} := \{R_1, R_2, \dots, R_m\}.$$

We further decomposed the sampling probability of 2-tables $\mathbb{P}[\mathcal{A} \mid \Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i)]$ into

$$\mathbb{P} \left[\mathcal{A} \mid \Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i) \wedge \mathcal{A}_{\exists} \right] \cdot \mathbb{P} \left[\mathcal{A}_{\exists} \mid \Gamma_T \wedge \bigwedge_{i \in [n]} \eta_i(e_i) \right],$$

where \mathcal{A}_{\exists} is a \mathcal{P}_{\exists} -structure over Δ . It decomposes the conditional sampling problem of \mathcal{A} into two subproblems—one is to sample \mathcal{A}_{\exists} and the other to sample the remaining substructures conditional on \mathcal{A}_{\exists} . The subproblem of sampling \mathcal{A} conditioned on \mathcal{A}_{\exists} can be reduced into a sampling problem on a UFO^2 sentence with evidence

$$\Gamma' = \forall x \forall y : \psi(x, y) \wedge \bigwedge_{i \in [m]} \tau_i(e_i) \wedge \mathcal{A}_{\exists}, \quad (\text{D.1})$$

since all existentially-quantified formulas have been satisfied with \mathcal{A}_{\exists} . Recall that for a UFO^2 sentence, when the 1-types τ_i of all elements has been fixed, the sampling problem of 2-tables can be decomposed into multiple independent sampling problems of the 2-table for each element tuple (see Section 4.2). Sampling 2-tables from (D.1) can be solved in a similar way, by decomposing the sampling probability into

$$\mathbb{P} \left[\bigwedge_{i, j \in [n] : i < j} \pi_{i,j}(e_i, e_j) \mid \Gamma' \right] = \prod_{i, j \in [n] : i < j} \mathbb{P} [\pi_{i,j}(e_i, e_j) \mid \psi'(e_i, e_j)],$$

where $\psi'(e_i, e_j)$ is the simplified formula of $\psi(e_i, e_j) \wedge \psi(e_j, e_i)$ obtained by replacing the ground 1-literals with their truth values by the 1-types $\tau_i(e_i)$ and $\tau_j(e_j)$, and the sampled structure \mathcal{A}_{\exists} . The 2-tables $\pi_{i,j}$ as well as the simplified formulas $\psi'(e_i, e_j)$ are independent of each other, and thus can be efficiently sampled in parallel. The other subproblem of sampling \mathcal{A}_{\exists} can be solved by an

algorithm similar to TwoTablesSamplerForFO2. In this algorithm, the 2-tables used to partition cells are now defined over \mathcal{P}_\exists , whose size is only 4^m . Compared to the number 4^B of 2-tables in the original algorithm TwoTablesSampler, where $B \geq m$ is the number of binary predicates in Γ , the number of 2-tables in the new algorithm can be exponentially smaller. As a result, the for-loop that enumerates all 2-tables configurations in Algorithm 4 is exponentially faster.

Cache We cached the weight $\langle w, \bar{w} \rangle(\tau^i)$ of all 1-types and the weight $\langle w, \bar{w} \rangle(\pi^i)$ of all 2-tables, which are often used in our sampler.

D.2. The submodule $\text{ExSat}(\cdot, \cdot)$

The pseudo-code of ExSat is presented in Algorithm 7.

Algorithm 7 $\text{ExSat}(\mathbf{g}, \eta)$.

```

1: Decompose  $\mathbf{g}$  into  $\{g_{\eta^i, \pi^j}\}_{i \in [N_c], j \in [N_b]}$ 
2:  $(\beta, \tau) \leftarrow (\eta)$ 
3: // Check the coherence of 2-tables
4: for  $i \in [N_c]$  do
5:   for  $j \in [N_b]$  do
6:     //  $\tau(\eta^i)$  is the 1-type in  $\eta^i$ 
7:     if  $\pi^j$  is not coherent with  $\tau(\eta^i)$  and  $\tau$  and  $g_{\eta^i, \pi^j} > 0$  then
8:       return False
9:     end if
10:   end for
11: end for
12: // Check the satisfaction of existentially-quantified formulas
13:  $\forall j \in [N_b], h_{\pi^j} \leftarrow \sum_{i \in [N_c]} g_{\eta^i, \pi^j}$ 
14: for  $Z_k(x) \in \beta$  do
15:   for  $j \in [N_b]$  do
16:     if  $R_k(x, y) \in \pi^j$  and  $h_{\pi^j} > 0$  then
17:       return True
18:     end if
19:   end for
20: end for
21: return False

```

Appendix E. Missing details of experiments

E.1. Experiment settings

Sampling combinatorial structures The corresponding WFOMS problems for the uniform generation of combinatorial structures used in our experiments are presented as follows. The weighting functions w and \bar{w} map all predicates to 1.

- Functions:

$$\forall x \exists_{=1} y : f(x, y).$$

- Functions w/o fix points:

$$(\forall x \exists_{=1} y : f(x, y)) \wedge (\forall x : \neg f(x, x)).$$

- Permutations:

$$(\forall x \exists_{=1} y : \text{Per}(x, y)) \wedge (\forall y \exists_{=1} x : \text{Per}(x, y)).$$

- Permutation without fix-points:

$$(\forall x \exists_{=1} y : \text{Per}(x, y)) \wedge (\forall y \exists_{=1} x : \text{Per}(x, y)) \wedge (\forall x : \neg \text{Per}(x, x)).$$

Sampling from MLNs An MLN is a finite set of weighted first-order formulas $\{(w_i, \alpha_i)\}_{i \in [m]}$, where each w_i is either a real-valued weight or ∞ , and α_i is a first-order formula. Let \mathcal{P} be the vocabulary of $\alpha_1, \alpha_2, \dots, \alpha_m$. An MLN Φ paired with a domain Δ induces a probability distribution over \mathcal{P} -structures (also called possible worlds):

$$p_{\Phi, \Delta}(\omega) := \begin{cases} \frac{1}{Z_{\Phi, \Delta}} \exp\left(\sum_{(\alpha, w) \in \Phi_{\mathbb{R}}} w \cdot \#(\alpha, \omega)\right) & \text{if } \omega \models \Phi_{\infty} \\ 0 & \text{otherwise} \end{cases}$$

where $\Phi_{\mathbb{R}}$ and Φ_{∞} are the real-valued and ∞ -valued formulas in Φ respectively, and $\#(\alpha, \omega)$ is the number of groundings of α satisfied in ω . The sampling problem on an MLN Φ over a domain Δ is to randomly generate a possible world ω according to the probability $p_{\Phi, \Delta}(\omega)$.

The reduction from the sampling problems on MLNs to WFOMS can be performed following the same idea as in [53]. For every real-valued formula $(\alpha_i, w_i) \in \Phi_{\mathbb{R}}$, where the free variables in α_i are \mathbf{x} , we introduce a novel auxiliary predicate ξ_i and create a new formula $\forall \mathbf{x} : \xi_i(\mathbf{x}) \Leftrightarrow \alpha_i(\mathbf{x})$. For formula α_i with infinity weight, we instead create a new formula $\forall \mathbf{x} : \alpha_i(\mathbf{x})$. Denote the conjunction of the resulting set of sentences by Γ , and set the weighting function to be $w(\xi_i) = \exp(w_i)$ and $\bar{w}(\xi_i) = 1$, and for all other predicates, we set both w and \bar{w} to be 1. Then the sampling problem on Φ over Δ is reduced to the WFOMS $(\Gamma, \Delta, w, \bar{w})$.

By the reduction above, we can write the two MLNs used in our experiments to WFOMS problems. The weights of predicates are all set to be 1 unless otherwise specified.

- Friends-smokers MLN: the reduced sentence is

$$\begin{aligned} \forall x : \neg fr(x, x) \wedge \\ \forall x \forall y : fr(x, y) \Leftrightarrow fr(y, x) \wedge \\ \forall x \exists y : fr(x, y) \wedge \\ \forall x \forall y : (\xi_1(x) \Leftrightarrow sm(x)) \wedge \\ \forall x \forall y : \xi_2(x, y) \Leftrightarrow (fr(x, y) \wedge sm(x) \Rightarrow sm(y)), \end{aligned}$$

and the weight of ξ_1 and ξ_2 is set to be $w(\xi) = \exp(0)$, $w(\xi_2) = \exp(0.2)$.

- Employment MLN: the corresponding sentence is

$$\forall x : \xi(x) \Leftrightarrow (\exists y : workfor(x, y) \vee boss(x)),$$

and the weight of ξ is set to be $\exp(1.3)$.

- Deskmate MLN: the corresponding sentence is

$$\begin{aligned} \forall x : \neg mate(x, x) \wedge \neg fr(x, x) \wedge \\ \forall x \forall y : mate(x, y) \Rightarrow mate(y, x) \wedge \\ \forall x \exists_{=1} y : mate(x, y) \wedge \\ \forall y \exists_{=1} x : mate(x, y) \wedge \\ \forall x : \xi(x, y) \Leftrightarrow (mate(x, y) \Rightarrow fr(x, y)), \end{aligned}$$

and the weight of ξ is set to be $\exp(1.0)$.

E.2. More experimental results

We use the Kolmogorov-Smirnov (KS) test [29] to validate the conformity of the (count) distributions produced by our algorithm to the reference distributions. The KS test used here is based on the multivariate Dvoretzky-Kiefer-Wolfowitz (DKW) inequality recently proved by [33].

Let $\mathbf{X}_1 = (X_{1i})_{i \in [k]}$, $\mathbf{X}_2 = (X_{2i})_{i \in [k]}$, \dots , $\mathbf{X}_n = (X_{ni})_{i \in [k]}$ be n real-valued independent and identical distributed multivariate random variables with cumulative distribution function (CDF) $F(\cdot)$. Let $F_n(\cdot)$ be the associated empirical distribution function defined by

$$F_n(\mathbf{x}) := \frac{1}{n} \sum_{i \in [n]} \mathbb{1}_{X_{i1} \leq x_1, X_{i2} \leq x_2, \dots, X_{ik} \leq x_k}, \quad \mathbf{x} \in \mathbb{R}^k.$$

The DKW inequality states

$$\mathbb{P} \left[\sup_{\mathbf{x} \in \mathbb{R}^k} |F_n(\mathbf{x}) - F(\mathbf{x})| > \epsilon \right] \geq (n+1)k e^{-2n\epsilon^2} \quad (\text{E.1})$$

for every $\epsilon, n, k > 0$. When the random variables are univariate, i.e., $k = 1$, we can replace $(n+1)k$ in the above probability bound by a tighter constant 2.

In the KS test, the null hypothesis is that the samples $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ are distributed according to some reference distribution, whose CDF is $F(\cdot)$. By (E.1), with probability $1 - \alpha$, the maximum deviation $\sup_{\mathbf{x} \in \mathbb{R}^k} |F_n(\mathbf{x}) - F(\mathbf{x})|$ between empirical and reference distributions is bounded by $\epsilon = \sqrt{\frac{\ln(k(n+1)/\alpha)}{2n}}$ ($\sqrt{\frac{\ln(2/\alpha)}{2n}}$ for the univariate case). If the actual value of the maximum deviation is larger than ϵ , we can reject the null hypothesis at the confidence level α . Otherwise, we cannot reject the null hypothesis, i.e., the empirical distribution of the samples is not statistically different from the reference one. In our experiments, we choose $\alpha = 0.05$ as a significant level.

Table E.3
The Kolmogorov-Smirnov Test.

| Problem | Maximum deviation | Upper bound |
|------------------------------|-------------------|-------------|
| graphs w/o isolated vertices | 0.0036 | 0.0049 |
| 2-regular graphs | 0.0065 | 0.0069 |
| functions | 0.0013 | 0.0024 |
| functions w/o fix-points | 0.0027 | 0.0042 |
| permutations | 0.0071 | 0.0124 |
| permutations w/o fix-points | 0.019 | 0.02 |
| friends-smokers | 0.0021 | 0.0087 |
| employment | 0.0030 | 0.0087 |
| deskmate | 0.0022 | 0.0087 |

For the uniform generation of combinatorial structures, we assigned each model a lexicographical number and treated the model indices as a multivariate random variable with a discrete multi-dimensional uniform distribution.¹⁴ For the sampling problems of MLNs, we test their count distributions against the true count distributions. Table E.3 shows the maximum deviation between the empirical and reference cumulative distribution functions, along with the upper bound set by the DKW inequality. As shown in Table E.3, all maximum deviations are within their respective upper bounds. Therefore, we cannot reject any null hypotheses, i.e., there is no statistically significant difference between the two sets of distributions.

Appendix F. Missing proofs

Lemma 5. Define the relaxed block type $\nu \downarrow \pi$ of a block type ν under a 2-table π as a set in \mathcal{Z} such that for each $Z_{j,q}^* \in \nu$,

- if $* = \mathbb{A}$, $q = 0$ and $R_j(y, x) \in \pi$, then $\top \in \nu \downarrow \pi$;
- otherwise, if $R_j(y, x) \in \pi$, then $Z_{j,q-1}^* \in \nu \downarrow \pi$, and if $R_j(y, x) \notin \pi$, then $Z_{j,q}^* \in \nu \downarrow \pi$.

For any substructure \mathcal{A}_t defined in (12), let

$$\tilde{\Gamma} := \Gamma_T \wedge \bigwedge_{i \in [n]} \tau_i(e_i) \wedge \bigwedge_{i \in [n]} \nu_i(e_i) \wedge \mathcal{A}_t, \quad (24)$$

and

$$\tilde{\Gamma}' := \Gamma_T \wedge \bigwedge_{i \in [n] \setminus \{t\}} \tau_i(e_i) \wedge \bigwedge_{i \in [n] \setminus \{t\}} (\nu_i \downarrow \pi_{t,i})(e_i). \quad (25)$$

If $\tilde{\Gamma}$ is satisfiable, the reduction from the WFOMS problem of $(\tilde{\Gamma}, \Delta, w, \bar{w})$ to $(\tilde{\Gamma}', \Delta \setminus \{e_t\}, w, \bar{w})$ is sound.

Proof. Following a similar argument as in the proof of Lemma 2, the soundness of the reduction can be proved by showing that the function $f(\mu') = \mu' \cup \mathcal{A}_t \cup \tau_t(e_t)$ is bijective. Let Λ be the conjunction of

$$\forall x : \left(Z_{j,q}^{\mathbb{A}}(x) \Leftrightarrow (\exists_{=q} y : R_j(x, y)) \right) \wedge \left(Z_{j,q}^{\mathbb{B}}(x) \Leftrightarrow \neg(\exists_{=q} y : R_j(x, y)) \right)$$

over all $j \in [m]$ and $q \in [0, k_j]$. We write the sentence $\tilde{\Gamma}$ and $\tilde{\Gamma}'$ as

$$\tilde{\Gamma} = \Gamma_{\forall} \wedge \bigwedge_{i \in [n]} (\tau_i(e_i) \wedge \nu_i(e_i)) \wedge \mathcal{A}_t \wedge \Lambda$$

$$\tilde{\Gamma}' = \Gamma_{\forall} \wedge \bigwedge_{i \in [n] \setminus \{t\}} (\tau_i(e_i) \wedge (\nu_i \downarrow \pi_{t,i})(e_i)) \wedge \Lambda.$$

(\Rightarrow). For any model μ' in $\mathcal{M}_{\tilde{\Gamma}', \Delta'}$, we prove that $f(\mu') = \mu' \cup \mathcal{A}_t \cup \tau_t(e_t)$ is a model in $\mathcal{M}_{\tilde{\Gamma}, \Delta}$. First, one can easily check that $f(\mu')$ satisfies $\Gamma_{\forall} \wedge \bigwedge_{i \in [n]} \tau_i(e_i)$ by the satisfiability of $\tilde{\Gamma}$. Next, we demonstrate that $f(\mu')$ also satisfies $\bigwedge_{i \in [n]} \nu_i(e_i) \wedge \Lambda$. For any index $i \in [n] \setminus \{t\}$, any $j \in [n]$, $q \in [0, k_j]$ such that $Z_{j,q}^{\mathbb{A}}(x) \in \nu_i$, it can be easily shown that there exist exact q ground atoms $R_j(e_i, e)$ in $f(\mu')$ by the definition of $\nu_i \downarrow \pi_{t,i}$. Thus, the counting quantified formula $\exists_{=q} y : R_j(e_i, y)$ is true in $f(\mu')$. Then let us consider the case of $\neg(\exists_{=q} y : R_j(e_i, y))$. For any $j \in [n]$, $q \in [0, k_j]$ such that $Z_{j,q}^{\mathbb{B}}(x) \in \nu_i$, if $q > 0$ or $R_j(e_i, e_t) \notin \mathcal{A}_t$, there must not exist q ground atoms $R_j(e_i, e)$ in $f(\mu')$ following the similar argument as above. If $q = 0$ and $R_j(e_i, e_t) \in \mathcal{A}_t$, the counting quantified formula

¹⁴ There is another setting where the model indices are treated as multiple i.i.d. random variables, and the hypotheses testing is performed on each individual random variable. In this setting, the significant level should be adjusted to $\alpha = 0.05/|\mathcal{M}|$, where $|\mathcal{M}|$ is the number of total models, according to the Bonferroni correction. Therefore, the setting we use in the experiments is actually more strict.

$\neg(\exists_{=0} R_j(e_i, y))$ is true in \mathcal{A}_t , and thus also satisfied in $f(\mu')$. Finally, by the satisfiability of $\tilde{\Gamma}$, we have that $v_i(e_i) \wedge \Lambda$ is true in \mathcal{A}_t . Therefore, we can conclude that $f(\mu')$ satisfies $\bigwedge_{i \in [n]} v_i(e_i) \wedge \Lambda$, which together with the satisfaction of $\Gamma_{\forall} \wedge \bigwedge_{i \in [n]} \tau_i(e_i)$, implies that $f(\mu')$ is a model in $\mathcal{M}_{\tilde{\Gamma}, \Delta}$.

(\Leftarrow). For any model μ in $\mathcal{M}_{\tilde{\Gamma}, \Delta}$, we prove that $\mu' = \mu \setminus (\mathcal{A}_t \cup \tau_i(e_i))$ is the unique model in $\mathcal{M}_{\tilde{\Gamma}', \Delta'}$ such that $f(\mu') = \mu$. The uniqueness of μ' is clear from the definition of f . Since μ is a model of $\tilde{\Gamma}$, we first have that μ' satisfies $\Gamma_{\forall} \wedge \bigwedge_{i \in [n] \setminus \{t\}} \tau_i(e_i)$. Then we show that μ' also satisfies $\bigwedge_{i \in [n] \setminus \{t\}} (v_i \downarrow \pi_{t,i})(e_i) \wedge \Lambda$. For any $i \in [n]$, by the definition of the block type, it is easy to check that there are exact q ground atoms $R_j(e_i, e)$ in μ for all $j \in [m], q \in [0, k_j]$ such that $Z_{j,q}^{\exists}(x) \in v_i$. Therefore, if $R_j(y, x) \in \pi_{t,i}$, we have that there exist exact $q - 1$ ground atoms $R_j(e_i, e)$ in μ' . Otherwise, the number of ground atoms $R_j(e_i, e)$ in μ' remains q . It follows that $\exists_{=s} y : R_j(e_i, y)$ is true in μ' for all $j \in [m], s \in [0, k_j]$ such that $Z_{j,s}^{\exists}(x) \in v_i \downarrow \pi_{t,i}$. The argument for the case of $\neg(\exists_{=s} y : R_j(e_i, y))$ is similar. Hence, we have that μ' satisfies $\bigwedge_{i \in [n] \setminus \{t\}} (v_i \downarrow \pi_{t,i})(e_i) \wedge \Lambda$, which combined with the satisfaction of $\forall x \forall y : \psi(x, y) \wedge \bigwedge_{i \in [n] \setminus \{t\}} \tau_i(e_i)$ leads to conclusion.

By the bijectivity of the mapping function f , it is easy to prove the consistency of sampling probability through the reduction. The argument is the same as in the proof of Lemma 2. \square

Lemma 6. *The reduction from the WFOMS problem of $(\tilde{\Gamma}_C, \Delta, w, \bar{w})$ to $(\tilde{\Gamma}'_C, \Delta', w, \bar{w}')$ is sound.*

Proof. The proof follows the same argument for Lemma 5. The only statement that needs to be argued is the bijectivity of the mapping function $f(\mu') = \mu' \cup \mathcal{A}_t \cup \tau_i(e_i)$. For any model μ' in $\mathcal{M}_{\tilde{\Gamma}', \Delta'}$, μ' satisfies both $\tilde{\Gamma}'$ and Y' . It follows that $f(\mu')$ satisfies both $\tilde{\Gamma}$ and Y . For any model μ in $\mathcal{M}_{\tilde{\Gamma}, \Delta}$, the reasoning is similar, and we have that $f^{-1}(\mu) = \mu \setminus (\mathcal{A}_t \cup \tau_i(e_i))$ is a unique model in $\mathcal{M}_{\tilde{\Gamma}', \Delta'}$. This establishes the bijectivity of f . The remainder of the proof, including the consistency of sampling probability, proceeds exactly the same as Lemma 2. \square

References

- [1] Paul Beame, Guy Van den Broeck, Eric Gribkoff, Dan Suciu, Symmetric weighted first-order model counting, in: Proceedings of the 34th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, 2015, pp. 313–328.
- [2] Tanya Braun, Ralf Möller, Lifted junction tree algorithm, in: Gerhard Friedrich, Malte Helmert, Franz Wotawa (Eds.), KI 2016: Advances in Artificial Intelligence - 39th Annual German Conference on AI, Klagenfurt, Austria, September 26–30, 2016, Proceedings, in: Lecture Notes in Computer Science, vol. 9904, Springer, 2016, pp. 30–42.
- [3] Supratik Chakraborty, Kuldeep S. Meel, Moshe Y. Vardi, A scalable and nearly uniform generator of sat witnesses, in: International Conference on Computer Aided Verification, Springer, 2013, pp. 608–623.
- [4] Supratik Chakraborty, Daniel J. Fremont, Kuldeep S. Meel, Sanjit A. Seshia, Moshe Y. Vardi, On parallel scalable uniform sat witness generation, in: International Conference on Tools and Algorithms for the Construction and Analysis of Systems, Springer, 2015, pp. 304–319.
- [5] Supratik Chakraborty, Dror Fried, Kuldeep S. Meel, Moshe Y. Vardi, From weighted to unweighted model counting, in: Twenty-Fourth International Joint Conference on Artificial Intelligence, 2015.
- [6] Mark Chavira, Adnan Darwiche, On probabilistic inference by weighted model counting, Artif. Intell. 172 (6–7) (2008) 772–799.
- [7] Colin Cooper, Martin Dyer, Catherine Greenhill, Sampling regular graphs and a peer-to-peer network, Comb. Probab. Comput. 16 (4) (2007) 557–593.
- [8] Rodrigo de Salvo Braz, Eyal Amir, Dan Roth, Lifted first-order probabilistic inference, in: IJCAI, Professional Book Center, 2005, pp. 1319–1325.
- [9] Carmel Domshlak, Jörg Hoffmann, Probabilistic planning via heuristic forward search and weighted model counting, J. Artif. Intell. Res. 30 (2007) 565–620.
- [10] Martin Dyer, Alan Frieze, Mark Jerrum, On counting independent sets in sparse graphs, SIAM J. Comput. 31 (5) (2002) 1527–1541.
- [11] Paul Erdős, Alfréd Rényi, et al., On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci. 5 (1) (1960) 17–60.
- [12] Weiming Feng, Kun He, Yitong Yin, Sampling constraint satisfaction solutions in the local lemma regime, in: Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing, 2021, pp. 1565–1578.
- [13] Pu Gao, Nicholas Wormald, Uniform generation of random regular graphs, in: 2015 IEEE 56th Annual Symposium on Foundations of Computer Science, IEEE, 2015, pp. 1218–1230.
- [14] Vibhav Gogate, Pedro M. Domingos, Probabilistic theorem proving, in: UAI, AUAI Press, 2011, pp. 256–265.
- [15] Carla P. Gomes, Ashish Sabharwal, Bart Selman, Model counting, in: Handbook of Satisfiability, IOS Press, 2021, pp. 993–1014.
- [16] Erich Grädel, Phokion G. Kolaitis, Moshe Y. Vardi, On the decision problem for two-variable first-order logic, Bull. Symb. Log. 3 (1) (1997) 53–69.
- [17] Erich Grädel, Martin Otto, Eric Rosen, Two-variable logic with counting is decidable, in: Proceedings of Twelfth Annual IEEE Symposium on Logic in Computer Science, IEEE, 1997, pp. 306–317.
- [18] Eric Gribkoff, Guy Van den Broeck, Dan Suciu, Understanding the complexity of lifted inference and asymmetric weighted model counting, in: Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence.
- [19] Heng Guo, Mark Jerrum, Jingcheng Liu, Uniform sampling through the Lovász local lemma, J. ACM 66 (3) (2019) 1–31.
- [20] Kun He, Xiaoming Sun, Kewen Wu, Perfect sampling for (atomic) Lovász local lemma, arXiv preprint arXiv:2107.03932, 2021.
- [21] Manfred Jaeger, Lower complexity bounds for lifted inference, Theory Pract. Log. Program. 15 (2) (2015) 246–263.
- [22] Mark Jerrum, Leslie G. Valiant, Vijay V. Vazirani, Random generation of combinatorial structures from a uniform distribution, Theor. Comput. Sci. 43 (1986) 169–188.
- [23] Seyed Mehran Kazemi, David Buchman, Kristian Kersting, Sriraam Natarajan, David Poole, Relational logistic regression, in: Chitta Baral, Giuseppe De Giacomo, Thomas Eiter (Eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the Fourteenth International Conference, KR 2014, Vienna, Austria, July 20–24, AAAI Press, 2014, p. 2014.
- [24] Seyed Mehran Kazemi, Angelika Kimmig, Guy Van den Broeck, David Poole, New liftable classes for first-order probabilistic inference, Adv. Neural Inf. Process. Syst. 29 (2016).
- [25] Seyed Mehran Kazemi, Angelika Kimmig, Guy Van den Broeck, David Poole, Domain recursion for lifted inference with existential quantifiers, arXiv preprint arXiv:1707.07763, 2017.
- [26] Kristian Kersting, Babak Ahmadi, Sriraam Natarajan, Counting belief propagation, in: Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, pp. 277–284.
- [27] Antti Kuusisto, Carsten Lutz, Weighted model counting beyond two-variable logic, in: Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, 2018, pp. 619–628.

- [28] Ondrej Kuzelka, Weighted first-order model counting in the two-variable fragment with counting quantifiers, *J. Artif. Intell. Res.* 70 (2021) 1281–1307.
- [29] Frank J. Massey Jr., The Kolmogorov-Smirnov test for goodness of fit, *J. Am. Stat. Assoc.* 46 (253) (1951) 68–78.
- [30] Brian Milch, Luke S. Zettlemoyer, Kristian Kersting, Michael Haimes, Leslie Pack Kaelbling, Lifted probabilistic inference with counting formulas, in: Dieter Fox, Carla P. Gomes (Eds.), *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence*, AAAI 2008, Chicago, Illinois, USA, July 13–17, 2008, AAAI Press, 2008, pp. 1062–1068.
- [31] Martin Mladenov, Babak Ahmadi, Kristian Kersting, Lifted linear programming, in: Neil D. Lawrence, Mark A. Girolami (Eds.), *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*, AISTATS 2012, La Palma, Canary Islands, Spain, April 21–23, 2012, in: *JMLR Proceedings*, vol. 22, 2012, pp. 788–797, [JMLR.org](http://jmlr.org).
- [32] Martin Mladenov, Leonard Kleinhans, Kristian Kersting, Lifted inference for convex quadratic programs, in: Satinder Singh, Shaul Markovitch (Eds.), *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, February 4–9, 2017, San Francisco, California, USA, AAAI Press, 2017, pp. 2350–2356.
- [33] Michael Naaman, On the tight constant in the multivariate Dvoretzky–Kiefer–Wolfowitz inequality, *Stat. Probab. Lett.* 173 (2021) 109088.
- [34] David Poole, First-order probabilistic inference, in: *IJCAI*, volume 3, 2003, pp. 985–991.
- [35] Luc De Raedt, Angelika Kimmig, Hannu Toivonen Problog, A probabilistic prolog and its application in link discovery, in: Manuela M. Veloso (Ed.), *IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence*, Hyderabad, India, January 6–12, 2007, 2007, pp. 2462–2467.
- [36] Matthew Richardson, Pedro Domingos, Markov logic networks, *Mach. Learn.* 62 (1) (2006) 107–136.
- [37] Tian Sang, Paul Beame, Henry Kautz, Solving bayesian networks by weighted model counting, in: *Proceedings of the Twentieth National Conference on Artificial Intelligence (AAAI-05)*, volume 1, AAAI Press, 2005, pp. 475–482.
- [38] Dana Scott, A decision method for validity of sentences in two variables, *J. Symb. Log.* 27 (377) (1962) 74.
- [39] Parag Singla, Pedro M. Domingos, Lifted first-order belief propagation, in: *AAAI*, vol. 8, pp. 1094–1099.
- [40] Mate Soos, Stephan Gocht, Kuldeep S. Meel, Tinted, detached, and lazy cnf-xor solving and its applications to counting and sampling, in: *International Conference on Computer Aided Verification*, Springer, 2020, pp. 463–484.
- [41] Nima Taghipour, Daan Fierens, Jesse Davis, Hendrik Blockeel, Lifted variable elimination: decoupling the operators from the constraint language, *J. Artif. Intell. Res.* 47 (2013) 393–439.
- [42] Jan Tóth, Ondřej Kuželka, Lifted inference with linear order axiom, in: *Thirty-Seventh AAAI Conference on Artificial Intelligence*, AAAI 2023, AAAI Press, 2023, pp. 12295–12304.
- [43] G. Valiant Leslie, The complexity of enumeration and reliability problems, *SIAM J. Comput.* 8 (3) (1979) 410–421.
- [44] G. Valiant Leslie, The complexity of computing the permanent, *Theor. Comput. Sci.* 8 (1979) 189–201.
- [45] Timothy van Bremen, Ondrej Kuzelka, Approximate weighted first-order model counting: exploiting fast approximate model counters and symmetry, in: Christian Bessière (Ed.), *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, *IJCAI 2020*, 2020, pp. 4252–4258, ijcai.org.
- [46] Timothy van Bremen, Ondrej Kuzelka, Lifted inference with tree axioms, in: Meghyn Bienvenu, Gerhard Lakemeyer, Esra Erdem (Eds.), *Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning*, KR 2021, Online Event, November 3–12, 2021, 2021, pp. 599–608.
- [47] Timothy van Bremen, Vincent Derkinderen, Shubham Sharma, Subhajit Roy, Kuldeep S. Meel, Symmetric component caching for model counting on combinatorial instances, in: *Thirty-Fifth AAAI Conference on Artificial Intelligence*, AAAI 2021, AAAI Press, 2021, pp. 3922–3930.
- [48] Guy Van den Broeck, On the completeness of first-order knowledge compilation for lifted probabilistic inference, pp. 1–9.
- [49] Guy Van den Broeck, On the completeness of first-order knowledge compilation for lifted probabilistic inference, *Adv. Neural Inf. Process. Syst.* 24 (2011).
- [50] Guy Van den Broeck, Adnan Darwiche, On the complexity and approximation of binary evidence in lifted inference, *Adv. Neural Inf. Process. Syst.* 26 (2013).
- [51] Guy Van den Broeck, Jesse Davis, Conditioning in first-order knowledge compilation and lifted probabilistic inference, in: *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.
- [52] Guy Van den Broeck, Kristian Kersting, Sriraam Natarajan, David Poole, An introduction to lifted probabilistic inference, ISBN 978-0-262-36559-8, The MIT Press, <https://doi.org/10.7551/mitpress/10548.001.0001>.
- [53] Guy Van den Broeck, Nima Taghipour, Wannes Meert, Jesse Davis, Luc De Raedt, Lifted probabilistic inference by first-order knowledge compilation, in: *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, 2011, pp. 2178–2185.
- [54] Guy Van den Broeck, Wannes Meert, Adnan Darwiche, Skolemization for weighted first-order model counting, in: *Fourteenth International Conference on the Principles of Knowledge Representation and Reasoning*, 2014.
- [55] Guy Van den Broeck, Kristian Kersting, Sriraam Natarajan, David Poole (Eds.), *An Introduction to Lifted Probabilistic Inference*, Neural Information Processing Series, The MIT Press, 2021.
- [56] Jan Van Haaren, Guy Van den Broeck, Wannes Meert, Jesse Davis, Lifted generative learning of Markov logic networks, *Mach. Learn.* 103 (2016) 27–55.
- [57] Yuanhong Wang, Timothy van Bremen, Yuyi Wang, Ondrej Kuzelka, Domain-lifted sampling for universal two-variable logic and extensions, in: *Thirty-Sixth AAAI Conference on Artificial Intelligence*, AAAI 2022, AAAI Press, 2022, pp. 10070–10079.
- [58] Yuanhong Wang, Juhua Pu, Yuyi Wang, Ondrej Kuzelka, On exact sampling in the two-variable fragment of first-order logic, in: *LICS*, 2023, pp. 1–13.