Optimal Union Probability Interval Is NP-Hard

Petteri Kaski
Department of Computer Science, Aalto University, Espoo, Finland

Heikki Mannila
Department of Computer Science, Aalto University, Espoo, Finland

Chandra Kanta Mohapatra
Department of Computer Science, Aalto University, Espoo, Finland

Abstract

A problem dating back to Boole [*Laws of Thought*, Walton & Maberly, 1854] is what can be computed about the probability of a finite union of events when given as input the probabilities of intersections of some of the events. The modern geometric study of the problem can be traced back to Hailperin [*Amer. Math. Monthly* 2 (1965) 343–359] who phrased the problem in the language of linear programming and generalized it to logical formulas of the events other than disjunction, heralding a substantial body of work in probabilistic logic [Nilsson, *Artif. Intell.* 28 (1986) 71–87], including the *probabilistic satisfiability problem* of Georgakopoulos, Kavvadis, and Papadimitriou [*J. Complexity* 4 (1988) 1–11], as well as fundamental connections to the geometry of metrics via cut and correlation polytopes [Deza and Laurent, *Geometry of Cuts and Metrics*, Springer, 1997] and to the study of marginal polytopes in graphical models of machine learning [Wainwright and Jordan, *Found. Trends Mach. Learn.* 1 (2008) 1–305]. This paper (i) describes the pertinent geometry of Boole's problem via coordinate projections of an elementary polytope arising essentially from Hailperin's linear program on the atoms of a Venn diagram, and (ii) shows that computing the optimal interval for the union probability is NP-hard, resolving an apparent gap in the literature highlighted by Pitowsky [*Math. Programming* 50 (1991) 395–414] and Boros *et al.* [*Math. Oper. Res.* 39 (2014) 1311–1329 and 51 (2026) 134–148].

2012 ACM Subject Classification Theory of computation → Problems, reductions and completeness; Mathematics of computing → Combinatorial optimization; Mathematics of computing → Probability and statistics; Theory of computation → Linear programming

Keywords and phrases combinatorial probability, computational complexity, correlation polytope, generalized Bonferroni inequalities, graphical models, inclusion-exclusion, intersection probability, linear programming, union polytope

Digital Object Identifier 10.4230/LIPIcs.ESA.2026.129

Funding *Heikki Mannila*: Heikki Mannila is supported by the Technology Industries of Finland Centennial Foundation.
Chandra Kanta Mohapatra: Chandra Kanta Mohapatra is supported by Helsinki Institute for Information Technology (HIIT) postdoctoral fellowship.

Acknowledgements We thank the anonymous reviewers for their insightful comments and careful reading of the manuscript.

1 Introduction

Suppose we have n events $B_1, B_2, \ldots, B_n$ in a probability space Ω and write $|B|$ for the probability of an event B. The classical inclusion-exclusion formula

$$|\cup_{i\in[n]} B_i| = \sum_{\emptyset\neq S\subseteq[n]} (-1)^{|S|+1} |\cap_{i\in S} B_i| \tag{1}$$

efficiently solves for the probability of the union event $\cup_{i\in[n]}B_i$, given as input the probabilities of all the intersection events $B_S := \cap_{i\in S}B_i$ for all $\emptyset \neq S \subseteq [n] := \{1, 2, \ldots, n\}$ and $B_\emptyset := \Omega$. But what can one compute about the union probability when one is given as input the probabilities of only some of the intersection events? More specifically, suppose we are given as input a set family $\emptyset \neq \mathcal{F} \subseteq 2^{[n]} \setminus \{\emptyset\}$ together with the intersection probability $|B_S|$ for each $S \in \mathcal{F}$, what can we say about $|\cup_{i\in[n]} B_i|$?

This problem originates to Boole [9], who observed that when only individual event probabilities are given, the union probability satisfies the inequalities

$$\max_{i\in[n]} |B_i| \leq |\cup_{i\in[n]} B_i| \leq \min\bigl(1, \sum_{i\in[n]} |B_i|\bigr)\,, \tag{2}$$

where the upper bound in particular is the familiar *union bound*, or *Boole's inequality.* These inequalities are also best possible, as was shown by Fréchet [23, 24]. Boole's inequality generalizes to the classical *Bonferroni inequalities* [8]

$$\begin{aligned} |\cup_{i\in[n]} B_i| \;&\leq \sum_{\substack{S\subseteq[n]\\ 1\leq|S|\leq k}} (-1)^{|S|+1}|B_S| \quad \text{for } k \in [n] \text{ odd, and}\\ |\cup_{i\in[n]} B_i| \;&\geq \sum_{\substack{S\subseteq[n]\\ 1\leq|S|\leq k}} (-1)^{|S|+1}|B_S| \quad \text{for } k \in [n] \text{ even}\,, \end{aligned}$$

which however are not best possible; a substantial literature exists on improved and generalized Bonferroni inequalities as well as approximate versions of the inclusion–exclusion formula (cf. our discussion of related work below). Our interest in this paper is the geometry and computational complexity of the *best-possible* inequalities for given $\emptyset \neq \mathcal{F} \subseteq 2^{[n]} \setminus \{\emptyset\}$ and $|B_S|$ for each $S \in \mathcal{F}$. In particular, we show that the best-possible inequalities are NP-hard to compute already when $\mathcal{F}$ consists of sets of size at most two. We refer to the task of computing the best-possible inequalities as *Boole's problem* in what follows.

A geometric representation of Boole's problem can be traced back to Hailperin [32], who observed that the best-possible inequalities can be derived by linear programming.[1] Hailperin's key combinatorial observation is that it suffices to study the probabilities of the *Venn atoms* (the 2^n cells of a Venn diagram easily visualized for $n = 2, 3$) of the n events $B_1, B_2, \ldots, B_n$. More precisely, associate with each subset $T \subseteq [n]$ the *atomic event* or *atom*

$$X_T = (\cap_{i\in T} B_i) \cap (\cap_{i\in[n]\setminus T}\, \Omega\setminus B_i)$$

and observe that the atoms form a set partition of Ω. Moreover, for each $S \in \mathcal{F}$, the intersection event $B_S = \cap_{i\in S} B_i$ partitions into atoms as $B_S = \cup_{S\subseteq T} X_T$. Geometrically, introducing real unknowns x_T and b_S for the pertinent probabilities, the partitioning immediately gives rise to the *Venn polytope* $\tau^{(\mathcal{F})} \subseteq \mathbb{Q}^{|\mathcal{F}|+2^n}$ defined by the hyperplanes and halfspaces

$$\sum_{T\subseteq[n]} x_T = 1\,, \quad \sum_{S\subseteq T} x_T = b_S \text{ for all } S \in \mathcal{F}\,, \quad \text{and} \quad x_T \geq 0 \text{ for all } T \subseteq [n]\,. \tag{3}$$

We stress that both the atom probabilities x_T as well as the intersection probabilities b_S are coordinates of the Venn polytope.

[1] Hailperin in fact observed that linear programming can be used to study the probability of any logical function of the events $B_1, B_2, \ldots, B_n$, not just the disjunction (union) of the events, and the given input can similarly consist of the probabilities of any logical functions of the events, not just conjunctions (intersections) of events. For simplicity, we restrict to study the complexity of the union probability with given intersection probabilities.

▶ **Example 1** (A Venn polytope). The Venn polytope $\tau^{(\mathcal{F})}$ for $\mathcal{F} = \{\{1\},\{2\},\{1,2\}\}$ in its halfspace-representation (left) and vertex-representation (right):

$$\left[\begin{array}{r} x_\emptyset + x_{\{1\}} + x_{\{2\}} + x_{\{1,2\}} = 1 \\ x_{\{1\}} + x_{\{1,2\}} = b_{\{1\}} \\ x_{\{2\}} + x_{\{1,2\}} = b_{\{2\}} \\ x_{\{1,2\}} = b_{\{1,2\}} \\ x_\emptyset\,, x_{\{1\}}\,, x_{\{2\}}\,, x_{\{1,2\}} \geq 0 \end{array}\right] \qquad \begin{array}{ccccccc} b_{\{1\}} & b_{\{2\}} & b_{\{1,2\}} & x_\emptyset & x_{\{1\}} & x_{\{2\}} & x_{\{1,2\}} \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}.$$

Now, following Hailperin, since the union event partitions as $\cup_{i\in[n]}B_i = \cup_{\emptyset\neq T\subseteq[n]}X_T = \Omega\setminus X_\emptyset$, the best-possible inequalities for the union probability $u = |\cup_{i\in[n]} B_i|$ are obtained from the linear program that minimizes/maximizes the objective function $u = \sum_{\emptyset\neq T\subseteq[n]} x_T = 1 - x_\emptyset$ over the constraints (3) and $b_S = |B_S|$ for each $S \in \mathcal{F}$ in the given input; we refer to this linear program as *Hailperin's* linear program. We say that an abstract vector $b = (b_S \in [0,1] : S \in \mathcal{F})$ of probabilities is *feasible* if Hailperin's linear program is feasible. In particular, b is feasible if and only if there exists a probability space Ω and n events $B_1, B_2, \ldots, B_n$ in the space that *realize* b with $b_S = |B_S|$ for all $S \in \mathcal{F}$.

While Hailperin's characterization of the optimal solution to Boole's problem by linear programming is concise and elegant, the exponential size of the linear program in the number of events n leaves open the question whether more efficient optimal characterizations are available. We show that computing the minimum/maximum union probability is NP-hard, thus making it unlikely that efficient optimal characterizations are available in general. Our main results is as follows.

▶ **Theorem 2** (Main; NP-hardness of Boole's union probability problem). *Given as input a set family $\emptyset \neq \mathcal{F} \subseteq 2^{[n]} \setminus \{\emptyset\}$ and a feasible rational vector $b = (b_S \in [0,1] : S \in \mathcal{F})$, each of the following two problems is NP-hard to solve:*
1. *determine the minimum union probability for a realization of b, and*
2. *determine the maximum union probability for a realization of b.*

Moreover, NP-hardness holds even when $\mathcal{F}$ consists of sets of size at most two.

▶ Remark 3 (Polynomial-time solvable special cases). In contrast to NP-hardness, we observe that the union probability problem is efficiently solvable in time polynomial in the size of the input $(b, \mathcal{F})$, for example, when all the intersection probabilities are given, by the principle of inclusion and exclusion (1). More generally, when $\mathcal{F}$ has size exponential in the number of events n, we can solve Hailperin's linear program in time polynomial in the assumed-exponential input size.

Issues related to the complexity of Boole's problem have been addressed in previous research. Pitowsky [53] highlights Boole's problem and observes its probable intractability due to its close relationship to the geometry of correlation polytopes. Boros, Scozzari, Tardella, and Veneziani [12] and Boros and Lee [10] highlight the open complexity status of the problem even in the case of sets of size at most two. They also observe that a closely related problem, deciding the feasibility of a given abstract vector b of probabilities, is NP-hard, by either direct proof [53] or reduction from the separation problem of either a cut polytope (cf. Deza and Laurent [18]) or a Boolean quadric polytope (cf. Jaumard, Hansen, and Poggi de Aragão [36]). Among further closely related NP-hard problems is the *probabilistic satisfiability* (PSAT) problem of Georgakopoulos, Kavvadias, and Papadimitriou [29], which, given as input m clauses on n Boolean variables, together with a given probability for each of the m clauses, asks whether there exists a distribution of probabilities on the 2^n possible truth assignments to the Boolean variables such that each clause is satisfied with exactly the given probability.

Moreover, Kavvadias and Papadimitriou [38] have shown that PSAT remains NP-hard even when all clauses consist of positive literals only and contain at most two literals each, which also directly establishes NP-hardness of deciding feasibility of a given abstract vector b in our present setting. Our present contribution thus is essentially the fine-grained result that, when given a *feasible* vector b as input, it yet remains NP-hard to solve Hailperin's linear program to optimality (minimum or maximum of the union probability $u = 1 - x_\emptyset$). This contribution is motivated by data analysis applications in particular, where the vector b is obtained from existing data, and hence the feasibility of b is known to hold.

Geometrically, we observe that deciding whether a given abstract vector b is feasible and deciding whether a given feasible vector b has a realization with union probability u both admit characterization via polytopes that are elementary coordinate projections of the Venn polytope $\tau^{(\mathcal{F})}$. We recall that coordinate projections of a polytope are easily obtained from the vertex representation by projecting the vertices and taking the convex hull of the projections; we also recall Example 1 for a Venn polytope in vertex representation.

The *union polytope* $\sigma^{(\mathcal{F})} \subseteq \mathbb{Q}^{|\mathcal{F}|+1}$ is obtained by projecting the Venn polytope $\tau^{(\mathcal{F})}$ to coordinates $x_\emptyset$ and b_S for all $S \in \mathcal{F}$. The (generalized[2]) *correlation polytope* $\rho^{(\mathcal{F})} \subseteq \mathbb{Q}^{|\mathcal{F}|}$ is obtained by projecting the Venn polytope $\tau^{(\mathcal{F})}$ to coordinates b_S for all $S \in \mathcal{F}$. We immediately observe the coordinate-projection chain $\rho^{(\mathcal{F})} \leq \sigma^{(\mathcal{F})} \leq \tau^{(\mathcal{F})}$. Correlation polytopes [53], also studied more recently under the name *marginal polytopes* in the machine-learning community [69], are a well-known and systematically studied class of polytopes via cut polytopes and cut cones in the geometry of metrics [18], whereas the union polytopes are apparently a less studied family of polytopes.

▶ **Example 4** (Union polytopes and correlation polytopes)**.** Halfspace-representations of union polytopes $\sigma^{(\mathcal{F})}$ correlation polytopes $\rho^{(\mathcal{F})}$, writing $u = 1 - x_\emptyset$ for the union probability:

$\mathcal{F}$	$\{1,2\}$	$\{1\},\{2\}$	$\{1\},\{1,2\}$	$\{1\},\{2\},\{1,2\}$
$\sigma^{(\mathcal{F})}$	$\begin{bmatrix} b_{\{1,2\}} \geq 0 \\ b_{\{1,2\}} \leq u \\ u \leq 1 \end{bmatrix}$	$\begin{bmatrix} b_{\{1\}} \leq u \\ b_{\{2\}} \leq u \\ u \leq b_{\{1\}} + b_{\{2\}} \\ u \leq 1 \end{bmatrix}$	$\begin{bmatrix} b_{\{1,2\}} \geq 0 \\ b_{\{1,2\}} \leq b_{\{1\}} \\ b_{\{1\}} \leq u \\ u \leq 1 \end{bmatrix}$	$\begin{bmatrix} b_{\{1\}} + b_{\{2\}} = u + b_{\{1,2\}} \\ b_{\{1,2\}} \geq 0 \\ b_{\{1,2\}} \leq b_{\{1\}} \\ b_{\{1,2\}} \leq b_{\{2\}} \\ b_{\{1\}} + b_{\{2\}} \leq 1 + b_{\{1,2\}} \end{bmatrix}$
$\rho^{(\mathcal{F})}$	$\begin{bmatrix} b_{\{1,2\}} \geq 0 \\ b_{\{1,2\}} \leq 1 \end{bmatrix}$	$\begin{bmatrix} b_{\{1\}} \geq 0 \\ b_{\{2\}} \geq 0 \\ b_{\{1\}} \leq 1 \\ b_{\{2\}} \leq 1 \end{bmatrix}$	$\begin{bmatrix} b_{\{1,2\}} \geq 0 \\ b_{\{1,2\}} \leq b_{\{1\}} \\ b_{\{1\}} \leq 1 \end{bmatrix}$	$\begin{bmatrix} b_{\{1,2\}} \geq 0 \\ b_{\{1,2\}} \leq b_{\{1\}} \\ b_{\{1,2\}} \leq b_{\{2\}} \\ b_{\{1\}} + b_{\{2\}} \leq 1 + b_{\{1,2\}} \end{bmatrix}$

In particular, we observe both the Boole–Frechét bound (2) in $\sigma^{(\{1\},\{2\})}$ and the inclusion–exclusion formula (1) for $n = 2$ in $\sigma^{(\{1\},\{2\},\{1,2\})}$ above. We postpone a short generic analysis of the geometry of the polytopes to Section 2.

Geometrically, an equivalent formulation of Boole's problem is now easily seen to be the following: given a point b in a correlation polytope $\rho^{(\mathcal{F})}$ as input, find the preimage of b under coordinate projection in the corresponding union polytope $\sigma^{(\mathcal{F})}$. This preimage is exactly the interval of the union probability $u = 1 - x_\emptyset$ and is particularly simple to find by substituting b into the halfspaces of $\sigma^{(\mathcal{F})}$; the reader is encouraged to try this out with the union polytopes $\sigma^{(\mathcal{F})}$ in Example 4 above.

[2] When $\mathcal{F} = \binom{[n]}{2} \cup \binom{[n]}{1}$, the polytope $\rho^{(\mathcal{F})} \subseteq \mathbb{Q}^{|\mathcal{F}|}$ is known as the *correlation polytope* [53]. Correlation polytopes and the associated *cut polytopes* are well-studied classes of polytopes related to each other by the so-called covariance mapping; also the generalized correlation polytopes appear in earlier work, cf. e.g. Deza and Laurent [18, Chap. I.5].

In spite of the appealing simplicity of the small examples above, it is however unlikely that the correlation and union polytopes in general admit a simple description for their halfspace-representations. Indeed, as discussed above, deciding the feasibility, or, what is the same, deciding membership in the correlation polytope, for a given abstract vector b is known to be NP-hard. For completeness, we state this fact as a theorem.

▶ **Theorem 5** (Hardness of membership in a correlation polytope; [18, 36, 38, 53]). *Given a set family $\emptyset \neq \mathcal{F} \subseteq 2^{[n]} \setminus \{\emptyset\}$ and a rational vector $b = (b_S \in [0, 1] : S \in \mathcal{F})$ as input, it is NP-hard to decide whether $b \in \rho^{(\mathcal{F})}$. Moreover, this holds even when $\mathcal{F}$ consists of sets of size at most two.*

Similarly, deciding membership in a union polytope is NP-hard; we prove the following theorem in the appendix as a corollary of the proof of Theorem 2.

▶ **Theorem 6** (Hardness of membership in a union polytope). *Given a set family $\emptyset \neq \mathcal{F} \subseteq 2^{[n]} \setminus \{\emptyset\}$, a rational vector $b = (b_S \in [0, 1] : S \in \mathcal{F})$, and a rational number $x_\emptyset \in [0, 1]$ as input, it is NP-hard to determine whether $(b, x_\emptyset) \in \sigma^{(\mathcal{F})}$. Moreover, this holds even when $\mathcal{F}$ consists of sets of size at most two and $b \in \rho^{(\mathcal{F})}$.*

In contrast to the complexity of the halfspace-representations of the union and correlation polytopes, the vertex-representations of the polytopes admit description by coordinate projection from the vertices of the Venn polytopes; we record, and prove in the appendix, basic geometric facts about the polytopes in Propositions 7 to 9 in Section 2 as preliminaries to the proof of our main theorem in Section 3. One immediate consequence of the zero-one vertex-representations observed in Propositions 7 to 9 is that the membership problems in Theorems 5 and 6 are in fact NP-complete by Carathéodory's theorem.

Let us now give a short overview of the proof of our main theorem (Theorem 2). For NP-hardness of the minimum union probability problem, we proceed with a straightforward reduction from the fractional graph coloring problem [30, 48, 61]. The NP-hardness of the maximum union probability problem requires somewhat more work. Our approach is to transform a restricted linear program for maximum union probability in steps to a variant of its dual linear program whose polyhedron ϕ_G has facets that enable eventual relation to the k-clique problem on a given graph G. We establish this relation with standard techniques for rational polyhedra [31] by transforming a validity oracle (obtainable from an oracle for maximum union probability) for ϕ_G via the polar polytope ϕ_G^* and its polar $\phi_G^{**} = \phi_G$ to a membership oracle for ϕ_G, which enables immediate solution of the k-clique problem on G.

We refer to the monographs of Grötschel, Lovász, and Schrijver [31] for pertinent terminology on rational polyhedra and Ziegler [73] for zero-one polytopes; see also Schrijver [62, 63, 64]. For fractional graph theory, we refer to Scheinerman and Ullman [61].

Related work

Generalized Bonferroni inequalities and local lemmata in probability and logic. A large body of work studies generalized Bonferroni [8] inequalities, with or without Hailperin's [32] linear programming formulation. Determining the best Bonferroni bounds for binomial moments $S_k = \sum_{K \in \binom{[n]}{k}} |B_K|$ given up to order $k \leq 2$ is studied by Chung and Erdős [14], Dawson and Sankoff [16], Galambos [26], Kounias and Marin [40, 41], Kwerel [44], Sathe et al. [60], Hunter [34], Worsley [70], de Caen [17], Kuai et al. [42, 43], Frolov [22], and Yang et al. [72]. More generalized versions for $k \geq 2$ were also studied by Gallot [28], Prékopa and Gao [57], Galambos and Mucci [25], and Kwerel [44] providing several sharp inequalities. A generalized version of the union problem was studied by Galambos [26], Platz [54], Prékopa [55, 56], Boros and Prékopa [11], Galambos and Simonelli [27], Min-Young [45], Sibya [66], Spencer [67], and Dohmen [19] for $\Pr(X \geq t)$; where X is a random variable that counts the number of events occur among n many events and $t \in [n]$. Hailperin's seminal study [32] was followed by a more systematic study of sentential probability logic [33]; Nilsson [52] models probabilities of logical sentences, where union probabilities arise from combining dependent events under consistency constraints. Jaeger's [35] probabilistic inference rules automates such tight bounds on the target probability as a function of the input constraints. In a related setting, bounding the probability of the union of n bad events, the Lovász Local Lemma [20] shows that if each bad event has probability at most p and depends on at most d others, then the union probability is strictly less than 1, hence a good outcome exists, provided that $p(d+1) \leq \exp(-1)$; a constructive randomized version was later given by Moser and Tardos [51]. Barvinok [5] recently showed that, under a stricter condition, intersections of order $O(\log(n/\epsilon))$ suffice to approximate the good event's probability up to relative error ϵ.

Approximate and exact inclusion–exclusion in algorithms and complexity. Linial and Nisan [46] present an *approximate* level-wise variant of the inclusion–exclusion formula (1) relying on Chebyshev polynomials and use it to study distributions that fool constant-depth circuits, followed by work of Luby and Veličković [47], Bazzi [6], Razborov [58] and Braverman [13]; a related problem is to approximate the number of satisfying assignments for a DNF with m clauses and n variables was studied by Luby and Veličkovi [47], Kahn et al. [37] show logarithmic-size intersections suffice for uniqueness of the DNF assignment, and later Melkman and Shimony [50] study its construction. Further works that improve the Linial-Nisan framework in specific applications include Klein and Metsch [39] (cryptography) and Shrestov [65] (boolean symmetric functions). The *exact* principle of inclusion and exclusion and its generalization to Möbius inversion on partially ordered sets [68] other than $(2^{[n]}, \subseteq)$ are fundamental primitives in the study fine-grained and parameterized algorithms, cf. Fomin and Kratsch [21] and Cygan et al. [15], including the study of canonical hard problems such as matrix permanent [59] and graph coloring [7].

Data mining. Downwards-closed set families arise naturally in data mining, in the task of finding association rules or frequent sets from zero-one-valued data. In that problem, the data is a (large) zero-one matrix D, and the task is to find all sets S of columns such that at least a fraction ϵ of the rows of the matrix have a 1-element in each column of S. In our notation, the task is to find the family $\mathcal{F} = \{S \subseteq [n] : b_S \geq \epsilon\}$. In applications, the matrix D is typically sparse, and the zeros and ones have an asymmetric role; hence, for suitably chosen ϵ, the family $\mathcal{F}$ will have a reasonable size. The algorithms for finding the collection $\mathcal{F}$ are based on a level-wise search: first find the singleton sets satisfying the criterion, and when the sets of size k are known, build all candidate sets C of size $k+1$ such that the subsets of C of size k satisfy the criterion, and check whether C satisfies the criterion. This so-called *Apriori* algorithm works well in many cases. The early papers and algorithms for this include Agrawal et al. [2], Agrawal and Srikant [4], Mannila et al. [49], Agrawal et al. [3], and have since been followed by many others.[3] Theoretical analyses of the question also exist, but to our knowledge, the question we study in this paper has not been considered in the frequent set literature. Note that in the setting of frequent sets, the vector b corresponding to the collection $\mathcal{F}$ is known to be feasible, so Boole's problem is significant exactly in the case where the feasibility of b is known.

[3] On March 26, 2026, a search on Scopus for "association rule mining" returned 20,581 documents.

Further polytopes and extension complexity. Wainwright and Jordan [69] derive the notion of a marginal polytope by starting from a set of sufficient statistics and the corresponding mean parameters (the expectations of the sufficient statistics). Then the set of all distributions that realize the mean parameters is convex and hence can be described by inequalities, giving the marginal polytope. It is mostly used in the case where the sufficient statistics correspond to singletons or pairs of variables in the underlying probability space, but one could also consider, e.g., the case where there is a sufficient statistic for each intersection probability. Then the marginal polytope for this set of sufficient statistics is the correlation polytope. Related to correlation and union polytopes, Yang, Alajaji, and Takahara [71] seek to bound the union probability from weighted-aggregate input of the form $\{|B_i|\}_{i=1}^n$ and $\left\{\sum_{j=1}^n c_j\,|B_i\cap B_j|\right\}_{i=1}^n$, which geometrically corresponds to intersecting the union polytope with the affine subspace defined by the input and studying the resulting polytope; whether our present NP-hardness results extend to exact optimization in such intersection polytopes remains an interesting direction for future work. A yet further line of work is the extension complexity of the correlation polytopes, cf. Aboulker et al. [1].

2 Geometry of the Venn, union, and correlation polytopes

The following propositions collect some immediate observations on the geometry of the Venn, union, and correlation polytopes. The proofs of the propositions appear in the appendix. We need short preliminaries. For a nonnegative integer n, we write $[n] = \{1, 2, \ldots, n\}$ and $2^{[n]}$ for the set of all subsets of $[n]$. We write $\mathbb{Q}$ for the set of rational numbers. For a logical proposition P, let us work with *Iverson's bracket notation* and set $[\![P]\!] = 1$ if P is true and $[\![P]\!] = 0$ if P is false. Let $\emptyset \neq \mathcal{F} \subseteq 2^{[n]}$. Define the set-inclusion indicator vectors $\zeta^{(\mathcal{F},T)} \in \mathbb{Q}^{|\mathcal{F}|}$ for all $S \in \mathcal{F}$ and $T \subseteq [n]$ by the rule

$$\zeta_S^{(\mathcal{F},T)} = [\![S \subseteq T]\!]\,. \tag{4}$$

Let us write $e^{(T)} \in \mathbb{Q}^{2^n}$ for the standard basis vector defined for all $T, U \subseteq [n]$ by the rule $e_U^{(T)} = [\![T = U]\!]$. Let us write $\uparrow S = \{T : S \subseteq T \subseteq [n]\}$ for the *up-set* of a set $S \in \mathcal{F}$. For a set $V \subseteq \mathbb{Q}^d$, we write $\mathrm{conv}(V)$ for the convex hull of V.

▶ **Proposition 7** (Geometry of Venn polytopes). *We have the following.*

1. *The Venn polytope $\tau^{(\mathcal{F})} \subseteq \mathbb{Q}^{|\mathcal{F}|+2^n}$ has dimension $2^n - 1$ and has only zero-one-valued vertices which are given by the representation $\tau^{(\mathcal{F})} = \mathrm{conv}((\zeta^{(\mathcal{F},T)}, e^{(T)}) : T \subseteq [n])$.*
2. *The number of vertices of $\tau^{(\mathcal{F})}$ is 2^n and a set of $1 + |\mathcal{F}|$ hyperplanes for $\tau^{(\mathcal{F})}$ is given by the equalities in* (3)*.*

▶ **Proposition 8** (Geometry of correlation polytopes). *We have the following.*

1. *The correlation polytope $\rho^{(\mathcal{F})} \subseteq \mathbb{Q}^{|\mathcal{F}|}$ is full-dimensional and has only zero-one-valued vertices which are given by the representation $\rho^{(\mathcal{F})} = \mathrm{conv}(\zeta^{(\mathcal{F},T)} : T \subseteq [n])$.*
2. *The number of vertices of $\rho^{(\mathcal{F})}$ is $|\bigwedge_{S\in\mathcal{F}} \{\uparrow S, 2^{[n]}\setminus\uparrow S\}|$, where the meet is taken in the set partition lattice of the set $2^{[n]}$. In particular, the number of vertices is 2^n if and only if $\{k\} \in \mathcal{F}$ for all $k \in [n]$.*
3. *A vector $b = (b_S \in [0, 1] : S \in \mathcal{F})$ is feasible if and only if $b \in \rho^{(\mathcal{F})}$.*

▶ **Proposition 9** (Geometry of union polytopes). *We have the following.*

1. *The union polytope $\sigma^{(\mathcal{F})} \subseteq \mathbb{Q}^{|\mathcal{F}|+1}$ has dimension $|\mathcal{F}| + [\![\mathcal{F} \neq 2^{[n]} \setminus \{\emptyset\}]\!]$ and has only zero-one-valued vertices which are given by the representation $\sigma^{(\mathcal{F})} = \mathrm{conv}((\zeta^{(\mathcal{F},T)}, [\![T = \emptyset]\!]) : T \subseteq [n])$.*

2. *When $\sigma^{(\mathcal{F})}$ has codimension 1 it has the hyperplane $x_\emptyset + \sum_{\emptyset \neq S \subseteq [n]} (-1)^{|S|+1} b_S = 1$.*
3. *The number of vertices of $\sigma^{(\mathcal{F})}$ either agrees with the number of vertices of $\rho^{(\mathcal{F})}$ or is one more than this; the latter case occurs if and only if there exists an $k \in [n]$ with $\{k\} \notin \mathcal{F}$.*
4. *For every $b = (b_S \in [0,1] : S \in \mathcal{F})$ and $x_\emptyset \in [0,1]$ we have that $(b, x_\emptyset) \in \sigma^{(\mathcal{F})}$ if and only if b is realizable and there exists a realization x of b with union probability $u = 1 - x_\emptyset$.*

3 Hardness of optimally solving Boole's problem

This section restates and proves our main theorem.

▶ **Theorem 2** (Main; NP-hardness of Boole's union probability problem)**.** *Given as input a set family $\emptyset \neq \mathcal{F} \subseteq 2^{[n]} \setminus \{\emptyset\}$ and a feasible rational vector $b = (b_S \in [0,1] : S \in \mathcal{F})$, each of the following two problems is NP-hard to solve:*
1. *determine the minimum union probability for a realization of b, and*
2. *determine the maximum union probability for a realization of b.*

Moreover, NP-hardness holds even when $\mathcal{F}$ consists of sets of size at most two.

For convenience, we present the proof in two parts, one part for the minimum union probability, and one part for the maximum union probability. Before proceeding with the first part, let us recall the fractional chromatic number problem (e.g. [61, Chap. 3]), which is known to be NP-hard [48]. Let us write $V(G)$ for the vertex set and $\mathcal{I}(G)$ for the set of all nonempty independent sets of a graph G. Given a graph G as input, we are to compute the minimum value of $\sum_{I \in \mathcal{I}(G)} y_I$ subject to the constraints $\sum_{u \in I \in \mathcal{I}(G)} y_I \geq 1$ for all $u \in V(G)$ and $y_I \geq 0$ for all $I \in \mathcal{I}(G)$. This minimum value is the *fractional chromatic number* $\chi_f(G)$ of G. A map $y : \mathcal{I}(G) \to \mathbb{Q}$ that satisfies the constraints is a *fractional coloring* of G.

Proof of Theorem 2, Item 1 (minimum union probability). We reduce from the fractional chromatic number problem. Let G be a graph given as input. Without loss of generality we may assume that the vertex set $V(G) = [n]$ and that the edge set $E(G)$ consists of 2-element subsets of $V(G)$. Set

$$\mathcal{F} = \{\{u\} : u \in V(G)\} \cup E(G) \subseteq 2^{[n]}\,.$$

Define $b : \mathcal{F} \to \mathbb{Q}$ by setting

$$b_{\{u\}} = \frac{1}{n} \quad \text{for all } u \in V(G), \text{ and} \tag{5}$$

$$b_{\{u,v\}} = 0 \quad \text{for all } \{u,v\} \in E(G). \tag{6}$$

We immediately observe that $(b, \mathcal{F})$ is feasible; indeed, to obtain a realization x, for each $T \subseteq V(G) = [n]$ we can set

$$x_T = \begin{cases} \frac{1}{n} & \text{if } |T| = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Suppose now that $x^* = (x^*_T \in [0,1] : T \subseteq [n])$ is a realization of b that minimizes the union probability $\sum_{\emptyset \neq T \subseteq [n]} x^*_T$. Define $y = (y_T \in [0,1] : \emptyset \neq T \subseteq V(G))$ for all $\emptyset \neq T \subseteq V(G)$ by setting $y_T = n \cdot x^*_T$. We claim that y is zero outside of $\mathcal{I}(G)$ and a fractional coloring of G when restricted to $\mathcal{I}(G)$; thus $\sum_{\emptyset \neq T \subseteq V(G)} y_T \geq \chi_f(G)$ by definition of fractional coloring. Indeed, from (6) and x^* realizing b we observe that $y_T = 0$ unless $T \in \mathcal{I}(G)$. Moreover, from (5) and x^* realizing b we observe that $\sum_{u \in T \subseteq V(G)} y_T = 1$ holds for all $u \in V(G)$. Thus, the minimum union probability of $(b, \mathcal{F})$ is at least $\chi_f(G)/n$. Conversely, suppose that y^* is a fractional coloring of G with $\sum_{I\in\mathcal{I}(G)} y_I = \chi_f(G)$. Extend this fractional coloring to all subsets $\emptyset \neq T \subseteq [n]$ by setting $y^*_T = 0$ if $T \notin \mathcal{I}(G)$. By definition of fractional coloring, we have that $\sum_{u\in T\subseteq V(G)} y^*_T \geq 1$. Since all nonempty subsets of independent sets in G are independent, without loss of generality we can assume that $\sum_{u\in T\subseteq V(G)} y^*_T = 1$ holds for all $u \in V(G)$; indeed, to ensure this we can iterate over all $u_0 \in V(G)$ and do the following: while the slack $\left(\sum_{u_0\in T\in\mathcal{I}(G)} y^*_T\right) - 1$ is positive, greedily transfer mass from y^*_T to $y^*_{T\setminus\{u_0\}}$ for sets $u_0 \in T \in \mathcal{I}(G)$ with $|T| \geq 2$ and greedily reduce mass in $y^*_{\{u_0\}}$, until the slack is zero. It is immediate that the greedy process zeroes slack at u_0 only and leaves other $u \neq u_0$ untouched in terms of their slack; furthermore, $\sum_{T\in\mathcal{I}(G)} y^*_T$ does not increase in the process.

Now define $x = (x_T \in [0,1] : T \subseteq [n])$ by setting $x_\emptyset = 1 - \frac{1}{n}\sum_{\emptyset\neq T\subseteq[n]} y^*_T$ and $x_T = \frac{1}{n}\cdot y^*_T$ for all $\emptyset \neq T \subseteq [n]$; since fractional chromatic number always satisfies $\chi_f(G) \leq n$, in particular we conclude that $x_\emptyset \geq 0$. From (5) and (6) we observe that x realizes $(b, \mathcal{F})$ with union probability that equals $\chi_f(G)/n$. We conclude that the minimum union probability of $(b, \mathcal{F})$ equals $\chi_f(G)/n$. ◀

Let us now turn to the proof of Theorem 2(2), where the NP-hardness reduction will eventually be from the k-clique problem. We start with a lemma that shows an oracle for the linear program for the maximum union probabability problem on given input $(b, \mathcal{F})$ can be used to solve a graph-based linear program, which the subsequent proof will then use to solve k-clique.

▶ **Lemma 10** (A transformed special case of the dual linear program). *Assume access to an oracle that solves the maximum union probability problem for a set family $\emptyset \neq \mathcal{F} \subseteq 2^{[n]}$ and a feasible rational $b : \mathcal{F} \to [0,1]$ given as query. Then, given a nonempty n-vertex graph G and edge weights $w : E(G) \to \mathbb{Q}_{\geq 0}$ as input, we can solve the linear program to*

$$\begin{aligned} &\textit{maximize} && \sum_{\{u,v\}\in E(G)} w_{\{u,v\}} y_{\{u,v\}} \\ &\textit{subject to} && \sum_{E(G)\ni\{u,v\}\subseteq T} y_{\{u,v\}} \leq |T| - 1 \ \textit{for all cliques}\ \emptyset \neq T \subseteq V(G) \end{aligned} \tag{8}$$

for its optimum value in time polynomial in n and the rational encoding length of w, by making a single query to the maximum union probability oracle, with $\mathcal{F}$ consisting of sets of size at most two.

Proof. Let us recall the maximum union probability problem as a linear program on a given query $(b, \mathcal{F})$ with b feasible. Namely, we are to

$$\begin{aligned} &\text{maximize} && \sum_{\emptyset\neq T\subseteq[n]} x_T \\ &\text{subject to} && \sum_{T\subseteq[n]} x_T = 1,\ \sum_{S\subseteq T} x_T = b_S \text{ for all } S \in \mathcal{F}, \text{ and } x_T \geq 0 \text{ for all } T \subseteq [n]\,. \end{aligned} \tag{9}$$

We assume we have an oracle subroutine that on query $(b, \mathcal{F})$ with b feasible outputs the value of (9).

Let G be a graph and $w : E(G) \to \mathbb{Q}_{\geq 0}$ a function given as input. (We will discuss the input w in detail only towards the end of the proof of this lemma.) We may assume that $n \geq 2$ and that G has at least one edge. We may also assume that the vertex set $V(G) = [n]$ and that the edge set $E(G)$ consists of 2-element subsets of $V(G)$. Fix

$$\mathcal{F} = \{\{u\} : u \in V(G)\} \cup \{\{u,v\} : u,v \in V(G)\}\,.$$

We will also fix $b_{\{u\}} = \frac{1}{n}$ and $b_{\{u,v\}} = 0$ for all distinct $u, v \in V(G)$ with $\{u, v\} \notin E(G)$ in our query to (9) in what follows. With this fixing, the query $(b, \mathcal{F})$ to (9) that we will make in what follows is defined by a function $c : E(G) \to [0, \frac{1}{n^2}]$ when we set $b_{\{u,v\}} = c_{\{u,v\}}$ for all $\{u, v\} \in E(G)$; the upper bound $\frac{1}{n^2}$ for the values of c is to ensure feasibility of the resulting query b. Indeed, to see that such a function c results in a feasible query b to (9), observe that b is realized by the atom probabilities x defined for all $T \subseteq [n]$ by

$$x_T = \begin{cases} 0 & \text{if } |T| > 2, \text{ or } |T| = 2 \text{ and } T \notin E(G), \\ c_T & \text{if } |T| = 2 \text{ and } T \in E(G), \\ \frac{1}{n} - \sum_{T \subsetneq U \subseteq [n]} x_U & \text{if } |T| = 1, \\ 1 - \sum_{\emptyset \neq U \subseteq [n]} x_U & \text{if } T = \emptyset. \end{cases}$$

Now let us assume a function $c : E(G) \to [0, \frac{1}{n^2}]$ and thus a query $(b, \mathcal{F})$ has been fixed. We will proceed along a sequence of linear programs equivalent to (9). Since we have fixed $b_{\{u\}} = \frac{1}{n}$ for all $u \in V(G)$ and $x_T \geq 0$ holds for all $\emptyset \neq T \subseteq [n]$, we observe that

$$\sum_{\emptyset \neq T \subseteq [n]} x_T \leq \sum_{u \in V(G)} \sum_{u \in T \subseteq V(G)} x_T \leq n \cdot \frac{1}{n} = 1\,.$$

This inequality in particular implies that we can remove the constraint $\sum_{T \subseteq [n]} x_T = 1$ and the variable $x_\emptyset$ from (9); that is, an equivalent linear program is to

$$\begin{aligned} &\text{maximize} \sum_{\emptyset \neq T \subseteq [n]} x_T \\ &\text{subject to} \sum_{S \subseteq T} x_T = b_S \text{ for all } S \in \mathcal{F} \text{ and } x_T \geq 0 \text{ for all } \emptyset \neq T \subseteq [n]. \end{aligned} \tag{10}$$

The dual of (10) is to

$$\begin{aligned} &\text{minimize} \sum_{S \in \mathcal{F}} b_S y_S \\ &\text{subject to} \sum_{\mathcal{F} \ni S \subseteq T} y_S \geq 1 \text{ for all } \emptyset \neq T \subseteq [n], \end{aligned} \tag{11}$$

where each variable y_S for $S \in \mathcal{F}$ can take arbitrary rational values. Since $b_{\{u,v\}} = 0$ whenever $\{u, v\} \notin E(G)$, we observe that in (11) each constraint for $\emptyset \neq T \subseteq [n]$ that is not a clique in G can be trivially satisfied by assigning an arbitrarily large value to a variable $y_{\{u,v\}}$ for $T \supseteq \{u, v\} \notin E(G)$, without affecting the value of the objective function since $b_{\{u,v\}} = 0$. In particular, we can remove all the non-clique constraints and all the variables $y_{\{u,v\}}$ with $\{u, v\} \notin E(G)$ from (11) to obtain the equivalent linear program to

$$\begin{aligned} &\text{minimize } \frac{1}{n} \sum_{u \in V(G)} y_{\{u\}} + \sum_{\{u,v\} \in E(G)} c_{\{u,v\}} y_{\{u,v\}} \\ &\text{subject to} \sum_{V(G) \ni u \in T} y_{\{u\}} + \sum_{E(G) \ni \{u,v\} \subseteq T} y_{\{u,v\}} \geq 1 \text{ for all cliques } \emptyset \neq T \subseteq V(G). \end{aligned} \tag{12}$$

Now fix an arbitrary optimum solution y^* to (12). Without loss of generality we can assume that $y^*_{\{u\}} = 1$ holds for all $u \in V(G)$; indeed, $y^*_{\{u\}} \geq 1$ is immediate by the clique $T = \{u\}$, and the assumption $y^*_{\{u_0\}} = 1 + \epsilon$ for some $u_0 \in V(G)$ and $\epsilon > 0$ yields a contradiction to the optimality of y^* by decreasing $y^*_{\{u_0\}}$ (with objective coefficient $\frac{1}{n}$) by ϵ and increasing each $y^*_{\{u_0,v\}}$ (with objective coefficient $c_{\{u_0,v\}} \leq \frac{1}{n^2}$) by ϵ for each edge $\{u_0, v\} \in E(G)$ to obtain an overall decrease in the minimization objective of (12) without violating any of the clique constraints. Thus, a linear program equivalent to (12) is to

$$\begin{aligned} &\text{minimize } 1 + \sum_{\{u,v\}\in E(G)} c_{\{u,v\}} y_{\{u,v\}} \\ &\text{subject to } |T| + \sum_{E(G)\ni\{u,v\}\subseteq T} y_{\{u,v\}} \geq 1 \text{ for all cliques } \emptyset \neq T \subseteq V(G). \end{aligned} \tag{13}$$

Recalling that the variables $y_{\{u,v\}}$ can take arbitrary rational values, changing the sign of the variables and removing the known offset 1 from the objective, we observe that an offset-1-equivalent linear program to (13) is to

$$\begin{aligned} &\text{maximize } \sum_{\{u,v\}\in E(G)} c_{\{u,v\}} y_{\{u,v\}} \\ &\text{subject to } \sum_{E(G)\ni\{u,v\}\subseteq T} y_{\{u,v\}} \leq |T| - 1 \text{ for all cliques } \emptyset \neq T \subseteq V(G). \end{aligned} \tag{14}$$

It follows immediately from this sequence (9)–(14) of offset-equivalent linear programs that we can use our assumed maximum union probability oracle to obtain the maximum value of (8) for an arbitrary $c : E(G) \to [0, \frac{1}{n^2}]$.

Now consider the given rational input $w : E(G) \to \mathbb{Q}_{\geq 0}$ and recall that we are to solve (8) for its optimum value. When w is identically zero, the optimum value of (8) is zero, so we can assume that $\max_{e\in E(G)} w(e) > 0$ in what follows. Set $\alpha = (n^2 \max_{e\in E(G)} w_e)^{-1}$ and $c = \alpha w$, observing that all values of c are rational in the interval $[0, \frac{1}{n^2}]$. Furthermore, c has rational encoding length at most polynomial in n and the rational encoding length of w. Now invoke the maximum union probability oracle on input $(b, \mathcal{F})$ derived from c and offset the oracle output by 1 to obtain the optimum value t for (14), and return the value t/α, observing that the returned value is the optimum value of (8) on input w. This process runs in time bounded by a polynomial in n and the rational encoding length of w, making a single query $(b, \mathcal{F})$ to the maximum union probability oracle with $\mathcal{F}$ consisting of sets of size at most two. ◀

We are now ready for the main reduction from the k-clique problem.

Proof of Theorem 2, Item 2 (maximum union probability). We reduce from the k-clique problem. Let G be a graph and k be a nonnegative integer given as input. Without loss of generality we may assume that $n \geq k \geq 2$ and that G has at least one edge. We may also assume that the vertex set $V(G) = [n]$ and that the edge set $E(G)$ consists of 2-element subsets of $V(G)$. It now follows immediately from Lemma 10 that we have a subroutine to find the maximum value of any objective $w : E(G) \to \mathbb{Q}_{\geq 0}$ over the polyhedron

$$\phi_G = \left\{ y : E(G) \to \mathbb{Q} \ : \ \sum_{E(G)\ni\{u,v\}\subseteq T} y_{\{u,v\}} \leq |T| - 1 \text{ for all cliques } \emptyset \neq T \subseteq V(G) \right\} \tag{15}$$

We observe in particular that ϕ_G is unbounded, in fact with cone $\mathbb{Q}^{|E(G)|}_{\leq 0}$ as we will see, so it will be convenient to proceed via the polar to arrive at the k-clique problem.

We first construct a *validity oracle* oracle for ϕ_G; namely, given $\mu : E(G) \to \mathbb{Q}$ and $\nu \in \mathbb{Q}$ as input, we are to assert whether $\mu^\top y \leq \nu$ holds for all $y \in \phi_G$. On given input μ, ν, the validity oracle proceeds as follows. Unless μ takes only nonnegative values, assert that $\mu^\top y \le \nu$ does not hold for all $y \in \phi_G$; indeed, if $\mu(\{u,v\}) < 0$ for some $\{u,v\} \in E(G)$, construct a vector $y_0 : E(G) \to \mathbb{Q}$ that satisfies $y_0(\{u,v\}) < \min(0, \nu/\mu(\{u,v\}))$ and is zero elsewhere to witness $\mu^\top y_0 > \nu$ as well as $y_0 \in \phi_G$. When μ takes only nonnegative values, use the subroutine with $w = \mu$ to obtain the maximum value t of the objective $w^\top y$ for $y \in \phi_G$, and assert that $\mu^\top y \le \nu$ holds for all $y \in \phi_G$ if $t \le \nu$; otherwise assert that $\mu^\top y \le \nu$ does not hold for all $y \in \phi_G$. This completes the description of the validity oracle for ϕ_G.

The validity oracle for ϕ_G immediately gives a *membership oracle* for the *polar*

$$\phi_G^* = \left\{ z : E(G) \to \mathbb{Q} \;:\; z^\top y \le 1 \text{ for all } y \in \phi_G \right\} \tag{16}$$

of ϕ_G; indeed, by definition of the polar, for $\mu \in E(G) \to \mathbb{Q}$ given as input we have $\mu \in \phi_G^*$ if and only if $\mu^\top y \le 1$ holds for all $y \in \phi_G$.

We next turn this membership oracle for ϕ_G^* into a validity oracle for ϕ_G^*. Towards this end, we collect some geometric prerequisites about ϕ_G and ϕ_G^*. Recall the standard norm inequalities $\|y\|_2 \le \|y\|_1 \le \sqrt{|E(G)|}\|y\|_2$ for $y : E(G) \to \mathbb{Q}$. Since $\|y\|_2 \le \frac{1}{n}$ implies $\|y\|_1 \le n\|y\|_2 \le 1$, from (15) we observe that ϕ_G contains the 2-norm ball of radius $\frac{1}{n}$ centered at the origin; thus, the polar ϕ_G^* is a polytope contained in the 2-norm ball of radius n centered at the origin. We also observe that $\phi_G^* \subseteq \mathbb{Q}_{\ge 0}^{|E(G)|}$; indeed, any $z : E(G) \to \mathbb{Q}$ with $z(\{u,v\}) < 0$ for some $\{u,v\} \in E(G)$ has $z^\top y_0 > 1$ for $y_0 : E(G) \to \mathbb{Q}$ with $y_0(\{u,v\}) < 1/z(\{u,v\})$ and zero elsewhere, with $y_0 \in \phi_G$. Consider an arbitrary $z : E(G) \to \mathbb{Q}_{\ge 0}$ with $\|z\|_2 \le \frac{1}{n}$. From (15) it follows that for all $y \in \phi_G$ we have $z^\top y \le \|z\|_1 \le n\|z\|_2 \le 1$ and thus $z \in \phi_G^*$ by (16). It follows that ϕ_G^* contains the nonnegative part of the 2-norm ball of radius $\frac{1}{n}$ centered at the origin and that $z_{\{u,v\}} \ge 0$ is a face of ϕ_G^* for each $\{u,v\} \in E(G)$; so we have in particular that ϕ_G^* contains the 2-norm ball of radius $\frac{1}{2n^2}$ centered at $(\frac{1}{2n^2}, \frac{1}{2n^2}, \cdots, \frac{1}{2n^2})$. Furthermore, it is immediate that the cone of ϕ_G contains $\mathbb{Q}_{\le 0}^{|E(G)|}$, and we claim that equality holds; indeed, suppose that $y : E(G) \to \mathbb{Q}$ belongs to the cone of ϕ_G with $y(\{u,v\}) > 0$ for some $\{u,v\} \in E(G)$, and obtain a contradiction to $z \in \phi_G^*$ for $z(\{u,v\}) = \frac{1}{n}$ and zero elsewhere.

From the halfspace-representation (15) we observe that the facet-complexity [31, Definition 6.2.2] of $\phi_G \subseteq \mathbb{Q}^{|E(G)|}$ is bounded by $O(n^2)$, implying [31, Lemma 6.2.3] that the vertex-complexity [31, Definition 6.2.2] of ϕ_G is bounded by $O(n^6)$. Since the polar ϕ_G^* admits halfspace-representation with each inequality obtained from the vertex-representation[4] of ϕ_G, it immediately follows that the facet-complexity of ϕ_G^* is bounded by $O(n^6)$.

Since we know an explicit positive-radius ball inside ϕ_G^* as well as an explicit positive-radius ball containing ϕ_G^*, and the facet complexity of ϕ_G^* is bounded by a polynomial in n, the membership oracle for ϕ_G^* can be turned into a weak violation and thus weak validity oracle for ϕ_G^* [31, Theorem 4.3.2], which in turn can be turned into a strong validity oracle for ϕ_G^* [31, Theorem 6.3.2], which in turn gives a membership oracle for the polar of the polar ϕ_G^{**}. Since ϕ_G is closed and convex as well as contains the origin, we have $\phi_G^{**} = \phi_G$ [31, Sect. 0.1]) and thus a membership oracle for ϕ_G.

Finally, we will use the membership oracle for ϕ_G to solve the k-clique problem on G. To decide whether G has a clique of size k, set up the constant function $y^{(\kappa)} : E(G) \to \mathbb{Q}$ with $y^{(\kappa)}_{\{u,v\}} = \kappa$ for all $\{u,v\} \in E(G)$ for some constant $\kappa > 0$ yet to be determined. Observe

---

[4] We recall that any nonempty polyhedron $\pi \subseteq \mathbb{Q}^d$ admits representation as the Minkowski sum $\pi = \text{conv}(P) + \text{cone}(R)$ of a convex hull of a finite set of points $P \subseteq \mathbb{Q}^d$ and the convex cone of a finite set of rays $R \subseteq \mathbb{Q}^d$ so that the polar $\pi^* = \{z \in \mathbb{Q}^d : y^\top z \le 1 \text{ for all } y \in \pi\} \subseteq \mathbb{Q}^d$ of π has the halfspace-representation $\pi^* = \{z \in \mathbb{Q}^d : p^\top z \le 1 \text{ for all } p \in P \text{ and } r^\top z \le 0 \text{ for all } r \in R\}$, cf. [31, Sect. 0.1].

that we can use the membership oracle for ϕ_G to test whether $y^{(\kappa)} \in \phi_G$. For every clique $T \subseteq V(G)$ of G, we observe that $\kappa\binom{|T|}{2} > |T| - 1$ if and only if $|T| > \frac{2}{\kappa}$. Thus, using the membership oracle for ϕ_G, we can decide whether G has a clique of size k by setting $\kappa = \frac{2}{k-1}$; indeed, from (15) we have that the constant-function query $y^{(\kappa)} \notin \phi_G$ if and only if G has a clique of size of at least k, equivalently a k-clique. ◀

References

1. Pierre Aboulker, Samuel Fiorini, Tony Huynh, Marco Macchia, and Johanna Seif. Extension complexity of the correlation polytope. *Operations Research Letters*, 47(1):47–51, 2019. `doi:10.1016/j.orl.2018.12.001`.
2. Rakesh Agrawal, Tomasz Imielinski, and Arun N. Swami. Mining association rules between sets of items in large databases. In Peter Buneman and Sushil Jajodia, editors, *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, Washington, DC, USA, May 26-28, 1993*, pages 207–216. ACM Press, 1993. `doi:10.1145/170035.170072`.
3. Rakesh Agrawal, Heikki Mannila, Ramakrishnan Srikant, Hannu Toivonen, and A. Inkeri Verkamo. Fast discovery of association rules. In Usama M. Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth, and Ramasamy Uthurusamy, editors, *Advances in Knowledge Discovery and Data Mining*, pages 307–328. AAAI/MIT Press, 1996.
4. Rakesh Agrawal and Ramakrishnan Srikant. Fast algorithms for mining association rules in large databases. In Jorge B. Bocca, Matthias Jarke, and Carlo Zaniolo, editors, *VLDB'94, Proceedings of 20th International Conference on Very Large Data Bases, September 12-15, 1994, Santiago de Chile, Chile*, pages 487–499. Morgan Kaufmann, 1994. URL: `http://www.vldb.org/conf/1994/P487.PDF`.
5. Alexander Barvinok. Computing the probability of intersection, 2025. `doi:10.48550/arXiv.2507.10329`.
6. Louay M. J. Bazzi. Polylogarithmic independence can fool DNF formulas. *SIAM J. Comput.*, 38(6):2220–2272, 2009. `doi:10.1137/070691954`.
7. Andreas Björklund, Thore Husfeldt, and Mikko Koivisto. Set partitioning via inclusion-exclusion. *SIAM J. Comput.*, 39(2):546–563, 2009. `doi:10.1137/070683933`.
8. Carlo E. Bonferroni. Teoria statistica delle classi e calcolo delle probabilità. (Pubbl. d. R. Ist. Super. di Sci. Econom. e Commerciali di Firenze. 8) Firenze: Libr. Internaz. Seeber. 62 S., 1936. `doi:10.4135/9781412961288.n455`.
9. George Boole. *An Investigation of the Laws of Thought, on Which Are Founded the Mathematical Theories of Logic and Probabilities.* Walton and Maberly, London, 1854. `doi:10.5962/bhl.title.29413`.
10. Endre Boros and Joonhee Lee. Boole's probability bounding problem, linear programming aggregations, and nonnegative quadratic pseudo-Boolean functions. *Math. Oper. Res.*, 51(1):134–148, 2026. `doi:10.1287/moor.2023.0019`.
11. Endre Boros and András Prékopa. Closed form two-sided bounds for probabilities that at least r and exactly r out of n events occur. *Math. Oper. Res.*, 14(2):317–342, 1989. `doi:10.1287/moor.14.2.317`.
12. Endre Boros, Andrea Scozzari, Fabio Tardella, and Pierangela Veneziani. Polynomially computable bounds for the probability of the union of events. *Math. Oper. Res.*, 39(4):1311–1329, 2014. `doi:10.1287/moor.2014.0657`.
13. Mark Braverman. Polylogarithmic independence fools AC^0 circuits. *J. ACM*, 57(5):Art. 28, 10, 2010. `doi:10.1145/1754399.1754401`.
14. K. L. Chung and P. Erdös. On the application of the Borel-Cantelli lemma. *Trans. Amer. Math. Soc.*, 72:179–186, 1952. `doi:10.2307/1990661`.
15. Marek Cygan, Fedor V. Fomin, Ł ukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michał Pilipczuk, and Saket Saurabh. *Parameterized Algorithms.* Springer, Cham, 2015. `doi:10.1007/978-3-319-21275-3`.

16 D. A. Dawson and D. Sankoff. An inequality for probabilities. *Proc. Amer. Math. Soc.*, 18:504–507, 1967. doi:10.2307/2035487.

17 D. de Caen. A lower bound on the probability of a union. *Discrete Math.*, 169(1-3):217–220, 1997. doi:10.1016/S0012-365X(96)00107-0.

18 Michel Marie Deza and Monique Laurent. *Geometry of Cuts and Metrics*, volume 15 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 1997. doi:10.1007/978-3-642-04295-9.

19 Klaus Dohmen. *Improved Bonferroni Inequalities via Abstract Tubes*, volume 1826 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2003. Inequalities and identities of inclusion-exclusion type. doi:10.1007/b13785.

20 Paul Erdős and László Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. In *Infinite and Finite Sets (Colloq., Keszthely, 1973)*, volume 10 of *Colloq. Math. Soc. János Bolyai*, pages 609–627. North-Holland, Amsterdam, 1975.

21 Fedor V. Fomin and Dieter Kratsch. *Exact Exponential Algorithms*. Texts in Theoretical Computer Science. An EATCS Series. Springer, Heidelberg, 2010. doi:10.1007/978-3-642-16533-7.

22 Andrei N. Frolov. Bounds for probabilities of unions of events and the Borel-Cantelli lemma. *Statist. Probab. Lett.*, 82(12):2189–2197, 2012. doi:10.1016/j.spl.2012.08.002.

23 Maurice Fréchet. *Les probabilités associées à un système d'événements compatibles et dépendants*, volume 859 of *Actualités Scientifiques et Industrielles*. Hermann, Paris, 1940.

24 Maurice Fréchet. *Les probabilités associées à un système d'événements compatibles et dépendants*, volume 942 of *Actualités Scientifiques et Industrielles*. Hermann, Paris, 1943.

25 J. Galambos and R. Mucci. Inequalities for linear combinations of binomial moments. *Publicationes Mathematicae Debrecen*, 27(3-4):263–268, 1980. doi:10.5486/pmd.1980.27.3-4.09.

26 Janos Galambos. Bonferroni inequalities. *The Annals of Probability*, 5(4):577–581, 1977. doi:10.1214/aop/1176995765.

27 Janos Galambos and Italo Simonelli. *Bonferroni-Type Inequalities with Applications*. Probability and its Applications (New York). Springer-Verlag, New York, 1996.

28 S. Gallot. A bound for the maximum of a number of random variables. *J. Appl. Probability*, 3:556–558, 1966. doi:10.2307/3212139.

29 George Georgakopoulos, Dimitris Kavvadias, and Christos H. Papadimitriou. Probabilistic satisfiability. *J. Complexity*, 4(1):1–11, 1988. doi:10.1016/0885-064X(88)90006-4.

30 M. Grötschel, L. Lovász, and A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981. doi:10.1007/BF02579273.

31 Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric Algorithms and Combinatorial Optimization*, volume 2 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, second edition, 1993. doi:10.1007/978-3-642-78240-4.

32 Theodore Hailperin. Best possible inequalities for the probability of a logical function of events. *Amer. Math. Monthly*, 72:343–359, 1965. doi:10.2307/2313491.

33 Theodore Hailperin. *Sentential Probability Logic*. Lehigh University Press, Bethlehem, PA; Associated University Presses, London, 1996. Origins, development, current status, and technical applications.

34 David Hunter. An upper bound for the probability of a union. *Journal of Applied Probability*, 13(3):597–603, 1976. doi:10.1017/s0021900200104164.

35 Manfred Jaeger. Automatic derivation of probabilistic inference rules. *International Journal of Approximate Reasoning*, 28(1):1–22, 2001. doi:10.1016/S0888-613X(01)00044-5.

36 Brigitte Jaumard, Pierre Hansen, and Marcus Poggi de Aragão. Column generation methods for probabilistic logic. *INFORMS J. Comput.*, 3(2):135–148, 1991. doi:10.1287/IJOC.3.2.135.

37 Jeff Kahn, Nathan Linial, and Alex Samorodnitsky. Inclusion-exclusion: exact and approximate. *Combinatorica*, 16(4):465–477, 1996. doi:10.1007/BF01271266.

38 Dimitris J. Kavvadias and Christos H. Papadimitriou. A linear programming approach to reasoning about probabilities. *Ann. Math. Artif. Intell.*, 1:189–205, 1990. `doi:10.1007/BF01531078`.

39 Andreas Klein and Klaus Metsch. On approximate inclusion-exclusion. *Innov. Incidence Geom.*, 6/7:249–270, 2007/08.

40 Eustratios G. Kounias. Bounds for the probability of a union, with applications. *Ann. Math. Statist.*, 39:2154–2158, 1968. `doi:10.1214/aoms/1177698049`.

41 Stratis Kounias and Jacqueline Marin. Best linear Bonferroni bounds. *SIAM J. Appl. Math.*, 30(2):307–323, 1976. `doi:10.1137/0130031`.

42 H. Kuai, F. Alajaji, and G. Takahara. A lower bound on the probability of a finite union of events. *Discrete Math.*, 215(1-3):147–158, 2000. `doi:10.1016/S0012-365X(99)00246-0`.

43 Hongyan Kuai, Fady Alajaji, and Glen Takahara. Tight error bounds for nonuniform signaling over AWGN channels. *IEEE Trans. Inform. Theory*, 46(7):2712–2718, 2000. `doi:10.1109/18.887886`.

44 Seymour M. Kwerel. Bounds on the probability of the union and intersection of m events. *Advances in Applied Probability*, 7:431–448, 1975. `doi:10.2307/1426084`.

45 Min-Young Lee. *Bonferroni-Type Inequalities.* ProQuest LLC, Ann Arbor, MI, 1991. Thesis (Ph.D.), Temple University. URL: `http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:9134967`.

46 Nathan Linial and Noam Nisan. Approximate inclusion-exclusion. *Combinatorica*, 10(4):349–365, 1990. `doi:10.1007/BF02128670`.

47 M. Luby and B. Veličković. On Deterministic Approximation of DNF. *Algorithmica*, 16(4-5):415–433, 1996. `doi:10.1007/s004539900054`.

48 Carsten Lund and Mihalis Yannakakis. On the hardness of approximating minimization problems. *J. ACM*, 41(5):960–981, 1994. `doi:10.1145/185675.306789`.

49 Heikki Mannila, Hannu Toivonen, and A. Inkeri Verkamo. Efficient algorithms for discovering association rules. In Usama M. Fayyad and Ramasamy Uthurusamy, editors, *Knowledge Discovery in Databases: Papers from the 1994 AAAI Workshop, Seattle, Washington, USA, July 1994. Technical Report WS-94-03*, pages 181–192. AAAI Press, 1994.

50 Avraham A. Melkman and Solomon E. Shimony. A note on approximate inclusion-exclusion. *Discrete Appl. Math.*, 73(1):23–26, 1997. `doi:10.1016/S0166-218X(96)00048-0`.

51 Robin A. Moser and Gábor Tardos. A constructive proof of the general Lovász local lemma. *J. ACM*, 57(2):Art. 11, 15, 2010. `doi:10.1145/1667053.1667060`.

52 Nils J. Nilsson. Probabilistic logic. *Artificial Intelligence*, 28(1):71–87, 1986. `doi:10.1016/0004-3702(86)90031-7`.

53 Itamar Pitowsky. Correlation polytopes: their geometry and complexity. *Math. Programming*, 50(3):395–414, 1991. `doi:10.1007/BF01594946`.

54 Ole Platz. A sharp upper probability bound for the occurrence of at least m out of n events. *J. Appl. Probab.*, 22(4):978–981, 1985. `doi:10.1017/s0021900200108241`.

55 András Prékopa. Boole-Bonferroni inequalities and linear programming. *Operations Research*, 36(1):145–162, 1988. `doi:10.1287/opre.36.1.145`.

56 András Prékopa. Inequalities on expectations based on the knowledge of multivariate moments. In *Stochastic inequalities (Seattle, WA, 1991)*, volume 22 of *IMS Lecture Notes Monogr. Ser.*, pages 309–331. Inst. Math. Statist., Hayward, CA, 1992. `doi:10.1214/lnms/1215461959`.

57 András Prékopa and Linchun Gao. Bounding the probability of the union of events by aggregation and disaggregation in linear programs. *Discrete Appl. Math.*, 145(3):444–454, 2005. `doi:10.1016/j.dam.2004.03.003`.

58 Alexander Razborov. A simple proof of Bazzi's theorem. *ACM Trans. Comput. Theory*, 1(1), February 2009. `doi:10.1145/1490270.1490273`.

59 Herbert John Ryser. *Combinatorial Mathematics*, volume No. 14 of *The Carus Mathematical Monographs*. Mathematical Association of America, ; distributed by John Wiley and Sons, Inc., New York, 1963.

60 Y. S. Sathe, Meena Pradhan, and S. P. Shah. Inequalities for the probability of the occurrence of at least m out of n events. *Journal of Applied Probability*, 17(4):1127–1132, 1980. doi:10.1017/s002190020009745x.

61 Edward R. Scheinerman and Daniel H. Ullman. *Fractional Graph Theory*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., New York, 1997. A rational approach to the theory of graphs, With a foreword by Claude Berge, A Wiley-Interscience Publication.

62 Alexander Schrijver. *Combinatorial Optimization. Polyhedra and Efficiency. Vol. A*, volume 24,A of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003. Paths, flows, matchings, Chapters 1–38.

63 Alexander Schrijver. *Combinatorial Optimization. Polyhedra and Efficiency. Vol. B*, volume 24,B of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003. Matroids, trees, stable sets, Chapters 39–69.

64 Alexander Schrijver. *Combinatorial Optimization. Polyhedra and Efficiency. Vol. C*, volume 24,C of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003. Disjoint paths, hypergraphs, Chapters 70–83.

65 Alexander A. Sherstov. Approximate inclusion-exclusion for arbitrary symmetric functions. *Comput. Complexity*, 18(2):219–247, 2009. doi:10.1007/s00037-009-0274-4.

66 Masaaki Sibuya. Sharp Bonferroni-type inequalities in explicit forms. In *Probability Theory and Applications*, volume 80 of *Math. Appl.*, pages 165–194. Kluwer Acad. Publ., Dordrecht, 1992.

67 Thomas H. Spencer. Generalized Bonferroni inequalities. *J. Appl. Probab.*, 31(2):409–417, 1994. doi:10.2307/3215034.

68 Richard P. Stanley. *Enumerative Combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.

69 Martin J. Wainwright and Michael I. Jordan. Graphical models, exponential families, and variational inference. *Found. Trends Mach. Learn.*, 1(1-2):1–305, 2008. doi:10.1561/2200000001.

70 K. J. Worsley. An improved Bonferroni inequality and applications. *Biometrika*, 69(2):297–302, 1982. doi:10.1093/biomet/69.2.297.

71 Jun Yang, Fady Alajaji, and Glen Takahara. On bounding the union probability. In *IEEE International Symposium on Information Theory, ISIT 2015, Hong Kong, China, June 14-19, 2015*, pages 1761–1765. IEEE, 2015. doi:10.1109/ISIT.2015.7282758.

72 Jun Yang, Fady Alajaji, and Glen Takahara. Lower bounds on the probability of a finite union of events. *SIAM J. Discrete Math.*, 30(3):1437–1452, 2016. doi:10.1137/15M100866X.

73 Günter M. Ziegler. Lectures on 0/1-polytopes. In *Polytopes—combinatorics and computation (Oberwolfach, 1997)*, volume 29 of *DMV Sem.*, pages 1–41. Birkhäuser, Basel, 2000.

A Omitted proofs

This appendix restates and proves the claims whose proofs were omitted in the paper proper.

▶ **Proposition 7** (Geometry of Venn polytopes). *We have the following.*

1. *The Venn polytope $\tau^{(\mathcal{F})} \subseteq \mathbb{Q}^{|\mathcal{F}|+2^n}$ has dimension $2^n - 1$ and has only zero-one-valued vertices which are given by the representation $\tau^{(\mathcal{F})} = \operatorname{conv}((\zeta^{(\mathcal{F},T)}, e^{(T)}) : T \subseteq [n])$.*
2. *The number of vertices of $\tau^{(\mathcal{F})}$ is 2^n and a set of $1 + |\mathcal{F}|$ hyperplanes for $\tau^{(\mathcal{F})}$ is given by the equalities in* (3)*.*

Proof. The vertices $(\zeta^{(\mathcal{F},T)}, e^{(T)}) \in \{0,1\}^{|\mathcal{F}|+2^n}$ for $T \subseteq [n]$ of $\tau^{(\mathcal{F})}$ are easily solved from the systems of equations obtained by leaving one constraint out from the system (3) in all possible ways, and observing that the resulting system gives a vertex satisfying all the constraints in all other cases except when the leftmost constraint in (3) is left out. The 2^n vertices are affinely independent so the dimension of $\tau^{(\mathcal{F})}$ is at least $2^n - 1$. The dimension of $\tau^{(\mathcal{F})}$ is at most $2^n - 1$ because the $1 + |\mathcal{F}|$ hyperplanes in (3) are independent. ◀

▶ **Proposition 8** (Geometry of correlation polytopes). *We have the following.*

1. *The correlation polytope $\rho^{(\mathcal{F})} \subseteq \mathbb{Q}^{|\mathcal{F}|}$ is full-dimensional and has only zero-one-valued vertices which are given by the representation $\rho^{(\mathcal{F})} = \operatorname{conv}(\zeta^{(\mathcal{F},T)} : T \subseteq [n])$.*
2. *The number of vertices of $\rho^{(\mathcal{F})}$ is $|\bigwedge_{S\in\mathcal{F}} \{\uparrow S, 2^{[n]} \setminus \uparrow S\}|$, where the meet is taken in the set partition lattice of the set $2^{[n]}$. In particular, the number of vertices is 2^n if and only if $\{k\} \in \mathcal{F}$ for all $k \in [n]$.*
3. *A vector $b = (b_S \in [0,1] : S \in \mathcal{F})$ is feasible if and only if $b \in \rho^{(\mathcal{F})}$.*

Proof. For Item 1, the zero-one-valued vertices property is immediate from (4) and Proposition 7(1). From (4) we observe that the vectors $\zeta^{(\mathcal{F},T)}$ for $T \in \{\emptyset\} \cup \mathcal{F}$ are affinely independent; indeed, $\zeta^{(\mathcal{F},\emptyset)}$ is the zero vector, and the vectors $\zeta^{(\mathcal{F},T)}$ for $T \in \mathcal{F}$ listed in a linearization of the partial order $(\mathcal{F}, \subseteq)$ form a triangular matrix with nonzero diagonal. Thus, $\rho^{(\mathcal{F})}$ is full-dimensional. For Item 2, observe that for all $T, U \subseteq [n]$ we have from (4) that $\zeta^{(\mathcal{F},T)} \neq \zeta^{(\mathcal{F},U)}$ if and only if there exists an $S \in \mathcal{F}$ with $[\![S \subseteq T]\!] \neq [\![S \subseteq U]\!]$; this inequation holds if and only if the sets T and U lie in distinct cells of the set partition $\{\uparrow S, 2^{[n]} \setminus \uparrow S\}$. The cells of the meet $\bigwedge_{S\in\mathcal{F}} \{\uparrow S, 2^{[n]} \setminus \uparrow S\}$ are thus in bijective correspondence with the vectors in the set $\{\zeta^{(\mathcal{F},T)} : T \subseteq [n]\}$. Observe also that T and U lie in distinct cells of the set partition $\{\uparrow\{k\}, 2^{[n]} \setminus \uparrow\{k\}\}$ if and only if exactly one of T and U contains $k \in [n]$. Thus, $\bigwedge_{S\in\mathcal{F}} \{\uparrow S, 2^{[n]} \setminus \uparrow S\}$ is the discrete partition with 2^n cells when $\{k\} \in \mathcal{F}$ holds for all $k \in [n]$. Conversely, suppose that $\{k\} \notin \mathcal{F}$ for some $k \in [n]$, implying that $\zeta^{(\mathcal{F},\emptyset)} = \zeta^{(\mathcal{F},\{k\})}$ and thus $|\{\zeta^{(\mathcal{F},T)} : T \subseteq [n]\}| < 2^n$. For Item 3, recall our discussion of Hailperin's linear program in the introduction. ◀

▶ **Proposition 9** (Geometry of union polytopes). *We have the following.*

1. *The union polytope $\sigma^{(\mathcal{F})} \subseteq \mathbb{Q}^{|\mathcal{F}|+1}$ has dimension $|\mathcal{F}| + [\![\mathcal{F} \neq 2^{[n]} \setminus \{\emptyset\}]\!]$ and has only zero-one-valued vertices which are given by the representation $\sigma^{(\mathcal{F})} = \operatorname{conv}((\zeta^{(\mathcal{F},T)}, [\![T = \emptyset]\!]) : T \subseteq [n])$.*
2. *When $\sigma^{(\mathcal{F})}$ has codimension 1 it has the hyperplane $x_\emptyset + \sum_{\emptyset \neq S \subseteq [n]} (-1)^{|S|+1} b_S = 1$.*
3. *The number of vertices of $\sigma^{(\mathcal{F})}$ either agrees with the number of vertices of $\rho^{(\mathcal{F})}$ or is one more than this; the latter case occurs if and only if there exists an $k \in [n]$ with $\{k\} \notin \mathcal{F}$.*
4. *For every $b = (b_S \in [0,1] : S \in \mathcal{F})$ and $x_\emptyset \in [0,1]$ we have that $(b, x_\emptyset) \in \sigma^{(\mathcal{F})}$ if and only if b is realizable and there exists a realization x of b with union probability $u = 1 - x_\emptyset$.*

Proof. For Item 1, the zero-one-valued vertices property is immediate from (4) and Proposition 7(1). Since $\sigma^{(\mathcal{F})}$ coordinate-restricts to $\rho^{(\mathcal{F})}$, we have that $\sigma^{(\mathcal{F})}$ has dimension at least $|\mathcal{F}|$ by full-dimensionality of $\rho^{(\mathcal{F})}$, cf. Proposition 8(1). It suffices to show that $\sigma^{(\mathcal{F})}$ has dimension $|\mathcal{F}| + 1$ when $\mathcal{F} \neq 2^{[n]} \setminus \{\emptyset\}$; indeed, Item 2 shows that the dimension is $|\mathcal{F}|$ when $\mathcal{F} = 2^{[n]} \setminus \{\emptyset\}$. So let us assume $\mathcal{F} \neq 2^{[n]} \setminus \{\emptyset\}$. We split into two cases. First, suppose that there exists an $k \in [n]$ with $\{k\} \notin \mathcal{F}$. From (4) we observe that the vectors $(\zeta^{(\mathcal{F},T)}, [\![T = \emptyset]\!])$ for $T \in \{\emptyset, \{k\}\} \cup \mathcal{F}$ are affinely independent; indeed, $(\zeta^{(\mathcal{F},\{k\})}, 0)$ is the zero vector, and the vectors $(\zeta^{(\mathcal{F},T)}, [\![T = \emptyset]\!])$ for $T \in \{\emptyset\} \cup \mathcal{F}$ listed in a reverse linearization of the partial order $(\{\emptyset\} \cup \mathcal{F}, \subseteq)$, placing the empty set last, form a triangular matrix with nonzero diagonal. Thus, $\sigma^{(\mathcal{F})}$ is full-dimensional. Second, suppose that $\{k\} \in \mathcal{F}$ holds for all $k \in [n]$. Since $\mathcal{F} \neq 2^{[n]} \setminus \{\emptyset\}$, there thus exists a set $\mathcal{F} \not\ni U \subseteq [n]$ with $|U| \geq 2$ such that all nonempty proper subsets of U are in $\mathcal{F}$. We show that the vectors $(\zeta^{(\mathcal{F},T)}, [\![T = \emptyset]\!])$ for $T \in \{\emptyset, U\} \cup \mathcal{F}$ are affinely independent. Linearize the partial order $(\{\emptyset, U\} \cup \mathcal{F}, \subseteq)$ so that the sets in

$2^U \subseteq \{\emptyset, U\} \cup \mathcal{F}$ come before all the other sets. Place the vectors $(\zeta^{(\mathcal{F},T)}, [\![T = \emptyset]\!])$ into the columns of a matrix from left to right in this linearization order for $T \in \{\emptyset, U\} \cup \mathcal{F}$ so that $T = \emptyset$ gives the leftmost column. We can index the rows of this matrix, again in linearization order, by $S \in \{\emptyset\} \cup \mathcal{F}$, so that row $S = \emptyset$ has the entry $[\![T = \emptyset]\!]$ in column T, and row $S \in \mathcal{F}$ has the entry $[\![S \subseteq T]\!]$ in column T by (4). The matrix has shape $(|\mathcal{F}| + 1) \times (|\mathcal{F}| + 2)$. Now subtract the leftmost column ($T = \emptyset$) from all the other columns, delete the leftmost column, and negate the first row. It now suffices to show that the resulting $(|\mathcal{F}| + 1) \times (|\mathcal{F}| + 1)$ matrix has full rank to conclude affine independence. Recalling the linearization order, we observe that the matrix is upper triangular with nonzero diagonal, apart from the top-left $(2^{|U|} - 1) \times (2^{|U|} - 1)$ submatrix, with rows indexed by $S \in 2^U \setminus \{U\}$, and columns indexed by $T \in 2^U \setminus \{\emptyset\}$, with entries $[\![S \subseteq T]\!]$. This top-left submatrix has full rank; indeed, the submatrix equals, assuming lexicographic order is used to linearize 2^U, the $\bar{Z}$-matrix obtained from the base case $d = 1$, $Z = M = (1)$, and $f = g = (1)$ by applying the construction in the following immediate claim $|U| - 1$ times, yielding $\bar{Z}$ of shape $(2^{|U|} - 1) \times (2^{|U|} - 1)$.

▷ Claim 11. Let Z and M be mutually inverse $d \times d$ matrices and let f and g be $d \times 1$ vectors such that $\gamma = g^\top M f \neq 0$. Define the $(2d + 1) \times (2d + 1)$ matrices and $(2d + 1) \times 1$ vectors

$$\bar{Z} = \left[\begin{array}{c|c|c} Z & f & Z \\ \hline g^\top & 0 & g^\top \\ \hline 0 & f & Z \end{array}\right], \quad \bar{M} = \left[\begin{array}{c|c|c} M & 0 & -M \\ \hline \gamma^{-1} g^\top M & -\gamma^{-1} & 0 \\ \hline -\gamma^{-1} M f g^\top M & \gamma^{-1} M f & M \end{array}\right], \quad \bar{f} = \left[\begin{array}{c} f \\ \hline 0 \\ \hline 0 \end{array}\right], \quad \bar{g} = \left[\begin{array}{c} 0 \\ \hline 0 \\ \hline g \end{array}\right].$$

Then, $\bar{Z}$ and $\bar{M}$ are mutually inverse and $\bar{\gamma} = \bar{g}^\top \bar{M} \bar{f} = -\gamma \neq 0$.

For Item 2, from the proof of Item 1 above we observe that $\sigma^{(\mathcal{F})}$ is full-dimensional unless $\mathcal{F} = 2^{[n]} \setminus \{\emptyset\}$. When $\mathcal{F} = 2^{[n]} \setminus \{\emptyset\}$, that is, when we know all intersection probabilities, codimension 1 with the stated hyperplane follows from the inclusion-exclusion formula (1). For Item 3, recall the proof of Proposition 8(2); observe that the cell of the set partition $P = \bigwedge_{S \in \mathcal{F}} \{\uparrow S, 2^{[n]} \setminus \uparrow S\}$ that contains the empty set equals $\{\emptyset\}$ if and only if $\{k\} \in \mathcal{F}$ for all $k \in [n]$; otherwise, the meet $P \wedge \{\{\emptyset\}, 2^{[n]} \setminus \{\emptyset\}\}$ splits this cell into exactly two cells. We have that $|P \wedge \{\{\emptyset\}, 2^{[n]} \setminus \{\emptyset\}\}|$ is the number of vertices in $\sigma^{(\mathcal{F})}$. For Item 4, recall our discussion of Hailperin's linear program in the introduction. ◀

▶ **Theorem 6** (Hardness of membership in a union polytope). *Given a set family $\emptyset \neq \mathcal{F} \subseteq 2^{[n]} \setminus \{\emptyset\}$, a rational vector $b = (b_S \in [0, 1] : S \in \mathcal{F})$, and a rational number $x_\emptyset \in [0, 1]$ as input, it is NP-hard to determine whether $(b, x_\emptyset) \in \sigma^{(\mathcal{F})}$. Moreover, this holds even when $\mathcal{F}$ consists of sets of size at most two and $b \in \rho^{(\mathcal{F})}$.*

Proof. We reduce from the fractional chromatic number problem. Let G be a graph with n vertices given as input and assume the setting in the proof of Theorem 2(1) as regards the instance $(b, \mathcal{F})$ of minimum union probability constructed in that proof. Let us write $[u_{\min}, u_{\max}]$ for the union probability interval of the instance $(b, \mathcal{F})$, and recall that $(b, x_\emptyset) \in \sigma^{(\mathcal{F})}$ if and only if $u = 1 - x_\emptyset \in [u_{\min}, u_{\max}]$. By construction, the fractional chromatic number $\chi_f(G)$ satisfisfies $u_{\min} = \chi_f(G)/n$, and from (7) we have $u_{\max} = 1$, where both $u_{\min}$ and $u_{\max}$ as well as the vector b have rational encoding length bounded by a polynomial in n.

Now suppose we have available an oracle that given as input the instance $(b, \mathcal{F})$ and a rational $x_\emptyset$ decides whether $(b, x_\emptyset) \in \sigma^{(\mathcal{F})}$; that is, whether $u = 1 - x_\emptyset \in [u_{\min}, u_{\max}] = [\chi_f(G)/n, 1]$. Using the oracle a polynomial number of times in n, starting from $u = 1/2$ we can in time polynomial in n use binary search to find a rational number u with $|u - u_{\min}| < \frac{1}{2D^2}$, where D is an upper bound for the denominator of $u_{\min}$; in particular, we can assume that D has encoding length bounded by a polynomial in n. This rational u thus enables us to compute the exact value $u_{\min} = \chi_f(G)/n$ from u in time polynomial in n with continued fractions [31, Theorem 5.1.9], and thus solve the fractional chromatic number of G in time and number of oracle calls bounded by a polynomial in n. ◀