Combinatorics, Probability and Computing (2015) 24, 438–456.
doi:10.1017/S096354831400042X

Simplifying Inclusion–Exclusion Formulas

XAVIER GOAOC,[1†] JIŘÍ MATOUŠEK,[2,3 ‡§¶] PAVEL PATÁK,[4§‖]
ZUZANA SAFERNOVÁ[2‡§] and MARTIN TANCER[2‡§]

[1]Université Paris–Est Marne-la-Vallée, France
(e-mail: goaoc@univ-mlv.fr)

[2]Department of Applied Mathematics, Charles University,
Malostranské nám. 25, 118 00 Praha 1, Czech Republic
(e-mail: {matousek,zuzka,tancer}@kam.mff.cuni.cz)

[3]Institute of Theoretical Computer Science, ETH Zurich, 8092 Zurich, Switzerland

[4]Department of Algebra, Charles University, Sokolovská 83, 186 75 Praha 8, Czech Republic
(e-mail: patak@kam.mff.cuni.cz)

Received 13 December 2012; revised 17 April 2014; first published online 14 October 2014

Let $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ be a family of n sets on a ground set S, such as a family of balls in $\mathbb{R}^d$. For every finite measure μ on S, such that the sets of $\mathcal{F}$ are measurable, the classical *inclusion–exclusion formula* asserts that

$$\mu(F_1 \cup F_2 \cup \cdots \cup F_n) = \sum_{I:\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} \mu\left(\bigcap_{i\in I} F_i\right),$$

that is, the measure of the union is expressed using measures of various intersections. The number of terms in this formula is exponential in n, and a significant amount of research, originating in applied areas, has been devoted to constructing simpler formulas for particular families $\mathcal{F}$. We provide an upper bound valid for an arbitrary $\mathcal{F}$: we show that every system $\mathcal{F}$ of n sets with m non-empty fields in the Venn diagram admits an inclusion–exclusion formula with $m^{O(\log^2 n)}$ terms and with ± 1 coefficients, and that such a formula can be computed in $m^{O(\log^2 n)}$ expected time. For every $\varepsilon > 0$ we also construct systems with Venn diagram of size m for which every valid inclusion–exclusion formula has the sum of absolute values of the coefficients at least $\Omega(m^{2-\varepsilon})$.

† This research was done while the author was affiliated with INRIA, project team Vegas. A visit to Prague was partially supported by grant GRADR Eurogiga GIG/11/E023.
‡ Supported by ERC advanced grant 267165.
§ Partially supported by Charles University grant GAUK 421511.
¶ Partially supported by grant GRADR Eurogiga GIG/11/E023.
‖ Partially supported by Charles University grant SVV-2012-265317.

2010 *Mathematics subject classification*: Primary 05A19
Secondary 05A15, 05E45

1. Introduction

One of the basic topics in introductory courses of discrete mathematics is the *inclusion–exclusion principle* (also called the *sieve formula*), which allows one to compute the number of elements of a union $F_1 \cup F_2 \cup \cdots \cup F_n$ of n sets from the knowledge of the sizes of all intersections of the F_i.

We will consider a slightly more general setting, where we have a ground set S and a (finite) *measure* μ on S; then the inclusion–exclusion principle asserts that, for every collection $F_1, F_2, \ldots, F_n$ of μ-measurable sets, we have

$$\mu\Big(\bigcup_{i=1}^{n} F_i\Big) = \sum_{I:\emptyset\neq I\subseteq[n]} (-1)^{|I|+1}\mu\Big(\bigcap_{i\in I} F_i\Big). \tag{1.1}$$

(Here, as usual, $[n] = \{1,2,\ldots,n\}$ and $|I|$ denotes the cardinality of the set I.) This principle not only plays a fundamental role in various areas of mathematics such as probability theory and combinatorics, but it also has important algorithmic applications. For instance, it provides simple methods for the computation of volume or surface area of molecules in computational biology [21] and underlies, through efficient computation of Möbius transforms [12, Section 4.3.4], the best known algorithms for several NP-hard problems including graph k-colouring [3], the travelling salesman problem on bounded-degree graphs [2], the dominating set problem [22], and partial dominating set and set splitting problems [19].

The inclusion–exclusion principle involves a number of summands that is exponential in n, the number of sets. In general this cannot be avoided if one wants an *exact* formula valid for *every* family $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$; see Example 2.2 below for a family for which equation (1.1) is the only solution. However, since this is a serious obstacle to efficient use of inclusion–exclusion, much effort has been devoted to finding 'smaller' formulas. These efforts are essentially organized along two lines of research.

The first approach gives up on exactness and tries to *approximate* efficiently the measure of the union using the measure of only *some* of the intersections. The first results of this flavour are the classical *Bonferroni inequalities* [4].[1] It turns out that better approximations can be obtained by replacing the coefficients $(-1)^{|I|+1}$ by other suitable numbers, and such Bonferroni-type inequalities have been studied extensively; see, *e.g.*, [8]. Linial and Nisan [14] and Kahn, Linial and Samorodnitsky [11] have investigated how well $\mu(F_1 \cup \cdots \cup F_n)$ can be approximated if we know the measure of all intersections $\bigcap_{i\in I} F_i$ for all $I \subseteq [n]$ of size at most r. Their main finding is that having r at least of order $\sqrt{n}$ is both necessary and sufficient for a reasonable approximation in the worst case. This still leaves us with about $2^{\sqrt{n}}$ terms in approximate inclusion–exclusion formulas.

[1] These assert that if we omit all terms with $|I| > r$ on the right-hand side of (1.1), then we get an upper bound for the left-hand side for r odd, and a lower bound for the left-hand side for r even. The case $r = 1$ is the often-used *union bound* in probability theory.

https://doi.org/10.1017/S096354831400042X Published online by Cambridge University Press

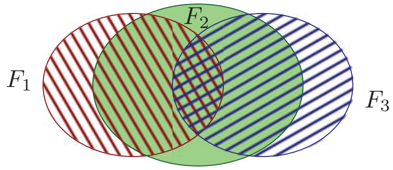

Figure 1. Three subsets of $\mathbb{R}^2$ admitting a simpler inclusion–exclusion formula. The ground set $F_1 \cup F_2 \cup F_3$ splits into six non-empty regions recognizable by the filling pattern.

The second line of research looks for 'small' inclusion–exclusion formulas valid for *specific* families of sets. To illustrate the type of simplifications afforded by fixing the sets, consider the family $\mathcal{F} = \{F_1, F_2, F_3\}$ of Figure 1. Since $F_1 \cap F_3 = F_1 \cap F_2 \cap F_3$, formula (1.1) can be simplified to

$$\mu(F_1 \cup F_2 \cup F_3) = \mu(F_1) + \mu(F_2) + \mu(F_3) - \mu(F_1 \cap F_2) - \mu(F_2 \cap F_3).$$

More generally, let us consider a family $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$, and let us say that a coefficient vector

$$\boldsymbol{\alpha} = (\alpha_I)_{\emptyset \neq I \subseteq [n]} \in \mathbb{R}^{2^n - 1}$$

is an *IE-vector for* $\mathcal{F}$ if we have

$$\mu\left(\bigcup_{i=1}^{n} F_i\right) = \sum_{I : \emptyset \neq I \subseteq [n]} \alpha_I \, \mu\left(\bigcap_{i \in I} F_i\right) \tag{1.2}$$

for every finite measure μ on the ground set of $\mathcal{F}$ (with all the F_i measurable). Given $\mathcal{F}$, we would like to find an IE-vector for $\mathcal{F}$ such that both the number of non-zero coefficients is small and the coefficients themselves are not too large. This idea, which we originally learned from [1], seems to originate in the work of Kratky [13] on families of disks in the plane, and a systematic study of such simplifications was initiated by Naiman and Wynn [17, 18]. A simplified inclusion–exclusion formula was also successfully used in an algorithm of Björklund, Husfeldt, Kaski and Koivisto [2]. We refer to the monograph of Dohmen [6] for an overview of this line of research.

Given a specific family $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ of sets, how small can we expect an inclusion–exclusion formula to be? This is, roughly speaking, the question we tackle in this paper. To formalize the problem, we should specify how $\mathcal{F}$ is given. Let us consider the *Venn diagram* of $\mathcal{F}$, which is the partition of the ground set S into equivalence classes according to the membership in the sets of $\mathcal{F}$. For each non-empty index set $\tau \subseteq [n]$, we define the *region* of τ, denoted by $\mathrm{reg}(\tau)$, as the set of all points that belong to the sets F_i with $i \in \tau$ and no others (see Figure 1):

$$\mathrm{reg}(\tau) = \left(\bigcap_{i \in \tau} F_i\right) \setminus \left(\bigcup_{i \notin \tau} F_i\right).$$

The *Venn diagram* of $\mathcal{F}$ is then the collection of all subsets of $[n]$ with non-empty regions, that is,

$$\mathcal{V} = \mathcal{V}(\mathcal{F}) := \{\tau \subseteq [n] : \mathrm{reg}(\tau) \neq \emptyset\}.$$

https://doi.org/10.1017/S096354831400042X Published online by Cambridge University Press

We regard the Venn diagram as a set system on the ground set $[n]$; it is a 'dual' of the set system $\mathcal{F}$.

We say that $\mathcal{F}$ is *standardized* if the ground set equals the union of the F_i and each non-empty region has exactly one point. It is easy to see that, as far as inclusion–exclusion formulas are concerned, all points in a single region are equivalent; it only matters which of the regions are non-empty. Therefore assuming that $\mathcal{F}$ is standardized does not mean a loss of generality. We will use this assumption in the algorithmic part of our main result, Theorem 1.1. For general $\mathcal{F}$ this requires a preprocessing step for $\mathcal{F}$, in which the part of the ground set S in each non-empty region is contracted to a single point.

Let $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ be a family of sets and let m denote the size of $\mathcal{V}$ (which equals the size of the ground set for $\mathcal{F}$ standardized). A linear-algebraic argument shows that *every* (finite) family $\mathcal{F}$ has an inclusion–exclusion formula with at most m terms (see Corollary 2.3) and m terms are sometimes necessary (see the beginning of Section 4). The question of how small a formula $\mathcal{F}$ admits may thus seem settled. There is, however, a caveat: this linear-algebraic argument may yield *exponentially large* coefficients (see Example 2.4). If we wanted to use such a formula, we would need to compute with very high precision, and perhaps more seriously, we would have to know the measures of the various intersections with enormous precision, in order to obtain a meaningful result. This may be totally impractical, for example in geometric settings where some physical measurements are involved, or where the measures of the intersections are computed with limited precision. Thus, we prefer inclusion–exclusion formulas where not only is the number of terms small but the coefficients are also small.

Our main result is the following general upper bound. To our knowledge, it is the first upper bound applicable to an arbitrary family.

Theorem 1.1. *Let n and m be integers and let*

$$D = \lceil 2e \ln m \rceil \left\lceil 2 + \ln \frac{n}{\ln m} \right\rceil.$$

Then, for every family $\mathcal{F}$ of n sets with Venn diagram of size m, there is an IE-vector $\boldsymbol{\alpha}$ for $\mathcal{F}$ that has at most

$$\sum_{i=1}^{D} \binom{n}{i} \leqslant m^{O(\ln^2 n)}$$

non-zero coefficients, and in which all non-zero coefficients are ± 1. Such an $\boldsymbol{\alpha}$ can be computed in $m^{O(\ln^2 n)}$ expected time if $\mathcal{F}$ is standardized.

The bound in this theorem is quasi-polynomial, but not polynomial, in m and n. We do not know if a polynomial bound can be achieved with ± 1 coefficients. We have at least the following lower bound, proved in Section 4, showing that inclusion–exclusion formulas of *linear* size are impossible in general.

Theorem 1.2. *For any $\varepsilon > 0$, for arbitrarily large values of m, there exists a family of sets with Venn diagram of size m for which any IE-vector has ℓ_1-norm at least $\Omega(m^{2-\varepsilon})$.*

https://doi.org/10.1017/S096354831400042X Published online by Cambridge University Press

We recall that the ℓ_1-norm of a real vector $\boldsymbol{x} \in \mathbb{R}^d$ is

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{d} |x_i|.$$

The ℓ_1-norm gives a lower bound on the tradeoff between the number of non-zero coefficients and their orders of magnitude (we recall that a formula with $O(m)$ non-zero coefficients is always attainable, the problem being that the coefficients may be too large).

Remark on ℓ_1-norm minimization. A useful heuristic for finding 'small' IE-vectors might be to look for an IE-vector of minimum ℓ_1-norm. In the linear-algebraic formulation, this means finding a solution of $A\boldsymbol{x} = \mathbf{1}$ of minimum ℓ_1-norm.

It is well known that finding a solution of minimum ℓ_1-norm of a linear system can be done in polynomial time, via linear programming. Several specialized algorithms for this problem have also been developed, with better performance than direct application of general-purpose linear programming solvers (see, *e.g.*, [24] for a recent overview). However, in our setting the number of columns of the matrix A may be exponential in m and n, and so even the input for an ℓ_1-norm minimizing algorithm would be too large.

There are linear programs with exponentially many variables (and polynomially many constraints) that can still be solved in polynomial time. For example, one may attempt, at least for theoretical purposes, to solve the dual linear program by the ellipsoid method, provided that a separation oracle is available.

In our setting, the task of the separation oracle can be formulated as follows in the setting of the original (standardized) set system $\mathcal{F} = \{F_1, \ldots, F_n\}$: *given weights* $w_1, \ldots, w_m \in \mathbb{Z}$ *of the points and threshold* c*, find a subset* $I \subseteq [n]$*, if one exists, such that the sum of weights of the points in* $\bigcap_{i \in I} F_i$ *is at least* c*.* Unfortunately, as was shown by Hoffmann, Okamoto, Ruiz-Vargas, Scheder and Solymosi [10], this problem is NP-complete not only for arbitrary set systems, but also, for example, for the case where each F_i is the complement of a hexagon in the plane. Thus, this approach does not seem to lead to a polynomial-time algorithm for finding an IE-vector of minimum ℓ_1-norm, even for rather simple geometric settings.

Topological background. In order to prove Theorem 1.1 we need several basic notions from topological combinatorics. We aim at a self-contained exposition that should make the proof accessible even to a reader who is not familiar with topological methods (we use the topological background mostly indirectly). For further reading we refer the reader to sources such as [9, 15, 16].

2. Preliminaries

We consider a family $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ of sets on a ground set S, and assume that the F_i are all distinct. Besides the Venn diagram $\mathcal{V}$, we associate yet another set system with

https://doi.org/10.1017/S096354831400042X Published online by Cambridge University Press

$\mathcal{F}$, namely, the *nerve*[2] $\mathcal{N}$ of $\mathcal{F}$:

$$\mathcal{N} = \mathcal{N}(\mathcal{F}) := \left\{\sigma \subseteq [n] : \sigma \neq \emptyset, \bigcap_{i\in\sigma} F_i \neq \emptyset\right\}.$$

So both of $\mathcal{N}$ and $\mathcal{V}$ have ground set $[n]$, and we have $\mathcal{V} \subseteq \mathcal{N}$.

Let us enumerate the elements of $\mathcal{V}$ as $\mathcal{V} = \{\tau_1, \tau_2, \ldots, \tau_m\}$ in such a way that $|\tau_i| \leqslant |\tau_j|$ for $i < j$, and let us enumerate $\mathcal{N} = \{\sigma_1, \sigma_2, \ldots, \sigma_{|\mathcal{N}|}\}$ so that the sets of $\mathcal{V}$ come first, *i.e.*, $\sigma_i = \tau_i$ for $i = 1, 2, \ldots, m$.

In the Introduction, we were indexing IE-vectors for $\mathcal{F}$ by all possible subsets $I \subseteq [n]$. But if I is not in the nerve, the corresponding intersection is empty, and thus without loss of generality we may assume that its coefficient is zero. Thus, from now on, we will index IE-vectors $\boldsymbol{x}$ as $(x_1, \ldots, x_{|\mathcal{N}|})$, where x_j is the coefficient of $\mu\big(\bigcap_{i\in\sigma_j} F_i\big)$.

IE-vectors from linear algebra. Let $A = (a_{jk})$ denote the 0–1 matrix with m rows and $|\mathcal{N}|$ columns such that $a_{jk} = 1$ if $\tau_j \supseteq \sigma_k$ and $a_{jk} = 0$ otherwise. Let $\mathbf{1}$ denote the m-dimensional vector with all entries equal to 1.

Lemma 2.1. *$\boldsymbol{x} \in \mathbb{R}^{|\mathcal{N}|}$ is an IE-vector for $\mathcal{F}$ if and only if $A\boldsymbol{x} = \mathbf{1}$.*

Proof. A vector $\boldsymbol{x} \in \mathbb{R}^{|\mathcal{N}|}$ is an IE-vector for $\mathcal{F}$ if and only if for every finite measure μ on S we have

$$\mu\left(\bigcup_{i=1}^{n} F_i\right) = \sum_{k=1}^{|\mathcal{N}|} x_k \mu\left(\bigcap_{i\in\sigma_k} F_i\right). \tag{2.1}$$

We first reformulate equation (2.1) using the regions of $\mathcal{F}$. The regions decompose $\bigcup_{i=1}^{n} F_i$ in a way that is compatible with the regions $\bigcap_{i\in\sigma} F_i$:

$$\bigcup_{i=1}^{n} F_i = \bigcup_{\tau\in\mathcal{V}} \mathrm{reg}(\tau) \quad \text{and for all } \sigma \in \mathcal{N}, \quad \bigcap_{i\in\sigma} F_i = \bigcup_{\tau\in\mathcal{V}:\tau\supseteq\sigma} \mathrm{reg}(\tau).$$

Moreover, the regions are pairwise disjoint. Thus, for every finite measure μ on S we have

$$\mu\left(\bigcup_{i=1}^{n} F_i\right) = \sum_{\tau\in\mathcal{V}} \mu\big(\mathrm{reg}(\tau)\big) \quad \text{and for all } \sigma \in \mathcal{N}, \quad \mu\left(\bigcap_{i\in\sigma} F_i\right) = \sum_{\tau\in\mathcal{V}:\tau\supseteq\sigma} \mu\big(\mathrm{reg}(\tau)\big),$$

and equation (2.1) is equivalent to

$$\sum_{\tau\in\mathcal{V}} \mu\big(\mathrm{reg}(\tau)\big) = \sum_{k=1}^{|\mathcal{N}|} x_k \left(\sum_{\tau\in\mathcal{V}:\tau\supseteq\sigma_k} \mu\big(\mathrm{reg}(\tau)\big)\right).$$

[2] This is the first notion from topological combinatorics that we need. Usually, a nerve also comes with an associated topological space that captures some of the properties of the underlying family $\mathcal{F}$. In our case, a purely combinatorial description of the nerve is sufficient. We also emphasize that the condition $\sigma \neq \emptyset$ in the definition of $\mathcal{N}(\mathcal{F})$ is not a standard one but it is convenient for our purposes.

Using the orderings on $\mathcal{V}$ and $\mathcal{N}$ and the definition of A, we obtain that $\boldsymbol{x} \in \mathbb{R}^{|\mathcal{N}|}$ is an IE-vector for $\mathcal{F}$ if and only if for every finite measure μ on S we have

$$\sum_{j=1}^{m} \mu\big(\mathrm{reg}(\tau_j)\big) = \sum_{k=1}^{|\mathcal{N}|} x_k \left(\sum_{j=1}^{m} a_{j,k} \mu\big(\mathrm{reg}(\tau_j)\big) \right) = \sum_{j=1}^{m} \left(\sum_{k=1}^{|\mathcal{N}|} a_{j,k} x_k \right) \mu\big(\mathrm{reg}(\tau_j)\big). \tag{2.2}$$

Now, if $A\boldsymbol{x} = \mathbf{1}$ then equation (2.2) trivially holds for all μ and $\boldsymbol{x}$ is an IE-vector for $\mathcal{F}$. Conversely, assume that $\boldsymbol{x}$ is an IE-vector for $\mathcal{F}$ and thus that equation (2.2) holds for all μ. For $1 \leqslant j \leqslant m$ we pick $p_j \in \mathrm{reg}(\tau_j)$ and define the measure $\mu_j : 2^S \to \mathbb{R}$ by $\mu_j(T) = 1$ if $p_j \in T$ and 0 otherwise. Equation (2.2) then specializes to

$$1 = \mu_j\big(\mathrm{reg}(\tau_j)\big) = \sum_{k=1}^{|\mathcal{N}|} x_k a_{j,k} \mu_j\big(\mathrm{reg}(\tau_j)\big) = \sum_{k=1}^{|\mathcal{N}|} a_{j,k} x_k.$$

This implies that $(A\boldsymbol{x})_j = 1$. The statement follows. □

Remark 1. In our definition a vector $\boldsymbol{x}$ is an IE-vector for $\mathcal{F}$ if and only if equation (1.2) is valid for every finite measure. As it follows from the proof of Lemma 2.1 this definition is equivalent to extending this requirement to every (finitely additive) signed measure. (A signed measure satisfies the classical axioms of a measure with the exception that it may take negative values.)

Example 2.2. Let $S = 2^{[n]} \setminus \{[n]\}$ and $F_i = 2^{[n]\setminus\{i\}}$ for $i \in [n]$. It is easy to see that here $\mathcal{N} = \mathcal{V}$ and A is a lower-triangular square matrix with ones on the diagonal. Hence A is invertible and, by Lemma 2.1, $\mathcal{F}$ has a unique IE-vector, namely, the one from the standard inclusion–exclusion formula.

Corollary 2.3. *For every finite family $\mathcal{F}$, there is a unique IE-vector $\boldsymbol{\alpha}$ supported on $\mathcal{V}$ (that is, such that $\alpha_I = 0$ for $I \notin \mathcal{V}$), and this $\boldsymbol{\alpha}$ has all entries integral.*

Proof. Let B be the $m \times m$ submatrix of A consisting of the first m columns of A. The IE-vectors for $\mathcal{F}$ supported on $\mathcal{V}$ are in one-to-one correspondence with the solutions of $B\boldsymbol{y} = \mathbf{1}$. Since B is lower-triangular and has ones on the main diagonal, it is non-singular, and hence $B\boldsymbol{y} = \mathbf{1}$ has exactly one solution. Moreover, since B is a lower-triangular 0–1 matrix, this solution is integral. □

Remark 2. The matrix B from the proof above can be regarded as the zeta-matrix of $\mathcal{V}$ ordered by inclusion. The vector $\boldsymbol{\alpha}$ from Corollary 2.3 can therefore be obtained via the Möbius inversion formula: see [23, Chapter 3].

This description also yields a recursive formula for $\boldsymbol{\alpha}$ which we use in Section 4. The condition $(B\boldsymbol{y})_j = 1$ becomes $\sum \alpha_\tau = 1$ where the sum is taken over all $\tau \in \mathcal{V}$ with $\tau \subseteq \tau_j$. That is, $\alpha_{\tau_j} = 1 - \sum \alpha_\tau$ where the sum is taken over all $\tau \in \mathcal{V}$ properly contained in τ_j.

Unfortunately, the IE-vector with small support given by Corollary 2.3 might have exponentially large coefficients, as the following example shows.

https://doi.org/10.1017/S096354831400042X Published online by Cambridge University Press

Example 2.4. Let $S = [5\ell]$ for some positive integer ℓ, and for $i \leqslant \ell$, let $g(i)$ stand for the smallest integer $j \geqslant i$ divisible by 5, that is, $g(i) = 5\lceil i/5 \rceil$. We consider the set system $\mathcal{F} = \{F_1, F_2, \dots, F_{5\ell}\}$ on S given by $F_i = \{i\} \cup \{g(i)+1, \dots, 5\ell\}$. Now $j \in F_i$ if and only if $i = j$ or $j > g(i)$. In particular, no two elements of S belong to the same region and the number of regions of $\mathcal{F}$ is $m = |S| = 5\ell$, which is also equal to the number n of sets in $\mathcal{F}$: $n = m = 5\ell$. The lower-triangular matrix B from the proof of Corollary 2.3 has a simple structure in terms of 5×5 blocks: the blocks on the diagonal are identity blocks, and the blocks below the diagonal are filled with ones. Let $\hat{\boldsymbol{x}}$ denote the solution of $B\boldsymbol{x} = \mathbf{1}$. The first five rows yield $\hat{x}_1 = \hat{x}_2 = \cdots = \hat{x}_5 = 1$. The next five rows imply that for $j = 6, 7, \dots, 10$ we have

$$\hat{x}_1 + \hat{x}_2 + \cdots + \hat{x}_5 + \hat{x}_j = 1,$$

and so $\hat{x}_6 = \hat{x}_7 = \cdots = \hat{x}_{10} = -4$. A simple induction yields $\hat{x}_i = (-4)^{(g(i)/5)-1}$. Altogether, the largest coefficient is of order $4^{n/5}$. (Replacing the constant 5 by another constant y yields a similar exponential growth with basis $(y-1)^{1/y}$; the choice $y = 5$ maximizes the basis of the exponent.)

Abstract tubes. Naiman and Wynn [17, 18] started their study of simplified inclusion–exclusion formulas with families $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$ that were tube-like in the sense that $F_i \cap F_j \subseteq F_k$ for all $i \leqslant k \leqslant j$ (as in our Figure 1). They then realized that the simplifications found for these ‘simple tubes’ hold in a broader setting, leading them to introduce the more general notion of an abstract tube. This notion will also play an important role in our considerations.

Definition 1. An (*abstract*) *simplicial complex* with vertex set $[n]$ is a hereditary system of non-empty subsets of $[n]$.[3] An *abstract tube* is a pair $(\mathcal{F}, \mathcal{K})$, where $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$ is a family of sets and $\mathcal{K}$ is a simplicial complex with vertex set $[n]$, such that for every non-empty region τ of the Venn diagram of $\mathcal{F}$, the subcomplex *induced* on $\mathcal{K}$ by τ, $\mathcal{K}[\tau] := \{\vartheta \in \mathcal{K} : \vartheta \subseteq \tau\}$, is contractible.[4]

As first noted by Naiman and Wynn [17, 18], if $(\mathcal{F}, \mathcal{K})$ is an abstract tube, then

$$\mu\left(\bigcup_{i=1}^{n} F_i\right) = \sum_{I \in \mathcal{K}} (-1)^{|I|+1} \mu\left(\bigcap_{i \in I} F_i\right). \tag{2.3}$$

Moreover, truncating the sum yields upper and lower bounds in the spirit of the Bonferroni inequalities ([18]; also see [6, Theorem 3.1.9]).

[3] As in the definition of the nerve, we exclude the empty set from the definition of a simplicial complex. This is again non-standard but convenient.

[4] By *contractible* we mean contractibility in the sense of topology; there is a topological space defined by $\mathcal{K}[\tau]$ and, roughly speaking, ‘contractible’ means that this space can be continuously shrunk to a point. Readers not at ease with this notion may want to look at Remark 3.

Remark 3. An earlier, more permissive definition of abstract tubes by [17] had the weaker condition '$\chi(\mathcal{K}[\tau]) = 1$' instead of '$\mathcal{K}[\tau]$ contractible', where χ is the *Euler characteristic*.[5] We recall that for a simplicial complex $\mathcal{L}$ in our sense, the Euler characteristic is defined as $\chi(\mathcal{L}) := \sum_{\sigma\in\mathcal{L}}(-1)^{|\sigma|+1}$. In this setting, if $(\mathcal{F},\mathcal{K})$ satisfies $\chi(\mathcal{K}[\tau]) = 1$ for every τ, then (2.3) can be proved in a few lines, using Lemma 2.1. Indeed, consider a simplicial complex $\mathcal{K}$ with vertex set $[n]$ and let $\boldsymbol{x} \in \mathbb{R}^{|\mathcal{N}|}$ stand for the vector with $x_k = (-1)^{|\sigma_k|+1}$ if $\sigma_k \in \mathcal{K}$ and $x_k = 0$ otherwise. Since

$$(A\boldsymbol{x})_j = \sum_{k:\sigma_k\subseteq\tau_j} x_k = \sum_{\sigma_k:\sigma_k\in\mathcal{K}[\tau_j]} (-1)^{|\sigma_k|+1},$$

we have $(A\boldsymbol{x})_j = \chi(\mathcal{K}[\tau_j])$. Thus, if all the $\mathcal{K}[\tau_j]$ have Euler characteristic 1, then $\boldsymbol{x}$ is an IE-vector, and (2.3) follows.

The stronger definition of abstract tubes involving contractibility, as opposed to the Euler characteristic, was needed in order to guarantee that truncations of equation (2.3) also yield Bonferroni-type inequalities [6, Theorem 3.1.9].

Small abstract tubes have been identified for families of balls [17, 18, 1] or half-spaces [18] in $\mathbb{R}^d$, and similar structures were found for families of pseudodisks [7]. We establish Theorem 1.1 by proving that for every family of sets there exists an abstract tube with 'small' size that, in addition, can be computed efficiently. We will use the following sufficient condition guaranteeing that $(\mathcal{F},\mathcal{K})$ is an abstract tube; it is a reformulation of [6, Theorem 4.2.5] (for the reader's convenience we include a simple proof). Let $\mathrm{MNF}(\mathcal{K})$ denote the system of all inclusion-minimal non-faces of $\mathcal{K}$, that is, of all non-empty sets $I \subseteq [n]$ with $I \notin \mathcal{K}$ but with $I' \in \mathcal{K}$ for every proper subset $I' \subset I$.

Proposition 2.5. *Let $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ be a family of sets with Venn diagram $\mathcal{V}$ and let $\mathcal{K}$ be a non-empty simplicial complex with vertex set $[n]$. If no set of $\mathcal{V}$ can be expressed as a union of sets in* $\mathrm{MNF}(\mathcal{K})$*, then $(\mathcal{F},\mathcal{K})$ is an abstract tube.* □

Proof. Let $\tau \in \mathcal{V}$ and let $a \in \tau$ such that a belongs to no element of $\mathrm{MNF}(\mathcal{K})$ contained in τ. Our task is to show that for every simplex $\vartheta \in \mathcal{K}[\tau]$ or $\vartheta = \emptyset$, we have $\vartheta \cup \{a\} \in \mathcal{K}[\tau]$. A simplicial complex $\mathcal{K}[\tau]$ satisfying the stated condition is known as a *cone* with *apex* a. Since every cone is contractible, it remains to show the condition.

If $\vartheta \cup \{a\} \notin \mathcal{K}[\tau]$, then $\vartheta \cup \{a\}$ contains some $\beta \in \mathrm{MNF}(\mathcal{K})$; since $\vartheta \in \mathcal{K}[\tau]$, the face β contains a, a contradiction. □

3. The upper bound: proof of Theorem 1.1

Abstract tubes from selectors. Let $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ be a family of sets, and let $\mathcal{V}$ be the Venn diagram of $\mathcal{F}$. A *selector* for $\mathcal{V}$ is a map $w : \mathcal{V} \to [n]$ such that $w(\tau) \in \tau$ for

[5] The fact that all contractible complexes have the same Euler characteristic follows from [9, Theorem 2.44]. The fact that it equals 1 can be verified on a point.

https://doi.org/10.1017/S096354831400042X Published online by Cambridge University Press

every $\tau \in \mathcal{V}$. For any selector w for $\mathcal{V}$ we define the simplicial complex

$$\mathcal{K}_w = \{\sigma \in \mathcal{N}(\mathcal{F}) : \text{for all non-empty } \vartheta \subseteq \sigma \text{ there exists } \tau \in \mathcal{V} \text{ such that } w(\tau) \in \vartheta \subseteq \tau\}.$$

We observe that $(\mathcal{F}, \mathcal{K}_w)$ is an abstract tube since the complex $\mathcal{K}_w$ satisfies the sufficient condition of Proposition 2.5.

Lemma 3.1. *For any selector w for $\mathcal{V}$, $(\mathcal{F}, \mathcal{K}_w)$ is an abstract tube.*

Proof. This is simple once the idea behind the definition of $\mathcal{K}_w$ is explained. Namely, in the condition of Proposition 2.5 we want to prevent each set $\tau \in \mathcal{V}$ from being a union of minimal non-faces of the simplicial complex $\mathcal{K}$. Our way of achieving that is to insist that every minimal non-face I contained in τ avoids the point $w(\tau)$; thus, we consider the set system of 'admissible minimal non-faces'

$$\mathcal{B}_w := \{I \subseteq [n], I \neq \emptyset : \text{ if } I \subseteq \tau \in \mathcal{V}, \text{ then } w(\tau) \notin I\}.$$

Then the above definition of $\mathcal{K}_w$ can be interpreted as follows: a simplex $\sigma \in \mathcal{N}$ belongs to $\mathcal{K}_w$ if it contains no $I \in \mathcal{B}_w$.[6] (Simplices outside $\mathcal{N}$ can be ignored, since their supersets cannot be contained in a set $\tau \in \mathcal{V}$.) Therefore, all minimal non-faces of $\mathcal{K}_w$ belong to $\mathcal{B}_w$ or lie outside $\mathcal{N}$, and hence $(\mathcal{F}, \mathcal{K}_w)$ is an abstract tube by Proposition 2.5. □

Let us remark that there is no loss of generality in passing from the abstract tubes as in Proposition 2.5 to those of the form $\mathcal{K}_w$. Indeed, if $\mathcal{K}$ satisfies the condition of Proposition 2.5, then every $\tau \in \mathcal{V}$ contains at least one point that is not contained in any minimal non-face I of $\mathcal{K}$ with $I \subseteq \tau$, and such a point can be chosen as $w(\tau)$; then we can easily check that $\mathcal{K}_w \subseteq \mathcal{K}$. (It is sufficient to check that if I is a minimal non-face of $\mathcal{K}$, then it is also a non-face of $\mathcal{K}_w$. For this we point out that such a minimal non-face I of $\mathcal{K}$ belongs to the set $\mathcal{B}_w$ defined above. Therefore it is a non-face of $\mathcal{K}_w$, possibly not a minimal one.)

No large simplices in random $\mathcal{K}_w$. Let ρ be a permutation of $[n]$. We define a selector w_ρ for $\mathcal{V}$ by taking $w(\tau)$ as the smallest element of τ in the linear ordering $\prec$ on $[n]$ given by $\rho(1) \prec \rho(2) \prec \cdots \prec \rho(n)$.

For better readability we write $\mathcal{K}_\rho$ instead of $\mathcal{K}_{w_\rho}$. We want to show that for random ρ, $\mathcal{K}_\rho$ is unlikely to contain too large simplices, and thus leads to a small inclusion–exclusion formula.

Let Γ denote the incidence matrix of $\mathcal{V}$, that is, the 0–1 matrix with m rows and n columns where $\Gamma_{ij} = 1$ if and only if $j \in \tau_i$ (if the original system $\mathcal{F}$ was standardized, then Γ is the transposition of the usual incidence matrix of $\mathcal{F}$). We also denote by Γ_ρ the matrix obtained by applying the permutation ρ to the columns of Γ: the $\rho(i)$th column of

[6] Note that for the formal verification, the condition σ contains no $I \in \mathcal{B}_w$ can be written, in symbols, as follows: $\forall I \subseteq [n], I \neq \emptyset : ((\forall \tau \in \mathcal{V} : I \subseteq \tau \Rightarrow w(\tau) \notin I) \Rightarrow I \not\subseteq \sigma)$. This is equivalent to $\forall I \subseteq [n], I \neq \emptyset : I \subseteq \sigma \Rightarrow (\exists \tau \in \mathcal{V} : I \subseteq \tau \wedge w(\tau) \in I)$ which is just a transcription of $\sigma \in \mathcal{K}_w$.

https://doi.org/10.1017/S096354831400042X Published online by Cambridge University Press

		i_1	i_2	i_3		i_4		i_5	
$\vdots$									
j_3	$0\cdots0$	0	0	1	$*\cdots*$	1	$*\cdots*$	1	$*\cdots*$
$\vdots$									
j_1	$0\cdots0$	1	1	1	$*\cdots*$	1	$*\cdots*$	1	$*\cdots*$
j_2	$0\cdots0$	0	1	1	$*\cdots*$	1	$*\cdots*$	1	$*\cdots*$
j_4	$0\cdots0$	0	0	0	$0\cdots0$	1	$*\cdots*$	1	$*\cdots*$
j_5	$0\cdots0$	0	0	0	$0\cdots0$	0	$0\cdots0$	1	$*\cdots*$
$\vdots$									

Figure 2. Illustration of Lemma 3.2: If $\rho(\tau) = \{i_1, i_2, \ldots, i_5\}$ for a simplex τ of $\mathcal{K}_\rho$, then Γ_ρ must contain a row j_s compatible with $\{i_s, i_{s+1}, \ldots, i_5\}$ for $s = 1, 2, \ldots, 5$. The j_3 row is emphasized, constrained values appearing in grey; rows j_s for other values of s are represented consecutively for clarity, but they can appear in any order and non-consecutively.

Γ_ρ is the ith column of Γ and represents the incidences between permuted $[n]$ and $\mathcal{V}$. We now argue that if $\mathcal{K}_\rho$ contains a large simplex, then Γ_ρ contains a particular substructure.

We say that a row R of Γ_ρ is *compatible* with a subset $I \subseteq [n]$ if R contains ones in all columns with index in I and zeros in all columns with index smaller than $\min(I)$.

Lemma 3.2. *If $\rho(\tau) = \{i_1, i_2, \ldots, i_k\}$ for a simplex τ in $\mathcal{K}_\rho$, with $i_1 < i_2 < \cdots < i_k$, then for every $s \in \{1, 2, \ldots, k\}$ the matrix Γ_ρ contains a row compatible with $\{i_s, i_{s+1}, \ldots, i_k\}$.*

Proof. Let $s \in \{1, 2, \ldots, k\}$, let $I_s = \{i_s, i_{s+1}, \ldots, i_k\}$, and let $\vartheta_s = \rho^{-1}(I_s)$. We refer to Figure 2. Since ϑ_s is a simplex of $\mathcal{K}_\rho$, there exists $\tau_{j_s} \in \mathcal{V}$ such that $w_\rho(\tau_{j_s}) \in \vartheta_s \subseteq \tau_{j_s}$ by definition of $\mathcal{K}_\rho$. Since $\vartheta_s \subseteq \tau_{j_s}$, we have $I_s = \rho(\vartheta_s) \subseteq \rho(\tau_{j_s})$, and hence the j_sth row of Γ_ρ has ones in all columns with index in I_s. Since $w_\rho(\tau_{j_s}) \in \vartheta_s$, the set $\rho(\tau_{j_s})$ contains no i with $i < i_s$ and the j_sth row of Γ_ρ has zeros in all columns with index smaller than $i_s = \min(I_s)$. It follows that the j_sth row of Γ_ρ is compatible with I_s. □

We will need the following inequality.

Lemma 3.3. *Let $x_1, \ldots, x_r$ be positive real numbers with $x_1 + \cdots + x_r \leqslant n$. Then*

$$\frac{x_1}{x_1 + \cdots + x_r} \cdot \frac{x_2}{x_2 + \cdots + x_r} \cdots \frac{x_{r-1}}{x_{r-1} + x_r} \leqslant \left(1 - \sqrt[r-1]{\frac{x_r}{n}}\right)^{r-1}.$$

https://doi.org/10.1017/S096354831400042X Published online by Cambridge University Press

Proof. Let us set $y_\ell := x_\ell + x_{\ell+1} + \cdots + x_r$. Then we have

$$\begin{aligned}
&\frac{x_1}{x_1+\cdots+x_r}\cdot\frac{x_2}{x_2+\cdots+x_r}\cdots\frac{x_{r-1}}{x_{r-1}+x_r}\\
&\quad=\frac{y_1-y_2}{y_1}\cdot\frac{y_2-y_3}{y_2}\cdots\frac{y_{r-1}-y_r}{y_{r-1}}\\
&\quad=\left(1-\frac{y_2}{y_1}\right)\cdot\left(1-\frac{y_3}{y_2}\right)\cdots\left(1-\frac{y_r}{y_{r-1}}\right)\\
&\quad\leqslant\left(\frac{1-y_2/y_1+1-y_3/y_2+\cdots+1-y_r/y_{r-1}}{r-1}\right)^{r-1}\\
&\quad=\left(1-\frac{y_2/y_1+y_3/y_2+\cdots+y_r/y_{r-1}}{r-1}\right)^{r-1}\\
&\quad\leqslant\left(1-\sqrt[r-1]{\frac{y_r}{y_1}}\right)^{r-1}\\
&\quad\leqslant\left(1-\sqrt[r-1]{\frac{x_r}{n}}\right)^{r-1}.
\end{aligned}$$

□

Now we aim at showing that for a random ρ, the condition in Lemma 3.2 is unlikely to be satisfied for large k. That condition prescribes the existence of k rows in Γ_ρ with a certain pattern. In order to get a good bound for k, we will not actually look for all of these k rows, but rather we will consider only each bth of them, for a suitable integer parameter b, and ignore the rest.

Namely, we fix two parameters r and b with $1 < b < n$ and set $k = rb$ (we think of $r \approx \ln n$ and $b \approx \ln m$). For an r-element index set $J \subseteq [m]$, let $\Gamma_\rho[J]$ denote the submatrix obtained from Γ_ρ by considering only the rows with indices in J. We say that a permutation ρ is *bad* for J if there exists a k-element set of column indices $I = \{i_1, i_2, \ldots, i_k\}$ with $i_1 < i_2 < \cdots < i_k$ such that for every $s \in \{1, b+1, \ldots, (r-1)b+1\}$, the matrix $\Gamma_\rho[J]$ contains a row compatible with $\{i_s, i_{s+1}, \ldots, i_k\}$. Finally, we define p_J as the probability that a random permutation ρ is bad for J.

Lemma 3.4. *We have* $p_J \leqslant (1-(b/n)^{1/(r-1)})^{b(r-1)}$.

Proof. Let ρ be a bad permutation for J, and let $I = \{i_s, i_{s+1}, \ldots, i_k\}$ be the corresponding set of column indices.

Let $\ell \in \{0, 1, \ldots, r-1\}$. By the compatibility conditions we have that for $i < i_{\ell\cdot b+1}$ the ith column of $\Gamma_\rho[J]$ contains at most ℓ entries 1: see Figure 3. Moreover, for

$$i \in \{i_{\ell\cdot b+1}, i_{\ell\cdot b+2}, \ldots, i_{(\ell+1)\cdot b}\},$$

the ith column of $\Gamma_\rho[J]$ contains exactly $\ell+1$ entries 1.

We now partition $[n]$ into $[n] = Q_0 \cup Q_1 \cup \cdots \cup Q_r$, where Q_ℓ consists of the indices of those columns of $\Gamma_\rho[J]$ that contain exactly ℓ entries 1 (and $r-\ell$ entries 0). In particular, from the discussion above, $|Q_\ell| \geqslant b$ for $\ell \in [r]$. For $\ell \in [r]$ and $p \in [b]$, let $g_\ell^{(p)}$ denote the

https://doi.org/10.1017/S096354831400042X Published online by Cambridge University Press

	i_1		i_2	$\cdots$	i_b		i_{b+1}	$\cdots$	i_{2b}		i_{2b+1}	$\cdots\cdots$	$i_{(r-1)b+1}$	$\cdots$	i_{rb+1}	
$0\cdots0$	1	$*$	1	$\cdots$	1	$*$	1	$\cdots$	1	$*$	1	$\cdots\cdots$	1	$\cdots$	1	$*$
0			$\cdots$			0	1	$\cdots$	1	$*$	1	$\cdots\cdots$	1	$\cdots$	1	$*$
0					$\cdots$					0	1	$\cdots\cdots$	1	$\cdots$	1	$*$
							$\vdots$									
0							$\cdots$					0	1	$\cdots$	1	$*$

Figure 3. Compatibility conditions in Lemma 3.4. Only the rows of J are shown and similarly as before, and their order can be arbitrary.

pth smallest element of $\rho(Q_\ell)$. A necessary condition on ρ is

$$g_1^{(b)} < g_2^{(1)} < g_2^{(b)} < g_3^{(1)} < \cdots < g_{r-1}^{(b)} < g_r^{(1)}.$$

Now, let us assume that ρ is a random permutation (uniformly chosen). For $\ell \in [r]$, let E_ℓ denote the event $E_\ell := \{g_\ell^{(b)} < \min(g_{\ell+1}^{(1)}, g_{\ell+2}^{(1)}, \ldots, g_r^{(1)})\}$, and we bound p_J by the conditional probability

$$p_J \leqslant \mathbb{P}(E_1) \cdot \mathbb{P}(E_2|E_1) \cdot \mathbb{P}(E_3|E_1 \cap E_2) \cdots \mathbb{P}(E_{r-1}|E_1 \cap \cdots \cap E_{r-2}). \tag{3.1}$$

For $\ell \in [r-1]$, $\mathbb{P}(E_\ell|E_1 \cap \cdots \cap E_{\ell-1})$ is the probability that the b smallest elements of $\rho(Q_\ell) \cup \rho(Q_{\ell+1}) \cup \cdots \cup \rho(Q_r)$ belong to $\rho(Q_\ell)$. This probability is equal to

$$\binom{|Q_\ell|}{b} \Big/ \binom{|Q_\ell| + |Q_{\ell+1}| + \cdots + |Q_r|}{b} \leqslant \left(\frac{|Q_\ell|}{|Q_\ell| + |Q_{\ell+1}| + \cdots + |Q_r|}\right)^b.$$

So, letting $x_\ell = |Q_\ell|$, inequality (3.1) implies

$$\begin{aligned} p_J &\leqslant \left(\frac{x_1}{x_1 + x_2 + \cdots + x_r} \cdot \frac{x_2}{x_2 + x_3 + \cdots + x_r} \cdot \ldots \cdot \frac{x_{r-1}}{x_{r-1} + x_r}\right)^b \\ &\leqslant \left(1 - \sqrt[r-1]{\frac{|Q_r|}{n}}\right)^{b(r-1)}, \end{aligned}$$

the last inequality being Lemma 3.3. Then the lemma follows using $|Q_r| \geqslant b$. □

Proof of Theorem 1.1. Let n and $m \geqslant 2$ be integers.[7] Let $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ be a family of n sets whose Venn diagram $\mathcal{V}$ has size m. Let $p(k)$ denote the probability that $\mathcal{K}_\rho$ contains at least one simplex of size k, where ρ is chosen uniformly at random among all permutations of $[n]$. From Lemmas 3.2 and 3.4, for every $r > 2$ and $b \geqslant 2$ we have

$$p(rb) \leqslant \binom{m}{r} \left(1 - \sqrt[r-1]{b/n}\right)^{b(r-1)} \leqslant m^r e^{b(r-1)\ln(1 - \sqrt[r-1]{b/n})} \leqslant m^r e^{-b(r-1)\sqrt[r-1]{b/n}}.$$

Assuming that $b \geqslant 2e \ln m$, we get $p(rb) \leqslant m^{r - 2e(r-1)\sqrt[r-1]{b/n}}$, and choosing $r \geqslant 1 + \ln(n/b)$, we obtain

$$\sqrt[r-1]{b/n} = \exp\left(-\frac{1}{r-1}\ln\frac{n}{b}\right) \geqslant e^{-1} \quad \text{and} \quad p(rb) \leqslant m^{2-r} \leqslant \frac{1}{2}.$$

[7] Note that the case $m = 1$ is somewhat trivial since every maximal face of $\mathcal{N}$ belongs to $\mathcal{V}$, and thus there is an IE-vector with a single non-zero coefficient, namely 1, in this case.

Thus, with

$$D = \lceil 2e \ln m \rceil \left\lceil 2 + \ln \frac{n}{\ln m} \right\rceil$$

as in the theorem, we have $p(D) \leqslant 1/2$ (note that setting $r = \lceil 2 + \ln(n/\ln m) \rceil$ implies $r > 2$ as required since $m \leqslant 2^n$). So there exists a permutation ρ^* of $[n]$ such that $\mathcal{K}_{\rho^*}$ contains no simplex of size D (or larger). By Lemma 3.1, $(\mathcal{F}, \mathcal{K}_{\rho^*})$ is an abstract tube and $\mathcal{K}_{\rho^*}$ has at most $\sum_{i=1}^{D} \binom{n}{i}$ simplices. The IE-vector obtained from the abstract tube $(\mathcal{F}, \mathcal{K}_{\rho^*})$ as in equation (2.3) is as claimed in the theorem.

In order to actually compute a suitable coefficient vector, we choose a random permutation ρ and compute $\mathcal{K}_\rho$ by the following incremental algorithm. We use two auxiliary set systems $\mathcal{A}$ and $\mathcal{B}$, initialized to $\mathcal{A} = \mathcal{B} = \{\emptyset\}$ (the idea is that $\mathcal{B}$ contains all the simplices of $\mathcal{K}_\rho$ found so far, and $\mathcal{A} \subseteq \mathcal{B}$ contains those for which we still need to test one-element extensions). In each step, we take some $\sigma \in \mathcal{A}$, remove it from $\mathcal{A}$, and for each $i \notin \sigma$, we test whether $\sigma \cup \{i\}$ belongs to $\mathcal{K}_\rho$ (for this, we just check if there is $\tau \in \mathcal{V}$ such that $w_\rho(\tau) \in \sigma \cup \{i\} \subseteq \tau$; note that we have a direct access to $\mathcal{V}$ in $O(m)$ time since $\mathcal{F}$ is standardized). Those $\sigma \cup \{i\}$ that pass this test are added to both $\mathcal{A}$ and $\mathcal{B}$. The algorithm finishes either when $\mathcal{A} = \emptyset$ (in this case we set $\mathcal{K}_\rho = \mathcal{B} \setminus \{\emptyset\}$ and return the corresponding IE-vector), or when we first discover a simplex $\sigma \in \mathcal{K}_\rho$ of size larger than D. In the latter case, we discard the current permutation ρ, choose a new one, and repeat the algorithm.

The choice of a random permutation ρ takes $O(n \ln n)$ time and n random bits. Accepting or rejecting a new simplex by brute-force testing takes $O(mn)$ time. The expected number of times we have to start over with a new permutation ρ is $O(1)$. Altogether, the expected running time of this algorithm is

$$O\left(\binom{n}{D} mn \right) = m^{O(\ln^2 n)}. \qquad \square$$

4. The lower bound: proof of Theorem 1.2

For every m between n and 2^n there exists a system of n sets with Venn diagram of size m whose only IE-vector has m non-zero entries. Indeed, let $\mathcal{K} = \{\vartheta_1, \vartheta_2, \ldots, \vartheta_m\}$ be a simplicial complex over $[n]$ such that $[n] = \bigcup \mathcal{K}$ and $|\mathcal{K}| = m$. We define $F_i = \{t \in [m] : i \in \vartheta_t\}$ for $1 \leqslant i \leqslant n$ and put $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$. It can easily be checked that $\mathcal{V}(\mathcal{F}) = \mathcal{N}(\mathcal{F}) = K$ and so, as observed in Example 2.2, the matrix A is square, lower-triangular, and has ones on the diagonal; thus, there is a unique IE-vector for $\mathcal{F}$ and it has m non-zero entries. In this section we improve on this lower bound.

We recall that by Corollary 2.3, every set system $\mathcal{F}$ has a unique IE-vector with support in the Venn diagram $\mathcal{V}(\mathcal{F})$. We first argue that for some set systems constructed from lattices, this IE-vector is the one with minimal ℓ_1-norm. We then provide an explicit construction, based on projective spaces over finite fields, where the ℓ_1-norm is near-quadratic in m.

Set systems from lattices. We need to work with (finite) lattices as order-theoretic notions. A finite partially ordered set L is a *lattice* if, for every subset S of L, there is the least upper bound for S called the *join* of S and the greatest lower bound called the *meet* of S. A finite lattice always contains the least element 0. An *atom* is an element $a \in L$ such that 0 is the only element less than a. A lattice is *atomistic* if each element is a join of some subset of atoms.

Given a finite atomistic lattice L we construct the following set system $\mathcal{F} = \mathcal{F}(L)$. Up to a relabelling, we can assume that the set of atoms of L is $\mathbf{At} = \{1, 2, \ldots, n\}$. For every atom $a \in \mathbf{At}$ we define $F_a := \{x \in L : x \geqslant a\}$, and for every $x \in L$ we set $\mathbf{At}_x := \{a \in \mathbf{At} : a \leqslant x\}$. For $\mathcal{F} = \{F_a : a \in \mathbf{At}\}$ we have $\mathcal{V}(\mathcal{F}) = \{\mathbf{At}_x : x \in L \setminus \{0\}\}$. In particular, $\mathcal{V}(\mathcal{F})$ equipped with the inclusion relation is isomorphic to $L \setminus \{0\}$. Also note that x is the join of $\mathbf{At}_x$ since L is atomistic.

Lemma 4.1. *Let L be a finite atomistic lattice and let $\mathcal{F} = \mathcal{F}(L)$ be the set system described above. Then, among all IE-vectors for $\mathcal{F}$, the one with support in $\mathcal{V}(\mathcal{F})$ has minimal ℓ_1-norm.*

Proof. Let A be the matrix with rows indexed by $\mathcal{V}$ and columns indexed by $\mathcal{N} = \mathcal{N}(\mathcal{F})$, as defined before Lemma 2.1, and let B be the $m \times m$ submatrix consisting of the first m columns of A.

We want to show that every column of A is equal to a column of B. By the definition of A, this means that for every $\sigma \in \mathcal{N}$ we need to find some $v \in \mathcal{V}$ such that

$$\{\tau \in \mathcal{V} : \sigma \subseteq \tau\} = \{\tau \in \mathcal{V} : v \subseteq \tau\}.$$

We set s to be the join of σ. (Note that σ is a subset of $[n] = \mathbf{At}$ and, therefore, of L.) We aim to show that $\mathbf{At}_s$ is the required v. This way, we have obtained a $v \in \mathcal{V}$ such that the join of v equals the join of σ since s is also the join of the atoms contained in $\mathbf{At}_s$. A set $\tau \in \mathcal{V}$ can also be described as $\mathbf{At}_x$ for some $x \in L \setminus \{0\}$ due to our description of $\mathcal{V}$. Then the condition $\sigma \subseteq \tau$ translates to $x \geqslant a$ for every $a \in \sigma$. This is equivalent to $x \geqslant s$ since s is the join of σ. Similarly, $v \subseteq \tau$ translates to $x \geqslant a$ for every $a \in v$, which is again equivalent to $x \geqslant s$. Therefore, $\sigma \subseteq \tau$ if and only if $v \subseteq \tau$, as required.

Hence every column of A occurs in B as asserted. It follows that every solution of $A\boldsymbol{x} = \mathbf{1}$ can be transformed to a solution of $B\boldsymbol{y} = \mathbf{1}$ with the same or smaller ℓ_1-norm (if k is the index of a column outside B with $x_k \neq 0$, and that kth column equals the jth column of B, then we can zero out x_k while replacing x_j with $x_j + x_k$). Since $B\boldsymbol{y} = \mathbf{1}$ has a unique solution, it has to be a solution of minimum ℓ_1-norm as claimed. □

Construction based on projective spaces. Let q be a power of a prime number. Let P be a projective space of dimension d over the finite field F_q. That is, the points of P are all one-dimensional subspaces of the vector space F_q^{d+1}, and k-dimensional subspaces of P correspond to $(k+1)$-dimensional linear subspaces of F_q^{d+1}. We let L be the lattice of all subspaces of P (including the zero one, of projective dimension -1, as zero), where the join of subspaces of P corresponds to the (projective) span and the meet corresponds to the intersection. It is easy to check (and well known) that L is an atomistic lattice.

https://doi.org/10.1017/S096354831400042X Published online by Cambridge University Press

We obtain our lower bound from the family $\mathcal{F} = \mathcal{F}(L)$ and so, according to Lemma 4.1, we need only compute the size of $\mathcal{V}(\mathcal{F})$ and the ℓ_1-norm for the IE-vector with support in $\mathcal{V}$ to provide a lower bound. In order to do so, we need to work with q-binomial coefficients.

Definition 2 (q-binomial coefficients).

(i) Given a positive integer k, we define $[k]_q := 1 + q + q^2 + \cdots + q^{k-1}$.
(ii) Given non-negative integers n and k with $n \geqslant k$, we define

$$\binom{n}{k}_q := \frac{[n]_q[n-1]_q[n-2]_q \cdots [n-k+1]_q}{[1]_q[2]_q[3]_q \cdots [k]_q}.$$

We remark that it is well known that $\binom{n}{k}_q$ is actually a polynomial in q since the division is exact. From the definition above we deduce that the leading term of $\binom{n}{k}_q$ is $q^{k(n-k)}$. We also need the following facts regarding q-binomial coefficients to finish the calculations. See, for example, [5] and [20].

Lemma 4.2.

(i) *The number of k-dimensional subspaces of a d-dimensional projective space over F_q is $\binom{d+1}{k+1}_q$.*
(ii) (*The Cauchy binomial theorem*)

$$\sum_{i=0}^{k} q^{(i(i-1))/2} \binom{k}{i}_q t^i = \prod_{i=0}^{k-1} (1 + tq^i).$$

Now we can finally estimate the size of $|\mathcal{V}(\mathcal{F})|$ and the ℓ_1-norm of the resulting IE-formula.

Lemma 4.3.

(i) *The number of non-empty subspaces of P, that is, the size of $\mathcal{V}(\mathcal{F})$ is $\Theta\big(q^{\lfloor (d+1)^2/4 \rfloor}\big)$.*
(ii) *In the (unique) IE formula for $\mathcal{F}$, the coefficients of the subspaces of dimension k are all equal to $(-1)^k q^{(k(k+1))/2}$.*
(iii) *The ℓ_1-norm of the resulting IE-formula is $\Theta\big(q^{(d(d+1))}/2\big)$.*

Proof. Concerning statement (i), Lemma 4.2(i) implies that

$$|\mathcal{V}(\mathcal{F})| = |L \setminus \{0\}| = \sum_{k=0}^{d} \binom{d+1}{k+1}_q,$$

which is a polynomial in q. Since we know that the leading term of $\binom{d+1}{k+1}_q$ is

$$q^{(k+1)((d+1)-(k+1))},$$

https://doi.org/10.1017/S096354831400042X Published online by Cambridge University Press

we deduce that the middle q-binomial coefficient(s) has/have the leading term of the highest power. That is, the leading term of the polynomial above equals $q^{\lfloor (d+1)^2/4 \rfloor}$ or $2q^{\lfloor (d+1)^2/4 \rfloor}$ (depending on the parity of d), as required.

We prove statement (ii) by induction. The statement clearly holds for $k = 0$. Suppose that it is valid for all $i < k$. Using Lemma 4.2(i) again, we see that every subspace of dimension k has $\binom{k+1}{i+1}_q$ subspaces of dimension i. Therefore, using the recursive formula from Remark 2, the coefficient of this subspace has to be

$$1 - \sum_{i=0}^{k-1} (-1)^i q^{(i(i+1))/2} \binom{k+1}{i+1}_q = \sum_{j=0}^{k} (-1)^j q^{(j(j-1))/2} \binom{k+1}{j}_q.$$

However, using the Cauchy binomial theorem for the second equality below, this sum equals

$$\begin{aligned} \sum_{j=0}^{k} (-1)^j q^{(j(j-1))/2} \binom{k+1}{j}_q &= (-1)^k q^{(k(k+1))/2} + \sum_{j=0}^{k+1} (-1)^j q^{(j(j-1))/2} \binom{k+1}{j}_q \\ &= (-1)^k q^{(k(k+1))/2} + \prod_{j=0}^{k} (1 - q^j) \\ &= (-1)^k q^{(k(k+1))/2} + 0, \end{aligned}$$

which concludes the induction.

It remains to prove statement (iii). Using statement (ii), we deduce that the ℓ_1-norm of the resulting formula equals

$$\sum_{k=0}^{d} q^{(k(k+1))/2} \binom{d+1}{k+1}_q.$$

The leading term of this polynomial (in q) is $2q^{(d(d+1))/2}$. Indeed, the leading term of

$$q^{(k(k+1))/2} \binom{d+1}{k+1}_q$$

equals

$$q^{(k(k+1))/2+(k+1)(d-k)},$$

and is greatest for $k = d$ and $k = d - 1$. □

Proof of Theorem 1.2. Fix $\varepsilon > 0$ and let $d > 2/\varepsilon$ be some integer, chosen to be odd for simplicity. Recall that the above analysis holds for any q that is a prime power, so q can be chosen arbitrarily large. The set system $\mathcal{F}(L)$ consists of $n = [d+1]_q = \Theta(q^d)$ sets. The Venn diagram $\mathcal{V}(\mathcal{F}(L))$ has size $m = \Theta\big(q^{((d+1)^2)/4}\big)$ and the ℓ_1-norm of the formula supported by the Venn diagram is

$$\Theta\big(q^{(d(d+1))/2}\big) = \Theta\big(m^{4/((d+1)^2)\cdot(d(d+1))/2}\big) = \Theta\big(m^{2-2/(d+1)}\big) \geqslant \Omega\big(m^{2-\varepsilon}\big).$$

Lemma 4.1 ensures that this formula minimizes the ℓ_1-norm. □

https://doi.org/10.1017/S096354831400042X Published online by Cambridge University Press

5. Open problems

For several NP-hard problems, the best exponential-time algorithms rely on the inclusion–exclusion principle [3, 22, 19]. Whether these algorithms can be improved using Theorem 1.1 is an open problem that is perhaps best illustrated on an example.

Consider for instance the question of counting the number of *k-covers*: given a family $\mathcal{X} = \{X_1, X_2, \ldots, X_p\}$ of subsets of $[n]$, we want to determine how many k-element subsets of $\mathcal{X}$ have their union equal to $[n]$. Björklund, Husfeldt and Koivisto [3, Section 3.1] proposed the following approach. For $i \in [n]$, let F_i denote the set of k-element subsets of $\mathcal{X}$ whose union does not contain i:

$$F_i = \{(Y_1, Y_2, \ldots, Y_k) \in \mathcal{X} : i \notin Y_1 \cup Y_2 \cup \cdots \cup Y_k\}.$$

For a subset $\sigma \subseteq [n]$ let $\mathrm{av}(\sigma) = \{X \in \mathcal{X} : X \cap \sigma = \emptyset\}$. The number c_k of k-covers of $\mathcal{X}$ can be written, using the inclusion–exclusion principle, as

$$c_k = |\mathcal{X}|^k - \left|\bigcup_{i=1}^{n} F_i\right| = |\mathcal{X}|^k - \sum_{\emptyset \subsetneq \sigma \subseteq [n]} (-1)^{|\sigma|+1} \left|\bigcap_{i\in\sigma} F_i\right| = |\mathcal{X}|^k - \sum_{\emptyset \subsetneq \sigma \subseteq [n]} (-1)^{|\sigma|+1} |\,\mathrm{av}(\sigma)|^k. \tag{5.1}$$

Let f denote the indicator function of X and let $\tilde{f}$ be its *Möbius* transform: for $I \subseteq [n]$,

$$\tilde{f}(I) = \sum_{S \subseteq I} f(S)$$

($\tilde{f}$ is sometimes also called the *Zeta* transform). Since

$$|\,\mathrm{av}(\sigma)| = \sum_{S \subseteq [n] \setminus \sigma} f(S) = \tilde{f}([n] \setminus \sigma),$$

c_k can be deduced from the Möbius transform of f by summing its kth powers.

If $\mathcal{K}$ is a simplicial complex with n vertices and $|\mathcal{K}|$ simplices, then the values of $\tilde{f}(\sigma)$ for all $\sigma \in \mathcal{K}$ can be computed in $O(n|\mathcal{K}|)$ time by Yates' algorithm [12, Section 4.3.4]. The above method for counting k-covers therefore runs in time $O(n2^n)$. Simplifying the inclusion–exclusion formula (5.1) while keeping its support hereditary, as Theorem 1.1 does, improves the running time to $O(ns)$, where s is the size of the formula ($s = m^{O(\log^2 n)}$ in Theorem 1.1). When the Venn diagram of the F_i has size $m = 2^{o(n)}$, this complexity becomes sub-exponential in n. However, the catch is that, in the above example and many other problems [3, 22, 19], the family $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ is not standardized, which is a crucial assumption for the computational statement in Theorem 1.1. Whether a simplified formula can be computed efficiently in this context is an open problem.

References

[1] Attali, D. and Edelsbrunner, H. (2007) Inclusion–exclusion formulas from independent complexes. *Discrete Comput. Geom.* **37** 59–77.

[2] Björklund, A., Husfeldt, T., Kaski, P. and Koivisto, M. (2008) The travelling salesman problem in bounded degree graphs. In *Automata, Languages and Programming I*, Vol. 5125 of *Lecture Notes in Computer Science*, Springer, pp. 198–209.

[3] Björklund, A., Husfeldt, T. and Koivisto, M. (2009) Set partitioning via inclusion–exclusion. *SIAM J. Comput.* **39** 546–563.
[4] Bonferroni, C. E. (1936) Teoria statistica delle classi e calcolo delle probabilità. *Pubbl. d. R. Ist. Super. di Sci. Econom. e Commerciali di Firenze* **8** 1–62.
[5] Cohn, H. (2004) Projective geometry over $\mathbb{F}_1$ and the Gaussian binomial coefficients. *Amer. Math. Monthly* **111** 487–495.
[6] Dohmen, K. (2003) *Improved Bonferroni Inequalities via Abstract Tubes*, Vol. 1826 of *Lecture Notes in Mathematics*, Springer.
[7] Edelsbrunner, H. and Ramos, E. A. (1997) Inclusion–exclusion complexes for pseudodisk collections. *Discrete Comput. Geom.* **17** 287–306.
[8] Galambos, J. (1996) *Bonferroni-Type Inequalities with Applications*, Springer.
[9] Hatcher, A. (2001) *Algebraic Topology*, Cambridge University Press.
[10] Hoffmann, M., Okamoto, Y., Ruiz-Vargas, A., Scheder, D. and Solymosi, J. (2012) Solution to GWOP problem 17 'A Regional Oracle'. Oral presentation, *Tenth Gremo Workshop on Open Problems*, Bergün (GR), Switzerland.
[11] Kahn, J., Linial, N. and Samorodnitsky, A. (1996) Inclusion–exclusion: Exact and approximate. *Combinatorica* **16** 465–477.
[12] Knuth, D. E. (1997) *The Art of Computer Programming 2*, Addison-Wesley.
[13] Kratky, K. W. (1978) The area of intersection of n equal circular disks. *J. Phys. A* **11** 1017–1024.
[14] Linial, N. and Nisan, N. (1990) Approximate inclusion–exclusion. *Combinatorica* **10** 349–365.
[15] Matoušek, J. (2003) *Using the Borsuk–Ulam Theorem*, Universitext, Springer.
[16] Munkres, J. R. (1984) *Elements of Algebraic Topology*, Addison-Wesley.
[17] Naiman, D. Q. and Wynn, H. P. (1992) Inclusion–exclusion–Bonferroni identities and inequalities for discrete tube-like problems via Euler characteristics. *Ann. Statist.* **20** 43–76.
[18] Naiman, D. Q. and Wynn, H. P. (1997) Abstract tubes, improved inclusion–exclusion identities and inequalities and importance sampling. *Ann. Statist.* **25** 1954–1983.
[19] Nederlof, J. and van Rooij, J. M. M. (2010) Inclusion/exclusion branching for partial dominating set and set splitting. In *Parameterized and Exact Computation*, Vol. 6478 of *Lecture Notes in Computer Science*, Springer, pp. 204–215.
[20] Pólya, G. and Alexanderson, G. L. (1971) Gaussian binomial coefficients. *Elem. Math.* **26** 102–109.
[21] Perrot, G., Cheng, B., Gibson, K. D., Vila, J., Palmer, K. A., Nayeem, A., Maigret, B. and Scheraga, H. A. (1992) MSEED: A program for the rapid analytical determination of accessible surface areas and their derivatives. *J. Comput. Chem.* **13** 1–11.
[22] van Rooij, J. M. M., Nederlof, J. and van Dijk, T. C. (2009) Inclusion/exclusion meets measure and conquer: Exact algorithms for counting dominating sets. In *Algorithms: ESA 2009*, Vol. 5757 of *Lecture Notes in Computer Science*, Springer, pp. 554–565.
[23] Stanley, R. P. (1997) *Enumerative Combinatorics 1*, Vol. 49 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press. Corrected reprint of the 1986 original.
[24] Yang, A. Y., Ganesh, A., Zhou, Z., Sastry, S. S. and Ma, Y. (2010) Fast L1-Minimization algorithms for robust face recognition. `arXiv:1007.3753`