TIGHT PROBABILITY BOUNDS WITH PAIRWISE INDEPENDENCE*

ARJUN KODAGEHALLI RAMACHANDRA† AND KARTHIK NATARAJAN‡

Abstract. While useful probability bounds for n pairwise independent Bernoulli random variables adding up to at least an integer k have been proposed in the literature, none of these bounds are tight in general. In this paper, we provide several results in this direction. Firstly, when $k = 1$, the tightest upper bound on the probability of the union of n pairwise independent events is provided in closed-form for any input marginal probability vector $\boldsymbol{p} \in [0, 1]^n$. To prove the result, we show the existence of a positively correlated Bernoulli random vector with transformed bivariate probabilities, which is of independent interest. Building on this, we show that the ratio of the Boole union bound and the tight pairwise independent bound is upper bounded by 4/3 and that the ratio is attained. Applications of the result in correlation gap analysis and distributionally robust bottleneck optimization are discussed. The result is extended to find the tightest lower bound on the probability of the intersection of n pairwise independent events. Secondly, for any $k \geq 2$ and input marginal probability vector $\boldsymbol{p} \in [0, 1]^n$, new upper bounds are derived by exploiting ordering of probabilities. Numerical examples are provided to illustrate when the bounds provide improvement over existing bounds. Lastly, we identify specific instances when the existing and the new bounds are tight, for example with identical marginal probabilities.

Key words. pairwise independence, probability bounds, linear programming

MSC codes. 60-08, 90C05

1. Introduction.

Probability bounds for sums of Bernoulli random variables have been extensively studied by researchers in various communities including probability and statistics, computer science, combinatorics and optimization. In this paper, our focus is on pairwise independent Bernoulli random variables. It is well known that while mutually independent random variables are pairwise independent, the reverse is not true. Feller [18] attributes Bernstein [4] with identifying one of the earliest examples of $n = 3$ pairwise independent random variables that are not mutually independent. For general n, constructions of pairwise independent Bernoulli random variables can be found in the works of Geisser and Mantel [24], Karloff and Mansour [30], Koller and Meggido [31], pairwise independent discrete random variables in Feller [17], Lancaster [36], Joffe [29], O'Brien [41] and pairwise independent normal random variables in Geisser and Mantel [24]. One of the motivations for studying constructions of pairwise independent random variables particularly in the computer science community is that the joint distribution can have a low cardinality support (polynomial in the number of random variables) in comparison to mutually independent random variables (exponential in the number of random variables). The reader is referred to Lancaster [36] and more recent papers of Babai [2] and Gavinsky and Pudlák [23] who provide precise lower bounds on the entropy of the joint distribution of pairwise independent random variables that only grow logarithmically with the number of random variables. The low cardinality of such distributions have important ramifications in the efficient derandomization of algorithms for NP-hard combinato-

*Submitted: March 2021, Revised: April 2022.

Funding: The research of the second author was partly supported by the MOE Academic Research Fund Tier 2 grant MOE2019-T2-2-138, "Enhancing Robustness of Networks to Dependence via Optimization".

†Engineering Systems and Design, Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372. arjun_ramachandra@sutd.edu.sg.

‡Engineering Systems and Design, Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372. karthik_natarajan@sutd.edu.sg

rial optimization problems (see the review article of Luby and Widgerson [37] and the references therein for results on pairwise independent and more generally t-wise independent random variables).

In this paper, we are interested in the problem of computing probability bounds for the sum of pairwise independent Bernoulli random variables adding up to at least an integer k. Given an integer $n \geq 2$, denote by $[n] = \{1, 2, \ldots, n\}$ and by $K_n = \{(i,j) : 1 \leq i < j \leq n\}$ (it can be viewed as a complete graph on n nodes). Given integers $i < j$, let $[i,j] = \{i, i+1, \ldots, j-1, j\}$. Consider a Bernoulli random vector $\tilde{\boldsymbol{c}} = (\tilde{c}_1, \ldots, \tilde{c}_n)$ with marginal probabilities given by $p_i = \mathbb{P}(\tilde{c}_i = 1)$ for $i \in [n]$. Denote by $\boldsymbol{p} = (p_1, \ldots, p_n) \in [0,1]^n$, the univariate marginal vector and by $\Theta(\{0,1\}^n)$, the set of all probability distributions supported on $\{0,1\}^n$. Consider the set of joint probability distributions of Bernoulli random variables consistent with the given marginal probabilities and pairwise independence:

$$\Theta(\boldsymbol{p}, p_i p_j; (i,j) \in K_n) = \Big\{\theta \in \Theta(\{0,1\}^n) \;\Big|\; \mathbb{P}_\theta\left(\tilde{c}_i = 1\right) = p_i, \forall i \in [n], \\ \mathbb{P}_\theta\left(\tilde{c}_i = 1, \tilde{c}_j = 1\right) = p_i p_j,\ \forall (i,j) \in K_n\Big\}.$$

This set of distributions is nonempty for any $\boldsymbol{p} \in [0,1]^n$, since the distribution of mutually independent random variables lies in the set. Our problem of interest is to compute the maximum probability that n random variables adds up to at least an integer $k \in [n]$ over all distributions in the set. Denote this tightest upper bound by $\overline{P}(n,k,\boldsymbol{p})$ (observe that the bivariate probabilities here are simply given by the product of the univariate probabilities). Then,

$$\overline{P}(n,k,\boldsymbol{p}) = \max_{\theta \in \Theta(\boldsymbol{p}, p_i p_j; (i,j) \in K_n)} \mathbb{P}_\theta\left(\sum_{i \in [n]} \tilde{c}_i \geq k\right). \tag{1.1}$$

Two useful bounds that have been proposed for this problem are discussed next:

(a) Chebyshev [10] bound: The one-sided version of the Chebyshev tail probability bound uses the first and second moments of the random variables. Since the Bernoulli random variables are assumed to be pairwise independent or equivalently uncorrelated, the variance of the sum is given by:

$$\text{Variance}\left(\sum_{i \in [n]} \tilde{c}_i\right) = \sum_{i \in [n]} p_i(1-p_i).$$

Applying the Chebyshev bound gives:

$$\overline{P}(n,k,\boldsymbol{p}) \leq \begin{cases} 1, & k < \displaystyle\sum_{i \in [n]} p_i, \\ \dfrac{\sum_{i \in [n]} p_i(1-p_i)}{\sum_{i \in [n]} p_i(1-p_i) + (k - \sum_{i \in [n]} p_i)^2}, & \displaystyle\sum_{i \in [n]} p_i \leq k \leq n. \end{cases} \tag{1.2}$$

(b) Schmidt, Siegel and Srinivasan [54] bound: The Schmidt, Siegel and Srinivasan bound is derived by bounding the tail probability using the moments of multilinear polynomials. This is in contrast to the Chernoff-Hoeffding bound (see Chernoff [11], Hoeffding [27]) which bounds the tail probability of the sum of independent random variables using the moment generating function. A multilinear polynomial of degree j in n variables is defined as:

$$S_j(\boldsymbol{c}) = \sum_{1 \leq i_1 < i_2 < \ldots < i_j \leq n} c_{i_1} c_{i_2} \ldots c_{i_j}.$$

At the crux of the analysis in [54] is the observation that all the higher moments of the sum of Bernoulli random variables can be generated using linear combinations of the expected values of multilinear polynomials of the random variables. The construction of the bound makes use of the equality:

$$\binom{\sum_{i \in [n]} c_i}{j} = S_j(\boldsymbol{c}), \quad \forall \boldsymbol{c} \in \{0,1\}^n, \forall j \in [0, \textstyle\sum_{i \in [n]} c_i], \tag{1.3}$$

where $S_0(\boldsymbol{c}) = 1$ and $\binom{r}{s} = r!/(s!(r-s)!)$ for any pair of integers $r \geq s \geq 0$. The bound derived in Schmidt et al. [54] (see Theorem 7, part (II) on page 239) for pairwise independent random variables is[1]:

$$\overline{P}(n, k, \boldsymbol{p}) \leq \min \left(1, \frac{\sum_{i \in [n]} p_i}{k}, \frac{\sum_{(i,j) \in K_n} p_i p_j}{\binom{k}{2}}\right). \tag{1.4}$$

While both the Chebyshev bound in (1.2) and the Schmidt, Siegel and Srinivasan bound in (1.4) are useful, neither of them are tight for general values of n, k and $\boldsymbol{p} \in [0,1]^n$. In this paper, we work towards tightening these bounds for pairwise independent random variables and identifying instances when the bounds are tight.

1.1. Other related bounds. Consider the set of joint distributions of Bernoulli random variables consistent with the marginal probability vector $\boldsymbol{p} \in [0,1]^n$ and general bivariate probabilities given by $p_{ij} = \mathbb{P}(\tilde{c}_i = 1, \tilde{c}_j = 1)$ for all $(i,j) \in K_n$:

$$\begin{aligned}\Theta(\boldsymbol{p}, p_{ij}; (i,j) \in K_n) = \Big\{ \theta \in \Theta(\{0,1\}^n) \ \Big| \ & \mathbb{P}_\theta(\tilde{c}_i = 1) = p_i, \forall i \in [n], \\ & \mathbb{P}_\theta(\tilde{c}_i = 1, \tilde{c}_j = 1) = p_{ij}, \ \forall (i,j) \in K_n \Big\}.\end{aligned}$$

Unlike the pairwise independent case, verifying if this set of distributions is nonempty is already known to be a NP-complete problem (see Pitowsky [45]). The tightest upper bound on the tail probability over all distributions in this set is given by:

$$\max_{\theta \in \Theta(\boldsymbol{p}, p_{ij}; (i,j) \in K_n)} \mathbb{P}_\theta \left(\sum_{i \in [n]} \tilde{c}_i \geq k \right),$$

[1] While the statement in the theorem in [54] is provided for $k > \sum_i p_i$, it is straightforward to see that their analysis would lead to the form provided here for general k.

where the bound is set to $-\infty$ if the set of feasible distributions is empty. The bound is given by the optimal value of the linear program (see Hailperin [26]):

$$(1.5)\quad \begin{aligned} \max \quad & \sum_{\boldsymbol{c}\in\{0,1\}^n:\sum_t c_t\geq k} \theta(\boldsymbol{c}) \\ \text{s.t} \quad & \sum_{\boldsymbol{c}\in\{0,1\}^n} \theta(\boldsymbol{c}) = 1, \\ & \sum_{\boldsymbol{c}\in\{0,1\}^n:c_i=1} \theta(\boldsymbol{c}) = p_i, && \forall i\in[n], \\ & \sum_{\boldsymbol{c}\in\{0,1\}^n:c_i=1,c_j=1} \theta(\boldsymbol{c}) = p_{ij}, && \forall (i,j)\in K_n, \\ & \theta(\boldsymbol{c}) \geq 0, && \forall \boldsymbol{c}\in\{0,1\}^n, \end{aligned}$$

where the decision variables are the joint probabilities $\theta(\boldsymbol{c}) = \mathbb{P}(\tilde{\boldsymbol{c}} = \boldsymbol{c})$ for all $\boldsymbol{c} \in \{0,1\}^n$. The number of decision variables in the formulation grows exponentially in the number of random variables n. The dual linear program is given by:

$$(1.6)\quad \begin{aligned} \min \quad & \sum_{(i,j)\in K_n} \lambda_{ij}p_{ij} + \sum_{i\in[n]} \lambda_i p_i + \lambda_0 \\ \text{s.t} \quad & \sum_{(i,j)\in K_n} \lambda_{ij}c_ic_j + \sum_{i\in[n]} \lambda_i c_i + \lambda_0 \geq 0, \quad \forall \boldsymbol{c}\in\{0,1\}^n, \\ & \sum_{(i,j)\in K_n} \lambda_{ij}c_ic_j + \sum_{i\in[n]} \lambda_i c_i + \lambda_0 \geq 1, \quad \forall \boldsymbol{c}\in\{0,1\}^n : \textstyle\sum_t c_t \geq k. \end{aligned}$$

The dual linear program in (1.6) has a polynomial number of decision variables but an exponential number of constraints. This linear program is always feasible (simply set $\lambda_0 = 1$ and remaining dual variables to be zero) and strong duality thus holds. Given the large size of the primal and dual linear programs that need to be solved, two main approaches have been studied in the literature:

(a) The first approach is to find closed-form bounds by generating simple dual feasible solutions (see Kounias [32], Kounias and Marin [33], Sathe et al. [53], Móri and Székely [40], Dawson and Sankoff [12], Galambos [20, 21], de Caen [13], Kuai et al. [34], Dohmen and Tittmann [14] and related graph-based bounds in Hunter [28], Worsley [59], Veneziani [56], Vizvári [58]). These bounds have shown to be tight in specific instances (in Section 2.1 we discuss some of these instances).

(b) The second approach is to reduce the size of the linear programs used and solve them numerically. As the number of random variables n increase, the linear programs quickly become intractable and thus many papers adopting this approach, aggregate the primal decision variables, thus obtaining weaker bounds as a trade-off for the reduced size. Formulations of linear programs using partially or fully aggregated univariate, bivariate or m-variate information for $2 \leq m < n$ have been proposed in Kwerel [35], Platz [46], Prékopa [47, 48], Boros and Prékopa [6], Prékopa and Gao [49], Qiu et al. [51], Yang et al. [61], Yoda and Prékopa [62]). Techniques to solve the dual formulation have been studied in Boros et al. [7].

Using the second approach, in some cases, closed-form bounds have been derived as solutions of the aggregated linear programs. One such bound which is of relevance to this paper is developed in Boros and Prékopa [6] when the first and second binomial moments of an integer random variable supported on $[0, n]$ are known. They computed the tightest upper bound on $\mathbb{P}(\tilde{\xi} \geq k)$ by considering all distributions ω of an integer random variable $\tilde{\xi}$ supported on $[0, n]$ given by the set:

$$\left\{\omega([0,n]) \;\middle|\; \mathbb{E}_\omega\left[\begin{pmatrix}\tilde{\xi}\\ j\end{pmatrix}\right] = S_j,\ j = 1,2\right\}.$$

Setting $\tilde{\xi} = \sum_i \tilde{c}_i$ with $S_1 = \mathbb{E}[S_1(\tilde{\boldsymbol{c}})]$ and $S_2 = \mathbb{E}[S_2(\tilde{\boldsymbol{c}})]$ gives a closed-form upper bound as follows:

(1.7)
$$\mathbb{P}\left(\sum_{i\in[n]} \tilde{c}_i \geq k\right) \leq \begin{cases} 1, & k < \dfrac{(n-1)S_1 - 2S_2}{n - S_1}, \\ \dfrac{(k+n-1)S_1 - 2S_2}{kn}, & \dfrac{(n-1)S_1 - 2S_2}{n-S_1} \leq k < 1 + \dfrac{2S_2}{S_1}, \\ \dfrac{(i-1)(i-2S_1) + 2S_2}{(k-i)^2 + (k-i)}, & k \geq 1 + \dfrac{2S_2}{S_1}, \end{cases}$$

where $i = \lceil ((k-1)S_1 - 2S_2)/(k - S_1)\rceil$ and the ceiling function $\lceil x \rceil$ maps x to the smallest integer greater than or equal to x. Similar to the Chebyshev bound and the Schmidt, Siegel and Srinivasan bound, the Boros and Prékopa bound in (1.7) is not generally tight since it uses aggregated moment information, rather than the specific marginal probabilities. Another useful upper bound derived under weaker assumptions is the Boole union bound [5] (see also Fréchet [19]) for $k = 1$. This bound is valid even with arbitrary dependence among the Bernoulli random variables. Let $\Theta(\boldsymbol{p})$ denotes the set of joint distributions supported on $\{0,1\}^n$ consistent with the univariate marginal probability vector $\boldsymbol{p} \in [0,1]^n$. The Boole union bound is given as:

$$\overline{P}_u(n, 1, \boldsymbol{p}) = \max_{\theta \in \Theta(\boldsymbol{p})} \mathbb{P}_\theta\left(\sum_{i\in[n]} \tilde{c}_i \geq 1\right) = \min\left(\sum_{i\in[n]} p_i, 1\right). \quad (1.8)$$

Clearly, $\overline{P}(n, 1, \boldsymbol{p}) \leq \overline{P}_u(n, 1, \boldsymbol{p})$. Extensions of this bound for $k \geq 2$ is provided in Rüger [52].

1.2. Contributions and structure. This brings us to the key contributions and the structure of the current paper:

(a) In Section 2, we establish (see Lemma 2.1) that a positively correlated Bernoulli random vector $\tilde{\boldsymbol{c}}$ with the univariate probability vector $\boldsymbol{p} \in [0,1]^n$ and transformed bivariate probabilities $p_i p_j / p$ where $\max_i p_i \leq p \leq 1$, always exists. The lemma helps us compute the tightest upper bound on the probability of the union of n pairwise independent events and is of independent interest. By a simple transformation, the results from Lemma 2.1 are extended to show the existence of an alternate positively correlated Bernoulli random vector (see Corollary 2.2). Feasibility is not guaranteed for arbitrary correlation structures with Bernoulli random vectors and hence these two results provide useful sufficient conditions.

(b) We then provide the tightest upper bound on the probability on the union of n pairwise independent events ($k = 1$) in closed-form (see Theorem 2.3). The contributions of Theorem 2.3 lie in:
   1. Establishing that when the random variables are pairwise independent, for any given marginal vector $\boldsymbol{p} \in [0,1]^n$, the upper bound proposed in Kounias

[32], Hunter [28] and Worsley [59] is tight. These bounds were initially developed for the sum of dependent Bernoulli random variables with arbitrary bivariate probabilities (using tree structures from graph theory) and are not tight in general (see Example 2.4 in Section 2.1). Interestingly for pairwise independent random variables, we prove that the bound is tight by using Lemma 2.1.

2. Providing an explicit construction of an extremal distribution (not unique) that attains this bound (see Table 2).
3. Proving that the ratio of the Boole union bound and the pairwise independent bound is upper bounded by 4/3 and that this is attained (see Proposition 2.5). Applications of the result in correlation gap analysis and distributionally robust bottleneck combinatorial optimization are discussed (see examples 2.6 and 2.7).
4. Deriving the tightest lower bound on the probability of the intersection of n pairwise independent events ($k = n$) in closed-form (see Corollary 2.9).

(c) In Section 3, we focus on $k \geq 2$ and present new bounds exploiting the ordering of probabilities (see Theorem 3.1). These ordered bounds improve on the closed-form bounds discussed in Section 1 and numerical examples are provided to illustrate this result.

(d) In Section 4, we provide instances where some of the existing bounds and the newly proposed ordered bounds are tight:

1. First, we identify a special case when the existing closed-form bounds are tight. When the random variables are identically distributed, in Section 4.1, we provide the tightest upper bound in closed-form (see Theorem 4.1) for any $k \in [n]$. The proof is based on showing an equivalence with a linear programming formulation of an aggregated moment bound for which closed-form solutions have been derived by Boros and Prékopa [6]. While the expression of the tight closed-form bound is complicated in form in comparison with the Chebyshev bound in (1.2) and the Schmidt, Siegel and Srinivasan bound in (1.4), it helps us identify conditions when the latter bounds are guaranteed to be tight (see Proposition 4.3).
2. This result with identical marginals is further extended to show tightness for more general t-wise independent variables (see Corollary 4.2). The tight bounds for $t \geq 4$ can be derived as the optimal solution to an aggregated linear program first proposed by Prékopa [48].
3. Next, when $n-1$ marginal probabilities are identical, Proposition 4.5 provides instances when the new ordered bounds are tight. Numerical examples are provided to illustrate this result.

(e) We conclude in Section 5 and identify some future research questions.

2. Tight upper bound for $k = 1$. The goal of this section is to provide the tightest upper bound on the probability of the union of pairwise independent events. Towards this, we start by generating a feasible solution to the dual linear program in (1.6) with $k = 1$, $p_{ij} = p_i p_j$ for all $(i,j) \in K_n$ and probabilities sorted in increasing value as $0 \leq p_1 \leq p_2 \leq \ldots \leq p_n \leq 1$. Consider the dual solution:

$$\lambda_0 = 0,\ \lambda_i = 1\ \forall i \in [n],\ \lambda_{in} = -1\ \forall i \in [n-1] \text{ and } \lambda_{ij} = 0 \text{ otherwise.}$$

The left hand side of the dual constraints in (1.6) then simplifies to:

$$\begin{aligned}\sum_{(i,j)\in K_n}\lambda_{ij}c_ic_j+\sum_{i\in[n]}\lambda_ic_i+\lambda_0 &= -\sum_{i\in[n-1]}c_ic_n+\sum_{i\in[n]}c_i\\ &= c_n+\sum_{i\in[n-1]}c_i(1-c_n).\end{aligned}$$

To verify that this solution is dual feasible, observe that with all $c_i = 0$, $c_n + \sum_{i\in[n-1]} c_i(1 - c_n) = 0$. When $c_n = 1$, regardless of the values of $c_1, \ldots, c_{n-1}$, we have $c_n + \sum_{i\in[n-1]} c_i(1-c_n) = 1$. Lastly, when $c_n = 0$ and at least one $c_i = 1$ for $i \in [n-1]$, we have $c_n + \sum_{i\in[n-1]} c_i(1-c_n) \geq 1$. This solution has an objective value of $\sum_{i\in[n]} p_i - p_n(\sum_{i\in[n-1]} p_i)$. From weak duality and using the trivial upper bound of 1, we have:

$$\overline{P}(n,1,\boldsymbol{p}) \leq \min\left(\sum_{i\in[n]}p_i - p_n\left(\sum_{i\in[n-1]}p_i\right),1\right).$$

Intuitively the first term in this expression is obtained using the probabilistic inequality:

$$\mathbb{P}\left(\sum_{i\in[n]}\tilde{c}_i \geq 1\right) \leq \sum_{j\in[n-1]}\mathbb{P}\left(\tilde{c}_j=1,\tilde{c}_n=0\right)+\mathbb{P}\left(\tilde{c}_n=1\right),$$

and is provided in the work of Kounias [32]. The key result we show is that it is always possible to construct a pairwise independent distribution which attains the upper bound. The proof involves showing that the problem can be transformed to proving the existence of a distribution of a Bernoulli random vector $\tilde{\boldsymbol{c}}$ with univariate probabilities given by $\mathbb{P}(\tilde{c}_i = 1) = p_i$ and transformed bivariate probabilities given by $\mathbb{P}(\tilde{c}_i = 1, \tilde{c}_j = 1) = p_ip_j/p_n$, where p_n is the largest univariate probability. In the following lemma, we prove a more general result on the existence of such a correlated Bernoulli random vector.

Lemma 2.1. *Given an arbitrary univariate probability vector $\boldsymbol{p} \in [0,1]^n$ and bivariate probabilities p_ip_j/p for $(i,j) \in K_n$ where $\max_i p_i \leq p \leq 1$, a Bernoulli random vector consistent with the given univariate and bivariate probabilities always exists.*

Proof. Sort the probabilities in increasing value as $0 \leq p_1 \leq p_2 \leq \ldots \leq p_n \leq 1$. We want to show that there always exists a distribution $\theta \in \Theta(\boldsymbol{p}, p_ip_j/p; (i,j) \in K_n)$ such that:

$$\begin{aligned}\sum_{\boldsymbol{c}\in\{0,1\}^n}\theta(\boldsymbol{c}) &= 1,\\ \sum_{\boldsymbol{c}\in\{0,1\}^n:c_i=1}\theta(\boldsymbol{c}) &= p_i, \qquad \forall i\in[n],\\ \sum_{\boldsymbol{c}\in\{0,1\}^n:c_i=1,c_j=1}\theta(\boldsymbol{c}) &= \frac{p_ip_j}{p}, \quad \forall (i,j)\in K_n,\end{aligned} \tag{2.1}$$

where $p_n \leq p \leq 1$. The proof is divided into two parts:

(1) We first argue that it is sufficient to verify the existence of joint probabilities

$\theta(\boldsymbol{c})$ for n Bernoulli random variables such that:

$$\begin{aligned} \sum_{\boldsymbol{c}\in\{0,1\}^n} \theta(\boldsymbol{c}) &= 1, \\ \sum_{\boldsymbol{c}\in\{0,1\}^n : c_i = 1} \theta(\boldsymbol{c}) &= p_i, \qquad \forall i \in [n], \\ \sum_{\boldsymbol{c}\in\{0,1\}^n : c_i = 1, c_j = 1} \theta(\boldsymbol{c}) &= \frac{p_i p_j}{p_n}, \quad \forall (i,j) \in K_n, \end{aligned} \tag{2.2}$$

where the bivariate probabilities are modified from $p_i p_j/p$ to $p_i p_j/p_n$. This is because with $1 \leq 1/p \leq 1/p_n$, we can find a $\lambda \in [0,1]$ such that:

$$\frac{1}{p} = \lambda \frac{1}{p_n} + (1-\lambda)1.$$

Then, we can create the convex combination of two distributions $\overline{\theta}$ and $\underline{\theta}$ as follows:

$$\theta = \lambda \overline{\theta} + (1-\lambda)\underline{\theta},$$

where $\overline{\theta}$ is a probability distribution which satisfies (2.2) and $\underline{\theta}$ is a pairwise independent joint distribution on n Bernoulli random variables with univariate probabilities given by p_i and bivariate probabilities given by $p_i p_j$. The distribution $\underline{\theta}$ always exists as we can simply choose the mutually independent distribution on n random variables with univariate probabilities p_i. The convex combination then guarantees the existence of a distribution θ which satisfies (2.1). In step (2), we prove the existence of such a $\overline{\theta}$.

(2) To show that (2.2) is feasible, observe that there always exists a feasible distribution on $n-1$ Bernoulli random variables with probabilities given by $\vartheta(\boldsymbol{c}_{-n}) = \mathbb{P}(\tilde{\boldsymbol{c}}_{-n} = \boldsymbol{c}_{-n})$ for all $\boldsymbol{c}_{-n} = (c_1, \ldots, c_{n-1}) \in \{0,1\}^{n-1}$ such that:

$$\begin{aligned} \sum_{\boldsymbol{c}_{-n}\in\{0,1\}^{n-1}} \vartheta(\boldsymbol{c}_{-n}) &= 1, \\ \sum_{\boldsymbol{c}_{-n}\in\{0,1\}^{n-1} : c_i = 1} \vartheta(\boldsymbol{c}_{-n}) &= \frac{p_i}{p_n}, \qquad \forall i \in [n-1], \\ \sum_{\boldsymbol{c}_{-n}\in\{0,1\}^{n-1} : c_i = 1, c_j = 1} \vartheta(\boldsymbol{c}_{-n}) &= \frac{p_i p_j}{p_n^2}, \quad \forall (i,j) \in K_{n-1}. \end{aligned} \tag{2.3}$$

Such a ϑ exists because we can simply choose the mutually independent distribution on $n-1$ random variables with univariate probabilities p_i/p_n where the bivariate probabilities are given by $(p_i/p_n)(p_j/p_n)$. Then, we construct the distribution on n random variables by setting the probability of the vector of all zeros to $1-p_n$, setting the probabilities of the scenarios $\mathbb{P}(\tilde{\boldsymbol{c}}_{-n} = \boldsymbol{c}_{-n}, \tilde{c}_n = 1)$ to $\vartheta(\boldsymbol{c}_{-n})p_n$ and setting all the remaining probabilities to zero. This creates a feasible distribution satisfying (2.2) as seen in the construction of Table 1. This completes the proof.

We remark that there are alternative approaches to construct distributions satisfying Lemma 2.1. An anonymous referee provided the following construction. Let $\tilde{\boldsymbol{d}}$ denote a Bernoulli random vector with mutually independent random variables with marginal probabilities given by $\mathbb{P}(\tilde{d}_i = 1) = p_i/p$ for $i \in [n]$ and a Bernoulli random variable $\tilde{z}$ constructed independently with $\mathbb{P}(\tilde{z} = 1) = p$. Define $\tilde{c}_i = \tilde{d}_i \tilde{z}$ for $i \in [n]$. Then $\mathbb{P}(\tilde{c}_i = 1) = p_i$ for $i \in [n]$ and $\mathbb{P}(\tilde{c}_i = 1, \tilde{c}_j = 1) = p_i p_j/p$ for $(i,j) \in K_n$. We next show that Lemma 2.1 can be extended to prove the existence of an alternative positively correlated Bernoulli random vector.

Table 1: Probabilities of the scenarios to create a feasible distribution $\overline{\theta}$ in (2.2).

Scenarios	c_1	c_2	$\ldots$	c_n	Probability
2^{n-1}	0	0	$\ldots$	0	$\theta(\boldsymbol{c}) = 1 - p_n$
	1	0	$\ldots$	0	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	1	1	$\ldots$	0	0
2^{n-1}	0	0	$\ldots$	1	$\theta(\boldsymbol{c}) = p_n\vartheta(\boldsymbol{c}_{-n})$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	1	1	$\ldots$	1	$\theta(\boldsymbol{c}) = p_n\vartheta(\boldsymbol{c}_{-n})$

□

Corollary 2.2. *Given an arbitrary univariate probability vector $\boldsymbol{p} \in [0,1]^n$ and bivariate probabilities $p_ip_j + \frac{p}{1-p}(1-p_i)(1-p_j)$ for $(i,j) \in K_n$ where $0 \leq p \leq \min_i p_i$, a Bernoulli random vector consistent with the given univariate and bivariate probabilities always exists.*

Proof. From Lemma 2.1, it is straightforward to see that there exists a feasible bivariate distribution ϑ with univariate probabilities $1-p_i$ and bivariate probabilities $(1-p_i)(1-p_j)/(1-p)$ where $0 \leq p \leq \min_i p_i$ (since $1 \geq 1-p \geq \max_i(1-p_i)$). Note that this distribution satisfies $\mathbb{P}_\vartheta(\tilde{\boldsymbol{c}}_i = 0) = p_i, \ \forall i \in [n]$ and

$$\begin{aligned}\mathbb{P}_\vartheta(\tilde{\boldsymbol{c}}_i = 0, \tilde{\boldsymbol{c}}_j = 0) &= \mathbb{P}_\vartheta(\tilde{\boldsymbol{c}}_i = 0) - [\mathbb{P}_\vartheta(\tilde{\boldsymbol{c}}_j = 1) - \mathbb{P}_\vartheta(\tilde{\boldsymbol{c}}_i = 1, \tilde{\boldsymbol{c}}_j = 1)] \\ &= p_i - [(1-p_j) + (1-p_i)(1-p_j)/(1-p)] \\ &= p_ip_j + \tfrac{p}{1-p}(1-p_i)(1-p_j),\end{aligned}$$

for all $(i,j) \in K_n$. By flipping the zeros and ones of the support of ϑ while retaining the same joint probabilities $\vartheta(\boldsymbol{c})$, we obtain the desired result. □

We note that Lemma 2.1 and Corollary 2.2 provide conditions on the bivariate probabilities which guarantee the feasibility of positively correlated Bernoulli random vectors. Feasibility is typically not guaranteed for arbitrary correlation structures with Bernoulli random vectors. While prior works have identified specific correlation structures that are compatible with Bernoulli random vectors (see Chaganty and Joe [9], Qaqish [50], Emrich and Piedmonte [16], Lunn and Davies [38]), the identified conditions in Lemma 2.1 and Corollary 2.2 appear to be new to the best of our knowledge. This brings us to the first theorem, which provides the tightest upper bound on the probability of the union of n pairwise independent events using Lemma 2.1.

Theorem 2.3. *Sort the probabilities in increasing value as $0 \leq p_1 \leq p_2 \leq \ldots \leq p_n \leq 1$. Then,*

$$\overline{P}(n,1,\boldsymbol{p}) = \min\left(\sum_{i\in[n]} p_i - p_n\left(\sum_{i\in[n-1]} p_i\right), 1\right). \tag{2.4}$$

Proof. With $p_{ij} = p_ip_j$ and $k=1$, the optimal value of the primal linear program in (1.5) is bounded since it is feasible and the objective function describes a probability value. The optimality conditions of linear programming states that $\{\theta(\boldsymbol{c}); \boldsymbol{c} \in \{0,1\}^n\}$ is primal optimal and $\{\lambda_{ij}; (i,j) \in K_n, \lambda_i; i \in [n], \lambda_0\}$ is dual optimal if and only if they satisfy: (i) the primal feasibility conditions in (1.5), (ii) the dual feasibility conditions in (1.6) and (iii) the complementary slackness conditions given by:

$$\begin{aligned}
\left(\sum_{(i,j)\in K_n} \lambda_{ij}c_ic_j + \sum_{i\in[n]} \lambda_i c_i + \lambda_0\right)\theta(\boldsymbol{c}) &= 0, \quad \forall \boldsymbol{c} \in \{0,1\}^n : \textstyle\sum_t c_t = 0,\\
\left(\sum_{(i,j)\in K_n} \lambda_{ij}c_ic_j + \sum_{i\in[n]} \lambda_i c_i + \lambda_0 - 1\right)\theta(\boldsymbol{c}) &= 0, \quad \forall \boldsymbol{c} \in \{0,1\}^n : \textstyle\sum_t c_t \geq 1.
\end{aligned}$$

(1) Proof of tightness of non-trivial bound in (2.4): We show that $\overline{P}(n,1,\boldsymbol{p}) = \sum_{i\in[n]} p_i - p_n(\sum_{i\in[n-1]} p_i)$ which is the non-trivial part of the upper bound in (2.4) when $\sum_{i\in[n-1]} p_i \leq 1$. Consider the dual feasible solution $\lambda_0 = 0$, $\lambda_i = 1\ \forall i \in [n]$, $\lambda_{in} = -1\ \forall i \in [n-1]$ and $\lambda_{ij} = 0$ otherwise. We verify the tightness of the bound, by showing there exists a primal solution (feasible distribution) which satisfies the complementary slackness conditions. Towards this, observe that from the complementary slackness conditions in (iii) for all values of $\boldsymbol{c} \in \{0,1\}^n$ with $\sum_{t\in[n-1]} c_t \geq 2$ and $c_n = 0$, we have:

$$c_n + \sum_{i\in[n-1]} c_i(1-c_n) - 1 > 0 \Longrightarrow \theta(\boldsymbol{c}) = 0.$$

This forces a total of $2^{n-1} - n$ scenarios to have zero probability. Building on this, we set the probabilities of the 2^n possible scenarios of $\tilde{\boldsymbol{c}}$ as shown in Table 2. The probability of the vector of all zeros (one scenario) is set to $1 - \sum_{i\in[n]} p_i + p_n(\sum_{i\in[n-1]} p_i)$. To match the bivariate probabilities $\mathbb{P}(\tilde{c}_i = 1, \tilde{c}_n = 0) = p_i(1-p_n)$, we have to then set the probability of the scenario where $c_i = 1, c_n = 0$ and all remaining $c_j = 0$ to $p_i(1-p_n)$. This corresponds to the $n-1$ scenarios in Table 2. Hence, to ensure

Table 2: Probabilities of 2^n scenarios.

Scenarios	c_1	c_2	$\dots$	c_{n-1}	c_n	Probability
1	0	0	$\dots$	0	0	$1 - \sum_{i\in[n]} p_i + p_n\left(\sum_{i\in[n-1]} p_i\right)$
$n-1$	1	0	$\dots$	0	0	$p_1(1-p_n)$
	0	1	$\dots$	0	0	$p_2(1-p_n)$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	0	0	$\dots$	1	0	$p_{n-1}(1-p_n)$
$2^{n-1}-n$	1	1	$\dots$	0	0	0
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	1	1	$\dots$	1	0	0
2^{n-1}	0	0	$\dots$	0	1	$\theta(\boldsymbol{c})$ } p_n
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	1	1	$\dots$	1	1	$\theta(\boldsymbol{c})$

feasibility of the distribution, we need to show that there exist nonnegative values of

$\theta(\boldsymbol{c})$ for the last 2^{n-1} scenarios such that:

$$\begin{aligned}\sum_{\boldsymbol{c}\in\{0,1\}^n:c_n=1} \theta(\boldsymbol{c}) &= p_n,\\ \sum_{\boldsymbol{c}\in\{0,1\}^n:c_i=1,c_n=1} \theta(\boldsymbol{c}) &= p_i p_n, \quad \forall i\in[n-1],\\ \sum_{\boldsymbol{c}\in\{0,1\}^n:c_i=1,c_j=1,c_n=1} \theta(\boldsymbol{c}) &= p_i p_j, \quad \forall (i,j)\in K_{n-1},\end{aligned}$$

or equivalently, by conditioning on $c_n = 1$, we need to show that there exists nonnegative values of $\vartheta(\boldsymbol{c}_{-n}) = \mathbb{P}(\tilde{\boldsymbol{c}}_{-n} = \boldsymbol{c}_{-n})$ for all $\boldsymbol{c}_{-n} = (c_1, \ldots, c_{n-1}) \in \{0,1\}^{n-1}$ such that:

$$\begin{aligned}\sum_{\boldsymbol{c}_{-n}\in\{0,1\}^{n-1}} \vartheta(\boldsymbol{c}_{-n}) &= 1,\\ \sum_{\boldsymbol{c}_{-n}\in\{0,1\}^{n-1}:c_i=1} \vartheta(\boldsymbol{c}_{-n}) &= p_i, \qquad \forall i\in[n-1],\\ \sum_{\boldsymbol{c}_{-n}\in\{0,1\}^{n-1}:c_i=1,c_j=1} \vartheta(\boldsymbol{c}_{-n}) &= \frac{p_i p_j}{p_n}, \quad \forall (i,j)\in K_{n-1}.\end{aligned} \tag{2.5}$$

This corresponds to verifying the existence of a probability distribution on $n-1$ Bernoulli random variables with univariate probabilities p_i and bivariate probabilities $p_i p_j / p_n$ where $p_1 \leq p_2 \leq \ldots \leq p_{n-1} \leq p_n$. Observe that in (2.5), the univariate probabilities remain the same but the random variables are no longer pairwise independent. Now we make use of Lemma 2.1 to claim that (2.5) is always feasible. By considering $n-1$ variables and setting $p = p_n \geq \max_{i\in[n-1]} p_i$, it is to easy to see from Lemma 2.1 that there exists a distribution which satisfies (2.5). An outline of the different distributions used in the construction is provided in Figure 1. This completes the proof for the case where $\sum_{i\in[n-1]} p_i \leq 1$ with:

$$\overline{P}(n,1,\boldsymbol{p}) = \sum_{i\in[n]} p_i - p_n \left(\sum_{i\in[n-1]} p_i\right).$$

(2) Proof of tightness of the trivial part of the bound in (2.4): To complete the proof, consider the case with $\sum_{i\in[n-1]} p_i > 1$. Then, there exists an index $t \in [2, n-1]$ such that $\sum_{i\in[t-1]} p_i \leq 1$ and $\sum_{i\in[t]} p_i > 1$. Let $\delta = 1 - \sum_{i\in[t-1]} p_i$. Clearly $0 \leq \delta < p_t$. From step (1), we know that there exists a distribution for $t+1$ pairwise independent random variables with marginal probabilities $p_1, p_2, \ldots, p_{t-1}, \delta, p_{t+1}$ such that the probability of the sum of the random variables being at least one is equal to one (since the sum of the first t probabilities in this case is equal to one). By increasing the marginal probability δ to p_t, we can only increase this probability. To see this, consider the distribution for $t+1$ mutually independent Bernoulli random variables with marginal probabilities $p_1, p_2, \ldots, p_{t-1}, 1, p_{t+1}$ where the probability of the sum of the random variables being at least one is equal to one. We can then find a $\lambda \in [0,1)$ such that $p_t = \lambda\delta + (1-\lambda)$ and construct a pairwise independent distribution for $t+1$ pairwise independent random variables with marginal probabilities $p_1, p_2, \ldots, p_{t-1}, p_t, p_{t+1}$ by using the convex combination of the two distributions with sum of the random variables taking a value at least one with probability one. We can generate the remaining random variables $\tilde{c}_{t+2}, \ldots, \tilde{c}_n$ independently with mar-

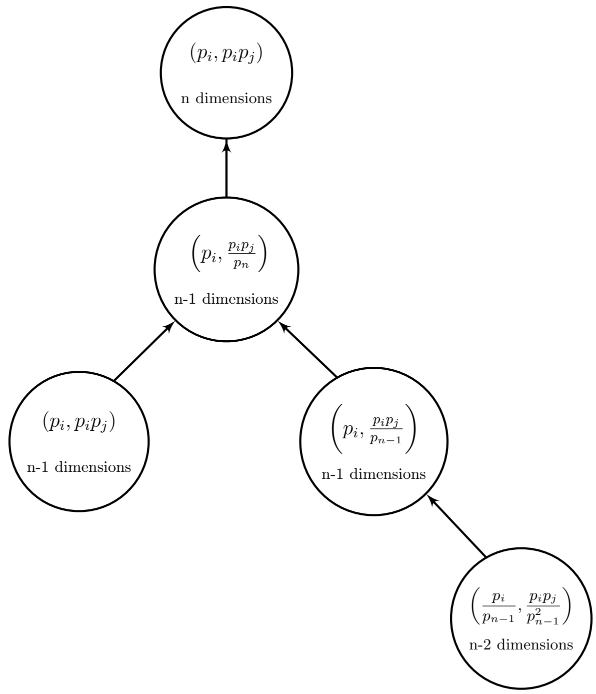


Fig. 1: Construction of the extremal distribution.

ginal probabilities $p_{t+2}, \ldots, p_n$. This provides a feasible distribution that attains the bound of one, thus completing the proof. ∎

2.1. Connection of Theorem 2.3 to existing results. Bounds on the probability that the sum of Bernoulli random variables is at least one has been extensively studied in the literature, under knowledge of general bivariate probabilities. Let A_i denote the event that $c_i = 1$ for each i, then, $k = 1$ simply corresponds to bounding the probability of the union of events. When the marginal probabilities $p_i = \mathbb{P}(A_i)$ for $i \in [n]$ and bivariate probabilities $p_{ij} = \mathbb{P}(A_i \cap A_j)$ for $(i, j) \in K_n$ are given, Hunter [28] and Worsley [59] derived the following bound by optimizing over spanning trees $\tau \in T$:

$$\mathbb{P}(\cup_i A_i) \leq \sum_{i \in [n]} p_i - \max_{\tau \in T} \sum_{(i,j) \in \tau} p_{ij}, \tag{2.6}$$

where T is the set of all spanning trees on the complete graph with n nodes with edge weights given by p_{ij}. A special case of the Hunter [28] bound was derived by Kounias [32]:

$$\mathbb{P}(\cup_i A_i) \leq \sum_{i \in [n]} p_i - \max_{j \in [n]} \sum_{i \neq j} p_{ij}, \tag{2.7}$$

which subtracts the maximum weight of a star spanning tree from the sum of the marginal probabilities. Tree bounds have been shown to be tight, in some special cases as outlined next:

(a) Zero bivariate probabilities for all pairs: When all the probabilities p_{ij} are zero, the bound reduces to the Boole union bound which is tight.

(b) Zero bivariate probabilities outside a given tree: Given a tree τ such that the bivariate probabilities p_{ij} are zero for edges $(i,j) \notin \tau$, Worsley [59] proved that the bound is tight (see Veneziani [57] for related results).

(c) Lower bounds on bivariate probabilities: Boros et al. [7] proved that by relaxing the equality of bivariate probabilities to lower bounds on bivariate probabilities:

$$\mathbb{P}\big(A_i \cap A_j\big) \geq p_{ij},\ \forall (i,j) \in K_n,$$

the tightest upper bound on the probability of the union is exactly the Hunter [28] and Worsley [59] bound (see Maurer [39] for related results).

(d) Pairwise independent variables (Theorem 2.3 in this paper): With pairwise independent random variables where $p_{ij} = p_i p_j$, the maximum weight spanning trees in (2.6) is exactly the star tree with the root at node n and edges (i,n) for all $i \in [n-1]$. In, this case, the Kounias [32], Hunter [28] and Worsley [59] bound reduce to the bound in (2.4) which is shown to be tight in Theorem 2.3 of this paper.

The next example illustrates that with general bivariate probabilities, even if a joint distribution exists, the Hunter [28], Worsley [59] bound and Kounias [32] bound are not guaranteed to be tight.

Example 2.4. Consider $n = 4$ Bernoulli random variables with univariate marginal probabilities:

$$p_1 = 0.35,\ p_2 = 0.19,\ p_3 = 0.13,\ p_4 = 0.2,$$

and bivariate probabilities:

$$p_{12} = 0.001,\ p_{13} = 0.022,\ p_{14} = 0.03,\ p_{23} = 0.017,\ p_{24} = 0.018,\ p_{34} = 0.019.$$

It can be verified using linear programming that a joint distribution with these given univariate and bivariate probabilities exists. The tight upper bound obtained by solving the linear program (1.5) is equal to:

$$\max_{\theta \in \Theta(\boldsymbol{p}, p_{ij};(i,j)\in K_4)} \mathbb{P}_\theta\left(\tilde{c}_1 + \tilde{c}_2 + \tilde{c}_3 + \tilde{c}_4 \geq 1\right) = 0.784.$$

Figure 2 displays the star spanning tree chosen by the Kounias [32] bound and the spanning tree chosen by the Hunter [28] and Worsley [59] bound. It is clear that none of these bounds are tight in this instance. Boros et al. [7] also provide randomly generated instances (see Table 1 of Section 4 in their paper) where the Hunter [28] and Worsley [59] bound is not tight, athough it provides the best performance among the upper bounds considered there.

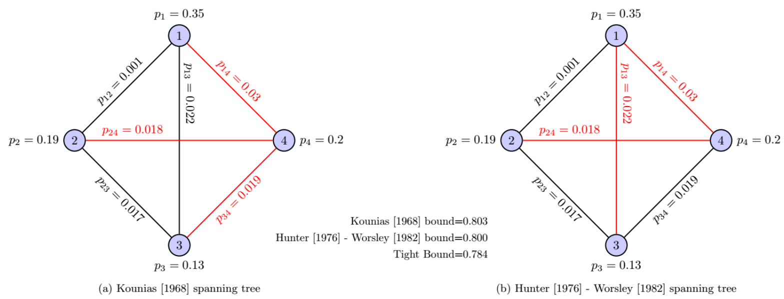


Fig. 2: Kounias [32], Hunter [28] and Worsley [59] spanning trees with general bivariates

Figure 3 demonstrates that with the same set of univariate marginals, when pairwise independence is enforced, the spanning trees obtained from all these approaches are identical and the bounds in (2.6) and (2.7) equal the tight bound 0.688 (from Theorem 2.3).

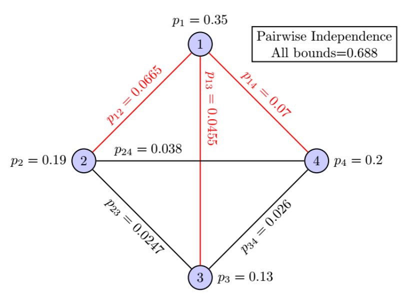


Fig. 3: Optimal spanning tree with pairwise independence when $\boldsymbol{p} = (0.35, 0.19, 0.13, 0.2)$.

2.2. Comparison with the union bound. The next proposition provides an upper bound on the ratio of the Boole union bound and the pairwise independent bound in (2.4) in Theorem 2.3.

Proposition 2.5. For all $\boldsymbol{p} \in [0,1]^n$, we have:

$$\frac{\overline{P}_u(n,1,\boldsymbol{p})}{\overline{P}(n,1,\boldsymbol{p})} \leq \frac{4}{3}.$$

The ratio of 4/3 is attained when $\sum_{i\in[n-1]} p_i = 1/2$ and $p_n = 1/2$.

Proof. Assume the probabilities are sorted in increasing value as $0 \leq p_1 \leq p_2 \leq \ldots \leq p_n \leq 1$. It is straightforward to see that if $\sum_{i\in[n-1]} p_i > 1$, both the bounds take the value of $\overline{P}(n,1,\boldsymbol{p}) = \overline{P}_u(n,1,\boldsymbol{p}) = 1$. Now assume, $\alpha = \sum_{i\in[n-1]} p_i \leq 1$.

The ratio is given as:

$$\begin{aligned}\frac{\overline{P}_u(n,1,\boldsymbol{p})}{\overline{P}(n,1,\boldsymbol{p})} &= \frac{\min\left(\sum_{i\in[n]} p_i, 1\right)}{\sum_{i\in[n]} p_i - p_n\left(\sum_{i\in[n-1]} p_i\right)} \\ &= \frac{\min(\alpha+p_n,1)}{\alpha+p_n-\alpha p_n}.\end{aligned}$$

If $\alpha + p_n \leq 1$, then we have:

$$\begin{aligned}\frac{\overline{P}_u(n,1,\boldsymbol{p})}{\overline{P}(n,1,\boldsymbol{p})} &= \frac{\alpha+p_n}{\alpha+p_n-\alpha p_n} \\ &= \frac{1}{1-\frac{1}{\frac{1}{\alpha}+\frac{1}{p_n}}} \\ &\leq \frac{4}{3} \\ &\text{[where the maximum is attained at } \alpha = 1-p_n \text{ and } p_n = 1/2].\end{aligned}$$

If $\alpha + p_n \geq 1$, then we have:

$$\begin{aligned}\frac{\overline{P}_u(n,1,\boldsymbol{p})}{\overline{P}(n,1,\boldsymbol{p})} &= \frac{1}{\alpha+p_n-\alpha p_n} \\ &= \frac{1}{\alpha(1-p_n)+p_n} \\ &\leq \frac{4}{3} \\ &\text{[where the maximum is attained at } \alpha = 1-p_n \text{ and } p_n = 1/2].\end{aligned}$$

This gives the bound of 4/3 when $p_n = 1/2$ and $\alpha = 1/2$. ∎

We next illustrate an application of Theorem 2.3 and Proposition 2.5 in comparing bounds with dependent and independent random variables in correlation gap analysis.

Example 2.6 (Correlation gap analysis). The notion of "correlation gap" was introduced by Agrawal et al. [1]. It is defined as the ratio of the worst-case expected cost for random variables with given univariate marginals to the expected cost when the random variables are independent. When $\tilde{\boldsymbol{c}}$ is a Bernoulli random vector and θ_{ind} denotes the independent distribution, the correlation gap is defined as:

$$\kappa_u(\boldsymbol{p}) = \sup_{\theta\in\Theta(\boldsymbol{p})} \frac{\mathbb{E}_\theta[f(\tilde{\boldsymbol{c}})]}{\mathbb{E}_{\theta_{ind}}[f(\tilde{\boldsymbol{c}})]}. \tag{2.8}$$

A function $f : \{0,1\}^n \to \mathbb{R}_+$ is: (i) submodular if $f(\boldsymbol{c}) + f(\boldsymbol{d}) \geq f(\boldsymbol{c}\wedge\boldsymbol{d}) + f(\boldsymbol{c}\vee\boldsymbol{d})$ for all $\boldsymbol{c},\boldsymbol{d} \in \{0,1\}^n$ with $\boldsymbol{c}\wedge\boldsymbol{d} = (\min(c_1,d_1),\ldots,\min(c_n,d_n))$ and $\boldsymbol{c}\vee\boldsymbol{d} = (\max(c_1,d_1),\ldots,\max(c_n,d_n))$ and (ii) nondecreasing if $f(\boldsymbol{c}) \geq f(\boldsymbol{d})$ for all $\boldsymbol{c} \geq \boldsymbol{d}$. A key result in this area is that for any nonnegative, nondecreasing, submodular function, the correlation gap is always upper bounded by $e/(e-1)$ (see Calinescu et al. [8], Agrawal et al. [1]). The example constructed in these papers show the bound is attained for the maximum of binary variables $f(\boldsymbol{c}) = \max_{i\in[n]} c_i$. For a given marginal vector $\boldsymbol{p}$, the correlation gap in (2.8) reduces to:

$$\begin{aligned}\kappa_u(\boldsymbol{p}) &= \frac{\max_{\theta\in\Theta(\boldsymbol{p})}\mathbb{E}_\theta[\max(\tilde{c}_1,\tilde{c}_2,...,\tilde{c}_n)]}{1-\prod_{i=1}^n(1-p_i)} \\ &= \frac{\max_{\theta\in\Theta(\boldsymbol{p})}\mathbb{P}_\theta\left(\sum_{i\in[n]}\tilde{c}_i\geq 1\right)}{1-\prod_{i=1}^n(1-p_i)} \\ &= \frac{\min\left(\sum_{i\in[n]}p_i,\ 1\right)}{1-\prod_{i=1}^n(1-p_i)}.\end{aligned} \tag{2.9}$$

We now provide an extension of this definition by considering the ratio of the worst-case expected cost when the random variables are pairwise independent to the expected cost when the random variables are independent. This is given as:

$$\kappa(\boldsymbol{p}) = \sup_{\theta\in\Theta(\boldsymbol{p},p_{ij};(i,j)\in K_n)} \frac{\mathbb{E}_\theta[f(\tilde{\boldsymbol{c}})]}{\mathbb{E}_{\theta_{ind}}[f(\tilde{\boldsymbol{c}})]},$$

which reduces in this specific case to:

$$\kappa(\boldsymbol{p}) = \frac{\min\left(\sum_{i\in[n]}p_i - p_n\left(\sum_{i\in[n-1]}p_i\right),1\right)}{1-\prod_{i=1}^n(1-p_i)}.$$

Clearly $\kappa(\boldsymbol{p})\leq\kappa_u(\boldsymbol{p})$. We next compare these two ratios.

(a) Worst-case analysis: Assume the marginal probability vector is given by $\boldsymbol{p} = (1/n,\ldots,1/n)$. For the independent distribution, the probability is given by $1-(1-1/n)^n$, while the Boole union bound is equal to one (attained by the distribution which assigns probability $1/n$ to each of n support points with $c_i = 1,\ c_j = 0, \forall j\neq i$ (for each $i\in[n]$) and zero otherwise). In this case, the limit of the ratio as n goes to infinity is given by:

$$\lim_{n\to\infty}\kappa_u(\boldsymbol{p}) = \frac{1}{1-(1-1/n)^n} = \frac{e}{e-1}\approx 1.5819.$$

Likewise it is easy to verify that with pairwise independence:

$$\lim_{n\to\infty}\kappa(\boldsymbol{p}) = \frac{1-1/n\,(1-1/n)}{1-(1-1/n)^n} = \frac{e}{e-1}\approx 1.5819.$$

Thus in the worst-case, both these bounds attain the ratio $e/(e-1)$.

(b) Instances where the correlation gap can be improved: On the other hand, Proposition 2.5 illustrates that for the probabilities $p_n = 1/2$ and $\sum_{i\in[n-1]}p_i = 1/2$, the pairwise independent bound is 3/4 and the Boole union bound is one. For example with $n = 2$ where $\boldsymbol{p} = (1/2,1/2)$, the Boole union bound is one, while both the pairwise independent bound and the independent probability is equal to 3/4. Then, we have $\kappa_u((1/2,1/2)) = 4/3$ while $\kappa((1/2,1/2)) = 1$. Thus in specific instances, the correlation gap can be tightened by considering pairwise independent random variables.

An application of the 4/3 bound in Proposition 2.5 in the context of distributionally robust optimization is discussed next.

Example 2.7 (Distributionally robust bottleneck combinatorial optimization). Consider a set of n elements indexed by $[n] = \{1, 2, \ldots, n\}$ where element i has a cost of c_i. Given a set of feasible solutions $\mathcal{X} \subseteq \{0,1\}^n$, the goal in the bottleneck combinatorial optimization problem is to find the solution $\boldsymbol{x} \in \mathcal{X}$ that minimizes the maximum cost among the selected elements (bottleneck cost). This is formulated as the bottleneck combinatorial optimization problem:

$$\min_{\boldsymbol{x}\in\mathcal{X}\subseteq\{0,1\}^n} \max_{i\in[n]} c_i x_i.$$

A threshold algorithm to solve this class of problems was developed by Edmonds and Fulkerson [15]. Consider a distributionally robust variant of this problem where the cost of the element i is a random variable $\tilde{c}_i$ and the joint distribution of $\tilde{\boldsymbol{c}}$ is not fully specified. The distributionally robust bottleneck optimization problem is formulated as:

$$\min_{\boldsymbol{x}\in\mathcal{X}\subseteq\{0,1\}^n} \max_{\theta\in\Theta} \mathbb{E}\left[\max_{i\in[n]} \tilde{c}_i x_i\right],$$

where Θ is the set of possible joint distributions and the goal is to find the solution $\boldsymbol{x} \in \mathcal{X}$ that minimizes the maximum expected bottleneck cost. Such problems have been studied in Agrawal et al. [1] where the distributions are specified up to marginal information and Xie et al. [60] where the distributions are assumed to lie in a ball around an empirical distribution specified by the Wasserstein distance. Here we consider the set of distributions with pairwise independent random variables where $\Theta = \Theta(\boldsymbol{p}, p_i p_j; (i,j) \in K_n)$. The next proposition provides a 4/3-approximation algorithm for this problem.

Proposition 2.8. Let OPT be the optimal value of the distributionally robust bottleneck combinatorial optimization problem:

$$\mathrm{OPT} = \min_{\boldsymbol{x}\in\mathcal{X}\subseteq\{0,1\}^n} \underbrace{\max_{\theta\in\Theta(\boldsymbol{p},p_ip_j;(i,j)\in K_n)} \mathbb{E}\left[\max_{i\in[n]} \tilde{c}_i x_i\right]}_{f(\boldsymbol{x})}.$$

Suppose we can optimize linear functions over the set $\mathcal{X} \subseteq \{0,1\}^n$ in polynomial time. Then, we can find $\hat{\boldsymbol{x}}$ in polynomial time such that:

$$\mathrm{OPT} \leq f(\hat{\boldsymbol{x}}) \leq \frac{4}{3}\mathrm{OPT}.$$

Proof. When $\boldsymbol{x} \in \mathcal{X} \subseteq \{0,1\}^n$, each $\tilde{c}_i x_i$ is a Bernoulli random variable with $\mathbb{P}(\tilde{c}_i x_i = 1) = p_i x_i$. Using the Boole union bound, we have:

$$\max_{\theta\in\Theta(\boldsymbol{p})} \mathbb{E}\left[\max_{i\in[n]} \tilde{c}_i x_i\right] = \min\left(1, \sum_{i\in[n]} p_i x_i\right).$$

Consider the solution $\hat{\boldsymbol{x}}$ which is computable in polynomial time by solving the minimum cost combinatorial optimization problem:

$$\hat{\boldsymbol{x}} \in \arg\min_{\boldsymbol{x}\in\mathcal{X}\subseteq\{0,1\}^n} \sum_{i\in[n]} p_i x_i.$$

Let $\boldsymbol{x}^*$ denote the optimal solution and θ^* denote the worst-case pairwise independent distribution in OPT. Then we have:

$\square$

$$\begin{aligned}
\frac{f(\hat{\boldsymbol{x}})}{\text{OPT}} &= \frac{\max_{\theta\in\Theta(\boldsymbol{p},p_ip_j;(i,j)\in K_n)} \mathbb{E}\left[\max_{i\in[n]} \tilde{c}_i\hat{x}_i\right]}{\mathbb{E}_{\theta^*}\left[\max_{i\in[n]} \tilde{c}_i x_i^*\right]} \\
&\leq \frac{\max_{\theta\in\Theta(\boldsymbol{p})} \mathbb{E}\left[\max_{i\in[n]} \tilde{c}_i\hat{x}_i\right]}{\mathbb{E}_{\theta^*}\left[\max_{i\in[n]} \tilde{c}_i x_i^*\right]} \\
&\quad [\text{since } \Theta(\boldsymbol{p}, p_ip_j; (i,j)\in K_n) \subseteq \Theta(\boldsymbol{p})] \\
&= \frac{\min\left(1, \sum_{i\in[n]} p_i\hat{x}_i\right)}{\mathbb{E}_{\theta^*}\left[\max_{i\in[n]} \tilde{c}_i x_i^*\right]} \\
&\leq \frac{\min\left(1, \sum_{i\in[n]} p_i x_i^*\right)}{\mathbb{E}_{\theta^*}\left[\max_{i\in[n]} \tilde{c}_i x_i^*\right]} \\
&\quad [\text{since } \boldsymbol{x}^* \text{ is only feasible for the sum objective}] \\
&= \frac{\overline{P}_u(n,1,\boldsymbol{p}\cdot\boldsymbol{x}^*)}{\overline{P}(n,1,\boldsymbol{p}\cdot\boldsymbol{x}^*)} \\
&\quad [\text{where } \boldsymbol{p}\cdot\boldsymbol{x}^* = (p_1x_1^*,\ldots,p_nx_n^*)] \\
&\leq \tfrac{4}{3} \\
&\quad [\text{from Proposition 2.5}].
\end{aligned}$$

Proposition 2.8 can be applied to instances such as the bottleneck assignment, bottleneck matching problem and bottleneck shortest path problems and provides a 4/3-approximation for these instances. The next result shows that Theorem 2.3 can be used to prove a tight lower bound on the probability of the intersection of pairwise independent events.

2.3. Tight lower bound for $k = n$. Denote the tightest lower bound on the probability of the intersection of pairwise independent events by $\underline{P}(n,n,\boldsymbol{p})$. Then,

$$\underline{P}(n,n,\boldsymbol{p}) = \min_{\theta\in\Theta(\boldsymbol{p},p_ip_j;(i,j)\in K_n)} \mathbb{P}_\theta\left(\sum_{i\in[n]} \tilde{c}_i = n\right).$$

COROLLARY 2.9. *Sort the probabilities in increasing value as* $0 \leq p_1 \leq p_2 \leq \ldots \leq p_n \leq 1$. *Then,*

$$\underline{P}(n,n,\boldsymbol{p}) = \max\left(p_1\left(\sum_{i=2}^{n} p_i - (n-2)\right), 0\right). \tag{2.10}$$

Proof. The proof follows from that of the union probability bound in Theorem 2.3. Define a complementary Bernoulli random variable $d_i = 1 - c_{n-i+1},\ i\in[n]$, with transformed probabilities $\mathbb{P}(\tilde{d}_i = 1) = q_i = 1 - p_{n-i+1},\ i\in[n]$ and thus $0 \leq q_1 \leq q_2 \leq \ldots \leq q_n \leq 1$. We first note that the maximum probability of the union of pairwise independent events can be expressed as an equivalent maximization problem defined on $\boldsymbol{d}$ as follows:

$$\overline{P}(n,1,\boldsymbol{p}) = \overline{Q}(n,n-1,\boldsymbol{q}) = \max_{\theta\in\Theta(\boldsymbol{q},q_iq_j;(i,j)\in K_n)} \mathbb{P}_\theta\left(\sum_{i\in[n]} \tilde{d}_i \leq n-1\right) \tag{2.11}$$

where $\overline{Q}(n, n-1, \boldsymbol{q})$ is the maximum probability that at most $n-1$ complimentary events occur. The proof is then completed by noting that the tight lower intersection bound $\underline{P}(n, n, \boldsymbol{q})$ can be expressed as

$$\begin{aligned}
\underline{P}(n,n,\boldsymbol{q}) &= 1-\overline{Q}(n,n-1,\boldsymbol{q}) \\
&= 1-\overline{P}(n,1,\boldsymbol{p}) \\
&= 1-\min\left(\sum\nolimits_{i\in[n]} p_i - p_n\left(\sum\nolimits_{i\in[n-1]} p_i\right), 1\right) \\
&= 1-\min\left(1-(1-p_n)\left(1-\sum\nolimits_{i\in[n-1]} p_i\right), 1\right) \\
&= \max\left(q_1\left(\sum_{i=2}^{n} q_i-(n-2)\right), 0\right).
\end{aligned}$$

and replacing $\boldsymbol{q}$ by $\boldsymbol{p}$. ∎

Extremal Distribution: The primal distribution which attains the non-trivial part of the tight intersection bound $\underline{P}(n, n, \boldsymbol{q})$ is shown in Table 3. It can be constructed from the union probability extremal distribution $\theta^\star$ in Table 2 by flipping the zeros and one's of the support, reversing the bits (to ensure ordering of the transfomed probabilities) and retaining the same joint probabilities $\theta^\star(\boldsymbol{c})$ but expressed in terms of $\boldsymbol{q}$ instead of $\boldsymbol{p}$.

Table 3: Probabilities of 2^n scenarios.

Scenarios	d_1	d_2	$\dots$	d_{n-1}	d_n	Probability	
2^{n-1}	0	0	$\dots$	0	0	$\theta(\boldsymbol{d})$	$1-q_1$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
	0	1	$\dots$	1	1	$\theta(\boldsymbol{d})$	
$2^{n-1}-n$	1	0	$\dots$	0	0	0	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
	1	1	$\dots$	1	0	0	
$n-1$	1	0	$\dots$	1	1	$q_1(1-q_2)$	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
	1	1	$\dots$	0	1	$q_1(1-q_{n-1})$	
	1	1	$\dots$	1	0	$q_1(1-q_n)$	
1	1	1	$\dots$	1	1	$q_1\left(\sum_{i=2}^{n} q_i-(n-2)\right)$	

Note that the feasibility of the joint distribution in Table 3 depends on the existence of nonnegative values $\theta(\boldsymbol{d})$ for the first 2^{n-1} scenarios or alternatively by conditioning on $d_1 = 0$, there exist nonnegative values of $\vartheta(\boldsymbol{d}_{-1}) = \mathbb{P}(\tilde{\boldsymbol{d}}_{-1} = \boldsymbol{d}_{-1})$ for all $\boldsymbol{d}_{-1} = (d_2 \dots, d_n) \in \{0,1\}^{n-1}$ such that:

(2.12)
$$\begin{aligned}
\sum_{\boldsymbol{d}_{-1}\in\{0,1\}^{n-1}} \vartheta(\boldsymbol{d}_{-1}) &= 1, \\
\sum_{\boldsymbol{d}_{-1}\in\{0,1\}^{n-1}:d_i=0} \vartheta(\boldsymbol{d}_{-1}) &= 1-q_i, && \forall i\in[2,n], \\
\sum_{\boldsymbol{d}_{-1}\in\{0,1\}^{n-1}:d_i=0,d_j=0} \vartheta(\boldsymbol{d}_{-1}) &= \frac{(1-q_i)(1-q_j)}{1-q_1}, && \forall (i,j)\in\{(i,j): 2\le i<j\le n\},
\end{aligned}$$

where the constraints in (2.12) is expressed in terms of non-occurence of the Bernoulli events represented by $\boldsymbol{d}$, *i.e.* $d_i = 0$ instead of $d_i = 1$. The existence of such a feasible bivariate distribution ϑ can be independently verified from Corollary 2.2 by noting that the Bernoulli random vector defined there satisfies $\mathbb{P}(\tilde{\boldsymbol{c}}_i = 0) = 1 - p_i, \ \forall i \in [n]$ and $\mathbb{P}(\tilde{\boldsymbol{c}}_i = 0, \tilde{\boldsymbol{c}}_j = 0) = (1-p_i)(1-p_j)/(1-p)$ for all $(i,j) \in K_n$, subsequently replacing p_i by q_i and setting $p = q_1 <= \min_{i \in [2,n]} q_i$ for $n-1$ variables instead of n.

2.3.1. Connection of Corollary 2.9 to existing results. The intersection bound $\underline{P}(n,n,\boldsymbol{p})$ derived in Corollary 2.9 is zero when $\sum_{i=2}^{n} p_i \leq n-2$. In related work with identical probabilities p, Benjamini et al. [3] compute that the minimum intersection probability for t-wise independent Bernoulli random variables and identify when it is zero. They prove that $\underline{P}(n,n,p) = 0$ for all $t < n$ and $p \leq 1/2$ which matches our result with pairwise independence ($t = 2$) since $p \leq (n-2)/(n-1) \leq 1/2$ for all $n \geq 3$. We will show in Section 4.1 that with pairwise independent identical Bernoulli's, it is possible to derive closed-form tight upper and lower bounds on the intersection probability and more generally $\overline{P}(n,k,\boldsymbol{p})$ and $\underline{P}(n,k,\boldsymbol{p})$ for any $k \in [n]$. With arbitrary dependence among the Bernoulli random variables, the Fréchet [19] lower intersection bound is given as:

$$\underline{P}_u(n,n,\boldsymbol{p}) = \min\nolimits_{\theta \in \Theta(\boldsymbol{p})} \mathbb{P}_\theta\left(\sum\nolimits_{i\in[n]} \tilde{c}_i = n\right) = \max\left(\sum\nolimits_{i \in [n]} p_i - (n-1), 0\right). \tag{2.13}$$

Clearly, $\underline{P}(n,n,\boldsymbol{p}) \geq \underline{P}_u(n,1,\boldsymbol{p})$ and the lower bound is thus improved with pairwise independence.

3. Improved bounds with non-identical marginals for $k \geq 2$. In the previous section, we resolved the question of finding the tightest bound on the probability of the union of pairwise independent events. We now shift attention to the case of at least k pairwise independent events occurring where $k \geq 2$. Deriving tight bounds for general k appears to be challenging. We exploit the ordering of the probabilities to provide new upper bounds by creating feasible solutions to the dual linear program in (1.6). We make use of the observation that all three bounds in (1.2), (1.4) and (1.7) can be expressed in terms of the first two aggregated (or equivalently binomial) moments of the sum of pairwise independent random variables with $S_1 = \sum_i p_i$ and $S_2 = \sum_{(i,j)\in K_n} p_i p_j$. The new ordered bounds improve on these three closed-form bounds. We will refer to the original bounds in (1.2), (1.4) and (1.7) as unordered bounds from this point onwards. The next theorem provides probability bounds for the sum of pairwise independent random variables with possibly non-identical marginals when $k \geq 2$.

THEOREM 3.1. *Sort the input probabilities in increasing order as $p_1 \leq \ldots \leq p_n$. Define the partial binomial moment $S_{1r} = \sum_{i \in [n-r]} p_i$ for $r \in [0, n-1]$ and $S_{2r} = \sum_{(i,j)\in K_{n-r}} p_i p_j$ for $r \in [0, n-2]$.*

(a) The ordered Schmidt, Siegel and Srinivasan bound is a valid upper bound on $\overline{P}(n,k,\boldsymbol{p})$:

$$\overline{P}(n,k,\boldsymbol{p}) \leq \min\left(1, \min_{r_1 \in [0,k-1]}\left(\frac{S_{1r_1}}{k - r_1}\right), \min_{r_2\in[0,k-2]}\left(\frac{S_{2r_2}}{\binom{k-r_2}{2}}\right)\right), \forall k \in [2,n]. \tag{3.1}$$

(b) The ordered Boros and Prékopa bound is a valid upper bound on $\overline{P}(n,k,\boldsymbol{p})$:

$$\overline{P}(n,k,\boldsymbol{p}) \leq \min_{r\in[0,k-1]} BP(n-r,k-r,\boldsymbol{p}), \quad \forall k \in [2,n], \tag{3.2}$$

where:

$$BP(n-r,k-r,\boldsymbol{p}) = \begin{cases} 1, & k < \dfrac{(n-r-1)S_{1r}-2S_{2r}}{n-r-S_{1r}}+r, \\[2ex] \dfrac{(k+n-2r-1)S_{1r}-2S_{2r}}{(k-r)(n-r)}, & \dfrac{(n-r-1)S_{1r}-2S_{2r}}{n-r-S_{1r}}+r \le k < 1+\dfrac{2S_{2r}}{S_{1r}}+r, \\[2ex] \dfrac{(i-1)(i-2S_{1r})+2S_{2r}}{(k-r-i)^2+(k-r-i)}, & k \ge 1+\dfrac{2S_{2r}}{S_{1r}}+r. \end{cases}$$

and $i = \lceil((k-r-1)S_{1r}-2S_{2r})/(k-r-S_{1r})\rceil$.

(c) The ordered Chebyshev bound is a valid upper bound on $\overline{P}(n,k,\boldsymbol{p})$*:*

$$\overline{P}(n,k,\boldsymbol{p}) \le \min_{r\in[0,k-1]} CH(n-r,k-r,\boldsymbol{p}), \forall k \in [2,n], \tag{3.3}$$

where:

$$CH(n-r,\ k-r,\ \boldsymbol{p}) = \begin{cases} 1, & k < S_{1r}+r, \\ \dfrac{S_{1r}-(S_{1r}^2-2S_{2r})}{S_{1r}-(S_{1r}^2-2S_{2r})+(k-r-S_{1r})^2}, & S_{1r}+r \le k \le n. \end{cases}$$

Proof.

(a) We observe that for any $r_1 \in [0,k-1]$ and any subset $S \subseteq [n]$ of the random variables of cardinality $n-r_1$, an upper bound is given by:

$$\begin{aligned} \mathbb{P}\left(\sum_{i\in[n]} \tilde{c}_i \ge k\right) &\le \mathbb{P}\left(\sum_{i\in S} \tilde{c}_i \ge k-r_1\right) \\ &\quad \left[\text{since } \textstyle\sum_{i\in[n]} c_i \ge k \text{ implies } \sum_{i\in S} c_i \ge k-r_1\right] \\ &\le \frac{\mathbb{E}\left[\sum_{i\in S}\tilde{c}_i\right]}{k-r_1} \\ &\quad [\text{using Markov inequality}] \\ &= \frac{\sum_{i\in S} p_i}{k-r_1}. \end{aligned}$$

The tightest upper bound of this form is obtained by minimizing over all $r_1 \in [0,k-1]$ and subsets $S \subseteq [n]$ with $|S| = n-r_1$:

$$\begin{aligned} \mathbb{P}\left(\sum_{i\in[n]} \tilde{c}_i \ge k\right) &\le \min_{r_1\in[0,k-1]} \min_{S:|S|=n-r_1} \frac{\sum_{i\in S} p_i}{k-r_1} \\ &= \min_{r_1\in[0,k-1]} \frac{\sum_{i\in[n-r_1]} p_i}{k-r_1} \\ &\quad [\text{using the } n-r_1 \text{ smallest probabilities}]. \end{aligned} \tag{3.4}$$

We derive the other term in (3.1) using a similar approach while accounting for pairwise independence. For any $r_2 \in [0,k-2]$ and any subset $S \subseteq [n]$ of the random variables of cardinality $n-r_2$, an upper bound is given by:

$$\begin{aligned}
\mathbb{P}\left(\sum_{i\in[n]}\tilde{c}_i \geq k\right) &\leq \mathbb{P}\left(\sum_{i\in S}\tilde{c}_i \geq k-r_2\right)\\
&= \mathbb{P}\left(\binom{\sum_{i\in S}\tilde{c}_i}{2} \geq \binom{k-r_2}{2}\right)\\
&\leq \frac{\mathbb{E}\left[\sum_{i\in S}\sum_{j\in S:j>i}\tilde{c}_i\tilde{c}_j\right]}{\binom{k-r_2}{2}}\\
&\quad \text{[using equation (1.3) and Markov inequality]}\\
&= \frac{\sum_{i\in S}\sum_{j\in S:j>i}\mathbb{E}[\tilde{c}_i]\mathbb{E}[\tilde{c}_j]}{\binom{k-r_2}{2}}\\
&\quad \text{[using pairwise independence]}\\
&= \frac{\sum_{i\in S}\sum_{j\in S:j>i}p_ip_j}{\binom{k-r_2}{2}}.
\end{aligned}$$

The tightest upper bound of this form is obtained by minimizing over $r_2 \in [0, k-2]$ and all sets S of size $n-r_2$. This gives:

$$\begin{aligned}
\mathbb{P}\left(\sum_{i\in[n]}\tilde{c}_i \geq k\right) &\leq \min_{r_2\in[0,k-2]}\ \min_{S:|S|=n-r_2} \frac{\sum_{i\in S}\sum_{j\in S:j>i}p_ip_j}{\binom{k-r_2}{2}}\\
&= \min_{r_2\in[0,k-2]}\left(\frac{\sum_{(i,j)\in K_{n-r_2}}p_ip_j}{\binom{k-r_2}{2}}\right)\\
&\quad \text{[using the } n-r_2 \text{ smallest probabilities].}
\end{aligned} \tag{3.5}$$

From the bounds (3.4) and (3.5), we get:

$$\overline{P}(n,k,\boldsymbol{p}) \leq \min\left(1, \min_{r_1\in[0,k-1]}\left(\frac{S_{1r_1}}{k-r_1}\right), \min_{r_2\in[0,k-2]}\left(\frac{S_{2r_2}}{\binom{k-r_2}{2}}\right)\right), \quad \forall k \in [2,n],$$

where $S_{1r_1} = \sum_{i\in[n-r_1]}p_i$ for $r_1 \in [0, n-1]$ and $S_{2r_2} = \sum_{(i,j)\in K_{n-r_2}}p_ip_j$ for $r_2 \in [0, n-2]$. One can interpret this bound as creating a set of dual feasible solutions and picking the best among them. The dual formulation is:

$$\begin{aligned}
\overline{P}(n,k,\boldsymbol{p}) = \min\ & \sum_{(i,j)\in K_n}\lambda_{ij}p_ip_j + \sum_{i\in[n]}\lambda_ip_i + \lambda_0\\
\text{s.t}\ & \sum_{(i,j)\in K_n}\lambda_{ij}c_ic_j + \sum_{i\in[n]}\lambda_ic_i + \lambda_0 \geq 0 \quad \forall \boldsymbol{c} \in \{0,1\}^n,\\
& \sum_{(i,j)\in K_n}\lambda_{ij}c_ic_j + \sum_{i\in[n]}\lambda_ic_i + \lambda_0 \geq 1, \quad \forall \boldsymbol{c} \in \{0,1\}^n : \textstyle\sum_t c_t \geq k.
\end{aligned}$$

The components of the second term in (3.1) are obtained by choosing dual feasible solutions with $\lambda_i = 1/(k-r_1)$ for $i \in [n-r_1]$ and setting all other dual variables to 0. Similarly, the components of the third term are obtained by choosing dual feasible solutions with $\lambda_{ij} = 1/\binom{k-r_2}{2}$ for $(i,j) \in K_{n-r_2}$ and setting all other dual variables to 0.

(b) The bound in (3.2) is obtained by using the inequality:

$$\mathbb{P}\left(\sum_{i\in[n]}\tilde{c}_i \geq k\right) \leq \mathbb{P}\left(\sum_{i\in[n-r]}\tilde{c}_i \geq k-r\right), \quad \forall r \in [0, k-1],$$

in conjunction with the bound in (1.7) computed from Boros and Prékopa [6]. We compute an upper bound on $\mathbb{P}\left(\sum_{i\in[n-r]}\tilde{c}_i \geq k-r\right)$ by using the aggregated moments S_{1r} and S_{2r} with the Boros and Prékopa bound from (1.7) as follows:

$$BP(n-r, k-r, \boldsymbol{p}) = \begin{cases} 1, & k < \dfrac{(n-r-1)S_{1r} - 2S_{2r}}{n-r-S_{1r}} + r, \\[2ex] \dfrac{(k+n-2r-1)S_{1r} - 2S_{2r}}{(k-r)(n-r)}, & \dfrac{(n-r-1)S_{1r} - 2S_{2r}}{n-r-S_{1r}} + r \leq k < 1 + \dfrac{2S_{2r}}{S_{1r}} + r, \\[2ex] \dfrac{(i-1)(i-2S_{1r}) + 2S_{2r}}{(k-r-i)^2 + (k-r-i)}, & k \geq 1 + \dfrac{2S_{2r}}{S_{1r}} + r, \end{cases}$$

where $i = \lceil((k-r-1)S_{1r} - 2S_{2r})/(k-r-S_{1r})\rceil$. Since the relation $\overline{P}(n, k, \boldsymbol{p}) \leq BP(n-r, k-r, \boldsymbol{p})$ is satisfied for every $0 \leq r \leq k-1$, the best upper bound on $\overline{P}(n, k, \boldsymbol{p})$ is obtained by taking the minimum over all possible values of r:

$$\overline{P}(n, k, \boldsymbol{p}) \leq \min_{r\in[0,k-1]} BP(n-r, k-r, \boldsymbol{p}), \quad \forall k \in [2, n].$$

(c) Proceeding in a similar manner as in (b), by using the aggregated moments S_{1r} and S_{2r} with Chebyshev bound, the upper bound for a given $r \in [0, k-1]$ can be written as follows:

$$CH(n-r,\ k-r,\ \boldsymbol{p}) = \begin{cases} 1, & k < S_{1r} + r, \\ \dfrac{S_{1r} - (S_{1r}^2 - 2S_{2r})}{S_{1r} - (S_{1r}^2 - 2S_{2r}) + (k-r-S_{1r})^2}, & S_{1r} + r \leq k \leq n. \end{cases}$$

The best upper bound on $\overline{P}(n, k, \boldsymbol{p})$ is obtained by taking the minimum over all possible values of r:

$$\overline{P}(n, k, \boldsymbol{p}) \leq \min_{r\in[0,k-1]} CH(n-r, k-r, \boldsymbol{p}), \quad \forall k \in [2, n].$$

□

3.1. Connection to existing results. Prior work in Rüger [52] shows that ordering of probabilities provides the tightest upper bound on the probability of n Bernoulli random variables adding up to at least k, when allowing for arbitrary dependence. Specifically, the bound derived there is:

$$\overline{P}_u(n, k, \boldsymbol{p}) = \max_{\theta\in\Theta(\boldsymbol{p})} \mathbb{P}_\theta\left(\sum_{i\in[n]}\tilde{c}_i \geq k\right) = \min\left(1, \min_{r\in[0,k-1]}\left(\frac{S_{1r}}{k-r}\right)\right).$$

However, this bound does not use pairwise independence information. Part (a) of Theorem 3.1 tightens the analysis in Rüger [52] for pairwise independent random

variables. It is also straightforward to see that the ordered Schmidt, Siegel and Srinivasan bound in (3.1) is at least as good as the bound in (1.4) (simply plug in $r = 0$). Building on the ordering of probabilities, the bound in (3.2) uses aggregated binomial moments for k ordered sets of random variables of size $n - r$ where $r \in [0, k-1]$. When $r = 0$, the bound in (3.2) reduces to the original aggregated moment bound of Boros and Prékopa in (1.7) and hence this bound is at least as tight. All the bounds in Theorem 3.1 are clearly efficiently computable.
It is easy to verify that the ordered Boros and Prékopa bound is at least as good as the other two ordered bounds, *i.e.*,

$$\text{Ordered bound (3.2)} \leq \min\left(\text{Ordered bound (3.1)}, \text{Ordered bound (3.3)}\right).$$

This is true since, each term of the ordered bounds are derived by finding upper bounds on the probability that the sum of the first $n - r$ random variables takes a value of at least $k - r$ using only the first two moments of the sum of these random variables. Since the Boros and Prékopa bound is the tightest upper bound possible when using only the first two moments of the sum, each term in the ordered Boros and Prékopa bound is at least as good as the corresponding term in the other two ordered bounds. Taking the minimum over all these terms implies that the ordered Boros and Prékopa bound must be at least as good as the other two bounds.

3.2. Further tightening of ordered bounds:. It is also worth mentioning that the bounds in Theorem 3.1 can in fact be strengthened further by using the tightest possible bound for $k = 1$ from Theorem 2.3. Specifically, we can tighten the ordered Schmidt, Siegel and Srinivasan bound in (3.1) as follows:

$$\min\left(1, \min_{r\in[0,k-2]} \min\left(\frac{S_{1r}}{k-r}, \frac{S_{2r}}{\binom{k-r}{2}}\right), \sum_{i\in[n-k+1]} p_i - p_{n-k+1} \sum_{i\in[n-k]} p_i\right).$$

where the last term corresponds to $r_1 = k - 1$ and is obtained by observing that:

$$\begin{aligned} \mathbb{P}\left(\sum_{i\in[n]} \tilde{c}_i \geq k\right) &\leq \mathbb{P}\left(\sum_{i\in[n-k+1]} \tilde{c}_i \geq 1\right) \\ &\leq \sum_{i\in[n-k+1]} p_i - p_{n-k+1} \sum_{i\in[n-k]} p_i \\ &\quad \text{[from Theorem 2.3]} \\ &\leq \textstyle\sum_{i\in[n-k+1]} p_i \\ &= S_{1(k-1)}/(k-(k-1)). \end{aligned}$$

The Boros and Prékopa bound and Chebyshev ordered bounds in (3.2) and (3.3) can be similarly tightened. Unlike the bounds in Theorem 3.1, these tightened bounds use partially disaggregated moment information. We next provide two numerical examples to illustrate the impact of ordering on the quality of the three bounds. We restrict attention, however, to the aggregated ordered moment bounds in Theorem 3.1 only.

3.3. Numerical illustrations.

Example 3.2 (Non-identical marginals). Consider an example with $n = 12$ random variables with the probabilities given by

$$p_1 = 0.0651, p_2 = 0.0977, p_3 = 0.1220, p_4 = 0.1705, p_5 = 0.3046, p_6 = 0.4402,$$
$$p_7 = 0.4952, p_8 = 0.6075, p_9 = 0.6842, p_{10} = 0.8084, p_{11} = 0.9489, p_{12} = 0.9656.$$

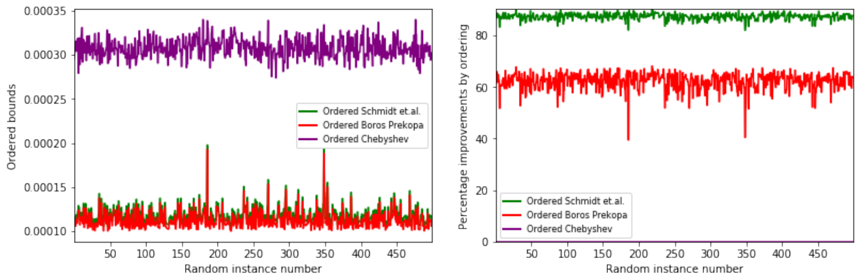
Table 4 compares the three ordered bounds with the three unordered bounds and the tight upper bound. Numerically, the ordered Boros and Prékopa bound (3.2) is found to be tight in this example for $k = 7, 8, 9, 12$ while the ordered Schmidt, Siegel and Srinivasan bound (3.1) is tight for $k = 12$. The ordered Boros and Prékopa bound is uniformly the best performing of the three bounds, while among the other two ordered bounds, none uniformly dominates the other. For example, comparing the ordered bounds when $7 \leq k \leq 9$, the Chebyshev bound outperforms the Schmidt, Siegel and Srinivasan bound, but when $k = 6$ or $10 \leq k \leq 12$, the Schmidt, Siegel and Srinivasan bound does better. Comparing the unordered bounds when $7 \leq k \leq 9$, the Schmidt, Siegel and Srinivasan bound (1.4) outperforms the Chebyshev bound (1.2) when $k = 6$ but for all $k \geq 7$, bound (1.2) does better. In terms of absolute difference between ordered and unordered bounds, ordering provides the maximum improvement to the Schmidt, Siegel and Srinivasan bound, followed by the Boros and Prékopa bound and the Chebyshev bound.

Table 4: Upper bound on the probability of sum of random variables equaling at least k for $n = 12$. For each value of k, the bottom row provides the tightest bound which can be computed in this example by solving an exponential sized linear program. The underlined instances illustrate cases when the other upper bounds are tight.

Bound	$k \in [1, 4]$	$k = 5$	$k = 6$	$k = 7$	$k = 8$	$k = 9$	$k = 10$	$k = 11$	$k = 12$
(1.2)	1	1	0.9553	0.5192	0.2552	0.1424	0.0889	0.0603	0.0434
(3.3)	1	1	0.9553	0.5192	0.2552	0.1424	0.0883	0.0549	0.0307
(1.4)	1	1	0.9517	0.6831	0.5123	0.3985	0.3188	0.2608	0.2173
(3.1)	1	1	0.9489	0.6162	0.3620	0.1827	0.0712	0.0250	0.0064
(1.7)	1	1	0.9497	0.5018	0.2509	0.1326	0.0795	0.0530	0.0379
(3.2)	1	1	0.9254	0.5018	0.2509	0.1290	0.0712	0.0249	0.0064
Tight	1	0.9957	0.8931	0.5018	0.2509	0.1290	0.0692	0.0230	0.0064

Example 3.3 (Non-identical marginals). In this example, we numerically compute the improvement of the new ordered bounds over the unordered bounds for $n = 100$ variables by creating 500 instances by randomly generating the probabilities $\boldsymbol{p} = (p_1, p_2, .., p_{100})$. First, we consider small marginal probabilities by uniformly and independently generating the entries of $\boldsymbol{p}$ between 0.01 and 0.05. When $k = n$, Figure 4a plots the three ordered bounds while Figure 4b shows the percentage improvement of the three bounds over their unordered counterparts. The percentage improvement is computed as $\big([\text{unordered-ordered}]/\text{unordered}\big) \times 100\%$. In this example with small marginals, the ordered Schmidt, Siegel and Srinivasan bound (3.1) is equal to the ordered Boros and Prékopa bound (3.2) as seen in Figure 4a. Ordering tends to improve the Schmidt, Siegel and Srinivasan bound significantly for smaller probabilities, since both the partial binomial moment terms S_{1r} and S_{2r} are smaller with smaller marginal probabilities for all $r \in [0, k-1]$.

The percentage improvement due to ordering in figure 4b is consistently above 80% for the Schmidt, Siegel and Srinivasan bound, while that of the Boros and Prékopa bound is around 60%. The ordered Chebyshev bound (3.3) shows an al-

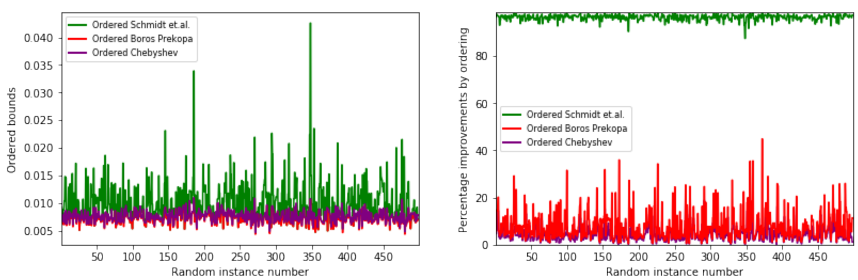


(a) Actual value of the ordered bounds (b) Percentage improvement of ordered bounds

Fig. 4: Smaller marginal probabilities p_i with $n = 100, k = 100$ and 500 instances.

most negligible improvement by ordering in this example.

Next, we consider similar plots when $k = n-1$ with larger marginal probabilities. The entries of $\boldsymbol{p}$ are generated uniformly and independently between 0.05 and 0.99. In Figure 5a, the ordered Chebyshev bound (3.3) performs better than the ordered



(a) Actual value of the ordered bounds (b) Percentage improvement of ordered bounds

Fig. 5: Larger marginal probabilities p_i with $n = 100, k = 99$ and 500 instances.

Schmidt, Siegel and Srinivasan bound (3.1). In Figure 5b, the percentage improvement due to ordering is again most significant for the Schmidt, Siegel and Srinivasan bound, being consistently above 90% while that of the Boros and Prékopa bound is less than 40% and that of the Chebyshev bound is less than 20%. It is also clear from Figures 4 and 5 that the ordered Boros and Prékopa bound (3.2) is the tightest of the three bounds across the instances, while among the other two bounds, none uniformly dominates the other.

4. Tightness in special cases. In this section, we identify two tight instances, one for the unordered bounds in (1.2), (1.4) and (1.7) and the other for the corresponding ordered bounds derived in Theorem 3.1. Firstly, in Section 4.1, for identical variables, the symmetry in the problem allows for closed-form tight bounds for any $k \in [2, n]$. We prove this by showing an equivalence of the exponential sized linear program (1.5) which computes the exact bound with a polynomial sized linear program analyzed in computing the Boros and Prékopa bound in (1.7). We use the exact bound to identify instances when the other two unordered bounds are tight.

The result with identical marginals is further extended to show tightness for t-wise independent variables. Secondly, in Section 4.2, we demonstrate the usefulness of the ordered bounds by identifying a special case when $n-1$ marginals are identical (with additional conditions on the probability and k), when the ordered bounds in (3.1) and (3.2) are tight.

4.1. Tightness of bounds with identical marginals. In this section, we provide probability bounds for n pairwise independent random variables adding up to at least $k \in [2, n]$ when their marginals are identical. The next theorem provides the tight bound with identical marginals, by applying the Boros and Prékopa bound in (1.7) to pairwise independent variables with $\tilde{\xi} = \sum_{i\in[n]} \tilde{c}_i$.

THEOREM 4.1. *Assume $p_i = p \in (0, 1)$ for $i \in [n]$. Let $\overline{P}(n, k, p)$ represent the tightest upper bound on the probability that n pairwise independent identical Bernoulli random variables add up to at least $k \in [n]$. Then,*

(4.1)
$$\overline{P}(n,k,p) = \begin{cases} 1, & k < (n-1)p, \quad (a) \\ \dfrac{((n-1)(1-p)+k)p}{k}, & (n-1)p \le k < 1+(n-1)p, \quad (b) \\ \dfrac{(i-1)(i-2np)+n(n-1)p^2}{(k-i)^2+(k-i)}, & k \ge 1+(n-1)p, \quad (c), \end{cases}$$

where $i = \lceil np(k-1-(n-1)p)/(k-np) \rceil$.

Proof. The tightest upper bound $\overline{P}(n, k, p)$ is the optimal value of the linear program:

$$(4.2) \qquad \begin{aligned} \overline{P}(n,k,p) = \max \quad & \sum_{\boldsymbol{c}\in\{0,1\}^n:\sum_i c_i \ge k} \theta(\boldsymbol{c}) \\ \text{s.t} \quad & \sum_{\boldsymbol{c}\in\{0,1\}^n} \theta(\boldsymbol{c}) = 1, \\ & \sum_{\boldsymbol{c}\in\{0,1\}^n:c_i=1} \theta(\boldsymbol{c}) = p, && \forall i \in [n], \\ & \sum_{\boldsymbol{c}\in\{0,1\}^n:c_i=1,c_j=1} \theta(\boldsymbol{c}) = p^2, && \forall (i,j) \in K_n, \\ & \theta(\boldsymbol{c}) \ge 0, && \forall \boldsymbol{c} \in \{0,1\}^n, \end{aligned}$$

where the decision variables are the joint probabilities $\theta(\boldsymbol{c}) = \mathbb{P}(\tilde{\boldsymbol{c}} = \boldsymbol{c})$ for $\boldsymbol{c} \in \{0, 1\}^n$. Consider the following linear program in $n+1$ variables which provides an upper bound on $\overline{P}(n, k, p)$:

$$(4.3) \qquad \begin{aligned} BP(n,k,p) = \max \quad & \sum_{\ell\in[k,n]} v_\ell \\ \text{s.t.} \quad & \sum_{\ell\in[0,n]} v_\ell = 1, \\ & \sum_{\ell\in[1,n]} \ell v_\ell = np, \\ & \sum_{\ell\in[2,n]} \binom{\ell}{2} v_\ell = \binom{n}{2} p^2, \\ & v_\ell \ge 0, && \forall \ell \in [0,n], \end{aligned}$$

where the decision variables are the probabilities $v_\ell = \mathbb{P}(\sum_{i\in[n]} \tilde{c}_i = \ell)$ for $l \in [0, n]$. Linear programs of the form (4.3) have been studied in Boros and Prékopa [6] in the context of aggregated binomial moment problems. As we shall see, these two formulations are equivalent with identical pairwise independent random variables.

(1) $\overline{P}(n,k,p) \leq BP(n,k,p)$: Given a feasible solution to (4.2) denoted by θ, construct a feasible solution to the linear program (4.3) as:

$$v_\ell = \sum_{\boldsymbol{c}\in\{0,1\}^n:\sum_i c_i = l} \theta(\boldsymbol{c}), \quad \forall l \in [0,n].$$

By taking expectations on both sides of the equality (1.3), we get:

$$\sum_{l\in[j,n]} \binom{l}{j} \mathbb{P}\left(\sum_{i\in[n]} \tilde{c}_i = l\right) = \mathbb{E}\left[S_j(\tilde{\boldsymbol{c}})\right], \quad \forall j \in [0,n].$$

Applying it for $j = 0,1,2$, we get the three equality constraints in (4.3):

$$\begin{aligned}
&\sum_{\ell\in[0,n]} v_\ell = 1,\\
&\sum_{\ell\in[1,n]} \ell v_\ell = \mathbb{E}\left[\sum_{i\in[n]} \tilde{c}_i\right] = np,\\
&\sum_{\ell\in[2,n]} \binom{\ell}{2} v_\ell = \mathbb{E}\left[\sum_{(i,j)\in K_n} \tilde{c}_i\tilde{c}_j\right] = n(n-1)p^2/2.
\end{aligned}$$

Lastly, the objective function value of this feasible solution satisfies:

$$\begin{aligned}
\sum_{\ell=k}^{n} v_\ell &= \sum_{\ell=k}^{n} \sum_{\boldsymbol{c}\in\{0,1\}^n:\sum_i c_i = l} \theta(\boldsymbol{c})\\
&= \sum_{\boldsymbol{c}\in\{0,1\}^n:\sum_i c_i \geq k} \theta(\boldsymbol{c}).
\end{aligned}$$

Hence, $\overline{P}(n,k,p) \leq BP(n,k,p)$.

(2) $\overline{P}(n,k,p) \geq BP(n,k,p)$: Given an optimal solution to (4.3) denoted by $\boldsymbol{v}$, construct a feasible solution to the linear program (4.2) by distributing v_ℓ equally among all the realizations in $\{0,1\}^n$ with exactly ℓ ones:

$$\theta(\boldsymbol{c}) = \frac{v_\ell}{\binom{n}{\ell}}, \quad \forall \boldsymbol{c} \in \{0,1\}^n : \textstyle\sum_{i\in[n]} c_i = \ell, \forall \ell \in [0,n].$$

The first constraint in (4.2) is satisfied since:

$$\begin{aligned}
\sum_{\boldsymbol{c}\in\{0,1\}^n} \theta(\boldsymbol{c}) &= \sum_{\ell\in[0,n]} \sum_{\boldsymbol{c}\in\{0,1\}^n:\sum_i c_i = \ell} \frac{v_\ell}{\binom{n}{\ell}}\\
&\quad \text{[since } \left|\{0,1\}^n : \textstyle\sum_{i\in[n]} c_i = \ell\right| = \tbinom{n}{\ell}\text{]}\\
&= \sum_{\ell\in[0,n]} v_\ell\\
&= 1.
\end{aligned}$$

The second constraint in (4.2) is satisfied since:

$$\begin{aligned}\sum_{\boldsymbol{c}\in\{0,1\}^n:c_j=1}\theta(\boldsymbol{c}) &= \sum_{\ell\in[1,n]}\frac{v_\ell}{\binom{n}{\ell}}\binom{n-1}{\ell-1}\\ &\quad \left[\text{since } \left|\{0,1\}^n : \textstyle\sum_{i\in[n]} c_i=\ell, c_j=1\right| = \binom{n-1}{\ell-1}\right]\\ &= \sum_{\ell\in[1,n]}\frac{\ell v_\ell}{n}\\ &= p.\end{aligned}$$

The third constraint in (4.2) satisfied since:

$$\begin{aligned}\sum_{\boldsymbol{c}\in\{0,1\}^n:c_i=1,c_j=1}\theta(\boldsymbol{c}) &= \sum_{\ell\in[2,n]}\frac{v_\ell}{\binom{n}{\ell}}\binom{n-2}{\ell-2}\\ &\quad \left[\text{since } \left|\{0,1\}^n : \textstyle\sum_{t\in[n]} c_t=\ell, c_i=1, c_j=1\right| = \binom{n-2}{\ell-2}\right]\\ &= \frac{2}{n(n-1)}\sum_{\ell\in[2,n]}\binom{\ell}{2}v_\ell\\ &= p^2.\end{aligned}$$

The objective function value of the feasible solution is given by:

$$\begin{aligned}\sum_{\boldsymbol{c}\in\{0,1\}^n:\sum_i c_i\geq k}\theta(\boldsymbol{c}) &= \sum_{\ell\in[k,n]}\ \sum_{\boldsymbol{c}\in\{0,1\}^n:\sum_i c_i=l}\theta(\boldsymbol{c})\\ &= \sum_{\ell\in[k,n]} v_\ell\\ &= BP(n,k,p).\end{aligned}$$

Hence, the optimal objective value of the two linear programs are equivalent. The formula for the tight bound in the theorem is then exactly the Boros and Prékopa bound in (1.7) (the bound $BP(n,k,p)$ is also derived in the work of [53], although tightness of the bound is not shown there). It is straightforward to verify that the following distributions attain the bounds for each of the cases (a)-(c) in the statement of the theorem:

(a) The probabilities are given as:

$$\theta(\boldsymbol{c}) = \begin{cases} \dfrac{(1-p)(j-(n-1)p)}{\binom{n-1}{j-1}}, & \text{if } \displaystyle\sum_{t\in[n]} c_t = j-1,\\ \dfrac{(1-p)(1+(n-1)p-j)}{\binom{n-1}{j}}, & \text{if } \displaystyle\sum_{t\in[n]} c_t = j,\\ \dfrac{n(n-1)p^2+(j-1)(j-2np)}{(n-j)^2+(n-j)}, & \text{if } \displaystyle\sum_{t\in[n]} c_t = n,\end{cases}$$

where $j = \lceil (n-1)p \rceil$ and all other support points have zero probability.

(b) The probabilities are given as:

$$\theta(\boldsymbol{c}) = \begin{cases} \dfrac{1-p}{k}(k-(n-1)p), & \text{if } \displaystyle\sum_{t\in[n]} c_t = 0, \\ \dfrac{p(1-p)}{\binom{n-2}{k-1}}, & \text{if } \displaystyle\sum_{t\in[n]} c_t = k, \\ \dfrac{p((n-1)p-(k-1))}{n-k}, & \text{if } \displaystyle\sum_{t\in[n]} c_t = n, \end{cases}$$

where all other support points have zero probability.

(c) The probabilities are given as:

$$\theta(\boldsymbol{c}) = \begin{cases} \dfrac{np[(n-1)p-(k+i-1)]+ik}{\binom{n}{i-1}(k-i+1)}, & \text{if } \displaystyle\sum_{t\in[n]} c_t = i-1, \\ \dfrac{np[(k+i-2)-(n-1)p]-k(i-1)}{\binom{n}{i}(k-i)}, & \text{if } \displaystyle\sum_{t\in[n]} c_t = i, \\ \dfrac{n(n-1)p^2+(i-1)(i-2np)}{\binom{n}{k}[(k-i)^2+(k-i)]}, & \text{if } \displaystyle\sum_{t\in[n]} c_t = k, \end{cases}$$

where all other support points have zero probability and the index i is evaluated as stated in equation (4.1)(c). It is straightforward to see that with identical marginals, the tight union bound in Theorem 2.3 reduces to the bound in case (b) of Theorem 4.1. ∎

4.1.1. Connection of Theorem 4.1 to existing results. Tightness results with identical Bernoulli random variables have been established in the literature in the context of occurence of at least and exactly k out of n events for specific regimes of the parameters n, k and p. Theorem 4.1 however, provides the tight bounds for all values of (n, k, p). Recent work by Garnett [22] provides the tight upper bound on the probability that the sum of pairwise independent Bernoulli random variables exceeds the mean by a small amount (this corresponds to case (b)). Pinelis [44] derives a closed-form tight lower bound on the probability of occurence of exactly one of out n events. Benjamini et al. [3] and Peled et al. [43] derived closed-form upper and lower bounds (not necessarily tight) on the maximal intersection probability of more general t-wise independent Bernoulli random variables (this corresponds to $k = n$ in case (c) for $t = 2$). These bounds were shown to match each other up to multiplicative factors of lower order in a large regime of the parameters n, p, t. The connection of the intersection probability with the linear program based approach of Boros and Prékopa [6] has been mentioned in these papers, although the equivalence for all values of k is not established. Corollary 4.2 in this paper, however, establishes the equivalence for all values of n, k, p, t. The usefulness of Theorem 4.1 lies in the fact that it can be extended to incorporate a wide variety of cases involving identical Bernoulli events by using the results from Boros and Prékopa [6] as follows:

i) Tight closed-form lower bounds on probability of occurrence of at least k out of n events

ii) Tight closed-form upper and lower bounds on the probability of occurence of exactly k out of n events

iii) Tight linear program based upper and lower bounds for t-wise independent variables ($t \geq 3$) from the symmetry assumptions (see Corollary 4.2).

We note that when $k \geq 1+(n-1)p$, the tight lower bound from [6] can be derived as:

$$\underline{P}(n,k,p) = \begin{cases} \frac{\left(2+(n-1)p-k\right)p}{n-k+1}, & 1+(n-1)p \leq k < 2+(n-1)p \\ 0, & k \geq 2+(n-1)p. \end{cases}$$

When $k = n \geq 1+(n-1)p$, this bound reduces to $\max\left(p\left((n-1)p-(n-2)\right),0\right)$ which is exactly the intersection bound computed in Corollary 2.9 with identical probabilities.

Corollary 4.2. *Consider identical t-wise independent Bernoulli random variables with probabilities $p \in (0,1)$ where $t \in [2,n]$. Then, the tightest upper bound on the probability of n such variables adding up to at least $k \in [n]$, denoted by $\overline{P}(n,k,p,t)$, can be computed as the optimal value of the aggregated linear program proposed by Prékopa [48]:*

$$\text{(4.4)} \quad \begin{aligned} \overline{P}(n,k,p,t) = \max \quad & \sum_{\ell=k}^{n} v_\ell \\ s.t. \quad & \sum_{\ell=m}^{n} \binom{\ell}{m} v_\ell = \binom{n}{m} p^m, \quad && \forall m \in [0,t], \\ & v_\ell \geq 0, && \forall \ell \in [0,n], \end{aligned}$$

where the decision variables are the probabilities $v_\ell = \mathbb{P}(\sum_{i=1}^n \tilde{c}_i = \ell)$ for $l \in [0,n]$.

Proof. The proof is straightforward from the proof of Theorem 4.1 which implies the equivalence of (4.4) with the large-sized linear program:

$$\text{(4.5)} \quad \begin{aligned} \overline{P}(n,k,p,t) = \max \quad & \sum_{\boldsymbol{c} \in \{0,1\}^n : \sum_i c_i \geq k} \mathbb{P}(\boldsymbol{c}) \\ \text{s.t.} \quad & \sum_{\boldsymbol{c} \in \{0,1\}^n} \mathbb{P}(\boldsymbol{c}) = 1, \\ & \sum_{\boldsymbol{c} \in \{0,1\}^n :\ c_i = 1,\ \forall i \in J} \mathbb{P}(\boldsymbol{c}) = p^m, \quad \forall J \in I_m,\ m \in [t], \\ & \mathbb{P}(\boldsymbol{c}) \geq 0, \quad \forall \boldsymbol{c} \in \{0,1\}^n, \end{aligned}$$

where $I_m = \{I \subseteq [n] : |I| = m\}$. In particular for any given feasible solution of (4.4), we can distribute the probability mass v_ℓ evenly across the $\binom{n}{\ell}$ scenarios for every $\ell \in [0,n]$ and satisfy all the constraints in (4.5) while for any given feasible solution of (4.5), we can aggregate the probabilities $\mathbb{P}(\boldsymbol{c})$ as

$$v_\ell = \sum_{\boldsymbol{c} \in \{0,1\}^n : \sum_i c_i = l} \mathbb{P}(\boldsymbol{c}), \quad \forall l \in [0,n].$$

and satisfy all constraints in (4.4). ∎

We note that for 3-wise independent variables, a closed-form expression for the optimal objective in (4.4) using the first three binomial moments has been provided in [6]. Further, the corresponding tight lower bound $\underline{P}(n,k,p,t)$ can be computed as the optimal value of the minimization version of the aggregated linear program in (4.4).

4.1.2. Tightness of alternative bounds. We next discuss an application of Theorem 4.1. Since the marginals are identical, it is easy to see that the ordered bounds in Theorem 3.1 reduce to the unordered bounds corresponding to $r = 0$. While the unordered Boros and Prékopa bound provides the tightest upper bound with identical marginals, the formula is more involved than the unordered Chebyshev bound which reduces to:

$$\overline{P}(n,k,p) \leq \begin{cases} 1, & k < np, \\ np(1-p)/\big(np(1-p)+(k-np)^2\big), & np \leq k \leq n. \end{cases} \tag{4.6}$$

and the unordered Schmidt, Siegel and Srinivasan bound which reduces to:

$$\overline{P}(n,k,p) \leq \min\left(1, \frac{np}{k}, \frac{n(n-1)p^2}{k(k-1)}\right). \tag{4.7}$$

It is possible to then use Theorem 4.1 to identify conditions on the parameters (n, k, p) for which the bounds in (4.6) and (4.7) are tight. We only focus on the non-trivial cases where the tight bound is strictly less than one and $n \geq 3$. Henceforth, the Chebyshev and Schmidt, Siegel and Srinivasan bounds referred to in this section are the unordered bounds.

Proposition 4.3.

(a) For $p = \alpha/(n-1)$ and any integer $\alpha \in [n-2]$, the Chebyshev bound in (4.6) is tight for the values of $k = \alpha + 1$ and $k = n$.

(b) For $p \leq 1/(n-1)$, the Schmidt, Siegel and Srinivasan bound in (4.7) is tight for all $k \in [2, n]$ while for $p > 1/(n-1)$, the bound is tight for all values of $k \in [\lceil 1 + (n-1)p \rceil, \lfloor n(n-1)p^2/(np-1) \rfloor]$.

Proof. Since Theorem 4.1 provides the tight bound, we simply need to show the equivalence with the bounds in (4.6) and (4.7) for the instances in the proposition.

(a) Consider $p = \alpha/(n-1)$ for any integer $\alpha \in [n-2]$.

1. Set $k = \alpha + 1$. This corresponds to case (c) in Theorem 4.1. Plugging in the values, the index i which is required for finding the tight bound is given by:

$$\begin{aligned} i &= \left\lceil \frac{n\alpha(\alpha+1-1-\alpha)/(n-1)}{\alpha+1-n\alpha/(n-1)} \right\rceil \\ &= 0. \end{aligned}$$

The corresponding tight bound in (4.1) gives:

$$\overline{P}(n,k,p) = \frac{n\alpha}{(n-1)(\alpha+1)} = \frac{np}{np+1-p}.$$

It is straightforward to verify by plugging in the values that the Chebyshev bound is exactly the same.

2. Set $k = n$. This corresponds to case (c) in Theorem 4.1. Plugging in the values, the index i in the tight bound is given by:

$$\begin{aligned} i &= \left\lceil \frac{n\alpha(n-1-\alpha)/(n-1)}{n-n\alpha/(n-1)} \right\rceil \\ &= \alpha. \end{aligned}$$

The tight bound in (4.1) gives:

$$\overline{P}(n,k,p) = \frac{\alpha}{(n-1)(n-\alpha)} = \frac{p}{p+n(1-p)}.$$

It is straightforward to verify by plugging in the values that the Chebyshev bound is exactly the same in this case.

(b) Observe that the last two terms in the Schmidt, Siegel and Srinivasan bound in (4.7) satisfy:

$$\frac{n(n-1)p^2}{k(k-1)} \leq \frac{np}{k} \text{ when } k \geq 1 + (n-1)p.$$

Since $k \geq 1+(n-1)p$ implies $1 \geq np/k$, the bound in (4.7) reduces to $n(n-1)p^2/k(k-1)$. The range of $k \geq 1+(n-1)p$ corresponds to case (c) in Theorem 4.1. If $k = 1+(n-1)p$, the index $i = \lceil np(k-(1+(n-1)p))/(k-np) \rceil = 0$ and the tight bound from (4.1) is:

$$\frac{np}{1+(n-1)p},$$

which is exactly the Schmidt, Siegel and Srinivasan bound. We can also verify that when the index $i = 1$ in case (c), then the tight bound in (4.1) reduces to:

$$\begin{aligned} \overline{P}(n,k,p) &= \frac{n(n-1)p^2 + (1-1)(1-2np)}{(k-1)^2 + (k-1)} \\ &= \frac{n(n-1)p^2}{k(k-1)}. \end{aligned}$$

We now identify conditions when $k > 1 + (n-1)p$ and the index i is equal to one.

1. Consider $0 < p \leq 1/(n-1)$. For the values of p in this interval, the valid range of k in case (c) corresponds to integer values of $k \geq 1 + (n-1)p$ which means $k \geq 2$. For the probability $0 < p \leq 1/n$, the index i satisfies:

$$\begin{aligned} i &= \left\lceil np\left(1 - \frac{1-p}{k-np}\right)\right\rceil \\ &= 1 \\ &\quad [\text{since } 0 < np \leq 1 \text{ and } 1-p \in (0,1) \text{ and } k - np > 1-p]. \end{aligned}$$

For the probability $1/n < p \leq 1/(n-1)$, the index i satisfies:

$$\begin{aligned} i &= \left\lceil (n-1)p\left(\frac{\frac{k-1}{n-1} - p}{\frac{k}{n} - p}\right)\right\rceil \\ &= 1 \\ &\quad [\text{since } 0 < (n-1)p \leq 1 \text{ and } 0 < \tfrac{k-1}{n-1} - p \leq \tfrac{k}{n} - p]. \end{aligned}$$

Hence, the bound in (4.7) is tight in this case for all integer values of $k \geq 2$.

2. For $p > 1/(n-1)$, the index $i = 1$ when $k(np-1) \leq n(n-1)p^2$. This corresponds to all integer values $k \in [\lceil 1+(n-1)p \rceil, \lfloor n(n-1)p^2/(np-1) \rfloor]$. ◻

A specific instance to show the tightness of the Chebyshev bound is to set $p = 1/2$, $k = n$ and $n = 2^m - 1$ where m is an integer. Using m independent Bernoulli random variables it is then possible to construct n pairwise independent Bernoulli random variables (see Tao [55], Goemans [25], Pass and Spektor [42] for this construction). Proposition 4.3(a) includes this instance (set $\alpha = (n-1)/2$, $k = n$ and $n = 2^m - 1$). In addition, Proposition 4.3(a) identifies other values of p and k where the Chebyshev bound is tight. Proposition 4.3(b) also shows that the Schmidt, Siegel and Srinivasan bound is tight for identical marginals for small probability values ($p \leq 1/(n-1)$), for all values of k, except $k = 1$. We now provide a numerical illustration of the results in Theorem 4.1 and Proposition 4.3.

Example 4.4 (Identical marginals). In Table 5, we provide a numerical comparison of the bounds for $n = 11$ for a set of values of p and k. The instances in Table 5 cover all the conditions identified in Proposition 4.3 when the Chebyshev and Schmidt, Siegel and Srinivasan bounds are tight. The instances when the Chebyshev bound is tight correspond to (i) $p = 0.1$ (here $\alpha = 1$ and the Chebyshev bound is tight for $k = 2$ and $k = 11$), (ii) $p = 0.2$ (here $\alpha = 2$ and the Chebyshev bound is tight for $k = 3$ and $k = 11$) and (iii) $p = 0.5$ (here $\alpha = 5$ and the Chebyshev bound is tight for $k = 6$ and $k = 11$). The Schmidt, Siegel and Srinivasan bound is tight for the small values of $p = 0.01, 0.05, 0.10$ (which are less than or equal to $1/(n-1) = 0.1$) and for all values of k, except $k = 1$.

Table 5: Upper bound on probability of sum of random variables for $n = 11$. For each value of p and k, the table provides the tight bound in (4.1) followed by the Chebyshev bound (4.6) and the Schmidt, Siegel and Srinivasan bound (4.7). The underlined instances illustrate nontrivial cases when the upper bounds in either (4.6) or (4.7) are tight.

p/k	1	2	3	4	5	6	7	8	9	10	11
0.01	0.1090	0.00550	0.00184	0.00092	0.00055	0.00037	0.00027	0.00020	0.00016	0.00013	0.00010
	0.1208	0.02959	0.01288	0.00715	0.00454	0.00313	0.00229	0.00175	0.00138	0.00112	0.00092
	0.11000	0.00550	0.00184	0.00092	0.00055	0.00037	0.00027	0.00020	0.00016	0.00013	0.00010
0.05	0.5250	0.13750	0.04583	0.02292	0.01375	0.00917	0.00655	0.00491	0.00382	0.00306	0.00250
	0.7206	0.19905	0.08008	0.04205	0.02571	0.01729	0.01240	0.00933	0.00726	0.00582	0.00477
	0.5500	0.13750	0.04583	0.02292	0.01375	0.00917	0.00655	0.00491	0.00382	0.00306	0.00250
0.10	1	0.55000	0.18333	0.09167	0.05500	0.03667	0.02620	0.01965	0.01528	0.01223	0.01000
	1	0.55000	0.21522	0.10532	0.06112	0.03960	0.02766	0.02038	0.01562	0.01235	0.01000
	1	0.55000	0.18333	0.09167	0.05500	0.03667	0.02620	0.01965	0.01528	0.01223	0.01000
0.11	1	0.59950	0.22184	0.11092	0.06655	0.04437	0.03037	0.02170	0.01627	0.01266	0.01013
	1	0.63310	0.25156	0.12154	0.06975	0.04484	0.03113	0.02283	0.01744	0.01375	0.01112
	1	0.60500	0.22184	0.11092	0.06655	0.04437	0.03170	0.02377	0.01849	0.01479	0.01210
0.15	1	0.78750	0.41250	0.19584	0.09792	0.05875	0.03916	0.02798	0.02098	0.01632	0.01306
	1	0.91968	0.43489	0.20253	0.11109	0.06901	0.04672	0.03362	0.02531	0.01972	0.01579
	1	0.82500	0.41250	0.20625	0.12375	0.08250	0.05893	0.04419	0.03437	0.02750	0.02250
0.20	1	1	0.73334	0.33334	0.16667	0.10000	0.06667	0.04762	0.03572	0.02778	0.02223
	1	1	0.73334	0.35200	0.18334	0.10865	0.07097	0.04972	0.03667	0.02812	0.02223
	1	1	0.73334	0.36667	0.22000	0.14667	0.10477	0.07858	0.06112	0.04889	0.04000
0.50	1	1	1	1	1	0.91667	0.54167	0.29167	0.17500	0.11667	0.08334
	1	1	1	1	1	0.91667	0.55000	0.30556	0.18334	0.11957	0.08334
	1	1	1	1	1	0.91667	0.65477	0.49108	0.38195	0.30556	0.25000

It is also clear why the Schmidt, Siegel and Srinivasan bound is not tight for $k = 1$, since it just reduces to the Markov bound np and does not exploit the pairwise independence information. For $k = 1$, the tight bound from Theorem 4.1 is given by $np - (n-1)p^2$ (see Theorem 2.3 which reduces to the same bound for $k = 1$). For larger values of p above 0.1, such as $p = 0.11$ in the table, from Proposition 4.3(b), the Schmidt, Siegel and Srinivasan bound is tight for $k \in [\lceil 2.1 \rceil, \lfloor 6.33 \rfloor]$ which corresponds to $k \in [3, 6]$. This can be similarly verified for the other probabilities $p = 0.15, 0.2, 0.5$ in the table.

4.2. Tightness of ordered bounds in a special case. In this section, we provide an instance when two of the ordered bounds derived in Section 3 are shown to be tight. While the ordered bounds in Theorem 3.1 are not tight in general, the next proposition identifies a special case with almost identical marginals when the bounds of Schmidt, Siegel and Srinivasan in (3.1) and Boros and Prékopa in (3.2) are shown to be attained.

Proposition 4.5. Suppose the marginal probabilities equal $p \in (0, 1/(n-1)]$ for $n-1$ random variables and $q \in (0, 1)$ for one random variable. Then, the ordered

bounds in (3.1) and (3.2) are tight for the following three instances and are given by:
(4.8)

$$\overline{P}(n,k,p,q) = \begin{cases} \dfrac{\binom{n-1}{2}p^2}{\binom{k-1}{2}}, & k \geq 3,\ q \geq (n-2)p, & \text{(a)},\\ \dfrac{\binom{n-1}{2}p^2}{\binom{k-1}{2}}, & k \in \big[\lceil 2+(n-2)p/q \rceil,\ n\big],\ p \leq q < (n-2)p, & \text{(b)},\\ pq, & k = n,\ 0 < q < p, & \text{(c)}. \end{cases}$$

Proof. We first prove that the ordered bounds of Schmidt, Siegel and Srinivasan and Boros and Prékopa reduce to the bound in (4.8) in each of the three cases and then show that the bound is tight.

(1) Show reduction of ordered bounds to the bound in (4.8): Let $\overline{P}(n,k,p,q)$ represent the tightest upper bound when $n-1$ probabilities are p and one is q. It can be observed that the bound in (4.8) is non-trivial for the three instances since:

$$\begin{aligned} \frac{\binom{n-1}{2}p^2}{\binom{k-1}{2}} &= \frac{(n-1)p(n-2)p}{(k-1)(k-2)} < 1, \\ & \quad [\text{since } (n-2)p < (n-1)p \leq 1 \text{ and } k \geq 3 \text{ for cases (a) and (b)}], \\ pq &< 1, \\ & \quad [\text{since } q < p < 1 \text{ for case (c)}]. \end{aligned}$$

It is easy to verify that the ordered Schmidt, Siegel and Srinivasan bound in (3.1) reduces to the bound in (4.8) for a specific parameter r_2 in each of the three cases:

(4.9)
$$\begin{aligned} r_2 &= 1, && \text{cases (a) and (b)},\\ r_2 &= n-2, && \text{case (c)}. \end{aligned}$$

It can be similarly verified that the ordered Boros and Prékopa bound in (3.2) reduces to the bound in (4.8) with the following parameters r and i in each of the three cases:

(4.10)
$$\begin{aligned} r &= 1,\ i = 0, && \text{cases (a) and (b)},\\ r &= n-2,\ i = 0, && \text{case (c)}. \end{aligned}$$

The effectiveness of ordering is demonstrated by (4.9) and (4.10) in that the ordered bounds of Schmidt, Siegel and Srinivasan and Boros and Prékopa correspond to $r > 0$ while their unordered counterparts in (1.4) and (1.7) correspond to $r = 0$ (considering all n variables). The unordered bounds are thus strictly weaker than the ordered bounds which in turn are tight as proved in the next step.

(2) Prove tightness of the bound in (4.8) by constructing extremal distributions:

Consider the linear program to compute $\overline{P}(n,k,p,q)$ which can be written as:

$$
\begin{aligned}
\overline{P}(n,k,p,q) = \max \quad & \sum_{\boldsymbol{c}\in\{0,1\}^n:\sum_t c_t \geq k} \theta(\boldsymbol{c}) \\
\text{s.t} \quad & \sum_{\boldsymbol{c}\in\{0,1\}^n} \theta(\boldsymbol{c}) = 1, \\
& \sum_{\boldsymbol{c}\in\{0,1\}^n:c_i=1} \theta(\boldsymbol{c}) = p, && \forall i \in [n-1], \\
& \sum_{\boldsymbol{c}\in\{0,1\}^n:c_n=1} \theta(\boldsymbol{c}) = q, \\
& \sum_{\boldsymbol{c}\in\{0,1\}^n:c_i=1,c_j=1} \theta(\boldsymbol{c}) = p^2, && \forall (i,j) \in K_{n-1}, \\
& \sum_{\boldsymbol{c}\in\{0,1\}^n:c_i=1,c_n=1} \theta(\boldsymbol{c}) = pq, && \forall i \in [n-1], \\
& \theta(\boldsymbol{c}) \geq 0, && \forall \boldsymbol{c} \in \{0,1\}^n.
\end{aligned}
\tag{4.11}
$$

We now proceed to prove tightness of the bound in (4.8) for each of the three instances of the (n,k,p,q) tuple by constructing feasible distributions of (4.11) which attain the bound.

1. $\overline{P}(n,k,p,q) = \dfrac{\binom{n-1}{2}p^2}{\binom{k-1}{2}}$ (cases (a) and (b)):

   The following distribution attains the tight bound:

$$
\theta(\boldsymbol{c}) =
\begin{cases}
(1-q)(1-(n-1)p), & \text{if } \sum_{t\in[n]} c_t = 0, & (x), \\
p(1-q), & \text{if } \sum_{t\in[n-1]} c_t = 1, c_n = 0, & (y), \\
q(1-(n-1)p) + \frac{(n-1)(n-2)p^2}{(k-1)}, & \text{if } \sum_{t\in[n-1]} c_t = 0, c_n = 1, & (z), \\
p(q - \frac{n-2}{k-2}p), & \text{if } \sum_{t\in[n-1]} c_t = 1, c_n = 1, & (u), \\
\frac{p^2}{\binom{n-3}{k-3}}, & \text{if } \sum_{t\in[n-1]} c_t = k-1, c_n = 1, & (v).
\end{cases}
\tag{4.12}
$$

   We use symbols x, y, z, u, v to denote the probability of the associated scenarios in (4.12). The constraints in (4.11) reduce to:

$$
\begin{aligned}
&\tbinom{n-2}{k-2}v + u + y = p \\
&\tbinom{n-1}{k-1}v + (n-1)u + z = q \\
&\tbinom{n-3}{k-3}v = p^2 \\
&\tbinom{n-2}{k-2}v + u = pq \\
&x + y + z + u + v = 1,
\end{aligned}
$$

   and using x, y, z, u, v from (4.12), it can be easily verified that all of the above constraints are satisfied. The non-negativity constraints for y, v are satisfied while $x \geq 0,\ z \geq 0$ is satisfied since $(n-1)p \leq 1$. Remaining case is u, for which we have:

$$
\begin{array}{llcl}
\text{case (a):} & u & = & p(q - \frac{n-2}{k-2}p) \\
 & & \geq & p(q - \frac{n-2}{3-2}p) \\
 & & & [\text{since } k \geq 3] \\
 & & = & p(q - (n-2)p) \\
 & & & [\text{since } q > (n-2)p] \\
 & & \geq & 0 \\
\text{case (b):} & u & = & p(q - \frac{n-2}{k-2}p) \\
 & & \geq & p(q - \frac{k-2}{k-2}q) \\
 & & & [\text{since } k \geq 2 + (n-2)p/q] \\
 & & = & 0.
\end{array}
$$

The only support points contributing to the objective function are the first set of $\binom{n-1}{k-1}$ scenarios, and so we have $\overline{P}(n,k,p,q) = \binom{n-1}{k-1}p^2/\binom{n-3}{k-3} = \binom{n-1}{2}p^2/\binom{k-1}{2}$.

2. $\overline{P}(n,k,p,q) = pq$ (case (c)):
The following distribution attains the tight bound pq:
(4.13)

$$
\theta(\boldsymbol{c}) = \begin{cases}
(1-p)(1-(n-2)p-q), & \text{if } \sum_{t\in[n]} c_t = 0, \quad (x), \\
p(1-p), & \text{if } \sum_{t\in[n-1]} c_t = 1, c_n = 0, \quad (y), \\
q(1-p), & \text{if } \sum_{t\in[n-1]} c_t = 0, c_n = 1, \quad (z), \\
p(p-q), & \text{if } \sum_{t\in[n-1]} c_t = n-1, c_n = 0, \quad (u), \\
pq, & \text{if } \sum_{t\in[n]} c_t = n, \quad (v).
\end{cases}
$$

The constraints in (4.11) reduce to:

$$
\begin{aligned}
& y + u + v = p \\
& z + v = q \\
& u + v = p^2 \\
& v = pq \\
& x + y + z + u + v = 1,
\end{aligned}
$$

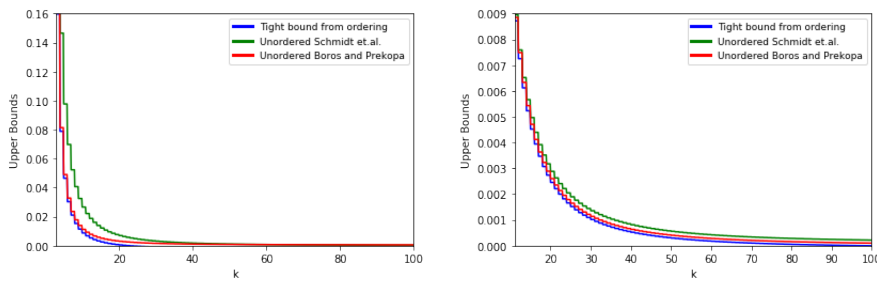
and using x, y, z, u, v from (4.13), it can be easily verified that all of the above constraints are satisfied. The non-negativity contraints for y, z, u, v are satisfied by $0 < q \leq p \leq 1$ while for x, we have:

$$
\begin{array}{lcl}
x & = & (1-p)(1-(n-2)p-q) \\
 & \geq & (1-p)(1-(n-2)p-p) \\
 & & [\text{since } q < p] \\
 & = & (1-p)(1-(n-1)p) \\
 & \geq & 0 \\
 & & [\text{since } (n-1)p \leq 1].
\end{array}
$$

The distribution in (4.13) attains the bound pq.

We have thus constructed two feasible probability distributions in (4.12) and (4.13) which attain the bound in (4.8) in each of the three instances defined by the (n, k, p, q) tuple. Hence the parameters r_2, r in (4.9) and (4.10) defined for each of the three cases must be the minimizers which exactly reduce the ordered bounds in (3.1) and (3.2) to the tight bound in (4.8). ∎

Example 4.6. This example demonstrates the usefulness of Proposition 4.5 when $n = 100$ and $p = 0.01$ where $(n-1)p \leq 1$. It compares the tight bounds computed from (4.8) with the unordered bounds of Schmidt, Siegel and Srinivasan from (1.4) and that of Boros and Prékopa from (1.7).



(a) $q = 0.99,\ q \geq (n-2)p,\ k \geq 3$

(b) $q = 0.1,\ p \leq q < (n-2)p,\ k \geq 12$

Fig. 6: Comparison of unordered bounds with tight bound when $n = 100,\ p = 0.01$

Figure 6a plots the two unordered bounds along with the tight bound when $q = 0.99$ (case (a) of Proposition 4.5), where the tight bound is valid for all k in $[3, n]$, while Figure 6b compares the bounds when $q = 0.1$ (case (b) of Proposition 4.5) for $k \geq 12$ as the tight bound is valid when $k \geq \lceil 2 + (n-2)p/q \rceil = \lceil 11.8 \rceil = 12$. The unordered Boros and Prékopa bound is much tighter than the unordered Schmidt, Siegel and Srinivasan bound in both figures. Hence, Figure 6 demonstrates that with ordering, the relative improvement of the Schmidt, Siegel and Srinivasan bound is much better than that of the Boros and Prékopa bound although both the ordered bounds reduce to the tight bound in (4.8).

5. Conclusion. In this paper we have provided results towards finding tight probability bounds for the sum of n pairwise independent random variables adding up to at least an integer k. In Section 2, we first established with Lemma 2.1 that a feasible correlated distribution of a Bernoulli random vector $\tilde{\boldsymbol{c}}$ with an arbitrary univariate probability vector $\boldsymbol{p} \in [0, 1]^n$ and transformed bivariate probabilities $p_i p_j / p$ where $\max_i p_i \leq p \leq 1$, always exists (this result was then extended to prove the existence of an alternate correlated Bernoulli random vector in Corollary 2.2). Theorem 2.3 then established that with pairwise independence, the Hunter [28] and Worsley [59] bound is tight for any $\boldsymbol{p} \in [0, 1]^n$, which, to the best of our knowledge, has not been shown thus far in the literature dedicated to this topic. In fact, paraphrasing from Boros [7] (Section 1.2), "As far as we know, in spite of the several studies dedicated to this problem, the complexity status of this problem, for feasible input, seems to be still open even for bivariate probabilities". With pairwise independent random variables, feasibility is guaranteed and Theorem 2.3 shows that the tightest upper bound is computable in polynomial time (in fact in a simple closed-form), thus providing a partial positive answer towards this question. The proof included the explicit construction of an extremal distribution (though not unique) in Table 2, that attains this bound. We then showed in Proposition 2.5 that the ratio of the Boole union bound and the pairwise independent bound is upper bounded by 4/3 and that this bound is attained. Applications of the result in correlation gap analysis and bottleneck optimization (in the distributionally robust optimization context) were discussed in examples 2.6 and 2.7. The tight upper bound on the union probability was then used to derive a closed-form expression for the tight lower bound on the intersection probability in Corollary 2.9, which, to the best of our knowledge, appears to be unknown in the literature. In Section 3, for $k \geq 2$, we proposed new bounds exploiting ordering of the probabilities (which are at least as good as the unordered bounds) and argued that the ordered Boros and Prékopa bound must be at least as good as the other two ordered bounds proposed in Theorem 3.1. To the best of our knowledge, this idea of ordering has not been exploited thus far to tighten probability bounds for pairwise independent random variables. We then showed in Section 3.2 that the ordered bounds can be further tightened by using the tight bound for $k = 1$ from Theorem 2.3. Numerical examples in Section 3.3 then demonstrated that while the Boros and Prékopa bound is uniformly the best performing of the three ordered bounds, the Schmidt, Siegel and Srinivasan bound shows the best improvement with ordering, in the examples considered. Section 4 provided instances when the unordered and ordered bounds are tight. In Section 4.1, for the special case of identical probabilities $p \in [0, 1]$ and any $k \in [n]$, we used a constructive proof exploiting the symmetry in the problem, to identify the best upper bound $\overline{P}(n, k, p)$ in closed-form and a corresponding extremal distribution. This result was further extended to provide tight bounds (not necessarily closed-form) for more general t-wise independent identical variables in Corollary 4.2. We then demonstrated the usefulness of this result by identifying instances when the existing unordered bounds are tight. Section 4.2 demonstrated the usefulness of the ordered bounds by identifying an instance with $n - 1$ identical probabilities (along with additional conditions on the identical probability and k), when the ordered bounds are tight.

We believe several interesting research questions arise from this work, two of which we list below:

(a) To the best of our knowledge, the computational complexity of evaluating (or approximating) the bound $\overline{P}(n, k, \boldsymbol{p})$ for general n, k and $\boldsymbol{p} \in [0, 1]^n$ is still unresolved. While we provide the answer in closed-form for $k = 1$, a natural question that arises is whether the tight bounds for general $k \geq 2$ with pairwise independent random variables are efficiently computable (or efficient to approximate)? We leave this for future research.

(b) The upper bound of 4/3 in Section 2.2 is derived for the ratio between the maximum probability for the union of arbitrarily dependent events and the probability of the union of pairwise independent events. We conjecture this upper bound is valid for the expected value of all non-decreasing, nonnegative submodular functions (of which the probability of the union is a special case) and leave it as an open question.

Acknowledgments. We would like to thank the Associate Editor Prasad Tetali and the reviewers for their careful reading of the paper and useful inputs.

REFERENCES

[1] S. Agrawal, Y. Ding, A. Saberi, and Y. Ye, *Price of correlations in stochastic optimization*, Operations Research, 1 (2012), pp. 150–162.
[2] L. Babai, *Entropy versus pairwise independence*, Available at http://people.cs.uchicago.edu/∼laci/papers/13augEntropy.pdf, (2013).
[3] I. Benjamini, O. Gurel-Gurevich, and R. Peled, *On k-wise independent distributions and boolean functions*, Working Paper, In: arXiv preprint:1201.3261, (2012).
[4] S. Bernstein, *Theory of probability, Moscow-Leningrad*, (1946).
[5] G. Boole, *The Laws of Thought (1916 reprint)*, 1854.
[6] E. Boros and A. Prékopa, *Closed form two-sided bounds for probabilities that at least r and exactly r out of n events occur*, Mathematics of Operations Research, 14 (1989), pp. 317–342.
[7] E. Boros, A. Scozzari, F. Tardella, and P. Veneziani, *Polynomially computable bounds for the probability of the union of events*, Mathematics of Operations Research, 39 (2014), pp. 1311–1329.
[8] G. Calinescu, C. Chekuri, M. Pál, and J. Vondrák, *Maximizing a monotone submodular function subject to a matroid constraint*, SIAM Journal on Computing, 40 (2007), pp. 1740–1766.
[9] N. R. Chaganty and H. Joe, *Range of correlation matrices for dependent bernoulli random variables*, Biometrika, 1 (2006), pp. 197–206.
[10] P. Chebyshev, *Des valeurs moyennes*, Journal de Mathématiques Pures et Appliquées, 2 (1867), pp. 177–184.
[11] H. Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Annals of Mathematical Statistics, 23 (1952), pp. 493–509.
[12] D. A. Dawson and D. Sankoff, *An inequality for probabilities*, Proceedings of the American Mathematical Society, 18 (1967), pp. 504–507.
[13] D. de Caen, *A lower bound on the probability of a union*, Discrete Mathematics, 169 (1997), pp. 217–220.
[14] K. Dohmen and P. Tittmann, *Improved Bonferroni inequalities and binomially bounded functions*, Electronic Notes in Discrete Mathematics, 28 (2007), pp. 91–93.
[15] J. Edmonds and D. R. Fulkerson, *Bottleneck extrema*, Journal of Combinatorial Theory, 3 (1970), pp. 299–306.
[16] L. J. Emrich and M. R. Piedmonte, *A method for generating high-dimensional multivariate binary variates*, The American Statistician, 45 (1991), pp. 302–304.
[17] W. Feller, *Non-Markovian processes with the semigroup property*, The Annals of Mathematical Statistics, 30 (1959), pp. 1252–1253.
[18] W. Feller, *An Introduction to Probability Theory and Its Applications: Volume I*, Wiley Series in Probability and Mathematical Statistics, 3 ed., 1968.
[19] M. Fréchet, *Généralisation du théoreme des probabilités totales*, Fundamenta mMthematicae, 1 (1935), pp. 379–387.
[20] J. Galambos, *Methods for proving Bonferroni type inequalities*, Journal of the London Mathematical Society, 2 (1975), pp. 561–564.
[21] J. Galambos, *Bonferroni inequalities*, The Annals of Probability, (1977), pp. 577–581.
[22] B. Garnett, *Small deviations of sums of independent random variables*, Journal of Combinatorial Theory, Series A, 169 (2020), pp. 105–119.
[23] D. Gavinsky and P. Pudlák, *On the joint entropy of d-wise-independent variables*, Commentationes Mathematicae Universitatis Carolinae, (2016), pp. 333–343.
[24] S. Geisser and N. Mantel, *Pairwise independence of jointly dependent variables*, The Annals of Mathematical Statistics, 33 (1962), pp. 290–291.
[25] M. Goemans, *Chernoff bounds, and some applications*, Lecture Notes, MIT, (2015).
[26] T. Hailperin, *Best possible inequalities for the probability of a logical function of events*, The American Mathematical Monthly, 72 (1965), pp. 343–359.
[27] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, Journal of the American Statistical Association, 58 (1963), pp. 13–30.
[28] D. Hunter, *An upper bound for the probability of a union*, Journal of Applied Probability, 13 (1976), pp. 597–603.
[29] A. Joffe, *On a set of almost deterministic k-independent random variables*, The Annals of Probability, 2 (1974), pp. 161–162.
[30] H. Karloff and Y. Mansour, *On construction of k-wise independent random variables*, in Proceedings of the 26th Annual ACM Symposium on Theory of Computing, 1994, pp. 564–573.
[31] D. Koller and N. Meggido, *Construcing small sample spaces satisfying given constraints*, SIAM Journal on Discrete Mathematics, 7 (1994), pp. 260–274.

[32] E. G. Kounias, *Bounds for the probability of a union, with applications*, The Annals of Mathematical Statistics, 39 (1968), pp. 2154–2158.
[33] S. Kounias and J. Marin, *Best linear Bonferroni bounds*, SIAM Journal on Applied Mathematics, 30 (1976), pp. 307–323.
[34] H. Kuai, F. Alajaji, and G. Takahara, *A lower bound on the probability of a finite union of events*, Discrete Mathematics, 215 (2000), pp. 147–158.
[35] S. M. Kwerel, *Most stringent bounds on aggregated probabilities of partially specified dependent probability systems*, Journal of the American Statistical Association, 70 (1975b), pp. 472–479.
[36] H. O. Lancaster, *Pairwise statistical independence*, The Annals of Mathematical Statistics, 36 (1965), pp. 1313–1317.
[37] M. Luby and A. Widgerson, *Pairwise independence and derandomization*, Foundations and Trends in Theoretical Computer Science, 1 (2005), pp. 239–201.
[38] A. D. Lunn and S. J. Davies, *A note on generating correlated binary variables*, Biometrika, 85 (1998), pp. 487–490.
[39] W. Maurer, *Bivalent trees and forests or upper bounds for the probability of a union revisited*, Discrete Applied Mathematics, 6 (1983), pp. 157–171.
[40] T. F. Móri and J. G. Székely, *A note on the background of several Bonferroni–Galambos-type inequalities*, Journal of Applied Probability, 22 (1985), pp. 836–843.
[41] G. L. O'Brien, *Pairwise independent random variables*, The Annals of Probability, 8 (1980), pp. 170–175.
[42] B. Pass and S. Spektor, *On Khintchine type inequalities for k-wise independent Rademacher random variables*, Statistics & Probability Letters, 132 (2018), pp. 35–39.
[43] R. Peled, A. Yadin, and A. Yehudayoff, *The maximal probability that k-wise independent bits are all 1*, Random Structures & Algorithms, 38 (2011), pp. 502–525.
[44] I. Pinelis, *Exact lower bound on an 'exactly one'probability*, Bulletin of the Australian Mathematical Society, 104 (2021), pp. 330–336.
[45] I. Pitowsky, *Correlation polytopes: Their geometry and complexity*, Mathematical Programming, 50 (1991), pp. 395–414.
[46] O. Platz, *A sharp upper probability bound for the occurrence of at least m out of n events*, Journal of Applied Probability, 22 (1985), pp. 978–981.
[47] A. Prékopa, *Boole-Bonferroni inequalities and linear programming*, Operations Research, 36 (1988), pp. 145–162.
[48] A. Prékopa, *Sharp bounds on probabilities using linear programming*, Operations Research, 38 (1990), pp. 227–239.
[49] A. Prékopa and L. Gao, *Bounding the probability of the union of events by aggregation and disaggregation in linear programs*, Discrete Applied Mathematics, 145 (2005), pp. 444–454.
[50] B. F. Qaqish, *A family of multivariate binary distributions for simulating correlated binary variables with specified marginal means and correlations*, Biometrika, 90 (2003), pp. 455–463.
[51] F. Qiu, S. Ahmed, and S. S. Dey, *Strengthened bounds for the probability of k-out-of-n events*, Discrete Applied Mathematics, 198 (2016), pp. 232–240.
[52] B. Rüger, *Das maximale signifikanzniveau des Tests: "Lehne H_0 ab, wennk untern gegebenen tests zur ablehnung führen"*, Metrika, 25 (1978), pp. 171–178.
[53] Y. S. Sathe, M. 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 (1980), pp. 1127–1132.
[54] J. Schmidt, A. Siegel, and A. Srinivasan, *Chernoff–Hoeffding bounds for applications with limited independence*, SIAM Journal on Discrete Mathematics, 8 (1995), pp. 223–250.
[55] T. Tao, *Topics in random matrix theory*, vol. 132, Graduate Studies in Mathematics, American Mathematical Society, 2012.
[56] P. Veneziani, *Graph-based upper bounds for the probability of the union of events*, The Electronic Journal of Combinatorics, 15 (2008).
[57] P. Veneziani, *Optimality conditions for Hunter's bound*, Discrete Mathematics, 308 (2008), pp. 6009–6014.
[58] B. Vizvári, *New upper bounds on the probability of events based on graph structures*, Mathematical Inequalities and Applications, 10 (2007), p. 217.
[59] K. J. Worsley, *An improved Bonferroni inequality and applications*, Biometrika, 69 (1982), pp. 297–302.
[60] W. Xie, J. Zhang, and S. Ahmed, *Distributionally robust bottleneck combinatorial problems: uncertainty quantification and robust decision making*, To appear in Mathematical Programming, (2021).
[61] J. Yang, F. Alajaji, and G. Takahara, *Lower bounds on the probability of a finite union of*

events, SIAM Journal on Discrete Mathematics, 30 (2016), pp. 1437–1452.
[62] K. Yoda and A. Prékopa, *Improved bounds on the probability of the union of events some of whose intersections are empty*, Operations Research Letters, 44 (2016), pp. 39–43.