Tight Probability Bounds with Pairwise Independence

Arjun Kodagehalli Ramachandra∗ Karthik Natarajan†

April 2020

Abstract

Probability bounds on the sum of \( n \) pairwise independent Bernoulli random variables exceeding an integer \( k \) have been extensively studied in the literature. However, these bounds are not tight in general. In this paper, we provide three results towards finding tight probability bounds for the sum of Bernoulli random variables with pairwise independence. First, for the sum of identically distributed random variables, we provide the tightest upper bound on the probability that the sum exceeds \( k \), for any \( k \). Second, when the distributions are non-identical, we provide the tightest upper bound on the probability that the sum exceeds \( k = 1 \). Lastly, we provide new upper bounds on the probability that the sum exceeds \( k \), for any \( k \geq 2 \), by exploiting ordering of the probabilities. These bounds improve on existing bounds and under certain conditions are shown to be tight. In each case, we prove tightness by identifying a distribution that attains the bound. Numerical examples are provided to illustrate when the bounds provide significant improvements.

1 Introduction

Pairwise independent Bernoulli random variables have been extensively studied by researchers in various communities including, but not limited to, probability and statistics, computer science and optimization. At the core of this analysis is the observation that while mutually independent random variables are pairwise independent, the reverse is not true. Feller [1968] attributes S. N. Bernstein with identifying one of the earliest examples of \( n = 3 \) random variables which are pairwise independent, but not mutually independent. For general \( n \), constructions of pairwise independent Bernoulli random variables can be found in Geisser and Mantel [1962], Karloff and Mansour [1994], Koller and Meggido [1994], pairwise independent discrete random variables in Feller [1959], Lancaster [1965], Joffe [1974], O’Brien [1980] and pairwise independent normal random variables in Geisser and Mantel [1962]. One of the motivations in studying the constructions of pairwise independent random variables 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. This has important ramifications in efficiently derandomizing

∗Engineering Systems and Design, Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372. Email: arjun_ramachandra@mymail.sutd.edu.sg
†Engineering Systems and Design, Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372. Email: karthik_natarajan@sutd.edu.sg
algorithms for several NP-hard optimization problems (see Luby and Widgerson [2005] and the references therein for constructions and applications with pairwise independent and more generally $t$-wise independent random variables).

In this paper, we are interested in computing probability bounds on the sum of pairwise independent Bernoulli random variables. Let $[n] = \{1, 2, \ldots, n\}$. For integers $i < j$, we use $[i, j]$ to denote $\{i, i + 1, \ldots, j - 1, j\}$. Let $I_n = \{(i, j) : 1 \leq i < j \leq n\}$ denote the set of pairwise indices in $[n]$. Consider a Bernoulli random vector $\tilde{c} = (\tilde{c}_1, \ldots, \tilde{c}_n)$ with marginal probabilities given by $p_i = P(\tilde{c}_i = 1)$ for $i \in [n]$. Let $\theta_{pw}$ denote the set of joint distributions supported on $\{0, 1\}^n$ consistent with the given marginal probabilities and pairwise independence:

$$\theta_{pw} = \left\{ \theta((0, 1)^n) \mid \forall i \in [n], \ P_{\theta}(\tilde{c}_i = 1) = p_i, \ \forall (i, j) \in I_n \right\}.$$ 

This set of distributions is clearly nonempty for any $p = (p_1, \ldots, p_n) \in (0, 1)^n$, since the mutually independent distribution is always feasible. We are interested in finding the maximum probability that the sum of the random variables exceeds an integer $k \in [n]$ for distributions in the set $\theta_{pw}$. Denoting this tightest upper bound by $P(n, k, p)$, we have

$$P(n, k, p) = \max_{\theta \in \theta_{pw}} P_{\theta} \left( \sum_{i=1}^{n} \tilde{c}_i \geq k \right).$$

Two upper bounds that have been proposed for this problem are the following:

(a) Chebyshev [1867] bound: To apply the Chebyshev one-sided tail probability bound, we make use of the first and the second moment of the sum of the random variables. Since the Bernoulli random variables are pairwise independent (or equivalently uncorrelated), applying the Chebyshev bound gives:

$$P(n, k, p) \leq \begin{cases} 1, & k < \sum_{i=1}^{n} p_i \\ \sum_{i=1}^{n} p_i (1 - p_i) / \left( \sum_{i=1}^{n} p_i (1 - p_i) + (k - \sum_{i=1}^{n} p_i)^2 \right), & \sum_{i=1}^{n} p_i \leq k \leq n. \end{cases}$$ (1.1)

(b) Schmidt et al. [1995] 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 Chernoff-Hoeffding bounds (see Chernoff [1952], Hoeffding [1963]) 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:

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

At the crux of the analysis is the observation that all the higher moments of the sum of Bernoulli random variables can be generated from linear combinations of the expected value of multilinear
polynomials of the random variables. The construction of their bound makes use of the equality:

\[
\left( \sum_{j=1}^{n} c_j \right) = S_j(c), \quad \forall c \in \{0, 1\}^n, \forall j \in [0, n],
\]

where \( S_0(c) = 1 \) and \( \binom{r}{s} = r!/(s!(r-s)!) \) for a pair of integers \( r \geq s \geq 0 \). The bound derived in Schmidt et al. [1995] for pairwise independent random variables is given by (see Theorem 7, part (II) on page 239):

\[
\mathcal{P}(n, k, p) \leq \min \left( 1, \sum_{i=1}^{n} p_i, \sum_{1 \leq i < j \leq n} p_i p_j \right).
\]

While both the bounds in (1.1) and (1.3) are very useful, neither of them are tight, in general, for the sum of pairwise independent Bernoulli random variables. We note that both the bounds can be expressed in terms of the first two aggregated (or equivalently binomial) moments for the sum of pairwise independent random variables, \( S_1(p) = \mathbb{E}[S_1(\tilde{c})] = \sum_i p_i \) and \( S_2(p) = \mathbb{E}[S_2(\tilde{c})] = \sum_{(i,j) \in I_n} p_i p_j \). The use of aggregated moments in developing probability bounds for sums of Bernoulli random variables has been analyzed in a series of papers by Prékopa [1988, 1990], Boros and Prékopa [1989], Prékopa and Gao [2005], Boros et al. [2014]. These papers use polynomial sized linear optimization formulations to find probability bounds and in some cases, closed form bounds are derived. Boros and Prékopa [1989] derived the tightest upper bound on \( \mathbb{P}(\tilde{\xi} \geq k) \) over all distributions \( v \) of an integer random variable \( \tilde{\xi} \) supported on \([0, n]\) which is assumed to lie in a set of distributions given by:

\[
\left\{ v([0, n]) \mid \mathbb{E}_v \left[ \left( \frac{\tilde{\xi}}{j} \right) \right] = S_j, j = 1, 2 \right\}.
\]

The upper bound is a closed form expression as follows:

\[
\mathbb{P}\left( \sum_{i=1}^{n} \tilde{c}_i \geq k \right) \leq \begin{cases} 1, & k < \frac{(n-1)S_1 - 2S_2}{n - S_1}, \\ \frac{(k + n - 1)S_1 - 2S_2}{kn}, & \frac{(n-1)S_1 - 2S_2}{n - S_1} \leq k < 1 + \frac{2S_2}{S_1}, \\ \frac{(i-1)(i-2S_1) + 2S_2}{(k-i)^2 + (k-i)}, & k \geq 1 + \frac{2S_2}{S_1}, \quad i = \left\lceil \frac{(k-1)S_1 - 2S_2}{k - S_1} \right\rceil, \end{cases}
\]

where the ceiling function \( \lceil x \rceil \) maps \( x \) to the smallest integer greater than or equal to \( x \). The bound in (1.4) can be applied to pairwise independent variables with \( \tilde{\xi} = \sum_{i \in [n]} \tilde{c}_i \). This brings us to the main results in this paper. In Section 2, we identify the tightest possible bound in a closed form when the marginals are identical. In Section 3, with non-identical marginals, we provide for \( k = 1 \), the tightest bound \( \mathcal{P}(n, 1, p) \) in closed form and then develop improved bounds for \( k \geq 2 \). Connections to existing results are discussed and numerical examples provided. In the rest of the paper, we will refer to the

---

1While the statement in the theorem in Schmidt et al. [1995] is for \( k > \sum_i p_i \), it is straightforward to see that the analysis would generalize to the form here. Schmidt et al. [1995] also consider \( t \)-wise independent random variables, which we do not focus on here.
three bounds in (1.1), (1.3) and (1.4) as the (a) Chebyshev, (b) Schmidt, Siegel and Srinivasan and (c) Boros and Prekopa bound respectively.

2 Identical marginals

The first theorem provides the tight bound for identical marginals. The proof of the theorem is based on showing an equivalence of the exponential sized linear program which computes the exact bound for identical marginals with a polynomial sized linear program in a new set of decision variables. The latter linear program admits a closed form solution.

**Theorem 2.1.** Assume \(p_i = p \in (0, 1)\) for \(i \in [n]\). Let \(\mathcal{P}(n, k, p)\) represent the tightest upper bound on the sum of \(n\) pairwise independent Bernoulli random variables exceeding an integer \(k \in [n]\). Then,

\[
\mathcal{P}(n, k, p) = \begin{cases} 
1, & k < (n - 1)p \\
\frac{(n - 1)(1 - p) + k}{k}, & (n - 1)p \leq k < 1 + (n - 1)p \\
\frac{n(n - 1)p^2 + (i - 1)(i - 2np)}{(k - i)^2 + (k - i)}, & k \geq 1 + (n - 1)p, \ i = \left\lceil np(k - 1 - (n - 1)p) - k^{-1}np \right\rceil 
\end{cases}
\]

(2.1)

**Proof.** Let \(\mathcal{C} = \{0, 1\}^n\). The tightest upper bound \(\mathcal{P}(n, k, p)\) is the optimal value of the linear program:

\[
\mathcal{P}(n, k, p) = \max \sum_{c \in \mathcal{C}, \sum_i c_i \geq k} \theta(c) \\
\text{s.t.} \sum_{c \in \mathcal{C}} \theta(c) = 1 \\
\sum_{c \in \mathcal{C}, c_i = 1} \theta(c) = p, \ \forall i \in [n] \\
\sum_{c \in \mathcal{C}, c_i = 1, c_j = 1} \theta(c) = p^2, \ \forall (i, j) \in I_n \\
\theta(c) \geq 0, \ \forall c \in \mathcal{C},
\]

(2.2)

where the decision variables are the joint probabilities \(\theta(c) = \mathbb{P}(\bar{c} = c)\) for \(c \in \mathcal{C}\). This linear program has an exponential number of decision variables with \(|\mathcal{C}| = 2^n\). Consider the following linear program
in $n + 1$ variables which provides an upper bound on $\mathcal{P}(n, k, p)$:

$$BP(n, k, p) = \max \sum_{\ell=k}^{n} v_\ell$$

s.t. $\sum_{\ell=0}^{n} v_\ell = 1$

$\sum_{\ell=1}^{n} \ell v_\ell = np$

$\sum_{\ell=2}^{n} \left( \frac{\ell}{2} \right) v_\ell = \left( \frac{n}{2} \right) p^2$

$v_\ell \geq 0, \quad \forall \ell \in [0, n],

(2.3)$

where the decision variables are the probabilities $v_\ell = \mathbb{P}(\sum_{i=1}^{n} \tilde{c}_i = \ell)$ for $\ell \in [0, n]$. Linear programs of the form (2.3) have been studied in Boros and Prékopa [1989] in the context of aggregated binomial moment problems. As we shall see, these two formulations are equivalent with identical pairwise independent random variables.

**Step (1):** $\mathcal{P}(n, k, p) \leq BP(n, k, p)$

Given a feasible solution to (2.2) denoted by $\theta$, construct a feasible solution to the linear program (2.3) as:

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

By taking expectations on both sides of the equality (1.2), we get:

$$\sum_{l=j}^{n} \binom{l}{j} \mathbb{P} \left( \sum_{i=1}^{n} \tilde{c}_i = l \right) = \mathbb{E} \left[ S_j(\tilde{c}) \right], \quad \forall j \in [0, n].$$

Applying it for $j = 0, 1, 2$, we get the three equality constraints in (2.3):

$$\sum_{\ell=0}^{n} v_\ell = 1$$

$$\sum_{\ell=1}^{n} \ell v_\ell = \mathbb{E} \left[ \sum_{i=1}^{n} \tilde{c}_i \right] = np$$

$$\sum_{\ell=2}^{n} \left( \frac{\ell}{2} \right) v_\ell = \mathbb{E} \left[ \sum_{(i,j) \in I_n} \tilde{c}_i \tilde{c}_j \right] = n(n-1)p^2/2.$$

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

$$\sum_{\ell=k}^{n} v_\ell = \sum_{\ell=k}^{n} \sum_{c \in \mathcal{C} : \sum_i c_i = \ell} \theta(c) = \sum_{c \in \mathcal{C} : \sum_i c_i \geq k} \theta(c).$$
Hence, \( P(n, k, p) \leq BP(n, k, p) \).

**Step (2):** \( P(n, k, p) \geq BP(n, k, p) \)

Given an optimal solution to (2.3) denoted by \( v \), construct a feasible solution to the linear program (2.2) by distributing \( v_\ell \) equally among all the realizations in \( C \) with exactly \( \ell \) ones:

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

The first constraint in (2.2) is satisfied since:

\[
\sum_{c \in C} \theta(c) = \sum_{\ell=0}^{n} \sum_{c \in C : \sum_{i=1}^{n} c_i = \ell} \frac{v_\ell}{\binom{n}{\ell}} \quad \text{[since \( |C : \sum_{i=1}^{n} c_i = \ell| = \binom{n}{\ell} \)]}
\]

\[
= \sum_{\ell=0}^{n} v_\ell = 1.
\]

The second constraint in (2.2) is satisfied since:

\[
\sum_{c \in C : c_j = 1} \theta(c) = \sum_{\ell=1}^{n} \frac{v_\ell}{\binom{n}{\ell}} \binom{n-1}{\ell-1} \quad \text{[since \( |C : \sum_{i=1}^{n} c_i = \ell, c_j = 1| = \binom{n-1}{\ell-1} \)]}
\]

\[
= \sum_{\ell=1}^{n} \frac{\ell v_\ell}{n} = p.
\]

The third constraint in (2.2) satisfied since:

\[
\sum_{c \in C : c_i = 1, c_j = 1} \theta(c) = \sum_{\ell=2}^{n} \frac{v_\ell}{\binom{n}{\ell}} \binom{n-2}{\ell-2} \quad \text{[since \( |C : \sum_{i=1}^{n} c_i = \ell, c_i = 1, c_j = 1| = \binom{n-2}{\ell-2} \)]}
\]

\[
= \frac{2}{n(n-1)} \sum_{\ell=2}^{n} \binom{\ell}{2} v_\ell = p^2.
\]

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

\[
\sum_{c \in C : \sum_{i=1}^{n} c_i \geq k} \theta(c) = \sum_{\ell=k}^{n} \sum_{c \in C : \sum_{i=1}^{n} c_i = \ell} \theta(c)
\]

\[
= \sum_{\ell=k}^{n} v_\ell = BP(n, k, p).
\]
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 Prekopa bound in (1.4) (the bound $BP(n, k, p)$ is also derived in the work of Sathe et al. [1980], though tightness of the bound is not shown there). It is also 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:

$$\theta(c) = \begin{cases} 
\frac{(1-p)(j - (n-1)p)}{\binom{n-1}{j-1}}, & \text{if } \sum_{t=1}^{n} c_t = j - 1 \\
\frac{(1-p)(1 + (n-1)p - j)}{\binom{n-1}{j}}, & \text{if } \sum_{t=1}^{n} c_t = j \\
n(n-1)p^2 + (j-1)(j-2np) \quad \frac{1}{(n-j)^2 + (n-j)}, & \text{if } \sum_{t=1}^{n} c_t = n,
\end{cases}$$

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

(b) The probabilities are given as:

$$\theta(c) = \begin{cases} 
\frac{1-p}{k} \left( (n-1)p \right), & \text{if } \sum_{t=1}^{n} c_t = 0 \\
p(1-p) \binom{n-2}{k-2}, & \text{if } \sum_{t=1}^{n} c_t = k \\
p(n-1)p - (k-1)) \quad \frac{1}{n-k}, & \text{if } \sum_{t=1}^{n} c_t = n,
\end{cases}$$

where all other support points have zero probability.

(c) The probabilities are given as:

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

where all other support points have zero probability and the index $i$ is evaluated as stated in equation (2.1)(c).

Connection to existing results: In related work, Benjamini et al. [2012] and Peled et al. [2011] derived probability bounds (not necessarily tight) on the sum of $t$-wise independent Bernoulli random variables with identical probabilities (as a special case, pairwise independent random variables are studied in these papers). For the specific case, where all the random variables take a value of 1 (this corresponds to $k = n$ in case (c)), the tight bound is provided in these works by making a connection...
with the Boros and Prekopa bound in (1.4). Recent work by Garnett [2020] provides the tight upper bound on the probability that the sum of pairwise independent Bernoulli random variables with identical marginals exceeds the mean by a small amount. This corresponds to case (b). Theorem 2.1 provides the equivalence for all values of \((n,k,p)\).

**Application:** We next discuss an application of Theorem 2.1. While the Boros and Prekopa bound provides the tightest upper bound with identical marginals, the formula is more complicated than the Chebyshev bound which reduces to:

\[
\Pr(n,k,p) \leq \begin{cases} 
1, & k < np \\
np(1-p)/(np(1-p)+(k-np)^2), & np \leq k \leq n.
\end{cases}
\]  

(2.4)

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

\[
\Pr(n,k,p) \leq \min\left(1,\frac{np}{k},\frac{n(n-1)p^2}{k(k-1)}\right).
\]  

(2.5)

It is possible to then use Theorem 2.1 to identify conditions on the parameters \((n,k,p)\) for which the bounds in (2.4) and (2.5) are tight. We only focus on the non-trivial cases where the tight bound is strictly less than 1 and \(n \geq 3\).

**Proposition 2.1.**

(a) For \(p = \alpha/(n-1)\) and any integer \(\alpha \in [n-2]\), the Chebyshev bound in (2.4) 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 (2.5) is tight for all \(k \in \{2, n\}\) while for \(p > 1/(n-1)\), the bound is tight for all \(k \in [\lceil 1 + (n-1)p \rceil, \lfloor n(n-1)p^2/(np-1) \rfloor]\).

**Proof.** Since Theorem 2.1 provides the tight bound, we simply need to show the equivalence with the bounds in (2.4) and (2.5) 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 2.1. Plugging in the values, the index \(i\) which is required for finding the tight bound is given by:

\[
i = \left\lceil \frac{n \alpha (\alpha + 1 - \alpha)/(n-1)}{\alpha + 1 - n \alpha/(n-1)} \right\rceil = 0.
\]

The corresponding tight bound in (2.1) gives:

\[
\Pr(n,k,p) = \frac{n \alpha}{(n-1)(\alpha + 1)} = \frac{np}{np + 1 - p}.
\]

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 2.1. Plugging in the values, the index $i$ in the tight bound is given by:

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

The tight bound in (2.1) gives:

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

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 (2.5) satisfy:

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

Since this implies $1 \geq np/k$, the bound in (2.5) 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 2.1. If $k = 1 + (n - 1)p$, the index $i = \left\lceil \frac{np(k - (1 + (n - 1)p))}{k - np} \right\rceil = 0$ and the tight bound 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 Theorem 2.1 reduces to:

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

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

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

$$i = \left\lceil \frac{np(k - np - (1 - p))}{k - np} \right\rceil = \left\lceil np\left(1 - \frac{1 - p}{k - np}\right) \right\rceil = 1 \text{ [since } 0 < np \leq 1 \text{ and } (1 - p) \in (0, 1) \text{ and } k - np \geq 1\text{].}$$
For the probability $1/n < p < 1/(n - 1)$, let $(n - 1)p = 1 - \delta$ where $\delta < 1$. Then, since $np > 1$, we have $n\frac{(1 - \delta)}{n - 1} > 1$ or equivalently $n\delta < 1$. The index $i$ satisfies:

$$i = \left\lceil \frac{np((n - 1)p - (k - 1))}{np - k} \right\rceil < \left\lceil \frac{np(1 - \delta - (k - 1))}{1 - k} \right\rceil \quad \text{[since } np > 1 \text{ and } (n - 1)p = 1 - \delta]\right.$$

$$= \left\lceil \frac{np(k - 2 + \delta)}{k - 1} \right\rceil < \left\lceil \frac{n(k - 2 + \delta)}{(n - 1)(k - 1)} \right\rceil \quad \text{[since } p < 1/(n - 1)\right]\right.$$

$$= \left\lceil \frac{n(k - 2 + \delta)}{nk - n - k + 1} \right\rceil \leq \left\lceil \frac{n(k - 2 + \delta)}{nk - 2n + 1} \right\rceil \quad \text{[since } k \leq \text{n]}\right.$$

$$= \left\lceil \frac{n(k - 2) + n\delta}{n(k - 2) + 1} \right\rceil = 1 \quad \text{[since } k \geq 2 \text{ and } 0 < n\delta < 1\right]\right.$$

Hence, the bound in (2.5) 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 $k \in \left[\left\lceil 1 + (n - 1)p \right\rceil, \left\lfloor n(n - 1)p^2/(np - 1) \right\rfloor \right]$.

Connection to existing results: A specific instance to show the tightness of the Chebyshev bound is to set $p = 1/2$, $k = n$ and $n = 2^m - 1$ using $m$ independent Bernoulli random variables to construct $n$ pairwise independent Bernoulli random variables (see Tao [2012], Goemans [2015], Pass and Spektor [2018] for this construction). Proposition 2.1(a) includes this instance (set $\alpha = (n - 1)/2$, $k = n$ and $n = 2^m - 1$). In addition, Proposition 2.1(a) identifies other values of $p$ and $k$ where the Chebyshev bound is tight. Proposition 2.1(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 $k = 1$. Interestingly, in the next section, we will provide the tightest upper bound for $k = 1$ for any set of marginal probabilities. We now provide a numerical illustration of the results in Theorem 2.1 and Proposition 2.1.
Example 1 (Identical marginals). In Table 1, we provide a numerical comparison of the bounds for $n = 11$ for a set of values of $p$ and $k$. The conditions identified in Proposition 2.1 covers all the instances in Table 1 where the Chebyshev bound and the Schmidt, Siegel and Srinivasan bound 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 $k = 10$), (ii) $p = 0.2$ (here $\alpha = 2$ and the Chebyshev bound is tight for $k = 3$ and $k = 10$) and (iii) $p = 0.5$ (here $\alpha = 5$ and the Chebyshev bound is tight for $k = 6$ and $k = 10$). 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 1.

Table 1: 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 (2.1) followed by the Chebyshev bound (2.4) and the Schmidt, Siegel and Srinivasan bound (2.5). The underlined instances illustrate cases when the upper bounds in either (2.4) or (2.5) are tight.

<table>
<thead>
<tr>
<th>$p/k$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.10900</td>
<td>0.00550</td>
<td>0.00184</td>
<td>0.00092</td>
<td>0.00055</td>
<td>0.00037</td>
<td>0.00027</td>
<td>0.00020</td>
<td>0.00016</td>
<td>0.00013</td>
<td>0.00010</td>
</tr>
<tr>
<td>0.02</td>
<td>0.10900</td>
<td>0.00550</td>
<td>0.00184</td>
<td>0.00092</td>
<td>0.00055</td>
<td>0.00037</td>
<td>0.00027</td>
<td>0.00020</td>
<td>0.00016</td>
<td>0.00013</td>
<td>0.00010</td>
</tr>
<tr>
<td>0.05</td>
<td>0.52500</td>
<td>0.13750</td>
<td>0.04583</td>
<td>0.02292</td>
<td>0.01375</td>
<td>0.00917</td>
<td>0.00655</td>
<td>0.00491</td>
<td>0.00382</td>
<td>0.00306</td>
<td>0.00250</td>
</tr>
<tr>
<td>0.10</td>
<td>0.55000</td>
<td>0.18333</td>
<td>0.09167</td>
<td>0.05500</td>
<td>0.03667</td>
<td>0.02620</td>
<td>0.01965</td>
<td>0.01528</td>
<td>0.01223</td>
<td>0.01000</td>
<td></td>
</tr>
<tr>
<td>0.11</td>
<td>0.59950</td>
<td>0.22184</td>
<td>0.11092</td>
<td>0.06655</td>
<td>0.04437</td>
<td>0.03037</td>
<td>0.02170</td>
<td>0.01827</td>
<td>0.01626</td>
<td>0.01013</td>
<td></td>
</tr>
<tr>
<td>0.15</td>
<td>0.78750</td>
<td>0.41250</td>
<td>0.19584</td>
<td>0.09792</td>
<td>0.05875</td>
<td>0.03916</td>
<td>0.02798</td>
<td>0.02098</td>
<td>0.01632</td>
<td>0.01306</td>
<td></td>
</tr>
<tr>
<td>0.20</td>
<td>0.73334</td>
<td>0.33334</td>
<td>0.16667</td>
<td>0.10000</td>
<td>0.06667</td>
<td>0.04762</td>
<td>0.03572</td>
<td>0.02778</td>
<td>0.02223</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>1.00000</td>
<td>0.54167</td>
<td>0.29167</td>
<td>0.17500</td>
<td>0.11667</td>
<td>0.08334</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.91667</td>
<td>0.55000</td>
<td>0.30556</td>
<td>0.18334</td>
<td>0.11957</td>
<td>0.08334</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.91667</td>
<td>0.55000</td>
<td>0.30556</td>
<td>0.18334</td>
<td>0.11957</td>
<td>0.08334</td>
<td></td>
</tr>
</tbody>
</table>

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 2.1 is given by $np - (n - 1)p^2$. The reader is referred to Problem 20 on page 26 of the book by Lovász [1979] which discusses other bounds for exactly the same problem.

For larger values of $p$ above 0.1, such as $p = 0.11$ in the table, from Proposition 2.1(b), the Schmidt, Siegel and Srinivasan bound is tight for $k \in \lceil 2.1 \rceil, [6.33]$ 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.
3 Non-identical marginals

In this section, we focus on probability bounds for the sum of pairwise independent random variables with non-identical marginals.

3.1 Tight upper bound for \( k = 1 \)

The next theorem provides the tight bound for \( k = 1 \). The proof of the theorem involves firstly constructing a feasible solution for the dual of the exponential sized linear program and secondly showing that there exists a feasible distribution which attains the bound. A key step in the proof of tightness is to show the existence of a distribution of a Bernoulli random vector with specified univariate probabilities and transformed bivariate probabilities, which should be of independent interest in itself.

**Theorem 3.1.** Given \( p \in (0, 1)^n \), let \( \overline{P}(n, 1, p) \) represent the tightest possible upper bound on the sum of \( n \) pairwise independent Bernoulli random variables exceeding \( k = 1 \). Sort the probabilities in increasing order as \( p_1 \leq p_2 \leq \ldots \leq p_n \). Then,

\[
\overline{P}(n, 1, p) = \begin{cases} 
1, & \sum_{i=1}^{n-1} p_i > 1 \\
\sum_{i=1}^{n} p_i - p_n \left( \sum_{i=1}^{n-1} p_i \right), & \sum_{i=1}^{n-1} p_i \leq 1 
\end{cases} 
\tag{3.1}
\]

**Proof.** Sort the variables in terms of the marginal probabilities \( p_1 \leq p_2 \leq \ldots \leq p_n \). The tightest upper bound is the optimal value of the linear program:

\[
\overline{P}(n, 1, p) = \max \sum_{c \in C : \sum_t c_t \geq 1} \theta(c) \\
\text{s.t.} \sum_{c \in C} \theta(c) = 1 \\
\sum_{c \in C : c_i = 1} \theta(c) = p_i, \quad \forall i \in [n] \\
\sum_{c \in C : c_i = 1, c_j = 1} \theta(c) = p_i p_j, \quad \forall (i, j) \in I_n \\
\theta(c) \geq 0, \quad \forall c \in C. 
\tag{3.2}
\]

The dual linear program is formulated as:

\[
\overline{P}_d(n, 1, p) = \min \sum_{(i,j) \in I_n} \lambda_{ij} p_i p_j + \sum_{i=1}^{n} \lambda_i p_i + \lambda_0 \\
\text{s.t.} \sum_{(i,j) \in I_n} \lambda_{ij} c_i c_j + \sum_{i=1}^{n} \lambda_i c_i + \lambda_0 \geq 0, \quad \forall c \in C : \sum_t c_t = 0 \\
\sum_{(i,j) \in I_n} \lambda_{ij} c_i c_j + \sum_{i=1}^{n} \lambda_i c_i + \lambda_0 \geq 1, \quad \forall c \in C : \sum_t c_t \geq 1. 
\tag{3.3}
\]
The optimality conditions of linear programming states that \( \{ \theta(c); c \in C \} \) is primal optimal and \( \{ \lambda_{ij}; (i,j) \in I_n, \lambda_i; i \in [n], \lambda_0 \} \) is dual optimal if and only if they satisfy (i) the primal feasibility conditions in (3.2), (ii) the dual feasibility conditions in (3.3) and (iii) the complementary slackness condition which is stated as:

\[
\begin{align*}
\left( \sum_{(i,j) \in I_n} \lambda_{ij}c_ic_j + \sum_{i=1}^{n} \lambda_ic_i + \lambda_0 \right) \theta(c) &= 0, \quad \forall c \in C : \sum_t c_t = 0 \\
\left( \sum_{(i,j) \in I_n} \lambda_{ij}c_ic_j + \sum_{i=1}^{n} \lambda_ic_i + \lambda_0 - 1 \right) \theta(c) &= 0, \quad \forall c \in C : \sum_t c_t \geq 1.
\end{align*}
\]

From strong duality, \( P(n, 1, p) = P_d(n, 1, p) \). We first prove the result in case (b) in (3.1), which is the non-trivial upper bound.

**Step (1): Create a dual feasible solution**

Step 1 of the proof is implicit in the work of Kounias [1968]. We reproduce it here for completeness. Construct a feasible solution to the dual linear program in (3.3) as follows:

\[
\begin{align*}
\lambda_0 &= 0, \\
\lambda_i &= +1, \forall i \in [n] \\
\lambda_{ij} &= \begin{cases} -1, & \text{if } j = n \\ 0, & \text{otherwise}, \forall i \in [n-1]. \end{cases}
\end{align*}
\]

The left hand side of the dual constraints in (3.2) simplifies to:

\[
\sum_{(i,j) \in I_n} \lambda_{ij}c_ic_j + \sum_{i=1}^{n} \lambda_ic_i + \lambda_0 = -\sum_{i=1}^{n-1} c_ic_n + \sum_{i=1}^{n} c_i = c_n + \sum_{i=1}^{n-1} c_i(1 - c_n).
\]

To verify that this solution is dual feasible, we observe that with all \( c_i = 0 \), \( c_n + \sum_{i=1}^{n-1} c_i(1 - c_n) = 0 \). When \( c_n = 1 \), regardless of the values of \( c_1, \ldots, c_{n-1} \), we have \( c_n + \sum_{i=1}^{n-1} c_i(1 - c_n) = 1 \). Lastly, when \( c_n = 0 \) and at least one \( c_i = 1 \) for \( i \in [n-1] \), we have \( c_n + \sum_{i=1}^{n-1} c_i(1 - c_n) \geq 1 \). This gives a dual feasible solution with the objective value in (3.1). From weak duality, this is an upper bound on \( P(n, 1, p) \).

**Step (2): Show tightness by constructing a pairwise independent distribution**

We verify the tightness of the bound, by showing there exists a primal solution (feasible distribution) which satisfies the complementary slackness conditions. Towards this, from the complementary slackness condition in (c), we have:

\[
\forall c \in C : \sum_t c_t \geq 2, c_n = 0, \quad \text{we have} \quad \left( c_n + \sum_{i=1}^{n-1} c_i(1 - c_n) - 1 \right) > 0 \quad \implies \quad \theta(c) = 0.
\]

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 $\tilde{c}$ as shown in Table 2. The probability of the vector of all zeros (one scenario) is set to $1 - \sum_{i=1}^{n} p_i - p_n \left( \sum_{i=1}^{n-1} p_i \right)$. To match the bivariate probabilities $P(\tilde{c}_i = 1, \tilde{c}_n = 0)$, we have to 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 feasibility of the distribution, we simply need to show that there exists nonnegative values $\theta(c)$ for the last $2^{n-1}$ scenarios such that:

$$
\sum_{c \in C: c_n = 1} \theta(c) = p_n, \\
\sum_{c \in C: c_i = 1, c_n = 1} \theta(c) = p_i p_n, \quad \forall i \in [n - 1] \\
\sum_{c \in C: c_i = 1, c_j = 1, c_n = 1} \theta(c) = p_i p_j, \quad \forall (i, j) \in I_{n-1}.
$$

or equivalently, by conditioning on $c_n = 1$, we need to show that there exists nonnegative values $\theta_{n-1}(c)$ where $c \in \{0, 1\}^{n-1}$ such that:

$$
\sum_{c \in \{0, 1\}^{n-1}} \theta_{n-1}(c) = 1, \\
\sum_{c \in \{0, 1\}^{n-1}; c_i = 1} \theta_{n-1}(c) = p_i, \quad \forall i \in [n - 1] \quad (3.4) \\
\sum_{c \in \{0, 1\}^{n-1}; c_i = 1, c_j = 1} \theta_{n-1}(c) = \frac{p_i p_j}{p_n}, \quad \forall (i, j) \in I_{n-1},
$$

This corresponds to verifying the existence of a probability distribution with $n - 1$ Bernoulli random variables, marginal probabilities $p_i$ and bivariate probabilities $p_i p_j / p_n$ where $p_n \geq p_{n-1} \geq p_{n-2} \geq \ldots \geq p_1$. Observe, that the distribution in (3.4) satisfies the original univariate marginal probabilities but

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>...</th>
<th>$c_{n-1}$</th>
<th>$c_n$</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 scenario</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>$1 - \sum_{i=1}^{n} p_i - p_n \left( \sum_{i=1}^{n-1} p_i \right)$</td>
</tr>
<tr>
<td>$n - 1$ scenarios</td>
<td>1</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>$p_1 (1 - p_n)$</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>$p_2 (1 - p_n)$</td>
</tr>
<tr>
<td></td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>...</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$p_{n-1} (1 - p_n)$</td>
</tr>
<tr>
<td>$2^{n-1} - n$ scenarios</td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$\theta(c)$</td>
</tr>
<tr>
<td>$2^{n-1}$ scenarios</td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
<td>:</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>1</td>
<td>$\theta(c)$</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\theta(c)$</td>
<td></td>
</tr>
</tbody>
</table>
is no longer pairwise independent. In the final step of the proof for case (b), we show that such a distribution always exists.

**Step (3): Complete proof of tightness by showing there exists a distribution that satisfies (3.4)**

1. We first argue that it is sufficient to prove the existence of a probability distribution with nonnegative values $\theta_{n-1}(c)$ and $n - 1$ Bernoulli random variables such that:

$$
\sum_{c \in \{0,1\}^{n-1}} \theta_{n-1}(c) = 1,
\sum_{c \in \{0,1\}^{n-1}; c_i=1} \theta_{n-1}(c) = p_i, \quad \forall i \in [n-1]
$$

$$
\sum_{c \in \{0,1\}^{n-1}; c_i=1, c_j=1} \theta_{n-1}(c) = \frac{p_ip_j}{p_{n-1}}, \quad \forall (i,j) \in I_{n-1},
$$

(3.5)

where the bivariate probabilities are modified from $p_ip_j/p_n$ to $p_ip_j/p_{n-1}$. To see this, observe that, since $1 \leq 1/p_n \leq 1/p_{n-1}$, we can always find a $\lambda \in [0, 1]$ such that:

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

Then, by considering a convex combination of the two distributions as follows:

$$
\theta_{n-1} = \lambda \overline{\theta}_{n-1} + (1 - \lambda) \theta_{n-1},
$$

where $\overline{\theta}_{n-1}$ is a probability distribution which satisfies (3.5) and $\theta_{n-1}$ is a joint distribution on $n - 1$ Bernoulli random variables with univariate marginals given by $p_i$ and bivariate probabilities given by $p_ip_j$, we guarantee (3.4) is feasible. Such a distribution $\overline{\theta}_{n-1}$ always exists (simply choose the mutually independent distribution on $n - 1$ random variables with marginal probabilities $p_i$). The distribution $\theta_{n-1}$ thus created has marginal probabilities given by $p_i$ and bivariate probabilities given by $p_ip_j/p_n$.

2. To show that (3.5) is feasible, by conditioning on $c_{n-1} = 1$, we simply need to show that there exists a probability distribution on $n - 2$ Bernoulli random variables with nonnegative values $\theta_{n-2}(c)$ such that:

$$
\sum_{c \in \{0,1\}^{n-2}} \theta_{n-2}(c) = 1,
\sum_{c \in \{0,1\}^{n-2}; c_i=1} \theta_{n-2}(c) = \frac{p_i}{p_{n-1}}, \quad \forall i \in [n-2]
$$

$$
\sum_{c \in \{0,1\}^{n-2}; c_i=1, c_j=1} \theta_{n-2}(c) = \frac{p_ip_j}{p_{n-1}}, \quad \forall (i,j) \in I_{n-2}.
$$

(3.6)

Then, by setting the vector of all zeros to $1 - p_{n-1}$ and scaling the probabilities when $c_{n-1} = 1$ (see the construction in Table 3), we obtain a feasible distribution. Such a distribution on $n - 2$ random variables with univariate probabilities given by $p_i/p_{n-1}$ and bivariate probabilities given by
$(p_i/p_{n-1})(p_j/p_{n-1})$ always exists (simply choose the mutually independent distribution on $n-2$ random variables with marginal probabilities $p_i/p_{n-1}$). The steps in (2) and (3) are shown in Figure 1. This completes the proof for part (b) and the tight bound is given by:

$$\overline{P}(n, 1, p) = \sum_{i=1}^{n} p_i - p_n (\sum_{i=1}^{n-1} p_i).$$

Table 3: Probabilities of the scenarios to create a feasible distribution in (3.5).

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$\ldots$</th>
<th>$c_{n-1}$</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^{n-2}$ scenarios</td>
<td>0</td>
<td>0</td>
<td>$\ldots$</td>
<td>0</td>
<td>$\theta_{n-1}(c) = 1 - p_{n-1}$</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>$\ldots$</td>
<td>0</td>
<td>$0$</td>
</tr>
<tr>
<td></td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>$\ldots$</td>
<td>1</td>
<td>0</td>
<td>$\theta_{n-1}(c) = p_{n-1}\theta_{n-2}(c)$</td>
</tr>
<tr>
<td>$2^{n-2}$ scenarios</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\theta_{n-1}(c) = p_{n-1}\theta_{n-2}(c)$</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Construction of the worst-case distribution with univariate and bivariate probabilities.

To complete the proof, we observe that for case (a) in the theorem with $\sum_{i=1}^{n-1} p_i > 1$, we can find an index $t \in [2, n-1]$ such that $\sum_{i=t}^{n-1} p_i \leq 1$ and $\sum_{i=1}^{t} p_i > 1$. Let $\delta = 1 - \sum_{i=t+1}^{n-1} p_i$. Clearly $0 \leq \delta < p_t$.

From the proof of case (b), we know 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 is greater than or equal to 1 is equal to one (since the sum of the first $t$ probabilities is equal to one). By increasing the marginal probability $\delta$ to $p_t$, we can only increase this probability. Hence, there exists a distribution for $t + 1$ random variables with probabilities $0 < p_1 \leq p_2 \leq \ldots \leq p_t \leq p_{t+1} < 1$ such that there is a zero probability for these random variables to simultaneously take a value of 0. We can generate the remaining random variables $\tilde{c}_{t+2}, \ldots, \tilde{c}_n$ independently with marginal probabilities $p_{t+2}, \ldots, p_n$. This provides a feasible distribution and attains the bound of 1 for case (a).

Connection to existing work: The problem of bounding 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 $\tilde{c}_i = 1$ for each $i$. Suppose 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 $(i, j) \in I_n$ are given. Then, $k = 1$ simply corresponds to the bounding the probability of the union of events. Kounias [1968] derived the following upper bound:

$$\mathbb{P}(\bigcup_i A_i) \leq \sum_{i=1}^{n} p_i - \max_{j \in [n]} \sum_{i \neq j} p_{ij},$$

which subtracts the maximum weight of a star spanning tree on a complete graph with $n$ nodes where the edge weights are given by $p_{ij}$ from the union bound. Hunter [1976] and Worsley [1982] tightened this bound by optimizing over the spanning trees $\tau \in T$:

$$\mathbb{P}(\bigcup_i A_i) \leq \sum_{i=1}^{n} p_i - \max_{\tau \in T} \sum_{(i, j) \in \tau} p_{ij},$$

where $T$ is the set of all spanning trees on the graph. Only in a few special cases has this tree bound been shown to be tight. For example, given a tree $\tau$ such that the bivariate probabilities $p_{ij}$ are 0 for all the edges $(i, j) \notin \tau$, the bound is known to be tight (see Worsley [1982]). It is straightforward to see that for pairwise independent random variables where $p_{ij} = p_i p_j$, the maximum weight spanning tree 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 bounds in Kounias [1968], Hunter [1976] and Worsley [1982] reduce to the bound in (3.1). The contribution of Theorem 3.1 is in showing that the bound is tight for all values of $n$ and $p \in (0, 1)^n$ with pairwise independence. In a more recent paper, Boros et al. [2014] derived polynomial time computable upper bounds (but not necessarily tight) for the union of events using bivariate probabilities. Specifically, they showed that by relaxing the equality of bivariate probabilities to lower bounds on bivariate probabilities as $\mathbb{P}(A_i \cap A_j) \geq p_{ij}, \forall (i, j) \in I_n$, the tightest upper bound on the probability of the union is exactly the Hunter-Worsley bound (see Maurer [1983] for related results). In fact paraphrasing from their paper (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.” Theorem 3.1 provides a partial positive answer towards this question for pairwise independent random variables.
Feasibility here is guaranteed from pairwise independence and Theorem 3.1 shows that the tightest upper bound is computable in polynomial time (in fact as a closed form). It is also straightforward to see that when the marginals are identical, the bound reduces to the bound derived in the previous section.

### 3.2 Improved bounds with non-identical marginals for $k \geq 2$

The next theorem provides new probability bounds for the sum of pairwise independent random variables with possibly non-identical marginals when $k \geq 2$. These bounds build on the previous results while exploiting the ordering of probabilities for the pairwise independent random variables.

**Theorem 3.2.** Sort the input probabilities in increasing order as $p_1 \leq p_2 \leq \ldots \leq p_n$. Define the partial binomial moment $S_{1r} = \sum_{i=1}^{n-r} p_i$ for $r \in [0, n-1]$ and $S_{2r} = \sum_{(i,j) \in I_{n-r}} p_i p_j$ for $r \in [0, n-2]$.

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

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

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

$$
\mathcal{P}(n, k, p) \leq \min_{0 \leq r \leq k-1} \text{BP}(n-r, k-r, p), \forall k \in [2, n]
$$

where:

$$
\text{BP}(n-r, k-r, p) = \begin{cases}
1, & k < \frac{(n-r-1)S_{1r} - 2S_{2r}}{n-r - S_{1r}} + r \\
\frac{(k-r + n-r-1)S_{1r} - 2S_{2r}}{(k-r)(n-r)}, & \frac{(n-r-1)S_{1r} - 2S_{2r}}{n-r - S_{1r}} + r \leq k < 1 + \frac{2S_{2r}}{S_{1r}} + r \\
\frac{(i-1)(i-2S_{1r}) + 2S_{2r}}{(k-r - i)^2 + (k-r - i)}, & 1 + \frac{2S_{2r}}{S_{1r}} + r \leq k \leq \left( \frac{(k-r-1)S_{1r} - 2S_{2r}}{k-r - S_{1r}} \right)
\end{cases}
$$

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

$$
\mathcal{P}(n, k, p) \leq \min_{0 \leq r \leq k-1} \text{CH}(n-r, k-r, p), \forall k \in [2, n]
$$

where:

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

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

$$
\mathbb{P} \left( \sum_{i=1}^{n} \tilde{c}_i \geq k \right) \leq \mathbb{P} \left( \sum_{i \in S} \tilde{c}_i \geq k - r_1 \right)
$$

[since $\sum_{i=1}^{n} c_i \geq k$ for $c \in \{0, 1\}^n$ implies $\sum_{i \in S} c_i \geq k - r_1$ for $c \in \{0, 1\}^n$]

$$
\leq \frac{\mathbb{E} \left[ \sum_{i \in S} \tilde{c}_i \right]}{k - r_1}
$$

[using Markov’s inequality]

$$
= \frac{\sum_{i \in S} p_i}{k - r_1}.
$$

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

$$
\mathbb{P} \left( \sum_{i=1}^{n} \tilde{c}_i \geq k \right) \leq \min_{0 \leq r_1 \leq k-1} \min_{S:|S|=n-r_1} \frac{\sum_{i \in S} p_i}{k - r_1}
$$

$$
= \min_{0 \leq r_1 \leq k-1} \frac{\sum_{i=1}^{n-r_1} p_i}{k - r_1}
$$

[using equation (1.2) and Markov’s inequality]

$$
= \sum_{i \in S} \sum_{j \in S: j > i} \frac{p_i p_j}{(k-r_1)^2}
$$

[using pairwise independence]

We derive the next term in (3.7) using a similar approach while accounting for pairwise independence. For any $0 \leq r_2 \leq k - 2$ and any subset $S \subseteq [n]$ of the random variables of cardinality $n - r_2$, an upper bound is given by:

$$
\mathbb{P} \left( \sum_{i=1}^{n} \tilde{c}_i \geq k \right) \leq \mathbb{P} \left( \sum_{i \in S} \tilde{c}_i \geq k - r_2 \right)
$$

$$
= \mathbb{P} \left( \left( \sum_{i \in S} \tilde{c}_i \right) \geq \binom{k - r_2}{2} \right)
$$

$$
\leq \frac{\mathbb{E} \left[ \sum_{i \in S} \sum_{j \in S: j > i} \tilde{c}_i \tilde{c}_j \right]}{\binom{k - r_2}{2}}
$$

[using equation (1.2) and Markov’s inequality]

$$
= \frac{\sum_{i \in S} \sum_{j \in S: j > i} \mathbb{E}[\tilde{c}_i] \mathbb{E}[\tilde{c}_j]}{\binom{k - r_2}{2}}
$$

[using pairwise independence]

$$
= \frac{\sum_{i \in S} \sum_{j \in S: j > i} p_i p_j}{\binom{k - r_2}{2}}.
$$
The tightest upper bound of this form is obtained by minimizing over $0 \leq r_2 \leq k - 2$ and all sets $S$ of size $n - r_2$. This gives:

$$
P\left(\sum_{i=1}^{n} \tilde{c}_i \geq k\right) \leq \min_{0 \leq r_2 \leq k - 2} \min_{S:|S|=n-r_2} \frac{\sum_{i \in S} \sum_{j \in S \setminus \{i\}} p_i p_j}{(k-r_2)}
$$

$$
= \min_{0 \leq r_2 \leq k - 2} \left(\frac{\sum_{(i,j) \in I_{n-r_2}} p_i p_j}{(k-r_2)}\right)
$$

[using the $n - r_2$ smallest probabilities].

From the bounds (3.10) and (3.11), we get:

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

where $S_{1r_1} = \sum_{i=1}^{n-r_1} p_i$ for $r_1 \in [0, n - 1]$ and $S_{2r_2} = \sum_{(i,j) \in I_{n-r_2}} p_i p_j$ for $r_2 \in [0, n - 2]$. It is straightforward to see that this approach is essentially creating a set of dual feasible solutions and picking the best among it. The dual formulation is:

$$
\mathcal{P}(n, k, p) = \min \sum_{(i,j) \in I_n} \lambda_{ij} p_i p_j + \sum_{i=1}^{n} \lambda_i p_i + \lambda_0
$$

s.t

$$
\sum_{(i,j) \in I_n} \lambda_{ij} c_i c_j + \sum_{i=1}^{n} \lambda_i c_i + \lambda_0 \geq 0 \quad \forall c \in C
$$

$$
\sum_{(i,j) \in I_n} \lambda_{ij} c_i c_j + \sum_{i=1}^{n} \lambda_i c_i + \lambda_0 \geq 1, \quad \forall c \in C : \sum_{t} c_t \geq k.
$$

Each component of the second term is 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, each component of the third term is obtained by choosing dual feasible solutions with $\lambda_{ij} = 1/(k-r_2)$ for $(i, j) \in I_{n-r_2}$ and setting all other dual variables to 0.

(b) The bound in (3.8) is obtained by using the inequality:

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

Then, we compute an upper bound on $P\left(\sum_{i=1}^{n-r} \tilde{c}_i \geq k - r\right)$ by using the aggregated moments $S_{1,}$
and $S_{2r}$ with the Boros and Prekopa bound from (1.4) as follows:

$$BP(n-r,k-r,p) = \begin{cases} 
1, & k < \frac{(n-r-1)S_{1r} - 2S_{2r}}{n-r} + r \\
\frac{(k-r + n-r-1)S_{1r} - 2S_{2r}}{(k-r)(n-r)}, & (n-r-1)S_{1r} - 2S_{2r} + r \leq k < 1 + \frac{2S_{2r}}{S_{1r}} + r \\
\frac{(i-1)(i-2S_{1r}) + 2S_{2r}}{(k-r-i)^2 + (k-r-i)}, & k \geq 1 + \frac{2S_{2r}}{S_{1r}} + r, \quad i = \left\lceil \frac{(k-r-1)S_{1r} - 2S_{2r}}{k-r-S_{1r}} \right\rceil
\end{cases}$$

Since the relation $P(n,k,p) \leq BP(n-r,k-r,p)$ is satisfied for every $0 \leq r \leq k-1$, the upper bound on $P(n,k,p)$ is obtained by taking the minimum over all possible values of $r$:

$$P(n,k,p) \leq \min_{0 \leq r \leq k-1} BP(n-r,k-r,p).$$

(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 $r$ ($0 \leq r \leq k-1$) can be written as follows:

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

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

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

Connection to existing work: Prior work in Rüger [1978] shows that ordering of probabilities provides the tightest upper bound on the probability of the sum of Bernoulli random variables exceeding $k$ while allowing for arbitrary dependence. Specifically, the bound derived there is:

$$\min \left( 1, \min_{0 \leq r \leq k-1} \left( \frac{S_{1r}}{k-r} \right) \right).$$

However, this bound does not use pairwise independence information. Part (a) of Theorem 3.2 tightens the analysis in Rüger [1978] for pairwise independent random variables. It is also straightforward to see that the ordered Schmidt, Siegel and Srinivasan bound in (3.7) is at least as good as the bound in (1.3) (simply plug in $r = 0$). Building on the ordering of probabilities, the bound in (3.8) uses aggregated binomial moments for $k$ ordered sets of random variables of size $n-r$ where $0 \leq r \leq k-1$. When $r = 0$, the bound in (3.8) reduces to the original aggregated moment bound of Boros and Prekopa in (1.4) and hence this bound is at least as tight. The bounds in Theorem 3.2 are clearly efficiently computable. We next provide two numerical examples to illustrate the impact of ordering on the quality of the three bounds.
Example 2 (Non-identical marginals). Consider an example with \( n = 12 \) random variables with the probabilities given by

\[
p = (0.0651, 0.0977, 0.1220, 0.1705, 0.3046, 0.4402, 0.4952, 0.6075, 0.6842, 0.8084, 0.9489, 0.9656).
\]

Table 4 compares the three ordered bounds with the three unordered bounds and the corresponding tight bound. Numerically, the ordered Boros and Prekopa bound is found to be tight in this example for \( k = 7, 8, 9, 12 \) while the ordered Schmidt, Siegel and Srinivasan bound is tight for \( k = 12 \). The Boros and Prekopa bound is uniformly the best performing of the three bounds, while among the other two 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 outperforms the Chebyshev bound when \( k = 6 \) but for all \( k \geq 7 \), the Chebyshev bound does better. In terms of absolute difference between ordered and unordered bounds, ordering appears to provide the maximum improvement to the Schmidt, Siegel and Srinivasan bound, followed by the Boros and Prekopa and the Chebyshev bound.

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

<table>
<thead>
<tr>
<th>Bounds</th>
<th>( k \in [1, 4] )</th>
<th>( k = 5 )</th>
<th>( k = 6 )</th>
<th>( k = 7 )</th>
<th>( k = 8 )</th>
<th>( k = 9 )</th>
<th>( k = 10 )</th>
<th>( k = 11 )</th>
<th>( k = 12 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schmidt, Siegel and Srinivasan</td>
<td>1</td>
<td>0.9517</td>
<td>0.6831</td>
<td>0.5123</td>
<td>0.3985</td>
<td>0.3188</td>
<td>0.2608</td>
<td>0.2173</td>
<td></td>
</tr>
<tr>
<td>Ordered Schmidt, Siegel and Srinivasan</td>
<td>1</td>
<td>0.9489</td>
<td>0.6162</td>
<td>0.3620</td>
<td>0.1827</td>
<td>0.0712</td>
<td>0.0250</td>
<td>0.0064</td>
<td></td>
</tr>
<tr>
<td>Chebyshev</td>
<td>1</td>
<td>0.9533</td>
<td>0.5192</td>
<td>0.2552</td>
<td>0.1424</td>
<td>0.0889</td>
<td>0.0603</td>
<td>0.0434</td>
<td></td>
</tr>
<tr>
<td>Ordered Chebyshev</td>
<td>1</td>
<td>0.9533</td>
<td>0.5192</td>
<td>0.2552</td>
<td>0.1424</td>
<td>0.0883</td>
<td>0.0549</td>
<td>0.0307</td>
<td></td>
</tr>
<tr>
<td>Boros and Prekopa</td>
<td>1</td>
<td>0.9497</td>
<td>0.5018</td>
<td>0.2509</td>
<td>0.1326</td>
<td>0.0795</td>
<td>0.0530</td>
<td>0.0379</td>
<td></td>
</tr>
<tr>
<td>Ordered Boros and Prekopa</td>
<td>1</td>
<td>0.9254</td>
<td>0.5018</td>
<td>0.2509</td>
<td>0.1290</td>
<td>0.0712</td>
<td>0.0249</td>
<td>0.0064</td>
<td></td>
</tr>
<tr>
<td>Tight bound</td>
<td>1</td>
<td>0.9957</td>
<td>0.8931</td>
<td>0.5018</td>
<td>0.2509</td>
<td>0.1290</td>
<td>0.0692</td>
<td>0.0230</td>
<td>0.0064</td>
</tr>
</tbody>
</table>

Example 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 \( p = (p_1, p_2, \ldots, p_{100}) \). First, we consider small marginal probabilities by uniformly and independently generating the entries of \( p \) between 0.01 and 0.05. When \( k = n \), Figure 2a plots the three ordered bounds while Figure 2b shows the percentage improvement of the three bounds over their unordered counterparts. The percentage improvement is computed as \((\text{unordered-ordered}/\text{unordered}) \times 100\%\).
In this example with small marginals, the ordered Schmidt, Siegel and Srinivasan bound is equal to the ordered Boros and Prekopa bound as seen in Figure 2a. Ordering tends to improve the Schmidt, Siegel and Srinivasan bound significantly for smaller probabilities, since both the partial binomial moment terms $S_1^r$ and $S_2^r$ are smaller with smaller marginal probabilities for all $r \in [0, k - 1]$. The percentage improvement due to ordering in figure 2b is consistently above 80% for the Schmidt, Siegel and Srinivasan bound, being while that of the Boros and Prekopa bound is around 60%. The ordered Chebyshev bound shows an almost negligible improvement by ordering in this example.

Next, we consider similar plots when $k = n - 1$ with larger marginal probabilities. The entries of $\mathbf{p}$ are generated uniformly and independently between 0.05 and 0.99.

In Figure 3a, the ordered Chebyshev bound from (3.9) performs better than the ordered Schmidt, Siegel and Srinivasan bound from (3.7). In Figure 3b, 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 Prekopa bound is less than 40% and that of the Chebyshev bound is less than 20%. It is also clear from Figures 2 and 3 that the ordered Boros and Prekopa bound from (3.8) is the tightest of the three bounds across the instances, while among the other two bounds, none uniformly dominates the other.

While the ordered bounds in Theorem 3.2 are not tight in general, the next proposition identifies a special case with almost identical marginals where the bound of Schmidt, Siegel and Srinivasan in (3.7) and Boros and Prekopa in (3.8) are shown to be attained.

**Proposition 3.1.** 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 bounds in (3.7) and (3.8) are tight for the following three instances and given by the bound:

\[
P(n, k, p, q) = \begin{cases} 
\binom{n-1}{k-2} p^2, & k \geq 3, \quad q \geq (n-2)p \quad \text{case (a)} \\
\binom{n-1}{k-2} p^2, & k \in \left[2 + \frac{(n-2)p}{q}, \ n\right], \quad p \leq q < (n-2)p \quad \text{case (b)} \\
pq, & k = n, \quad 0 < q < p \quad \text{case (c)}
\end{cases}
\]

**(3.12)**

**Proof.** We first prove that the ordered bounds of Schmidt, Siegel and Srinivasan and Boros and Prekopa reduce to the bound in (3.12) in each of the three cases and then show that the bound is tight.

**Step (1): Show reduction of ordered bounds to the bound in (3.12)**

Let \( P(n, k, p, q) \) represent the tightest upper bound when \( n-1 \) probabilities are \( p \) and one is \( q \). It can be observed that the bound in (3.12) is non-trivial for the three instances since:

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

It is easy to verify that the ordered Schmidt, Siegel and Srinivasan bound in (3.7) reduces to the bound in (3.12) for a specific parameter \( r_2 \) in each of the three cases:

\[
r_2 = 1, \quad \text{cases (a) and (b)} \\
r_2 = n - 2, \quad \text{case (c)}.
\]

**(3.13)**

It can be similarly verified that the ordered Boros and Prekopa bound in (3.8) reduces to the bound in (3.12) with the following parameters \( r \) and \( i \) in each of the three cases:

\[
r = 1, \quad i = 0 \quad \text{cases (a) and (b)} \\
r = n - 2, \quad i = 0 \quad \text{case (c)}.
\]

**(3.14)**

24
The effectiveness of ordering is demonstrated by (3.13) and (3.14) in that the ordered bounds of Schmidt, Siegel and Srinivasan and Boros and Prekopa correspond to \( r > 0 \) while their unordered counterparts in (1.3) and (1.4) 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.

**Step (2): Prove tightness of the bound in (3.12) by constructing worst-case distributions**

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

\[
P(n, k, p, q) = \max \sum_{c \in C} \theta(c) \sum_{t} c_t \geq k \quad \text{s.t} \quad \sum_{c \in C} \theta(c) = 1 \]

\[
\sum_{c \in C : c_t = 1} \theta(c) = p, \quad \forall i \in [n] \\
\sum_{c \in C : c_n = 1} \theta(c) = q \\
\sum_{c \in C : c_i = 1, c_j = 1} \theta(c) = p^2, \quad \forall (i, j) \in I_{n-1} \\
\sum_{c \in C : c_i = 1, c_n = 1} \theta(c) = p q, \quad \forall i \in [n-1] \\
\theta(c) \geq 0, \quad \forall c \in C.
\]

We now proceed to prove tightness of the bound in (3.12) for each of the three instances of the \((n, k, p, q)\) tuple by constructing feasible distributions of (3.15) which attain the bound.

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

The following distribution attains the tight bound:

\[
\theta(c) = \begin{cases} 
(1 - q)(1 - (n - 1)p), & \text{if } \sum_{t=1}^{n-1} c_t = 0 \\
p(1 - q), & \text{if } \sum_{t=1}^{n-1} c_t = 1, c_n = 0 \\
q(1 - (n - 1)p) + \frac{(n-1)(n-2)p^2}{(k-1)}, & \text{if } \sum_{t=1}^{n-1} c_t = 0, c_n = 1 \\
p(q - \frac{n-2}{k-2}p), & \text{if } \sum_{t=1}^{n-1} c_t = 1, c_n = 1 \\
\frac{p^2}{(n-3)}, & \text{if } \sum_{t=1}^{n-1} \tilde{c}_t = k - 1, c_n = 1 
\end{cases} 
\]

We use symbols \( x, y, z, u, v \) to denote the probability of the associated scenarios in (3.16). The
constraints in (3.15) reduce to:

\[
\begin{align*}
(n-2)\binom{k-2}{v} + u + y &= p \\
(n-1)\binom{k-1}{v} + (n-1)u + z &= q \\
(n-3)\binom{k-3}{v} &= p^2 \\
(n-2)\binom{k-2}{v} + u &= pq \\
x + y + z + u + v &= 1
\end{align*}
\]

and using \(x, y, z, u, v\) from (3.16), 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:

**case (a):** \(u = p(q - \frac{n-2}{k-2}p)\)

\[
\begin{align*}
&\geq y = p(q - \frac{n-2}{k-2}p) \\
&[\text{since } k \geq 3] \\
&= p(q - (n-2)p) \\
&[\text{since } q > (n-2)p] \\
&\geq 0
\end{align*}
\]

**case (b):** \(u = p(q - \frac{n-2}{k-2}p)\)

\[
\begin{align*}
&\geq p(q - \frac{k-2}{k-2}q) \\
&[\text{since } k \geq 2 + (n-2)p/q] \\
&= 0.
\end{align*}
\]

The only support points contributing to the objective function are the first set of \(\binom{n-1}{k-1}\) scenarios, and so we have

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

(b) \(\theta(c) = pq\) (case (c)):

The following distribution attains the tight bound \(pq\):

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

26
The constraints in (3.15) reduce to:

\[
\begin{align*}
y + u + v &= p \\
z + v &= q \\
u + v &= p^2 \\
v &= pq \\
x + y + z + u + v &= 1
\end{align*}
\]

and using \(x, y, z, u, v\) from (3.17), it can be easily verified that all of the above constraints are satisfied. The non-negativity constraints for \(y, z, u, v\) are satisfied by \(0 < q \leq p \leq 1\) while for \(x\), we have:

\[
x = (1 - p)(1 - (n - 2)p - q) \\
\geq (1 - p)(1 - (n - 2)p - p) \quad [\text{since } q < p] \\
= (1 - p)(1 - (n - 1)p) \\
\geq 0 \quad [\text{since } (n - 1)p \leq 1].
\]

The distribution in (3.17) attains the bound \(pq\). We have thus constructed two feasible probability distributions in (3.16) and (3.17) which attain the bound in (3.12) in each of the three instances defined by the \((n, k, p, q)\) tuple. Hence the parameters \(r_2, r\) in (3.13) and (3.14) defined for each of the three cases must be the minimizers which exactly reduce the ordered bounds in (3.7) and (3.8) to the tight bound in (3.12).

Example 4 (Almost identical marginals). This example demonstrates the usefulness of proposition 3.1 when \(n = 100\) and \(p = 0.01\) \(((n - 1)p \leq 1)\), by comparing the tight bounds computed from (3.12) with the unordered bounds of Schmidt, Siegel and Srinivasan from (1.3) and that of Boros and Prekopa from (1.4).
Figure 4: Comparison of unordered bounds with tight bound when \( n = 100, \ p = 0.01 \)

Figure 4a plots the two unordered bounds along with the tight bound when \( q = 0.99 \) (case (a) of proposition 3.1), where the tight bound is valid for all \( k \) in \([3, n]\), while figure 4b compares the bounds when \( q = 0.1 \) (case (b) of proposition 3.1) 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 Prekopa bound is much closer to the tight bound than the unordered Schmidt, Siegel and Srinivasan bound in both figures. Hence, Figure 4 demonstrates that with ordering, the relative improvement of the Schmidt, Siegel and Srinivasan bound is much better than that of the Boros and Prekopa bound although both the ordered bounds reduce to the tight bound in (3.12).

References


