NEW VERSIONS OF SUEN'S CORRELATION INEQUALITY

SVANTE JANSON

1. Introduction

Suen [8] found a remarkable correlation inequality, giving estimates for the probability that a collection of dependent random indicator variables vanish simultaneously, or in other words, for the probability that none of a collection of dependent events occurs.

The present author [4, 3] has found similar inequalities for a much more restricted situation; when applicable, these inequalities are somewhat better than Suen's, although the difference is negligible in many cases. (See Section 8 below.) Those inequalities have been used by several different authors for a variety of problems; there are, however, many situations where they are not applicable (see [8, 5] for two examples) and then Suen's inequality is a very attractive choice.

The purpose of the present note is to present some improvements and modifications of Suen's original inequality which (we hope) will be easy to apply in different situations.

The estimates considered here are exponential (unlike for example Chebyshev's inequality), in the sense that they typically are similar to the estimate $\exp(-\mu)$ for the independent case, where μ is the expected number of events. They are thus aimed at the case when the studied probability is very small, and has to be shown to be very small. In many applications, constants occuring in the estimates, even in the exponents, are immaterial; on the other hand, there are applications where very precise estimates are desired. For this reason, and because different versions of the inequality turn out to be useful in different situations, we will give several different versions of our estimates.

We give several upper bounds to the probability of simultaneous vanishing of a collection of indicator variables in Section 3; these are perhaps the main results of the paper. We give some corresponding lower bounds in Section 4, and in Section 5 an upper bound for the probability that only a few of the variables are non-zero. Section 6 contains the proofs of the results, while Section 7 contains three examples related to the sharpness of the results. Finally, Section 8 contains a short discussion of the results and some open problems.

Acknowledgement. This paper has benefitted from discussions with participants of the Eighth International Conference on Random Structures and Algorithms in Poznań in August 1997. In particular, I am grateful to Joel Spencer, both for his pertinent questions that inspired parts of this work, and for him generously allowing me to include his proof of Theorem 7.

Date: September 23, 1997.

2. Notation

We will throughout the paper use the following assumptions and notation. Recall that an indicator (or Bernoulli) random variable is a random variable taking only the values 0 and 1.

- $\{I_i\}_{i\in\mathcal{I}}$ is a finite family of indicator random variables (defined on a common probability space).
- Γ is a *dependency graph* for $\{I_i\}_{i\in\mathcal{I}}$, i.e. a graph with vertex set $\mathcal{I}$ such that if A and B are two disjoint subsets of $\mathcal{I}$, and Γ contains no edge between A and B, then the families $\{I_i\}_{i\in A}$ and $\{I_i\}_{i\in B}$ are independent.
- $S = \sum_i I_i$. In particular $S = 0$ if and only if all $I_i = 0$. The expression $\mathbb{P}(S = 0)$ in the theorems below may thus be regarded as shorthand for $\mathbb{P}(\text{every } I_i = 0)$.
- $i \sim j$, where $i, j \in \mathcal{I}$, if there is an edge in Γ between i and j. (In particular, $i \not\sim i$.)
- $i \sim A$, where $i \in \mathcal{I}$ and $A \subseteq \mathcal{I}$, if $i \sim j$ for some $j \in A$. ($i \in A$ is not excluded.)
- $p_i = \mathbb{P}(I_i = 1) = \mathbb{E}\, I_i$. Thus $\mathbb{P}(I_i = 0) = 1 - p_i$.
- $\mu = \mathbb{E}\, S = \sum_i p_i$.
- $\delta_i = \sum_{j\sim i} p_j$.
- $\delta = \max_{i\in\mathcal{I}} \delta_i$.
- $\Delta = \sum_{\{i,j\}:i\sim j} \mathbb{E}(I_i I_j)$, summing over *unordered* pairs $\{i, j\}$, i.e. over the edges in Γ. As a sum over ordered pairs we have $\Delta = \frac{1}{2}\sum_{i\in\mathcal{I}}\sum_{j\sim i} \mathbb{E}(I_i I_j)$.
- $\Delta_0 = \sum_{\{i,j\}:i\sim j} p_i p_j = \frac{1}{2}\sum_{i\in\mathcal{I}} p_i \delta_i$. Note that $\Delta_0 \le \Delta$ if the variables are positively correlated, while $\Delta_0 \ge \Delta$ if the variables are negatively correlated.
- $\varepsilon = \max_{i\in\mathcal{I}} p_i$.

Remark 1. In typical applications, there exists a natural dependency graph, but it should be observed that in general there is no unique choice, even if it is required to be minimal.

Remark 2. In the Lovász local lemma, see e.g. [1], a weaker notion of dependency graph is used, where the independence condition above is required only when A is a singleton $\{I_i\}$ (but B still is an arbitrary subset). These strong and weak versions are not equivalent, but we do not know any application where the difference matters.

We do not know whether the results below hold if Γ only is assumed to be a dependency graph in the weaker sense.

Remark 3. In particular, two variables I_i and I_j are independent unless there is an edge in Γ between i and j. Note, however, that this weaker condition (i.e. the condition above when both A and B are restricted to singletons) does not imply that Γ is a dependency graph, and that it is not sufficient for our results.

For a simple counter example to many of the results below under this weaker condition, colour the vertices of the complete graph K_n at random (independently and with equal probabilities) with two colours, and let, for each edge

ij, I_{ij} be the indicator that i and j have different colours. These $\binom{n}{2}$ indicator variables are pairwise independent, so we could take Γ to be the empty graph with no edges, and then the upper bounds below are all less than 2^{-cn^2}, for some $c > 0$; on the other hand, clearly $S = 0$ if and only if all vertices have the same colour, and thus $\mathbb{P}(S = 0) = 2^{1-n}$.

Remark 4. In several related papers, Δ is instead defined as a sum over ordered pairs, yielding twice our value. This should be kept in mind when comparing results.

3. Upper bounds

As explained in the introduction, we will in this section give several different upper bounds for $\mathbb{P}(S = 0)$, suitable for different (present or future) applications.

We begin with a slight improvement of Suen's original inequality [8], essentially following (but sharpening) the original proof. (Suen [8] has $2(\mathbb{E}(I_iI_j) + p_ip_j)$ instead of $\mathbb{E}(I_iI_j)$ below; the p_ip_j is really needed only for his lower bound, see Section 4, while the factor 2 is removed by a more careful estimate in the proof.) For convenience, all proofs are given in Section 6.

Theorem 1. *Let $\{I_i\}_{i\in\mathcal{I}}$ be a finite family of indicator random variables having a dependency graph Γ. Then, with notations as above,*

$$\mathbb{P}(S = 0) \le \exp\Big(\sum_{\{i,j\}:i\sim j} \mathbb{E}(I_iI_j) \prod_{k\sim\{i,j\}} (1-p_k)^{-1}\Big) \prod_{l\in\mathcal{I}} (1-p_l).$$

Large products are often less convenient than large sums. For example, it is often convenient to use the standard estimate $\prod(1-p_i) \le \exp\big(-\sum p_i\big) = e^{-\mu}$. In a similar spirit, the estimate in Theorem 1 can be modified as follows, often without significant loss.

Theorem 2. *With assumptions as in Theorem* 1*,*

$$\mathbb{P}(S = 0) \le \exp\Big(-\mu + \sum_{\{i,j\}:i\sim j} \mathbb{E}(I_iI_j) \exp\Big(\sum_{k\sim\{i,j\}} p_k\Big)\Big) \le e^{-\mu+\Delta e^{2\delta}}.$$

Theorems 1 and 2 are useful, and often quite sharp, when $\Delta < \mu$ and δ is small. If $\Delta \ge \mu$, Theorem 2 (and typically Theorem 1 too) is worthless, since the right hand side becomes greater than 1. In this case, the following version becomes useful.

Theorem 3. *With assumptions as in Theorem* 1*,*

$$\mathbb{P}(S = 0) \le \exp\Big(-\min\Big(\frac{\mu^2}{8\Delta}, \frac{\mu}{6\delta}, \frac{\mu}{2}\Big)\Big) = e^{-\mu^2/\max(8\Delta,2\mu,6\delta\mu)}.$$

The numerical constants in the exponents can be improved, but there is (with our proof, at least) a trade off between the different constants, and the optimal choice depends on the relations between μ, Δ and δ. The version given here is a compromise trying to be both simple and reasonably sharp. In an application where constant factors in the exponent are important, it is probably better to use the proof below, in particular (8), and choose q there according to the situation at hand.

Of the three quantities in the exponents in Theorem 3, the one involving δ is typically the least important. If we do not care about the constants, we can always replace it by a term involving Δ_0, which in many applications is dominated by Δ. (Again, the constants given here are not optimal.)

Theorem 4. *With assumptions as in Theorem* 1*,*

$$\mathbb{P}(S=0) \le e^{-\mu^2/\max(32\Delta, 48\Delta_0, 4\mu)}.$$

Theorem 5. *With assumptions as in Theorem* 1*, if furthermore the variables* $\{I_i\}$ *are positively correlated,*

$$\mathbb{P}(S=0) \le e^{-\mu^2/\max(48\Delta, 4\mu)}.$$

We do not know whether the terms involving δ really are needed in the estimates above, cf. Section 8. In most applications they are harmless, but in at least one [5] they affect the final result significantly. For that reason, we give two slightly stronger versions of Theorems 1 and 2, where this term is reduced as much as we have been able to achieve.

Define

$$\varphi_1(x) = 2\int_0^1 te^{tx}\,dt = 2\frac{xe^x - e^x + 1}{x^2} = e^x \frac{e^{-x} - 1 + x}{x^2/2}$$

and observe that $1 \le \varphi_1(x) \le e^x$ for $x \ge 0$. For small x, $\varphi_1(x) = 1 + \frac{2}{3}x + O(x^2)$.

Theorem 6. *With assumptions as in Theorem* 1*,*

$$\begin{aligned}\mathbb{P}(S=0) &\le \exp\Big(\tfrac12 \sum_i \sum_{j\sim i} \mathbb{E}(I_iI_j)(1-p_j)^{-1}\varphi_1\Big(\sum_{k\sim\{i,j\},\,k\neq i,j} p_k/(1-p_k)\Big)\Big) \\ &\qquad\qquad \times \prod_{l\in\mathcal{I}}(1-p_l) \\ &\le \exp\Big(\Delta(1-\varepsilon)^{-1}\varphi_1\big(2\delta/(1-\varepsilon)\big)\Big)\prod_{l\in\mathcal{I}}(1-p_l)\end{aligned}$$

and

$$\mathbb{P}(S=0) \le \exp\big(-\mu + e^{\varepsilon}\varphi_1(2e^{\varepsilon}\delta)\Delta\big).$$

Joel Spencer [6] has found a different proof of this type of inequalities, using only elementary probability calculations. This proof yields the following result, which sometimes is better than Theorem 6; when ε is negligible, Theorem 7 is better when $0 < \delta < 0.225\ldots$ while Theorem 6 is better for larger δ. (This strongly suggests that neither is the best possible, and that there is room for future improvements of these results.)

Define $\varphi_2(x)$, $0 \le x \le e^{-1}$, to be the smallest root of

$$\varphi_2(x) = e^{x\varphi_2(x)}. \tag{1}$$

It is well known that φ_2 is well defined in $[0, e^{-1}]$, with the Taylor series

$$\varphi_2(x) = \sum_0^\infty \frac{(n+1)^{n-1}}{n!} x^n, \qquad 0 \le x \le e^{-1};$$

in particular, $\varphi_2(x) = 1 + x + O(x^2)$.

Theorem 7 (Spencer [6])**.** *With assumptions as in Theorem* 1*, if furthermore* $\delta + \varepsilon \le e^{-1}$,

$$\mathbb{P}(S = 0) \le e^{\Delta\varphi_2(\delta+\varepsilon)} \prod_{k \in \mathcal{I}} (1 - p_k) \le e^{-\mu + \Delta\varphi_2(\delta+\varepsilon)}.$$

□

Remark 5. As is seen from the proof of Theorem 7, further improvements are possible if the dependency graph is such that the neigbourhoods of two adjacent vertices overlap significantly; cf. Theorem 6, where such an effect can be seen in the first upper bound.

4. Lower bounds

Suen [8] gave also a lower bound to $\mathbb{P}(S = 0)$ corresponding to his upper bound. Although the lower bound seems to have less applications, we for completeness give an improvement of it too corresponding to Theorem 1. (We ignore further improvements corresponding to Theorem 6.)

Theorem 8. *With assumptions as in Theorem* 1*, let*

$$\Delta^* = \sum_{\{i,j\}: i \sim j} \mathbb{E}(I_i I_j) \prod_{k \sim \{i,j\}} (1 - p_k)^{-1},$$

$$\Delta_0^* = \sum_{\{i,j\}: i \sim j} p_i p_j \prod_{k \sim \{i,j\}} (1 - p_k)^{-1}.$$

Then

$$\mathbb{P}(S = 0) \ge \Big(1 - \Delta_0^* \exp(\Delta^*)\Big) \prod_{k \in \mathcal{I}} (1 - p_k)$$

$$\ge \Big(1 - \Delta_0 e^{2\delta/(1-\varepsilon)} \exp\big(\Delta e^{2\delta/(1-\varepsilon)}\big)\Big) \prod_{k \in \mathcal{I}} (1 - p_k).$$

Remark 6. The careful reader may observe from the proof that the factor $\exp(\Delta^*)$ can be replaced by the somewhat smaller $\varphi_3(\Delta^*)$, where $\varphi_3(x) = (e^x - 1)/x$. (We do not know whether this factor depending on Δ^* really is needed at all, or whether it is an artefact of our proof.)

Moreover, since $\Delta_0^* \exp(\Delta^*) \le \exp(\Delta_0^* + \Delta^*) - \exp(\Delta^*) \le \exp(\Delta_0^* + \Delta^*) - 1$, the factor $1 - \Delta_0^* \exp(\Delta^*)$ may be replaced by the smaller $2 - \exp(\Delta^* + \Delta_0^*)$; Suen [8] has $2 - \exp(2\Delta^* + 2\Delta_0^*)$.

Theorem 8 is useful only when $\Delta_0 < 1$ and Δ is small, but even then it is often surpassed by the following quantitative version of the Lovász local lemma, cf. [1, Chapter 5].

Theorem 9. *With assumptions as in Theorem* 1*, suppose further that* $\delta+\varepsilon \le e^{-1}$*. Then, with* φ_2 *defined by* (1)*,*

$$\mathbb{P}(S=0) \ge \exp\bigl(-\mu\varphi_2(\delta+\varepsilon)\bigr).$$

5. Tail estimates

It is also possible to obtain exponential upper bounds for the lower tails $\mathbb{P}(S \le \lambda)$, $\lambda < \mu$, of the distribution of S, similar to the bounds given above for $\mathbb{P}(S=0)$. (There are no similar general bounds for the upper tails $\mathbb{P}(S \ge \lambda)$, $\lambda > \mu$, see [3, Example 2].)

Theorem 10. *With assumptions as in Theorem* 1*, and* $0 \le a \le 1$*,*

$$\mathbb{P}(S \le a\mu) \le \exp\Bigl(-\min\Bigl((1-a)^2\frac{\mu^2}{8\Delta+2\mu}, (1-a)\frac{\mu}{6\delta}\Bigr)\Bigr).$$

Note that the special case $a=0$ yields the corollary

$$\mathbb{P}(S=0) \le e^{-\mu^2/\max(8\Delta+2\mu, 6\delta\mu)},$$

which is only slightly weaker than Theorem 3.

Remark 7. The tail bounds given in [3] for a special case are of the same type as Theorem 10 but somewhat better; there 8Δ is replaced by 4Δ and there is no term with δ.

6. Proofs

We define for each index $i \in \mathcal{I}$ a partition $\mathcal{I} = \{i\} \cup N_i \cup U_i$ of the index set, with $N_i = \{j \in \mathcal{I} : j \sim i\}$ (the neighbours of i in Γ) and thus $U_i = \{j \ne i : j \not\sim i\}$; hence I_i is independent of $\{I_j : j \in U_i\}$.

Proof of Theorem 1. Define, for $0 \le t \le 1$, the random function

$$F(t) = \prod_{i \in \mathcal{I}} \bigl(1 - p_i - t(I_i - p_i)\bigr).$$

Thus $F(0) = \prod_i (1-p_i)$ and $F(1) = \prod_i (1-I_i)$, so $\mathbb{E}\,F(1) = \mathbb{P}(S=0)$. We thus want to compare $\mathbb{E}\,F(1)$ and $\mathbb{E}\,F(0) = F(0)$. (Note that in the case of independent indicators I_i, $\mathbb{E}\,F(t)$ is independent of t.)

We differentiate and obtain, introducing the notation $F_A(t) = \prod_{i\in A}\bigl(1-p_i-t(I_i-p_i)\bigr)$ for $A \subseteq \mathcal{I}$,

$$\begin{aligned} F'(t) &= -\sum_i (I_i - p_i)\prod_{j\ne i}\bigl(1-p_j-t(I_j-p_j)\bigr) = -\sum_i (I_i-p_i)F_{\mathcal{I}\setminus\{i\}}(t) \\ &= -\sum_i (I_i-p_i)F_{N_i}(t)F_{U_i}(t). \end{aligned} \tag{2}$$

Note that, for $0 \le t \le 1$ and each j,

$$0 \le 1 - p_j - t(I_j - p_j) \le 1 - p_j + tp_j \le 1. \tag{3}$$

Hence, for any set $A \subseteq \mathcal{I}$,

$$0 \leq \prod_{j\in A}(1-p_j+tp_j) - F_A(t) \leq \sum_{j\in A} tI_j, \tag{4}$$

and thus, considering the cases $I_i = 0$ and $I_i = 1$ separately,

$$(I_i - p_i)\Bigl(\prod_{j\in A}(1-p_j+tp_j) - F_A(t)\Bigr) \leq I_i \sum_{j\in A} tI_j. \tag{5}$$

Choosing $A = N_i$, this and (2) yield

$$F'(t) \leq -\sum_i F_{U_i}(t)(I_i - p_i)\prod_{j\in N_i}(1-p_j+tp_j) + \sum_i F_{U_i}(t) I_i \sum_{j\in N_i} tI_j.$$

Since, by the definition of a dependency graph, I_i and $F_{U_i}(t)$ are independent, $\mathbb{E}\bigl(F_{U_i}(t)(I_i - p_i)\bigr) = 0$. Moreover, using (3), $F_{U_i}(t) \leq F_{U_i\cap U_j}(t)$, which is independent of I_iI_j. Hence, for $0 \leq t \leq 1$,

$$\begin{aligned}\mathbb{E}\,F'(t) &\leq \sum_i\sum_{j\sim i} t\,\mathbb{E}(I_iI_jF_{U_i}(t)) \leq \sum_i\sum_{j\sim i} t\,\mathbb{E}(I_iI_jF_{U_i\cap U_j}(t)) \\ &= t\sum_i\sum_{j\sim i}\mathbb{E}(I_iI_j)\,\mathbb{E}\,F_{U_i\cap U_j}(t). \end{aligned} \tag{6}$$

Now, let $y_{ij} = \mathbb{E}(I_iI_j)\prod_{k\sim\{i,j\}}(1-p_k)^{-1}$ and $\Delta^* = \sum_{\{i,j\}:i\sim j} y_{ij}$. We claim that

$$\mathbb{E}\,F(t) \leq e^{t^2\Delta^*}\prod_{k\in\mathcal{I}}(1-p_k), \qquad 0 \leq t \leq 1. \tag{7}$$

In fact, we may by induction over $|\mathcal{I}|$, the number of variables, assume that the corresponding inequality holds for $\mathbb{E}\,F_A(t)$ for every proper subset A of $\mathcal{I}$. Since the corresponding values $y_{ij}^{(A)}$ for a subset A (and $i,j\in A$) satisfy $y_{ij}^{(A)} \leq y_{ij}$, it then follows from (6) that, using also that when $i\sim j$, $\mathcal{I}\setminus(U_i\cap U_j) = N_i\cup N_j = \{k : k\sim\{i,j\}\}$,

$$\begin{aligned}\mathbb{E}\,F'(t) &\leq t\sum_i\sum_{j\sim i}\mathbb{E}(I_iI_j)e^{t^2\Delta^*}\prod_{k\in U_i\cap U_j}(1-p_k) = t\sum_i\sum_{j\sim i} y_{ij}e^{t^2\Delta^*}\prod_{k\in\mathcal{I}}(1-p_k) \\ &= 2t\Delta^* e^{t^2\Delta^*}\prod_{k\in\mathcal{I}}(1-p_k) = \frac{d}{dt}e^{t^2\Delta^*}\prod_{k\in\mathcal{I}}(1-p_k).\end{aligned}$$

The estimate (7) now follows by integration, since it holds (with equality) for $t = 0$.

The theorem follows by choosing $t = 1$ in (7). □

Remark 8. If all p_i are equal, $p_i = p$ say, then $F(t) = (1-p+pt)^{|\mathcal{I}|}e^{-uS}$, with $u = \ln\bigl((1-p+tp)/(1-p)(1-t)\bigr)$; hence estimating $\mathbb{E}\,F(t)$, $0 \leq t \leq 1$, is equivalent to estimating the Laplace transform $\mathbb{E}\,e^{-uS}$, $u \geq 0$. In general, with unequal p_i, $\mathbb{E}\,F(t)$ is not equivalent to the Laplace transform; nevertheless, $\mathbb{E}\,e^{-uS}$ may be estimated by arguments similar to the ones above, which leads to results similar to the ones given here. It seems, however, that estimates of

F (as in Suen's paper [8]) lead to slightly stronger results. For applications of the Laplace transform, see the proof of Theorem 10 below, and also [4, 3] where it is used advantageously in a special case.

We derive Theorems 2–5 as corollaries to Theorem 1 by performing a random thinning of the family $\{I_i\}$. (This idea is used by Alon and Spencer [1] for a related inequality.) Let $q_i \in [0,1]$ and let $J_i \sim \mathrm{Be}(q_i)$ be Bernoulli variables that are independent of each other and of $\{I_i\}$. Define $I'_i = I_i J_i$ and $S' = \sum_{i\in\mathcal{I}} I'_i$. Clearly, $S' \le S$ and thus $\mathbb{P}(S=0) \le \mathbb{P}(S'=0)$. Moreover, Γ is a dependency graph for $\{I'_i\}$ too. Consequently, we may apply our estimates to the family $\{I'_i\}$ and obtain new estimates for $\mathbb{P}(S=0)$; with suitable choices of q_i, this may improve the original estimate.

Let $p'_i = \mathbb{E}\, I'_i = p_i q_i$, thus $0 \le p'_i \le p_i$.

Proof of Theorem 2. If we choose $q_i = (1-e^{-p_i})/p_i$, and thus $p'_i = 1-e^{-p_i}$, then Theorem 1 applied to $\{I'_i\}$ yields

$$\mathbb{P}(S=0) \le \mathbb{P}(S'=0) \le \exp\Big(\sum_{\{i,j\}:i\sim j} q_i q_j\, \mathbb{E}(I_i I_j) \prod_{k\sim\{i,j\}} e^{p_k}\Big) \prod_{l\in\mathcal{I}} e^{-p_l},$$

which implies the first inequality.

The final estimate follows because $\sum_{k\sim\{i,j\}} p_k \le \delta_i + \delta_j \le 2\delta$. □

Proof of Theorem 3. Using a new thinning $\{I'_i\}$, this time choosing all $q_i = q$ for some $q\in[0,1]$, we obtain from Theorem 2 applied to $\{I'_i\}$

$$\mathbb{P}(S=0) \le \mathbb{P}(S'=0) \le e^{-q\mu+q^2\Delta e^{2q\delta}}. \tag{8}$$

We choose $q = \min(\mu/4\Delta, 1/3\delta, 1)$; then $q\Delta \le \mu/4$ and $e^{2q\delta} \le e^{2/3} < 2$, and consequently (8) yields

$$\mathbb{P}(S=0) \le e^{-q\mu+\frac12 q\mu} = e^{-\frac12 q\mu},$$

which is the desired result. □

Proof of Theorem 4. This time we use a deterministic thinning $\{I_i\}_{i\in\mathcal{J}}$, where

$$\mathcal{J} = \{i\in\mathcal{I} : \delta_i \le 4\Delta_0/\mu\}.$$

We let μ', Δ' and δ' denote the quantities for our subfamily $\{I_i\}_{i\in\mathcal{J}}$ corresponding to μ, Δ and δ. By our choice of $\mathcal{J}$,

$$\sum_{i\notin\mathcal{J}} p_i \le \frac{\mu}{4\Delta_0}\sum_{i\in\mathcal{I}} p_i\delta_i = \frac{\mu}{4\Delta_0}2\Delta_0 = \frac{\mu}{2}$$

and thus

$$\mu' = \sum_{i\in\mathcal{J}} p_i = \sum_{i\in\mathcal{I}} p_i - \sum_{i\notin\mathcal{J}} p_i \ge \mu - \frac{\mu}{2} = \frac{\mu}{2}.$$

Moreover, clearly $\delta' \le 4\Delta_0/\mu$ and $\Delta' \le \Delta$.

Consequently, Theorem 3 applied to $S' = \sum_{i\in\mathcal{J}} I_i$ yields

$$\mathbb{P}(S=0) \le \mathbb{P}(S'=0) \le e^{-\min\left(\frac{\mu'^2}{8\Delta'},\frac{\mu'}{6\delta'},\frac{\mu'}{2}\right)} \le e^{-\min\left(\frac{\mu^2}{32\Delta},\frac{\mu}{48\Delta_0/\mu},\frac{\mu}{4}\right)}.$$

□

Proof of Theorem 5. When the variables are positively correlated, $\mathbb{E}(I_iI_j) \ge p_ip_j$ and thus $\Delta \ge \Delta_0$. □

Proof of Theorem 6. We follow the proof of Theorem 1, improving some estimates. We define, for $A \subseteq \mathcal{I}$, $\beta_A(t) = \prod_{i\in A}(1-p_i+tp_i)$. Using (3), we can improve (4) to

$$\prod_{j\in A}(1-p_j+tp_j) - F_A(t) \le \sum_{j\in A} tI_j\beta_{A\setminus\{j\}}(t),$$

and then (5) to

$$(I_i-p_i)\Bigl(\prod_{j\in A}(1-p_j+tp_j) - F_A(t)\Bigr) \le I_i(1-p_i)\sum_{j\in A} tI_j\beta_{A\setminus\{j\}}(t). \tag{9}$$

Moreover, $F_{U_i}(t) \le F_{U_i\cap U_j}(t)\beta_{U_i\setminus U_j}(t)$, which together with (9) yields as an improved version of (6), for $0\le t\le 1$,

$$\mathbb{E}\,F'(t) \le t\sum_i\sum_{j\sim i}\mathbb{E}(I_iI_j)\,\mathbb{E}\,F_{U_i\cap U_j}(t)(1-p_i)\beta_{N_i\setminus\{j\}}(t)\beta_{U_i\setminus U_j}(t). \tag{10}$$

We now define, when $i\sim j$, $N_{ij} = (N_i\setminus\{j\})\cup(U_i\setminus U_j) = N_i\cup N_j\setminus\{i,j\}$,

$$\psi(t) = t\sum_i\sum_{j\sim i}\mathbb{E}(I_iI_j)(1-p_j)^{-1}\prod_{k\in N_{ij}}\bigl(1+tp_k/(1-p_k)\bigr)$$

and

$$\Psi(t) = \int_0^t \psi(s)\,ds.$$

It then follows from (10) by induction as in the proof of Theorem 1 that

$$\mathbb{E}\,F(t) \le e^{\Psi(t)}\prod_{k\in\mathcal{I}}(1-p_k), \qquad 0\le t\le 1. \tag{11}$$

The first claim in the theorem now follows by choosing $t=1$ in (11), observing that

$$\psi(t) \le t\sum_i\sum_{j\sim i}\mathbb{E}(I_iI_j)(1-p_j)^{-1}\exp\Bigl(\sum_{k\in N_{ij}} tp_k/(1-p_k)\Bigr)$$

and thus

$$\Psi(1) = \int_0^1\psi(t)\,dt \le \sum_i\sum_{j\sim i}\mathbb{E}(I_iI_j)(1-p_j)^{-1}\tfrac12\varphi_1\Bigl(\sum_{k\in N_{ij}} p_k/(1-p_k)\Bigr).$$

The second inequality in the theorem is immediate.

For the final estimate, we perform a random thinning with $p_i' = 1-e^{-p_i}$ as in the proof of Theorem 2; we leave the details to the reader. □

For Theorem 7 we use the following form of the Lovász local lemma, which is easily proved by induction, see [1, Lemma 5.1.1 and its proof]. Let, for $A\subseteq\mathcal{I}$, $S_A = \sum_{i\in A} I_i$.

Lemma 1. *With assumptions as in Theorem* 1*, suppose further that* x_i*,* $i \in \mathcal{I}$*, are real numbers such that* $0 \le x_i < 1$ *and*

$$p_i \le x_i \prod_{j\sim i}(1-x_j), \qquad i \in \mathcal{I}. \tag{12}$$

Then, for any two subsets $A, B \subseteq \mathcal{I}$*,* $\mathbb{P}(S_A = 0 \mid S_B = 0) \ge \prod_{i\in A}(1-x_i)$*.* □

We will employ Lemma 1 with the following choice of x_i.

Lemma 2. *With assumptions as in Theorem* 1*, suppose further that* $\delta + \varepsilon \le e^{-1}$*. Let* $\varphi = \varphi_2(\delta+\varepsilon)$*, where* φ_2 *is defined by* (1)*, and let* $x_i = 1 - e^{-\varphi p_i}$*,* $i \in \mathcal{I}$*. Then* (12) *holds and, for any two subsets* $A, B \subseteq \mathcal{I}$*,*

$$\mathbb{P}(S_A = 0 \mid S_B = 0) \ge \prod_{i\in A}(1-x_i) = e^{-\varphi\sum_{i\in A}p_i}.$$

Proof of Lemma 2. By (1), $e^{(\delta+\varepsilon)\varphi} = \varphi$, and thus

$$x_i \prod_{j\sim i}(1-x_j) = (1-e^{-\varphi p_i})e^{-\varphi\delta_i} = (e^{\varphi p_i}-1)e^{-(\delta_i+p_i)\varphi} \ge \varphi p_i e^{-(\delta+\varepsilon)\varphi} = p_i.$$

Consequently (12) holds, and the final inequality follows by Lemma 1. □

Proof of Theorem 7 (*after Spencer* [6]). The main idea is to estimate the conditional probability $\mathbb{P}(I_i = 0 \mid S_{\mathcal{I}\setminus\{i\}} = 0)$. Recall that $S_{\mathcal{I}\setminus\{i\}} = S_{N_i} + S_{U_i}$, and that I_i and S_{U_i} are independent; furthermore, note the elementary identity $\mathbb{P}(E_1 \mid E_2 \text{ and } E_3) = \mathbb{P}(E_1 \text{ and } E_2 \mid E_3)/\mathbb{P}(E_2 \mid E_3)$ for any three events E_1, E_2, E_3. Consequently, we have the estimate

$$\begin{aligned}\mathbb{P}(I_i = 1 \mid S_{\mathcal{I}\setminus\{i\}} = 0) &= \frac{\mathbb{P}(I_i = 1 \text{ and } S_{N_i} = 0 \mid S_{U_i} = 0)}{\mathbb{P}(S_{N_i} = 0 \mid S_{U_i} = 0)}\\ &\ge \mathbb{P}(I_i = 1 \text{ and } S_{N_i} = 0 \mid S_{U_i} = 0)\\ &= \mathbb{P}(I_i = 1 \mid S_{U_i} = 0) - \mathbb{P}\Big(\bigcup_{j\sim i}\{I_i = I_j = 1\} \mid S_{U_i} = 0\Big)\\ &\ge \mathbb{P}(I_i = 1) - \sum_{j\sim i}\mathbb{P}(I_i = I_j = 1 \mid S_{U_i} = 0).\end{aligned} \tag{13}$$

Similarly, when $j \sim i \in \mathcal{I}$, since $\{I_i, I_j\}$ and $S_{U_i\cap U_j}$ are independent,

$$\begin{aligned}\mathbb{P}(I_i = I_j = 1 \mid S_{U_i} = 0) &= \frac{\mathbb{P}(I_i = I_j = 1 \text{ and } S_{U_i\cap N_j} = 0 \mid S_{U_i\cap U_j} = 0)}{\mathbb{P}(S_{U_i\cap N_j} = 0 \mid S_{U_i\cap U_j} = 0)}\\ &\le \frac{\mathbb{P}(I_i = I_j = 1 \mid S_{U_i\cap U_j} = 0)}{\mathbb{P}(S_{U_i\cap N_j} = 0 \mid S_{U_i\cap U_j} = 0)} = \frac{\mathbb{P}(I_i = I_j = 1)}{\mathbb{P}(S_{U_i\cap N_j} = 0 \mid S_{U_i\cap U_j} = 0)}.\end{aligned}$$

Applying Lemma 2, we thus obtain

$$\mathbb{P}(I_i = I_j = 1 \mid S_{U_i} = 0) \le \mathbb{P}(I_i = I_j = 1)e^{\varphi\sum_{k\in U_i\cap N_j}p_k}. \tag{14}$$

Note that $p_i \le x_i$ by (12), and thus $1 = (1-x_i)e^{\varphi p_i} \le (1-p_i)e^{\varphi p_i}$; moreover $(N_j \cap U_i) \cup \{i\} = N_j \setminus N_i$ provided $i \sim j$. By combining (13) and (14) we thus obtain the sought estimate

$$\begin{aligned}
\mathbb{P}(I_i = 0 \mid S_{\mathcal{I}\setminus\{i\}} = 0) &\le 1 - p_i + \sum_{j\sim i} \mathbb{P}(I_i = I_j = 1 \mid S_{U_i} = 0) \\
&\le 1 - p_i + (1-p_i)e^{\varphi p_i} \sum_{j\sim i} \mathbb{P}(I_i = I_j = 1) e^{\varphi \sum_{k\in U_i\cap N_j} p_k} \\
&= (1-p_i)\Big(1 + \sum_{j\sim i} \mathbb{E}(I_i I_j) e^{\varphi \sum_{k\in N_j\setminus N_i} p_k}\Big) \\
&\le (1-p_i)\exp\Big(\sum_{j\sim i} \mathbb{E}(I_i I_j) e^{\varphi \sum_{k\in N_j\setminus N_i} p_k}\Big).
\end{aligned} \tag{15}$$

For the final step, we may assume that the index set $\mathcal{I} = \{1,\dots,n\}$. We apply (15) to the subset $\{1,\dots,i\}$ of $\mathcal{I}$ and obtain, for $i = 1,\dots,n$,

$$\begin{aligned}
\mathbb{P}(I_i = 0 \mid I_1 = \cdots = I_{i-1} = 0) &\le (1-p_i)\exp\Big(\sum_{j\sim i,\, j<i} \mathbb{E}(I_i I_j) e^{\varphi \sum_{k\in N_j\setminus N_i} p_k}\Big) \\
&\le (1-p_i)\exp\Big(\sum_{j\sim i,\, j<i} \mathbb{E}(I_i I_j) e^{\varphi \delta_j}\Big).
\end{aligned}$$

The result follows by multiplying these inequalities for $i = 1,\dots,n$, and using $e^{\varphi\delta_j} \le e^{(\delta+\varepsilon)\varphi} = \varphi$. ∎

Proof of Theorem 8. We follow the proof of Theorem 1. The main idea is that there is a lower bound corresponding to (5), viz.

$$(I_i - p_i)\Big(\prod_{j\in A}(1 - p_j + tp_j) - F_A(t)\Big) \ge -p_i \sum_{j\in A} tI_j.$$

The same arguments that led to (6) then yield (we leave the details to the reader), for $0 \le t \le 1$,

$$\mathbb{E}\, F'(t) \ge -t \sum_i \sum_{j\sim i} p_i p_j\, \mathbb{E}\, F_{U_i\cap U_j}(t).$$

We now use the *upper* bound to $\mathbb{E}\, F_{U_i\cap U_j}(t)$ proved in (7) and obtain

$$\begin{aligned}
\mathbb{E}\, F'(t) &\ge -t \sum_i \sum_{j\sim i} p_i p_j e^{t^2\Delta^*} \prod_{k\in U_i\cap U_j} (1-p_k) \\
&= -2t\Delta_0^* e^{t^2\Delta^*} \prod_{k\in\mathcal{I}}(1-p_k) \ge -2t\Delta_0^* e^{\Delta^*} \prod_{k\in\mathcal{I}}(1-p_k).
\end{aligned}$$

Integrating this for $0 \le t \le 1$, we obtain

$$F(1) - F(0) \ge -\Delta_0^* e^{\Delta^*} \prod_{k\in\mathcal{I}}(1-p_k).$$

Since $\mathbb{P}(S=0)=F(1)$ and $F(0)=\prod_{k\in\mathcal{I}}(1-p_k)$, this yields the first inequality in the statement. The second follows by the estimate $(1-p_k)^{-1}\le \exp\bigl(p_k/(1-p_k)\bigr)\le \exp\bigl(p_k/(1-\varepsilon)\bigr)$. □

Proof of Theorem 9. This is just the special case $A=\mathcal{I}$ and $B=\emptyset$ of Lemma 2. □

Proof of Theorem 10. We use the standard method of first estimating the Laplace transform $\mathbb{E}\, e^{-tS}$ and then applying Markov's inequality. As remarked above, it is possible to estimate the Laplace transform directly by the method of the proof of Theorem 1. For simplicity, and to avoid repetitions, we will, however, instead use the results above together with another thinning.

Define thinned variables $I_i'=J_iI_i$ as above with $q_i=1-e^{-t}$ for every $i\in\mathcal{I}$. Since the thinned sum S' equals 0 if and only if J_iI_i equals 0 for all i, i.e., if and only if $J_i=0$ for the S indices i with $I_i=1$, the conditional probability $\mathbb{P}(S'=0\mid S)=e^{-tS}$ and thus

$$\mathbb{P}(S'=0)=\mathbb{E}\, e^{-tS}.$$

Now, for the thinned family $\{I_i'\}$, $\mu'=(1-e^{-t})\mu\ge (t-\frac{t^2}{2})\mu$, $\Delta'=(1-e^{-t})^2\Delta\le t^2\Delta$ and $\delta'=(1-e^{-t})\delta\le t\delta$. Consequently, Theorem 2 yields

$$\mathbb{E}\, e^{-tS}=\mathbb{P}(S'=0)\le e^{-\mu'+\Delta' e^{2\delta'}}\le e^{-t\mu+t^2(\frac{1}{2}\mu+\Delta e^{2t\delta})},$$

and thus by Markov's inequality, for any $t\ge 0$,

$$\mathbb{P}(S\le a\mu)\le \mathbb{E}\, e^{ta\mu-tS}\le e^{-t(1-a)\mu+t^2(\frac{1}{2}\mu+\Delta e^{2t\delta})}. \tag{16}$$

We choose here $t=\min\bigl((1-a)\mu/(4\Delta+\mu),1/3\delta\bigr)$; then $e^{2t\delta}\le e^{2/3}<2$, and thus

$$t^2(\tfrac{1}{2}\mu+\Delta e^{2t\delta})\le \tfrac{1}{2}t^2(\mu+4\Delta)\le \tfrac{1}{2}t(1-a)\mu.$$

Consequently, (16) yields

$$\mathbb{P}(S\le a\mu)\le e^{-t(1-a)\mu/2},$$

which is the claimed result. □

7. Examples

We give three examples showing that the estimates above are of the right order, and that no dramatic improvements are possible in general. We begin with two simple examples.

Example 1 (positive correlations, upper bounds)**.** Let $m,n\ge 1$ and $0<p<1$, and consider mn $\mathrm{Be}(p)$ variables consisting of n groups of m variables each, such that different groups are independent while the m variables in each group are *identical*. Let Γ be the obvious dependency graph consisting of n disjoint copies of the complete graph K_m. Then $\mu=mnp$, $\Delta=\binom{m}{2}np$, $\delta=(m-1)p$, $\varepsilon=p$, and $\mathbb{P}(S=0)=(1-p)^n$.

Consider first the case $m=2$, and thus $\mu=2np$, $\Delta=np$. Theorem 1 yields

$$\mathbb{P}(S=0)\le \exp\bigl(np(1-p)^{-2}\bigr)(1-p)^{2n};$$

the ratio between the right and left hand sides is (for $p \le 1/2$, say)

$$\begin{aligned}\exp\bigl(np(1-p)^{-2}\bigr)(1-p)^n &= \exp\bigl(np(1+2p+O(p^2)) - n(p+\tfrac12 p^2+O(p^3))\bigr)\\ &= \exp\bigl(\tfrac32 np^2 + O(np^3)\bigr).\end{aligned}$$

Note that the leading terms np in the exponent cancels, which shows that the estimate is quite sharp. Note also that the term $\prod_{k\sim\{i,j\}}(1-p_k)^{-1}$ in the exponent in Theorem 1 cannot be eliminated completely; in this example

$$\mathbb{P}(S=0) > \exp(\Delta)\prod_{l\in\mathcal{I}}(1-p_l). \tag{17}$$

Theorem 2 yields the slightly weaker

$$\mathbb{P}(S=0) \le \exp\bigl(-2np + npe^{2p}\bigr) = \exp\bigl(\tfrac52 np^2 + O(np^3)\bigr)\,\mathbb{P}(S=0).$$

The improved bound in Theorem 6 is

$$\exp\bigl(np(1-p)^{-1}\bigr)(1-p)^{2n} = \exp\bigl(\tfrac12 np^2 + O(np^3)\bigr)\,\mathbb{P}(S=0),$$

while Theorem 7 in this example yields (for $p \le 1/2e$) the weaker

$$\exp\bigl(np\varphi_2(2p)\bigr)(1-p)^{2n} = \exp\bigl(\tfrac32 np^2 + O(np^3)\bigr)\,\mathbb{P}(S=0).$$

Let us now consider $m \ge 3$. Then $\Delta \ge \mu$, and the bounds in Theorems 1 and 2 are > 1. For any $m \ge 2$, and $p \le 2/3$, Theorem 3 yields

$$\mathbb{P}(S=0) \le \exp(-\mu^2/8\Delta) = \exp\Bigl(-\frac{m}{4(m-1)}np\Bigr),$$

which for small p is off by a constant factor in the exponent. Theorems 4 and 5 yield similar results with even worse constants in the exponent.

Example 2 (negative correlations, lower bounds)**.** Let $m, n \ge 1$ and $0 < p \le 1/m$, and consider again mn $\mathrm{Be}(p)$ variables consisting of n independent groups of m variables each, this time assuming that the m variables in each group are such that never more than one of them equals 1. Then with Γ as in the preceding example, $I_iI_j = 0$ whenever $i \sim j$, and thus $\Delta = 0$; moreover, $\mu = mnp$, $\Delta_0 = \binom{m}{2}np^2$, $\delta = (m-1)p$, $\varepsilon = p$, and $\mathbb{P}(S=0) = (1-mp)^n$.

Theorem 8 yields

$$\mathbb{P}(S=0) \ge \Bigl(1 - \binom{m}{2}np^2(1-p)^{-m}\Bigr)(1-p)^{mn}$$

where the ratio between the right and left hand sides is (for $mp \le 1/2$ and $nm^2p^2 \le 1$, say)

$$\begin{aligned}\exp\Bigl(-\binom{m}{2}np^2(1+O(mp+nm^2p^2)) &- mn\bigl(p+\tfrac12 p^2+O(p^3)\bigr)\\ &+ n\bigl(mp+\tfrac12 m^2p^2+O(m^3p^3)\bigr)\Bigr)\\ &= \exp\bigl(O(nm^3p^3+n^2m^4p^4)\bigr).\end{aligned}$$

This is quite sharp for small mp provided n is not too large, but fails utterly when $nm(m-1)p^2 \geq 2$, and thus $\Delta_0 \geq 1$.

Theorem 9 yields, for $mp \leq 1/2e$,

$$\begin{aligned}\mathbb{P}(S=0) &\geq \exp\bigl(-mnp\varphi_2(mp)\bigr) = \exp\bigl(-mnp - nm^2p^2 + O(nm^3p^3)\bigr)\\ &= \exp\bigl(-\tfrac{1}{2}nm^2p^2 + O(nm^3p^3)\bigr)\,\mathbb{P}(S=0).\end{aligned}$$

This too is quite sharp when mp is small. It performs better than Theorem 9 when n is large, but is not as good when both mp and nm^2p^2 are small.

Example 3 (negative correlations, lower bound)**.** This more complicated example is due to Shearer [7], to which we refer for further details; it shows that the condition $\delta + \varepsilon \leq e^{-1}$ in Theorem 9 is best possible. (We make a minor modification since Shearer only considers the weak version of dependency graphs, cf. Remark 2.) More precisely:

Claim. *If $a > e^{-1}$, then there exists a finite family $\{I_i\}_{i\in\mathcal{I}}$ having a dependency graph Γ such that $\delta + \varepsilon < a$ and $\mathbb{P}(S=0) = 0$.*

Shearer's construction as is follows. Given a graph Γ and $0 \leq p \leq 1$, let V be the vertex set of Γ, say that a subset of V is dependent if it contains both endpoints of some edge in Γ (and independent otherwise), and define a signed measure $\mu_{\Gamma,p}$ on 2^V, the space of all subsets of V, such that for each $A \subseteq V$

$$\mu_{\Gamma,p}\{E \subseteq V : E \supseteq A\} = \begin{cases} p^{|A|} & A \text{ is independent},\\ 0 & A \text{ is dependent}.\end{cases}$$

It is easily seen that this defines a unique measure $\mu_{\Gamma,p}$, since $\mu_{\Gamma,p}(\{A\})$ may be expressed by inclusion-exclusion for every $A \subseteq V$. $\mu_{\Gamma,p}$ may assume negative values, but if it is a positive measure, i.e. if $\mu_{\Gamma,p}(\{A\}) \geq 0$ for every $A \subseteq V$, then $\mu_{\Gamma,p}$ is a probability measure on 2^V, and it is easily seen that with this measure, the indicators $I_i = \mathbf{1}(i \in E)$, $i \in V$, form a family of random indicator variables with dependency graph Γ; moreover, then $p_i = \mathbb{E}\, I_i = p$ for each i, and $\mathbb{E}(I_iI_j) = 0$ whenever $i \sim j$.

Let $a_m = (m+2)m^m/(m+1)^{m+1}$. Then $a_m \to e^{-1}$ as $m \to \infty$, and we may thus fix $m \geq 1$ such that $a_m < a$.

Let $b = a/(m+2) > m^m/(m+1)^{m+1}$. Shearer [7] showed that this implies that there exists an m-ary tree G such that $\mu_{G,b}(\emptyset) < 0$. We fix such a G; note that the root of G has degree m, the leaves degree 1 and all other vertices degree $m+1$.

On the other hand, it is easily seen that if p is small enough, then $\mu_{G,p}$ is a probability measure. (Indeed, if $p \leq m^m/(m+1)^{m+1}$, there is a simple probabilistic construction of $\mu_{G,p}$: Add external edges to the root and leaves of G such that every vertex has $m+1$ edges, 1 going 'down' and m going 'up'; colour all edges black or white independently at random with $\mathbb{P}(\text{white}) = q$, where $q(1-q)^m = p$; define $I_i = 1$ if the edge going down from vertex i is white and the m edges going up are black.)

By continuity of $p \mapsto \min\bigl\{\mu_{G,p}(\{A\}) : A \subseteq V \text{ is independent}\bigr\}$, and the fact that $\mu_{G,p}(\{A\}) = 0$ whenever A is dependent, it follows that for some p with

$0 < p < b$ we have $\mu_{G,p}(\{A\}) \geq 0$ for each A, and thus $\mu_{G,p}$ is a probability measure, but $\mu_{G,p}(\{A\}) = 0$ for at least one independent set A.

Fix such p and A, and let Γ be the (induced) subgraph of G obtained by deleting all vertices in A together with all their neighbours. (In fact, it seems likely that $A = \emptyset$, and thus $\Gamma = G$, but we have not verified this.) By a simple calculation [7], for every $B \subseteq V(\Gamma)$, $\mu_{\Gamma,p}(\{B\}) = \mu_{G,p}(\{A \cup B\})/p^{|A|}$; thus $\mu_{\Gamma,p}$ is a probability measure with $\mu_{\Gamma,p}(\emptyset) = 0$.

Let the indicators I_i be defined as above. For this family of indicators then $\mathbb{P}(S = 0) = \mu_{\Gamma,p}(\emptyset) = 0$ and $\delta + \varepsilon \leq (m+1)p + p < (m+2)b = a$, which verifies the claim.

8. Discussion

We may compare the upper bounds given in this paper with the results previously obtained in the special case when the indicators I_i all are of the type $\prod_{j \in Q_i} J_j$, where $\{J_j\}_{j \in Q}$ is a family of independent indicator variables and the Q_i are (arbitrary) subsets of the index set Q. In this case, using our notation, [4, 3] give the bounds

$$\mathbb{P}(S = 0) \leq \exp(-\mu + \Delta) \tag{18}$$

and

$$\mathbb{P}(S = 0) \leq \exp\bigl(-\mu^2/(2\Delta + \mu)\bigr); \tag{19}$$

Boppana and Spencer [2, 1] give, using a different proof and under somewhat more general conditions,

$$\mathbb{P}(S = 0) \leq \exp\bigl(\Delta/(1-\varepsilon)\bigr) \prod_{i \in \mathcal{I}} (1 - p_i) \leq \exp\bigl(-\mu + \Delta/(1-\varepsilon)\bigr). \tag{20}$$

(Alon and Spencer [1] also give, under the same conditions and provided $2\Delta \geq \mu(1-\varepsilon)$, the bound $\exp\bigl(-\mu^2(1-\varepsilon)/4\Delta\bigr)$; this is similar to but larger than the bound in (19).)

The bounds given here in Section 3 (for a much more general situation) are of a similar nature, but somewhat larger. The main difference is that δ plays no role in the bounds in the special case; for example, the bound in Theorem 2 differs from (18) only by the factor $e^{2\delta}$ multiplying Δ.

Furthermore, the constants in the exponent in the bounds in Theorems 3–5 are not as good as in (19) for the special case.

Example 1 shows that the Δ in the exponent in Theorem 2 and (18) is quite sharp; for example, it cannot in general be multiplied by a constant less than 1. However, we do not know whether the factor $e^{2\delta}$ multiplying Δ in Theorem 2 really is needed. As remarked above, the incomparable estimates in Theorems 6 and 7, which represent two different attempts to reduce this factor, suggest that none of the these results is the best possible. Indeed, it seems possible that the result (18) known for the special case above holds generally. (Recall, however, from Example 1 that the corresponding factor $\prod_{k \sim \{i,j\}} (1 - p_k)^{-1}$ in Theorem 1 cannot be completely eliminated, see (17). Similarly, by the same example, the factor $1/(1-\varepsilon)$ cannot be removed from the first bound in (20).)

Similarly, there seems to be room for improvements of the constants in Theorems 3–5. Example 1 shows that the best one could hope for is something like (19); for example, the constants in front of Δ and μ there are the best possible.

This leads to the following open problem.

Problem 1. *Do* (18)*,* (19) *and* (20) *hold under the assumptions of this paper?*

There are also other questions suggested by the results above.

Problem 2. [3] *gives a tail estimate similar to Theorem* 10*, but also a corresponding lower bound. Is there similar lower bound under the conditions of the present paper?*

As remarked in Remark 2, the Lovász local lemma (and our Theorem 9) requires only a weaker version of our condition that Γ is a dependency graph.

Problem 3. *Do the results of this paper hold if* Γ *only is assumed to be a dependency graph in the weaker sense?*

References

[1] N. Alon & J. Spencer, *The Probabilistic Method.* Wiley, New York 1992.
[2] R. Boppana & J. Spencer, A useful elementary correlation inequality. *J. Combin. Th. Ser. A* **50** (1989), 305–307.
[3] S. Janson, Poisson approximation for large deviations. *Random Struct. Alg.* **1** (1990), 221–229.
[4] S. Janson, T. Łuczak & A. Ruciński, An exponential bound for the probability of nonexistence of a specified subgraph in a random graph. In *Random Graphs '87* (Poznań 1987), Wiley, Chichester 1990, 73–87.
[5] L.M. Kirousis, E. Kranakis, D. Krizanc & Y.C. Stamatiou, Approximating the unsatisfiability threshold of random formulas. To appear.
[6] J. Spencer, personal communication, Poznań 1997.
[7] J. Shearer, On a problem of Spencer. *Combinatorica* **5** (1985), 241–245.
[8] W.C.S. Suen, A correlation inequality and a Poisson limit theorem for nonoverlapping balanced subgraphs of a random graph. *Random Struct. Alg.* **1** (1990), 231–242.

Department of Mathematics, Uppsala University, PO Box 480, S-751 06 Uppsala, Sweden
E-mail address: svante.janson@math.uu.se