Decoupling weakly dependent events

Stanislaw J. Szarek * Nicole Tomczak-Jaegermann †

In this note we discuss a probabilistic statement which conceptualizes arguments from recent papers [ST1, ST2]. Its framework appears to be sufficiently general to permit applications also beyond the original context.

The setting is as follows: we have a sequence of events the probability of each of which is rather small and we would like to deduce that the probability of their intersection is *very* small. This is of course straightforward if the events are independent; our approach allows to obtain comparable upper bounds on probabilities when the dependence is not "too strong." The precise formulation of the statement is somewhat technical, but its gist can be described as follows. Suppose that our probability space is a product space and that our events/sets are defined in terms of independent coordinates in a "local" way, i.e., while membership in each of the sets may depend on many or even all coordinates, it may be verified by checking a series of conditions each of which involves just a few coordinates (for example, to verify whether a sequence of vectors is orthogonal it is enough to look at just two elements of that sequence at a time). Then the probability of the intersection of these sets can "almost" be estimated as if they were independent. More precisely, the upper bound is not a product of their probabilities, but a homogeneous polynomial in the probabilities, the degree of which is high while the number of terms is controlled.

Problems similar in spirit if not in details were considered by many authors in probabilistic combinatorics and theoretical computer science. See, for example, [JR1, JR2] and their references; particularly [J2] seems to exhibit many formal similarities to our setting. (We thank M. Krivelevich for helping us navigate the combinatorics literature.) Let us note, however, that while the results cited above have, as a rule, a "large deviation feel," applications of our scheme go in the direction of "small ball" estimates, and

*Supported in part by a grant from the National Science Foundation (U.S.A.).
†This author holds the Canada Research Chair in Geometric Analysis.

no "dictionary" relating the other results to ours is apparent. On the other hand, all these statements can be considered as counterparts to Local Lovász Lemma (see e.g., [AS], Chapter 5) which, again under assumptions similar in spirit but different in detail, gives *lower* bounds on probabilities of intersections.

For $s \in \mathbb{N}$, we use the notation $[s]$ for the set $\{1, \ldots, s\}$. For a set J we denote by $|J|$ the cardinality of J.

We will present our result in two separate theorems. However, the first theorem is actually a special case of the second one and is stated here primarily for pedagogical reasons.

Theorem 1 *Let $d, N \in \mathbb{N}$. Consider a family of events $\{\Theta_{j,B} : j \in [N], B \subset [N]\}$ such that for any $j \in [N]$ and $B \subset [N]$ we have*

$$\Theta_{j,B} \subset \bigcup_{B' \subset B, |B'| \leq d} \Theta_{j,B'}. \tag{1}$$

For $j \in [N]$ set $\Theta_j := \Theta_{j,\{j\}^c}$, and for $\ell \in [N]$ set $\mathcal{J}_\ell = \{J \subset [N] : |J| = \ell\}$. Then for any $\ell \leq \lceil N/(2d+1) \rceil$, we have

$$\bigcap_{j=1}^{N} \Theta_j \subset \bigcup_{J \in \mathcal{J}_\ell} \bigcap_{j \in J} \Theta_{j,J^c}. \tag{2}$$

If additionally for any $I, J \subset [N]$ with $I \cap J = \emptyset$ the events $\{\Theta_{j,I} : j \in J\}$ are independent, then setting $p_j = \mathbb{P}(\Theta_j)$ for $j \in [N]$ we get

$$\mathbb{P}(\bigcap_{j=1}^{N} \Theta_j) \leq \sum_{J \in \mathcal{J}_\ell} \prod_{j \in J} \mathbb{P}(\Theta_{j,J^c}) \leq \sum_{J \in \mathcal{J}_\ell} \prod_{j \in J} p_j. \tag{3}$$

The following more general formulation – substituting conditional independence for independence in the hypothesis – appears to be more easily applicable to problems which come up naturally in convex geometry and combinatorics. To state it, we will use the following concept: a family $\{\Sigma_B : B \subset [N]\}$ of σ-algebras is nested if $B' \subset B$ implies $\Sigma_{B'} \subset \Sigma_B$.

Theorem 2 *In the notation of Theorem 1, assume that the family $\{\Theta_{j,B} : j \in [N], B \subset [N]\}$ satisfies condition* (1). *Let $\{\Sigma_B : B \subset [N]\}$ be a nested family of σ-algebras. Assume further that for any $I, J \subset [N]$ with*

$I \cap J = \emptyset$ the events $\{\Theta_{j,I} : j \in J\}$ are Σ_I-conditionally independent and that $\mathbb{P}(\Theta_j \,|\, \Sigma_{\{j\}^c}) \le p_j$ for $j \in [N]$. Then

$$\mathbb{P}(\bigcap_{j=1}^{N} \Theta_j) \le \sum_{J \in \mathcal{J}_\ell} \prod_{j \in J} p_j. \tag{4}$$

Proofs of the Theorems are based on the following combinatorial lemma.

Lemma 3 *Assume that a sequence $B_1, B_2, \dots, B_N$ of subsets of $[N]$ satisfies $|B_j| \le d$ and $j \notin B_j$ for $j = 1, 2, \dots, N$. Then there exists $J \subset [N]$ such that $|J| \ge N/(2d+1)$ and*

$$J \cap \bigcup_{j \in J} B_j = \emptyset.$$

Consider the $N \times N$ $\{0,1\}$-matrix $\Lambda = (\lambda_{i,j})$ defined by $\lambda_{i,j} = 1$ if $i \in B_j$ and $\lambda_{i,j} = 0$ if $i \notin B_j$, for $j \in [N]$. Then the lemma follows immediately from a result on suppression of matrices due to K. Ball (cf., [BT], Theorem 3.1). We recall the statement of this result as we feel it might be applicable in variety of contexts; for example in [ST2] it was used to control probabilities via analytic considerations rather than combinatorial ones.

Proposition 4 *Let $A = (a_{i,j})$ be an $N \times N$ matrix such that $a_{i,j} \ge 0$ for all i, j, $\sum_{i=1}^{N} a_{i,j} \le 1$ and $a_{j,j} = 0$ for all j. Then for every integer $t \ge 1$ there is a partition $\{J_s\}_{s=1}^{t}$ of $[N]$ such that for $s = 1, \dots, t$,*

$$\sum_{i \in J_s} a_{i,j} \le 2/t \qquad \text{for } j \in J_s.$$

In the setting of Lemma 3 we apply Proposition 4 to the matrix $A = (1/d)\Lambda$ and $t = 2d+1$.

We shall also provide an elementary proof of (a variant of) Lemma 3 that gives the estimate $|J| \ge N/(6d+2)$, which is sufficient for most applications. Alternatively, various variants of the lemma may be also derived from various forms of Turán's theorem, cf. [AS], p. 81–82. (We note, however, that Turán's theorem concerns *undirected* graphs, which correspond to symmetric matrices in the language of Proposition 4, and so the derivation requires some additional – even if not difficult – steps similar to the first part of the argument presented below.)

Proof Fix $a \geq 1$. Set $I := \{i : \sum_{j=1}^N \lambda_{i,j} \leq a\,d\}$ and let $m := |I|$. Since the sum of all entries of Λ is $\leq d\,N$, then $N - m = |I^c| \leq N/a$, and so $m \geq (1 - 1/a)N$.

Let $J \subset I$ be a maximal subset of I such that the corresponding $|J| \times |J|$ submatrix consists only of 0's, and set $|J| =: k$. To facilitate visualizing the argument the reader may think of $J = [k]$. By maximality of J we have that for each $i \in I \setminus J$ there is $j \in J$ such that $\lambda_{i,j} + \lambda_{j,i} \geq 1$. Summing up over $i \in I \setminus J$ we get

$$S := \sum_{i \in I \setminus J} \sum_{j \in J} (\lambda_{i,j} + \lambda_{j,i}) \geq |I \setminus J| = m - k.$$

We will now get two lower estimates on k from considerations in two separate rectangles. Set

$$t := \frac{1}{S} \sum_{i \in I \setminus J} \sum_{j \in J} \lambda_{j,i},$$

so that the number of 1's in the "upper-right" rectangle $J \times (I \setminus J)$ is equal to tS. On the other hand, for each $j \in J$, the number of 1's in the j'th row is less than or equal to $a\,d$, therefore $t\,(m-k) \leq tS \leq k(a\,d)$. Similarly, considering the "lower-left" rectangle $(I \setminus J) \times J$, and calculating the number of 1's in two different ways we get $(1-t)S \leq k\,d$, hence $(1-t)(m-k) \leq (1-t)S \leq k\,d$. Adding up the obtained inequalities we get $m - k \leq k\,d(a+1)$, which yields

$$k \geq \frac{a-1}{a\,(ad + d + 1)}\,N.$$

Setting, for example, $a = 2$, gives $k \geq N/(6d+2)$. $\square$

We are now ready for the proofs of the Theorems.

Proof of Theorem 1 Observe that by (1), $B' \subset B \subset [N]$ implies $\Theta_{j,B'} \subset \Theta_{j,B}$ for any $j \in [N]$. Fix $\ell \leq \lceil N/(2d+1) \rceil$. To show (2), let $\omega \in \bigcap_{j=1}^N \Theta_j$. Using (1) for each $j = 1, \ldots, N$ again we get sets $B_j \not\ni j$ (which may depend on ω) with $|B_j| \leq d$ such that $\omega \in \bigcap_{j=1}^N \Theta_{j,B_j}$. If $J \in \mathcal{J}_\ell$ is the set from Lemma 3 then, by the first observation above, we have $\omega \in \bigcap_{j=1}^N \Theta_{j,J^c}$. The set J may depend on ω as well, but since $J \in \mathcal{J}_\ell$, the inclusion (2) follows.

If the additional independence assumption is satisfied then the family $\{\Theta_{j,J^c} : j \in J\}$ is independent, hence $\mathbb{P}\Big(\bigcap_{j\in J} \Theta_{j,J^c}\Big) = \prod_{j\in J} \mathbb{P}(\Theta_{j,J^c})$. Since $\Theta_{j,J^c} \subset \Theta_j$, the last inequality follows as well. □

Proof of Theorem 2 By Theorem 1 the inclusion (2) holds.

Next, by the conditional independence assumption we have, for every $J \subset [N]$,

$$\mathbb{P}\Big(\bigcap_{j\in J} \Theta_{j,J^c} \,|\, \Sigma_{J^c}\Big) = \prod_{j\in J} \mathbb{P}\big(\Theta_{j,J^c} \,|\, \Sigma_{J^c}\big).$$

In turn, for $j \in J$,

$$\mathbb{P}\left(\Theta_{j,J^c} \,|\, \Sigma_{J^c}\right) \le \mathbb{P}\left(\Theta_j \,|\, \Sigma_{J^c}\right) = \mathbb{E}\Big(\mathbb{P}\left(\Theta_j \,|\, \Sigma_{\{j\}^c}\right) \,\Big|\, \Sigma_{J^c}\Big) \le p_j,$$

with the last estimate following from the (pointwise) upper bound on the random variable $\mathbb{P}\left(\Theta_j \,|\, \Sigma_{\{j\}^c}\right)$. Accordingly,

$$\mathbb{P}\Big(\bigcap_{j\in J} \Theta_{j,J^c}\Big) = \mathbb{E}\left(\mathbb{P}\Big(\bigcap_{j\in J} \Theta_{j,J^c} \,|\, \Sigma_{J^c}\Big)\right) \le \prod_{j\in J} p_j.$$

Therefore, by (2), we conclude that for $\ell = \lceil N/(2d+1) \rceil$,

$$\mathbb{P}\left(\bigcap_{j=1}^{N} \Theta_j\right) \le \sum_{J\in\mathcal{J}_\ell} \mathbb{P}\left(\bigcap_{j\in J} \Theta_{j,J^c}\right) \le \sum_{J\in\mathcal{J}_\ell} \prod_{j\in J} p_j,$$

that is, (4) holds. □

Even though the hypotheses of the theorems seem quite abstract, there exists a variety of natural probabilistic settings in which they are satisfied. We will now describe some such settings. While our examples make the appearance of conditional independence quite clear, securing uniform estimates for conditional probabilities often requires some additional technicalities which we will ignore here as they are only marginally related to the decoupling procedure.

Let $D_1, \dots, D_N, L_1, \dots, L_N$ be random convex subsets of $\mathbb{R}^n$ such that the family of pairs $\{(D_j, L_j)\}_{j=1}^N$ is independent. Set $E_j = \operatorname{span} L_j \subset \mathbb{R}^n$

and let P_{E_j} denote the orthogonal projection on E_j, for $j = 1, \dots, N$. For $B \subset [N]$, we let Σ_B be the σ-algebra generated by $\{D_i, L_i\}_{i \in B}$.

In a typical setting the sets D_i, L_i will be symmetric and $\operatorname{conv} D_i =: K$ will be a (random) symmetric convex body in $\mathbb{R}^n$. We will then define a normed space X as $\mathbb{R}^n$ endowed with the norm whose unit ball is K. We will be interested in particular in the character and complementability of subspaces of X, especially those determined by the E_j's.

For the first illustration of our scheme assume that there is $p \in (0,1)$ such that we have upper bounds for conditional probabilities

$$\mathbb{P}\left(E_j \cap \operatorname*{conv}_{i \neq j} D_i \not\subset L_j \,|\, \Sigma_{\{j\}^c} \right) \leq p \qquad \text{for all } j \in [N]. \tag{5}$$

In the simplest case when $D_i = L_i$ for $i \in [N]$, the complement of the event appearing in (5) can be alternatively described by the equality $E_j \cap K = D_j$, that is, the unit ball in E_j considered as a subspace of X (with the induced norm) being exactly D_j.

For $j \in [N]$ and $B \subset [N]$, let $\Theta_{j,B} = \{E_j \cap \operatorname{conv}_{i \in B} D_i \not\subset L_j\}$, then the corresponding Θ_j's are exactly the sets appearing in (5). By Caratheodory's theorem condition (1) is satisfied with $d = n+1$. (In fact, we do have here, and in examples that follow, equality of the sets. We note, however, that in actual applications one often needs the weaker hypothesis involving inclusion.) Also, if $I \cap J = \emptyset$ then events $\{\Theta_{j,I} : j \in J\}$ are Σ_I-conditionally independent. Therefore by Theorem 2 we get, with $\ell = \lceil N/(2n+3) \rceil$,

$$\mathbb{P}\left(\bigcap_{j=1}^{N} \Theta_j \right) \leq \binom{N}{\ell} p^{\ell} \leq \left(ep(2n+3) \right)^{\ell}. \tag{6}$$

For another illustration we assume the following upper bound on the conditional probabilities

$$\mathbb{P}\left(P_{E_j}\big(\operatorname*{conv}_{i \neq j} D_i\big) \not\subset L_j \,|\, \Sigma_{\{j\}^c} \right) \leq p \qquad \text{for all } j \in [N]. \tag{7}$$

Modulo some minor technicalities sets of this form were considered in [ST1]. Again, in the case when $D_i = L_i$ for $i = 1, \dots, N$, the complement of the event from (7) can be described as follows: the subspace E_j of X has the unit ball equal to D_j *and* is 1-complemented in X via the orthogonal projection.

Similarly as before, let $\Theta'_{j,B} = \left\{ P_{E_j}(\mathrm{conv}_{\, i\in B} D_i) \not\subset L_j \right\}$, for $j \in [N]$ and $B \subset [N]$. Then the corresponding Θ'_j's are exactly the sets appearing in (7). Now, condition (1) is clearly satisfied with $d = 1$, and if $I \cap J = \emptyset$ then events $\{\Theta'_{j,I} : j \in J\}$ are Σ_I-conditionally independent. Using Theorem 2 with $\ell = \lceil N/3 \rceil$ we then obtain

$$\mathbb{P}\left(\bigcap_{j=1}^{N} \Theta'_j \right) \leq \sum_{J \in \mathcal{J}_\ell} \mathbb{P}\left(\bigcap_{j\in J} \Theta'_{j,J^c} \,|\, \Sigma_{J^c} \right) \leq \binom{N}{\ell} p^\ell \leq (3ep)^\ell \, . \tag{8}$$

For more elaborated geometric interpretations of our scheme let us assume that $D_i \subset L_i$ for $i \in [N]$ and consider a random symmetric convex body $L \subset \mathbb{R}^n$. By Y denote the space $\mathbb{R}^n$ with the norm for which L is the unit ball; consider the formal identity operator $id_{X,Y} : X \to Y$ and let $k := \max \mathrm{codim}\, E_i$.

If $L_j = E_j \cap L$ for $j \in [N]$, then the complement of the set $\bigcap_{j=1}^N \Theta_j$ (appearing in (6)) is connected with an upper bound for the Gelfand numbers of $id_{X,Y}$. More precisely, $\omega \notin \bigcap_{j=1}^N \Theta_j$ implies that $c_k(id_{X,Y}) \leq 1$. Similarly, with L_j's of the same form, the complement of the set appearing in (8) relates to the approximation numbers of $id_{X,Y}$, namely $\omega \notin \bigcap_{j=1}^N \Theta'_j$ implies that $a_k(id_{X,Y}) \leq 1$.

Finally, if $L_j = P_{E_j} L$ for $j \in [N]$, then $\omega \notin \bigcap_{j=1}^N \Theta'_j$ implies that the k's Kolmogorov number of $id_{X,Y}$ satisfies $d_k(id_{X,Y}) \leq 1$.

Still another application of the present scheme can be found in [P] which provides a simpler and more structured proof of the result from [GLT] concerning highly asymmetric convex bodies. The sets corresponding to Θ_j's that appear in that paper are roughly of the form

$$\{g_j \in t \, \mathrm{conv}\, (\{g_i\}_{i \neq j}, 0) \quad \text{for all} \;\; j \in [N]\} \, ,$$

where the g_j's are i.i.d. Gaussian vectors and $t > 0$ is a constant. For more details we refer the reader to [P].

References

[AS] N. Alon & J. H. Spencer, *The probabilistic method. With an appendix by Paul Erdös.* John Wiley & Sons, Inc., New York, 1992.

[BT] J. Bourgain & L. Tzafriri, *Restricted invertibility of matrices and applications.* Analysis at Urbana, Vol. II (Urbana, IL, 1986–1987), 61–107, London Math. Soc. Lecture Note Ser., 138, Cambridge Univ. Press, Cambridge, 1989.

[GLT] E. D. Gluskin, A. Litvak & N. Tomczak-Jaegermann, *An example of a convex body without symmetric projections.* Israel J. of Math. 110 (2001), 267–277.

[J2] S. Janson, *Large Deviations for Sums of Partly Dependent Random Variables.* Random Struct. Alg. 24(3) (2004), 234–248

[JR1] S. Janson & A. Ruciński, *The infamous upper tail.* Random Struct. Alg. 20(3) (2002), 317–342.

[JR2] S. Janson & A. Ruciński, *The deletion method for upper tail estimates.* Combinatorica 24(4) (2004), 615–640

[P] P. Pivovarov *Random convex bodies lacking symmetric projections, revisited through decoupling.* Preprint.

[ST1] S. J. Szarek & N. Tomczak-Jaegermann, *Saturating Constructions for Normed Spaces.* Geometric and Functional Analysis, 14 (2004), 1352–1375.

[ST2] S. J. Szarek & N. Tomczak-Jaegermann, *Saturating Constructions for Normed Spaces II.* J. Funct. Anal., 221 (2005), 407–438.

Department of Mathematics
Case Western Reserve University
Cleveland, Ohio 44106-7058, U.S.A.
and
Equipe d'Analyse Fonctionnelle, BP 186
Université Pierre et Marie Curie
75252 Paris, France
szarek@cwru.edu

Department of Mathematical and Statistical Sciences,
University of Alberta,
Edmonton, Alberta, Canada T6G 2G1,
nicole@ellpspace.math.ualberta.ca