CONCENTRATION ESTIMATES FOR FUNCTIONS OF FINITE HIGH-DIMENSIONAL RANDOM ARRAYS

PANDELIS DODOS, KONSTANTINOS TYROS AND PETROS VALETTAS

Abstract. Let $\boldsymbol{X}$ be a d-dimensional random array on $[n]$ whose entries take values in a finite set $\mathcal{X}$, that is, $\boldsymbol{X} = \langle X_s : s \in \binom{[n]}{d} \rangle$ is an $\mathcal{X}$-valued stochastic process indexed by the set $\binom{[n]}{d}$ of all d-element subsets of $[n] \coloneqq \{1, \dots, n\}$. We give easily checked conditions on $\boldsymbol{X}$ which ensure, for instance, that for *every* function $f \colon \mathcal{X}^{\binom{[n]}{d}} \to \mathbb{R}$ which satisfies $\mathbb{E}[f(\boldsymbol{X})] = 0$ and $\|f(\boldsymbol{X})\|_{L_p} = 1$ for some $p > 1$, the random variable $f(\boldsymbol{X})$ becomes concentrated after conditioning it on a large subarray of $\boldsymbol{X}$. These conditions cover several classes of random arrays with not necessarily independent entries. Applications are given in combinatorics, and examples are also presented which show the optimality of various aspects of the results.

Contents

1. Introduction 2

Part 1. Proofs of the main results 9

2. From dissociativity to concentration 9
3. The box independence condition propagates 16
4. Proof of Theorem 3.2 20
5. Proof of Theorem 1.5 32
6. Proof of Theorem 1.4 and its higher-dimensional version 32
7. Extensions/Refinements 33

Part 2. Connection with combinatorics 35

8. Random arrays arising from combinatorial structures 35
9. Quasirandom families of graphs: proof of Theorem 1.8 40

Appendix A. Examples 48

References 56

2010 *Mathematics Subject Classification*: 05D10, 05D40, 60E15, 60G09, 60G42.

Key words: concentration inequalities, exchangeable random arrays, spreadable random arrays, martingale difference sequences, quasirandomness, density polynomial Hales–Jewett conjecture.

P.V. is supported by Simons Foundation grant 638224.

1. Introduction

1.1. **Motivation.** The concentration of measure refers to the powerful phenomenon asserting that a function which depends smoothly on its variables is essentially constant, as long as the number of the variables is large enough. There are various ways to quantify this "smooth dependence" (*e.g.*, Lipschitz conditions, bounds for the L_2 norm of the gradient, etc.). Detailed expositions can be found in [Le01, BLM13].

It is easy to see that this phenomenon is no longer valid if we drop the smoothness assumption. Nevertheless, one can still obtain some form of concentration under a much milder integrability condition (see [DKT16, Theorem 1′]).

Theorem. *For every $p > 1$ and every $0 < \varepsilon \leqslant 1$, there exists a constant $c > 0$ with the following property. If $n \geqslant 2/c$ is an integer, $\boldsymbol{X} = (X_1, \dots, X_n)$ is a random vector with independent entries which take values in a measurable space $\mathcal{X}$, and $f\colon \mathcal{X}^n \to \mathbb{R}$ is a measurable function with $\mathbb{E}[f(\boldsymbol{X})] = 0$ and $\|f(\boldsymbol{X})\|_{L_p} = 1$, then there exists an interval I of $[n]$ with $|I| \geqslant cn$ such that for every nonempty $J \subseteq I$ we have*

$$\mathbb{P}\big(\big|\mathbb{E}[f(\boldsymbol{X}) \,|\, \mathcal{F}_J]\big| \leqslant \varepsilon\big) \geqslant 1 - \varepsilon \tag{1.1}$$

where $\mathbb{E}[f(\boldsymbol{X}) \,|\, \mathcal{F}_J]$ stands for the conditional expectation of $f(\boldsymbol{X})$ with respect to the σ-algebra $\mathcal{F}_J \coloneqq \sigma(\{X_i : i \in J\})$.

(Here, and in what follows, $[n]$ denotes the discrete interval $\{1, \dots, n\}$.) Roughly speaking, this result asserts that if a function of several variables is sufficiently integrable, then, by integrating out some coordinates, it becomes essentially constant. It was motivated by—and it has found several applications in—problems in combinatorics (see [DK16]).

1.1.1. The goal of this paper is twofold: to develop workable tools in order to extend the conditional concentration estimate (1.1) to functions of random vectors $\boldsymbol{X}$ with not necessarily independent entries, and to present related applications. Of course, to this end some structural property of $\boldsymbol{X}$ is necessary. We focus on high-dimensional random arrays whose distribution is invariant under certain symmetries. Besides their intrinsic analytic and probabilistic interest, our choice to study functions of random arrays is connected to the *density polynomial Hales–Jewett conjecture*, an important combinatorial conjecture of Bergelson [Ber96]—see Subsection 1.5.

1.2. **Random arrays.** At this point it is useful to recall the definition of a random array.

Definition 1.1 (Random arrays, and their subarrays/sub-σ-algebras)**.** *Let d be a positive integer, and let I be a set with $|I| \geqslant d$. A d*-dimensional random array on *I is a stochastic process $\boldsymbol{X} = \langle X_s : s \in \binom{I}{d}\rangle$ indexed by the set $\binom{I}{d}$ of all d-element subsets of I. If J is a subset of I with $|J| \geqslant d$, then the* subarray of $\boldsymbol{X}$ determined by J *is the d-dimensional random array $\boldsymbol{X}_J \coloneqq \langle X_s : s \in \binom{J}{d}\rangle$; moreover, by $\mathcal{F}_J$ we shall denote the σ-algebra $\sigma(\{X_s : s \in \binom{J}{d}\})$ generated by $\boldsymbol{X}_J$.*

Of course, one-dimensional random arrays are just random vectors. On the other hand, two-dimensional random arrays are essentially the same as random symmetric matrices, and their subarrays correspond to principal submatrices; more generally, higher-dimensional random arrays correspond to random symmetric tensors. We employ the terminology of random arrays, however, since we are not using linear-algebraic tools.

1.2.1. *Notions of symmetry.* The study of random arrays with a symmetric distribution is a classical topic which goes back to the work of de Finetti; see [Au08, Au13, Kal05] for an exposition of this theory and its applications. Arguably, the most well-known notion of symmetry is exchangeability: a d-dimensional random array $\boldsymbol{X}$ on a (possibly infinite) set I is called *exchangeable* if for every finite permutation π of I, the random arrays $\boldsymbol{X}$ and $\boldsymbol{X}_\pi \coloneqq \langle X_{\pi(s)} : s \in \binom{I}{d}\rangle$ have the same distribution. Another well-known notion of symmetry, which is weaker than exchangeability, is spreadability: a d-dimensional random array $\boldsymbol{X}$ on a (possibly infinite) set I is called *spreadable*[1] if for every pair J, K of finite subsets of I with $|J| = |K| \geqslant d$, the subarrays $\boldsymbol{X}_J$ and $\boldsymbol{X}_K$ have the same distribution. Infinite, spreadable, two-dimensional random arrays have been studied by Fremlin and Talagrand [FT85], and—in greater generality—by Kallenberg [Kal92].

Beyond these notions, in this paper we will also consider the following approximate form of spreadability which naturally arises in combinatorial applications.

Definition 1.2 (Approximate spreadability)**.** *Let $\boldsymbol{X}$ be a d-dimensional random array on a* (*possibly infinite*) *set I, and let $\eta \geqslant 0$. We say that $\boldsymbol{X}$ is* η-spreadable (*or* approximately spreadable *if η is understood*)*, provided that for every pair J, K of finite subsets of I with $|J| = |K| \geqslant d$ we have*

$$\rho_{\mathrm{TV}}(P_J, P_K) \leqslant \eta \tag{1.2}$$

where P_J and P_K denote the laws of the random subarrays $\boldsymbol{X}_J$ and $\boldsymbol{X}_K$ respectively, and ρ_{TV} stands for the total variation distance.

The following proposition justifies Definition 1.2 and shows that approximately spreadable random arrays are the building blocks of arbitrary finite-valued, high-dimensional random arrays. The proof follows by a standard application of Ramsey's theorem [Ra30] taking into account the fact that the space of all probability measures on a finite set equipped with the total variation distance in compact.

Proposition 1.3. *For every triple m, n, d of positive integers with $n \geqslant d$, and every $\eta > 0$, there exists an integer $N \geqslant n$ with the following property. If $\mathcal{X}$ is a set with $|\mathcal{X}| = m$ and $\boldsymbol{X}$ is an $\mathcal{X}$-valued, d-dimensional random array on a set I with $|I| \geqslant N$, then there exists a subset J of I with $|J| = n$ such that the random array $\boldsymbol{X}_J$ is η-spreadable.*

[1] We point out that this is not standard terminology. In particular, in [FT85] spreadable random arrays are referred to as *deletion invariant*, while in [Kal05] they are called *contractable*.

1.3. **The concentration estimate.** We are ready to state one of the main extensions of (1.1) obtained in this paper; the question whether (1.1) could hold for random vectors with not independent entries, was asked by an anonymous reviewer of [DKT16] as well as by several colleagues in personal communication. In this introduction we restrict our discussion to boolean two-dimensional random arrays, mainly because this case is easier to grasp, but at the same time it is quite representative of the higher dimensional case. The general version is presented in Theorem 6.1 in Section 6; further extensions/refinements are given in Section 7.

Theorem 1.4. *Let $1 < p \leqslant 2$, let $0 < \varepsilon \leqslant 1$, let $k \geqslant 2$ be an integer, and set*

$$C = C(p,\varepsilon,k) \coloneqq \exp\Big(\frac{34}{\varepsilon^8 (p-1)^2}\, k^2\Big). \tag{1.3}$$

Also let $n \geqslant C$ be an integer, let $\boldsymbol{X} = \langle X_s : s \in \binom{[n]}{2}\rangle$ be a $\{0,1\}$-valued, $(1/C)$-spreadable, two-dimensional random array on $[n]$, and assume that

$$\big|\mathbb{E}[X_{\{1,3\}}X_{\{1,4\}}X_{\{2,3\}}X_{\{2,4\}}] - \mathbb{E}[X_{\{1,3\}}]\,\mathbb{E}[X_{\{1,4\}}]\,\mathbb{E}[X_{\{2,3\}}]\,\mathbb{E}[X_{\{2,4\}}]\big| \leqslant \frac{1}{C}. \tag{1.4}$$

Then for every function $f\colon \{0,1\}^{\binom{[n]}{2}} \to \mathbb{R}$ with $\mathbb{E}[f(\boldsymbol{X})] = 0$ and $\|f(\boldsymbol{X})\|_{L_p} = 1$ there exists an interval I of $[n]$ with $|I| = k$ such that for every $J \subseteq I$ with $|J| \geqslant 2$ we have

$$\mathbb{P}\big(\big|\mathbb{E}[f(\boldsymbol{X})\,|\,\mathcal{F}_J]\big| \leqslant \varepsilon\big) \geqslant 1-\varepsilon. \tag{1.5}$$

Recall that $\mathcal{F}_J$ denotes the σ-algebra generated by $\boldsymbol{X}_J$. (See Definition 1.1.) Thus, Theorem 1.4 asserts that the random variable $f(\boldsymbol{X})$ becomes concentrated after conditioning it on a *subarray* of $\boldsymbol{X}$. Also observe that (1.4) together with the $(1/C)$-spreadability of $\boldsymbol{X}$ imply that for every $i,j,k,\ell \in [n]$ with $i<j<k<\ell$ we have

$$\big|\mathbb{E}[X_{\{i,k\}}X_{\{i,\ell\}}X_{\{j,k\}}X_{\{j,\ell\}}] - \mathbb{E}[X_{\{i,k\}}]\,\mathbb{E}[X_{\{i,\ell\}}]\,\mathbb{E}[X_{\{j,k\}}]\,\mathbb{E}[X_{\{j,\ell\}}]\big| \leqslant \frac{6}{C}. \tag{1.6}$$

As we shall shortly see, as the parameter C gets bigger, the estimate (1.6) forces the random variables $X_{\{i,k\}}, X_{\{i,\ell\}}, X_{\{j,k\}}, X_{\{j,\ell\}}$ to behave independently. (It also implies that the correlation matrix of $\boldsymbol{X}$ is close to the identity.) Therefore, we may view (1.6) as an (*approximate*) *box independence* condition for $\boldsymbol{X}$. We present various examples of spreadable random arrays that satisfy the box independence condition in Section 8.

Finally we point out that (1.6) is essentially an optimal condition in the sense that for every integer $n \geqslant 4$ there exist

— a boolean, exchangeable, two-dimensional random array $\boldsymbol{X}$ on $[n]$, and
— a translated multilinear polynomial $f\colon \mathbb{R}^{\binom{[n]}{2}} \to \mathbb{R}$ of degree 4 with $\mathbb{E}[f(\boldsymbol{X})] = 0$ and $\|f(\boldsymbol{X})\|_{L_\infty} \leqslant 1$,

such that the correlation matrix of $\boldsymbol{X}$ is the identity, and for which (1.6) and (1.5) do not hold. (See Proposition A.1 in Appendix A; the case "$d \geqslant 3$" is treated in Proposition A.2.)

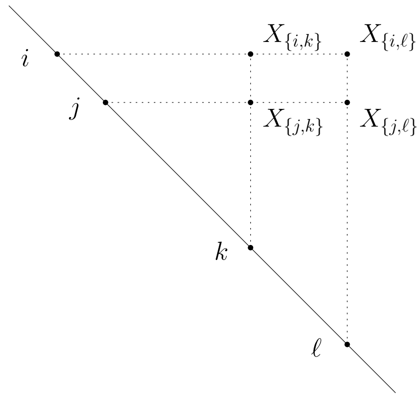

FIGURE 1. The box independence condition.

1.4. **Basic steps of the proof.** The first step of the proof of Theorem 1.4—which can be loosely described as its analytical part—is to show that the conditional concentration of $f(\boldsymbol{X})$ is equivalent to an approximate form of the dissociativity of $\boldsymbol{X}$; this is the content of Theorem 2.2 in Section 2. The proof of this step is based on estimates for martingale difference sequences in L_p spaces, and it applies to random arrays with arbitrary distributions (in particular, not necessarily approximately spreadable). The main advantage of this reduction is that it enables us to forget about the function f and focus exclusively on the random array $\boldsymbol{X}$.

The second—and more substantial—step is the verification of the approximate dissociativity of $\boldsymbol{X}$. This is a consequence of the following theorem which is one of the main results of this paper. (As before, at this point we restrict our discussion to boolean two-dimensional random arrays; the general version is given in Theorem 3.2.)

Theorem 1.5 (Propagation of randomness)**.** *Let $n \geqslant 4$ be an integer, let $0 < \eta, \vartheta \leqslant 1$, and set $\ell \coloneqq \binom{\lfloor n/2 \rfloor}{2}$. Also let $\boldsymbol{X} = \langle X_s : s \in \binom{[n]}{2} \rangle$ be a $\{0,1\}$-valued, η-spreadable, two-dimensional random array on $[n]$ such that for every $i,j,k,\ell \in [n]$ with $i<j<k<\ell$,*

$$\mathbb{E}[X_{\{i,k\}}X_{\{i,\ell\}}X_{\{j,k\}}X_{\{k,\ell\}}] \leqslant \mathbb{E}[X_{\{i,k\}}]\,\mathbb{E}[X_{\{i,\ell\}}]\,\mathbb{E}[X_{\{j,k\}}]\,\mathbb{E}[X_{\{j,\ell\}}] + \vartheta. \tag{1.7}$$

Then for every nonempty $\mathcal{F} \subseteq \binom{[n]}{2}$ such that $\cup\mathcal{F}$ has cardinality at most $n/2$, we have

$$\Big|\mathbb{E}\Big[\prod_{s\in\mathcal{F}} X_s\Big] - \prod_{s\in\mathcal{F}} \mathbb{E}[X_s]\Big| \leqslant 400|\mathcal{F}|\,\big(n^{-1/16} + \eta^{1/16} + \vartheta^{1/16}\big). \tag{1.8}$$

Theorem 1.5 shows that the box independence condition[2] propagates and forces all, not too large, subarrays of $\boldsymbol{X}$ to behave independently. Its proof is based on combinatorial

[2]Note that in Theorem 1.5 we only need the one-sided version (1.7) of (1.6). Of course, in retrospect, Theorem 1.5 yields that (1.7) is actually equivalent to (1.6) albeit with a slightly different constant.

and probabilistic ideas, and it is analogous[3] to the phenomenon—discovered in the theory of quasirandom graphs [CGW88, CGW89]—that a graph G that contains (roughly) the expected number of 4-cycles must also contain the expected number of any other, not too large, graph H. We comment further on the relation between the box independence condition and quasirandomness of graphs and hypergraphs in Subsection 8.1.

1.5. **Connection with combinatorics.** We proceed to discuss a representative combinatorial application of our main results.

1.5.1. *Families of graphs.* We start by observing that for every integer $n \geqslant 2$ we may identify a graph G on $[n]$ with an element of $\{0,1\}^{\binom{[n]}{2}}$ via its indicator function $\mathbf{1}_G$. (More generally, for every nonempty finite index set I we identify subsets of I with elements of $\{0,1\}^I$.) Thus, we view the set $\{0,1\}^{\binom{[n]}{2}}$ as the *space of all graphs on n vertices* and we denote by μ the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}}$. Our application is related to the following conjecture of Gowers [Go09, Conjecture 4].

Conjecture 1.6. *Let $0 < \delta \leqslant 1$ and assume that n is sufficiently large in terms of δ. Then for every family of graphs $\mathcal{A} \subseteq \{0,1\}^{\binom{[n]}{2}}$ with $\mu(\mathcal{A}) \geqslant \delta$ there exist $G, H \in \mathcal{A}$ with $H \subseteq G$ such that the difference $G \setminus H$ is a clique, that is, $G \setminus H = \binom{X}{2}$ for some $X \subseteq [n]$ with $|X| \geqslant 2$.*

Conjecture 1.6 is a special, but critical, case of the density polynomial Hales–Jewett conjecture [Ber96]; for a detailed discussion of its significance we refer to [Go09] where Conjecture 1.6 was proposed as a polymath project.

Despite the fact that there is considerable interest, there is nearly no information on Conjecture 1.6 in the literature. (See, however, the online discussion in [Go09].) This is partly due to the fact that, while the understanding of quasirandom graphs is very satisfactory, it is unclear what a quasirandom *family of graphs* actually is. Our results are pointing precisely in this direction[4].

1.5.2. *Quasirandom families of graphs.* In order to motivate the reader, let us say that a family of graphs $\mathcal{A} \subseteq \{0,1\}^{\binom{[n]}{2}}$ is *isomorphic invariant*[5] if for every permutation π of $[n]$ and every $G \subseteq \binom{[n]}{2}$ we have

$$G \in \mathcal{A} \quad \text{if and only if} \quad G_\pi \coloneqq \{\pi(e) : e \in G\} \in \mathcal{A}; \tag{1.9}$$

[3] In fact, this is more than an analogy; indeed, it is easy to see that Theorem 1.5 yields the aforementioned property of quasirandom graphs.

[4] Here, it is important to note that this is a rather basic step of the analysis of Conjecture 1.6; indeed, the combinatorial core of almost every problem in density Ramsey theory is to isolate its quasirandom and structure components—see, *e.g.*, [Tao08] for an exposition of this general philosophy.

[5] Isomorphic invariant families of graphs are also referred to as *graph properties*. It may be argued that Conjecture 1.6 is more natural for isomorphic invariant families of graphs, but we do not impose such a restriction in our results.

that is, G belongs to $\mathcal{A}$ if every isomorphic copy of G belongs to $\mathcal{A}$. As we shall see in Proposition 9.1, if $\mathcal{A} \subseteq \{0,1\}^{\binom{[n]}{2}}$ is an arbitrary isomorphic invariant family of graphs, then there exists a nonnegative parameter $\gamma(\mathcal{A})$ such that

— $\gamma(\mathcal{A}) \geqslant \mu(\mathcal{A})^4 - o_{n\to\infty}(1)$, and
— for every $U = \{i < j < k < \ell\} \in \binom{[n]}{4}$, denoting by $\mathbf{P}$ the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}\setminus\binom{U}{2}}$, we have

$$\mathbf{P}\big(W : W \cup \{i,k\}, W \cup \{i,\ell\}, W \cup \{j,k\}, W \cup \{j,\ell\} \in \mathcal{A}\big) = \gamma(\mathcal{A}).$$

On the other hand, if $\mathcal{A} \subseteq \{0,1\}^{\binom{[n]}{2}}$ is random, then clearly $\gamma(\mathcal{A}) = \mu(\mathcal{A})^4 + o_{n\to\infty}(1)$.

Keeping these observations in mind, we view as quasirandom those families of graphs $\mathcal{A}$ whose parameter $\gamma(\mathcal{A})$ is not significantly larger from the corresponding parameter of a random family of graphs with the same density. This is, essentially, the content of the following definition.

Definition 1.7 (Quasirandom families of graphs)**.** *Let $n \geqslant 2$ be an integer, let $\theta > 0$, and let $\mathcal{A} \subseteq \{0,1\}^{\binom{[n]}{2}}$ be a (not necessarily isomorphic invariant) family of graphs. We say that $\mathcal{A}$ is* θ-quasirandom *if there exists $\mathcal{U} \subseteq \binom{[n]}{4}$ with $|\mathcal{U}| \geqslant (1-\theta)\binom{n}{4}$ such that for every $U = \{i<j<k<\ell\} \in \mathcal{U}$ we have*

$$\mathbf{P}\big(W : W \cup \{i,k\}, W \cup \{i,\ell\}, W \cup \{j,k\}, W \cup \{j,\ell\} \in \mathcal{A}\big) \leqslant \mu(\mathcal{A})^4 + \theta. \tag{1.10}$$

Here, as above, $\mathbf{P}$ denotes the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}\setminus\binom{U}{2}}$. (Thus, if $\mathcal{A}$ is isomorphic invariant, then $\mathcal{A}$ is θ-quasirandom if $\gamma(\mathcal{A}) \leqslant \mu(\mathcal{A})^4 + \theta$.)

The reader might have already observed the similarity between Definition 1.7 and the classical 4-cycle condition of quasirandomness of graphs [CGW88, CGW89].

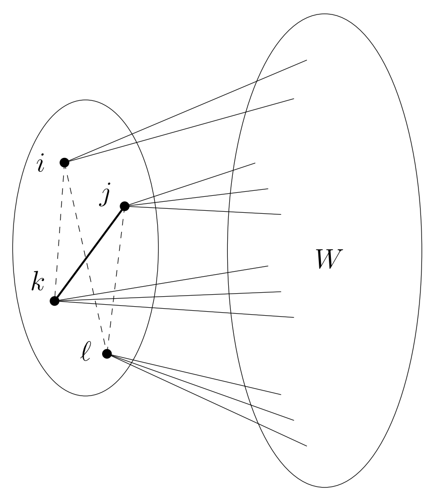

FIGURE 2. Quasirandom families of graphs.

1.5.3. The following theorem—which relies on both conditional concentration and Theorem 1.5, and whose proof is given in Section 9—shows that Definition 1.7 is indeed a sensible notion.

Theorem 1.8. *For every $0 < \delta \leqslant 1$ and every integer $k \geqslant 2$ there exist $\theta > 0$ and an integer $q_0 \geqslant k$ with the following property. Let $n \geqslant q_0$ be an integer, and let $\mathcal{A} \subseteq \{0,1\}^{\binom{[n]}{2}}$ be a θ-quasirandom family of graphs with $\mu(\mathcal{A}) \geqslant \delta$. Then, there exist $K \subseteq [n]$ with $|K| = k$ and $W \subseteq \binom{[n]}{2} \setminus \binom{K}{2}$ such that*

$$\{W\} \cup \Big\{W \cup e : e \in \binom{K}{2}\Big\} \subseteq \mathcal{A}. \tag{1.11}$$

In particular, there exist $G, H \in \mathcal{A}$ with $H \subseteq G$ such that $G \setminus H$ is a clique.

The proof of Theorem 1.8 is effective; see Remark 9.5 for its quantitative aspects.

1.6. **Related work.** Although Theorem 1.4 (as well as its higher dimensional extension, Theorem 6.1) is somewhat distinct from the traditional setting of concentration of smooth functions, it is related with several results which we are about to discuss.

Arguably, the one-dimensional case—that is, the case of random vectors—is the most heavily investigated. It is impossible to give here a comprehensive review; we only mention that concentration estimates for functions of finite exchangeable random vectors have been obtained in [Bob04, Ch06].

The two-dimensional case is also heavily investigated, in particular, in the literature around various random matrix models. However, closer to the spirit of this paper is the work of Latala [La06] and the subsequent papers [AdWo15, GSS19, V19] which obtain exponential concentration inequalities for smooth functions (*e.g.*, polynomials) of high-dimensional random arrays whose entries are of the form

$$X_s = \prod_{i \in s} \xi_i \tag{1.12}$$

where $(\xi_1, \dots, \xi_n)$ is a random vector with independent entries and a well-behaved distribution. Note that all these arrays are dissociated, and are additionally exchangeable if the random variables $\xi_1, \dots, \xi_n$ are identically distributed.

That said, the study of concentration inequalities for functions of more general finite high-dimensional random arrays is nearly not developed at all, mainly because the structure of finite high-dimensional[6] random arrays is quite complicated (see, also, [Au13, page 16] for a discussion on this issue). We make a step in this direction in the companion paper [DTV21].

[6] The understanding is better in the one-dimensional case—see [DF80].

1.7. **Organization of the paper.** We close this section by giving an outline of the contents of this paper. It is divided into two parts, Part 1 and Part 2, which are largely independent of each other and can be read separately.

Part 1 consists of Sections 3 up to 8. The main result in Section 2 is Theorem 2.2 which reduces conditional concentration to approximate dissociativity. The next two sections, Sections 3 and 4, are devoted to the proof of Theorem 1.5 and its higher-dimensional extension, Theorem 3.2. In Section 3 we introduce related definitions and we also present some consequences. The proof of Theorem 3.2 is given in Section 4; this is the most technically demanding part of the paper. In Section 5 we give the proof of Theorem 1.5, while in Section 6 we complete the proofs of Theorem 1.4 and its higher-dimensional extension, Theorem 6.1. Lastly, in Section 7 we pressent extensions/refinements of Theorems 1.4 and 6.1 for dissociated random arrays (Theorem 7.1), for vector-valued functions of random arrays (Theorem 7.3) and a simultaneous conditional concentration result (Theorem 7.4).

Part 2 consists of Sections 8 and 9 and it is entirely devoted to the connection of our results with combinatorics. In Section 8 we give examples of combinatorial structures for which our conditional concentration results are applicable, and in Section 9 we give the proof of Theorem 1.8.

Finally, in Appendix A we present examples which show the optimality of the box independence condition.

Part 1. Proofs of the main results

2. From dissociativity to concentration

2.1. **Main result.** Let d be a positive integer, and recall that a d-dimensional random array $\boldsymbol{X}$ on a (possibly infinite) subset I of $\mathbb{N}$ is called *dissociated* if for every $J,K\subseteq I$ with $|J|,|K|\geqslant d$ and $\max(J)<\min(K)$, the σ-algebras $\mathcal{F}_J$ and $\mathcal{F}_K$ are independent, that is, for every $A\in\mathcal{F}_J$ and $B\in\mathcal{F}_K$ we have $\mathbb{P}(A\cap B)=\mathbb{P}(A)\,\mathbb{P}(B)$. Dissociativity is a classical concept in probability (see [MS75]); we will need the following approximate version of this notion.

Definition 2.1 (Approximate dissociativity)**.** *Let n,ℓ,d be positive integers such that $n\geqslant \ell\geqslant 2d$, and let $0\leqslant\beta\leqslant 1$. We say that a d-dimensional random array $\boldsymbol{X}$ on $[n]$ is* (β,ℓ)-dissociated *provided that for every $J,K\subseteq[n]$ with $|J|,|K|\geqslant d$, $|J|+|K|\leqslant\ell$ and $\max(J)<\min(K)$, and every pair of events $A\in\mathcal{F}_J$ and $B\in\mathcal{F}_K$ we have*

$$\big|\mathbb{P}(A\cap B)-\mathbb{P}(A)\,\mathbb{P}(B)\big|\leqslant\beta. \tag{2.1}$$

The following theorem—which is the main result in this section—provides the link between conditional concentration and approximate dissociativity.

Theorem 2.2. *Let d be a positive integer, let $1 < p \leqslant 2$, let $0 < \varepsilon \leqslant 1$, let $k \geqslant d$ be an integer, and set*

$$\beta = \beta(p,\varepsilon) \coloneqq \Big(\frac{\varepsilon}{10}\Big)^{\frac{10}{p-1}} \tag{2.2}$$

$$\ell = \ell(p,\varepsilon,k) \coloneqq \Big\lceil \frac{4}{\varepsilon^4(p-1)}\, k\Big\rceil. \tag{2.3}$$

Also let $n \geqslant \ell$ be an integer, and let $\boldsymbol{X}$ be a (β,ℓ)-dissociated, d-dimensional random array on $[n]$ whose entries take values in a measurable space $\mathcal{X}$. Then for every measurable function $f\colon \mathcal{X}^{\binom{[n]}{d}} \to \mathbb{R}$ with $\mathbb{E}[f(\boldsymbol{X})] = 0$ and $\|f(\boldsymbol{X})\|_{L_p} = 1$ there exists an interval I of $[n]$ with $|I| = k$ such that for every $J \subseteq I$ with $|J| \geqslant d$ we have

$$\mathbb{P}\big(\big|\mathbb{E}[f(\boldsymbol{X})\,|\,\mathcal{F}_J]\big| \leqslant \varepsilon\big) \geqslant 1-\varepsilon. \tag{2.4}$$

We note that for spreadable random arrays there is a converse of Theorem 2.2, namely, approximate dissociativity is in fact necessary in order to have conditional concentration; see Proposition 2.8 in Subsection 2.6.

2.2. **Moment bound.** The following moment estimate is the main step of the proof of Theorem 2.2.

Theorem 2.3. *Let d,ℓ,n be positive integers with $n \geqslant \ell \geqslant 2d$, let $0 < \beta \leqslant 1$, and let $\boldsymbol{X}$ be a d-dimensional random array on $[n]$ which is (β,ℓ)-dissociated and whose entries take values in a measurable space $\mathcal{X}$. Then, for every $1 < p \leqslant 2$, every measurable function $f\colon \mathcal{X}^{\binom{[n]}{d}} \to \mathbb{R}$ with $f(\boldsymbol{X}) \in L_p$, every integer k with $d \leqslant k \leqslant \lfloor \ell/2 \rfloor$, and every $I \in \binom{[n]}{\ell}$, there exists $J \in \binom{I}{k}$ with the following property. For any $1 \leqslant r < p$, we have*

$$\big\|\mathbb{E}[f(\boldsymbol{X})\,|\,\mathcal{F}_J] - \mathbb{E}[f(\boldsymbol{X})]\big\|_{L_r} \leqslant \Big((p-1)^{-1/2}\sqrt{\frac{2k}{\ell}} + 10\beta^{\frac{1}{r}-\frac{1}{p}}\Big)\, \big\|f(\boldsymbol{X}) - \mathbb{E}[f(\boldsymbol{X})]\big\|_{L_p} \tag{2.5}$$

where $\mathcal{F}_J$ denotes the σ-algebra generated by the subarray $\boldsymbol{X}_J$. (See Definition 1.1.) Moreover, if I is an interval of $[n]$, then J is an interval too.

Theorem 2.3 easily yields Theorem 2.2. We present the details below.

Proof of Theorem 2.2 assuming Theorem 2.3. Set $r \coloneqq (p+1)/2$ and notice that with this choice we have $1 < r < p \leqslant 2$. Since $\mathbb{E}[f(\boldsymbol{X})] = 0$ and $\|f(\boldsymbol{X})\|_{L_p} = 1$, by Theorem 2.3 applied for the interval $I_1 \coloneqq [\ell]$, there exists an interval I_2 of $[\ell]$ with $|I_2| = k$ such that

$$\big\|\mathbb{E}[f(\boldsymbol{X})\,|\,\mathcal{F}_{I_2}]\big\|_{L_r} \leqslant (p-1)^{-1/2}\sqrt{\frac{2k}{\ell}} + 10\beta^{\frac{1}{r}-\frac{1}{p}}. \tag{2.6}$$

We claim that the interval I_2 is as desired. Indeed, fix a subset J of I_2 with $|J| \geqslant d$, and observe that $\mathcal{F}_J \subseteq \mathcal{F}_{I_2}$. Therefore, by (2.6) and the fact that the conditional expectation is a linear contraction on L_r, we obtain that

$$\big\|\mathbb{E}[f(\boldsymbol{X})\,|\,\mathcal{F}_J]\big\|_{L_r} \leqslant (p-1)^{-1/2}\sqrt{\frac{2k}{\ell}} + 10\beta^{\frac{1}{r}-\frac{1}{p}}.$$

By Markov's inequality, this estimate yields that

$$\mathbb{P}\big(\big|\mathbb{E}[f(\boldsymbol{X})\,|\,\mathcal{F}_J]\big| \geqslant \varepsilon\big) \leqslant (1/\varepsilon)^r \cdot \Big((p-1)^{-1/2}\sqrt{\frac{2k}{\ell}} + 10\beta^{\frac{1}{r}-\frac{1}{p}}\Big)^r. \tag{2.7}$$

By (2.7), the choice of r and the choice of β and ℓ in (2.2) and (2.3) respectively, we conclude that

$$\mathbb{P}\big(\big|\mathbb{E}[f(\boldsymbol{X})\,|\,\mathcal{F}_J]\big| \geqslant \varepsilon\big) \leqslant \varepsilon. \tag{2.8}$$

which clearly implies (2.4). The proof of Theorem 2.3 is completed. □

The rest of this section is devoted to the proof of Theorem 2.3 which is based on inequalities for martingales in L_p spaces. Martingales are, of course, standard tools in the proofs of concentration estimates. Typically, one decomposes a given random variable X into martingale increments, and then controls an appropriate norm of X by controlling the norm of the increments. In the proof of Theorem 2.3 we also decompose a given random variable into martingale increments but, in contrast, we seek to find one of the increments which has controlled norm. This method, known as the *energy increment strategy*, was introduced in the present probabilistic setting by Tao [Tao06] for "$p = 2$", and then extended in the full range of admissible p's in [DKT16]. Having said that, we also note that the main novelty of the present paper lies in the selection of the filtration.

We now briefly describe the contents of the rest of this section. In Subsection 2.3 we present the analytical estimate which is used[7] in the proof of Theorem 2.3. In Subsection 2.4 we prove an orthogonality result for pairs of σ-algebras which satisfy the estimate (2.1). The proof of Theorem 2.3 is completed in Subsection 2.5. Finally, in Subsection 2.6 we show that, for spreadable random arrays, the assumption of approximate dissociativity in Theorem 2.2 is necessary.

2.3. **Martingale difference sequences.** It is an elementary, though important, fact that martingale difference sequences are orthogonal in L_2. We will need the following extension of this fact.

Proposition 2.4. *Let $1 < p \leqslant 2$. Then for every martingale difference sequence $(d_i)_{i=1}^m$ in L_p we have*

$$\Big(\sum_{i=1}^m \|d_i\|_{L_p}^2\Big)^{1/2} \leqslant \big(p-1\big)^{-1/2}\,\Big\|\sum_{i=1}^m d_i\Big\|_{L_p}. \tag{2.9}$$

In particular,

$$\min_{1\leqslant i\leqslant m} \|d_i\|_{L_p} \leqslant \frac{1}{\sqrt{m(p-1)}}\,\Big\|\sum_{i=1}^m d_i\Big\|_{L_p}. \tag{2.10}$$

[7] Square-function estimates could also be used, but they do not yield optimal dependence with respect to the integrability parameter p.

We note that the constant $(p-1)^{-1/2}$ in (2.9) is optimal; this sharp estimate was proved by Ricard and Xu [RX16] who deduced it from a uniform convexity inequality for L_p spaces—see [Pi11, Lemma 4.32]. (See, also, [DKK16, Appendix A] for an exposition.)

2.4. **Mixing and orthogonality.** In what follows, it is convenient to introduce the following terminology. Let $(\Omega,\Sigma,\mathbb{P})$ be a probability space, and let $0\leqslant\beta\leqslant 1$; given two sub-$\sigma$-algebras $\mathcal{A},\mathcal{B}$ of Σ, we say that $\mathcal{A}$ and $\mathcal{B}$ are *β-mixing* provided that for every $A\in\mathcal{A}$ and every $B\in\mathcal{B}$ we have

$$\big|\mathbb{P}(A\cap B)-\mathbb{P}(A)\,\mathbb{P}(B)\big|\leqslant\beta. \tag{2.11}$$

Notice that in the extreme case "$\beta=0$" the estimate (2.11) is equivalent to saying that the σ-algebras $\mathcal{A}$ and $\mathcal{B}$ are independent, which in turn implies for every random variable X with $\mathbb{E}[X]=0$ we have $\mathbb{E}\big[\mathbb{E}[X\,|\,\mathcal{A}]\,|\,\mathcal{B}\big]=0$. The main result in this subsection (Proposition 2.7 below) is an approximate version of this fact.

We start with the following lemma.

Lemma 2.5. *Let $(\Omega,\Sigma,\mathbb{P})$ be a probability space, let $0<\beta\leqslant 1$, and let $\mathcal{A},\mathcal{B}$ be two sub-σ-algebras of Σ which are β-mixing. Then for every real-valued, bounded, random variable X and every $1\leqslant p\leqslant\infty$ we have*

$$\big\|\mathbb{E}\big[\mathbb{E}[X\,|\,\mathcal{A}]\,|\,\mathcal{B}\big]-\mathbb{E}[X]\big\|_{L_p}\leqslant(4\beta)^{1/p}\,\|X-\mathbb{E}[X]\|_{L_\infty}. \tag{2.12}$$

For the proof of Lemma 2.5 we need the following simple fact.

Fact 2.6. *Let (X,Σ,μ) be a measure space, and let $f\colon X\to\mathbb{R}$ be an integrable function. Then we have*

$$\|f\|_{L_1(\mu)}\leqslant 2\sup_{A\in\Sigma}\Big|\int_A f\,d\mu\Big|. \tag{2.13}$$

In particular, if $x_1,\dots,x_m\in\mathbb{R}$, then

$$\sum_{i=1}^m|x_i|\leqslant 2\max_{\emptyset\neq I\subseteq[m]}\Big|\sum_{i\in I}x_i\Big|. \tag{2.14}$$

Proof. Since $[f\geqslant 0],[f<0]\in\Sigma$, we have

$$\|f\|_{L_1(\mu)}=\Big|\int_{[f\geqslant 0]}f\,d\mu\Big|+\Big|\int_{[f<0]}f\,d\mu\Big|\leqslant 2\sup_{A\in\Sigma}\Big|\int_A f\,d\mu\Big|$$

as desired. □

We proceed to the proof of Lemma 2.5.

Proof of Lemma 2.5. We prove the L_1-estimate; the L_p-estimate for $p>1$ follows from the L_1-L_∞ bound, and the fact that the conditional expectation is a linear contraction on L_∞. Without loss of generality we may assume that $\mathbb{E}[X]=0$. (If not, then we work with the random variable $X':=X-\mathbb{E}[X]$ instead of X). Set $Z:=\mathbb{E}[X\,|\,\mathcal{A}]$, and observe that $\mathbb{E}[Z]=\mathbb{E}[X]=0$. Hence, by Fact 2.6, it suffices to obtain an upper bound for

$\big|\mathbb{E}[Z\mathbf{1}_B]\big|$ for arbitrary $B\in\mathcal{B}$. To this end, note that $\|Z\|_{L_\infty}\leqslant\|X\|_{L_\infty}$; therefore, by a standard approximation, we may assume that Z is of the form $\sum_{i=1}^N a_i\mathbf{1}_{A_i}$ where N is a positive integer, $|a_i|\leqslant\|Z\|_{L_\infty}$ for every $i\in[N]$, and the family $\{A_1,\dots,A_N\}$ forms a partition of Ω into measurable events. Let $B\in\mathcal{B}$ be arbitrary. Using the fact that $\sum_{i=1}^N a_i\,\mathbb{P}(A_i)=\mathbb{E}[Z]=0$ and the triangle inequality, we have

$$\big|\mathbb{E}[Z\mathbf{1}_B]\big|=\Big|\sum_{i=1}^N a_i\,\mathbb{P}(A_i\cap B)\Big|\leqslant\sum_{i=1}^N|a_i|\cdot|\mathbb{P}(A_i\cap B)-\mathbb{P}(A_i)\,\mathbb{P}(B)|. \tag{2.15}$$

If we set $x_i:=\mathbb{P}(A_i\cap B)-\mathbb{P}(A_i)\,\mathbb{P}(B)$, we obtain that

$$\big|\mathbb{E}[Z\mathbf{1}_B]\big|\leqslant\sum_{i=1}^N|a_i|\cdot|x_i|\leqslant 2\|Z\|_{L_\infty}\max_{\emptyset\neq I\subseteq[N]}\Big|\sum_{i\in I}x_i\Big| \tag{2.16}$$

where we have also used the pointwise bound $|a_i|\leqslant\|Z\|_{L_\infty}$ and Fact 2.6. Finally, setting $A_I:=\bigcup_{i\in I}A_i$ for every nonempty $I\subseteq[N]$, then we have

$$\Big|\sum_{i\in I}x_i\Big|=\big|\mathbb{P}(A_I\cap B)-\mathbb{P}(A_I)\,\mathbb{P}(B)\big|\leqslant\beta \tag{2.17}$$

since the sets $A_1,\dots,A_N$ are pairwise disjoint and $A_I\in\mathcal{A}$. We conclude that

$$\big|\mathbb{E}\big[\mathbb{E}[Z\,|\,\mathcal{B}]\mathbf{1}_B\big]\big|=\big|\mathbb{E}[Z\mathbf{1}_B]\big|\leqslant 2\beta\|X\|_{L_\infty}. \tag{2.18}$$

Since $B\in\mathcal{B}$ was arbitrary, the result follows. □

We are now ready to state the main result in this subsection.

Proposition 2.7. *Let $(\Omega,\Sigma,\mathbb{P})$ be a probability space, let $0<\beta\leqslant1$, and let $\mathcal{A},\mathcal{B}$ be two sub-σ-algebras of Σ which are β-mixing. Let $1\leqslant r<p\leqslant\infty$, and let $X\in L_p$. Then,*

$$\big\|\mathbb{E}\big[\mathbb{E}[X\,|\,\mathcal{A}]\,|\,\mathcal{B}\big]-\mathbb{E}[X]\big\|_{L_r}\leqslant 10\beta^{\frac1r-\frac1p}\,\|X-\mathbb{E}[X]\|_{L_p}. \tag{2.19}$$

Proof. We will obtain the estimate by truncating X and employing Lemma 2.5. We lay out the details. As in the proof of Lemma 2.5, we may assume that $\mathbb{E}[X]=0$. Let $t>0$ (to be chosen later) be the truncation level, and set $X_t:=X\mathbf{1}_{[|X|\leqslant t]}$. Markov's inequality yields that $\mathbb{P}(|X|>t)\leqslant t^{-p}\|X\|_{L_p}^p$, thus applying Hölder's inequality we obtain that

$$\|X_t-X\|_{L_r}^r=\mathbb{E}\big[|X|^r\,\mathbf{1}_{[|X|>t]}\big]\leqslant\|X\|_{L_p}^r\,\mathbb{P}(|X|>t)^{1-\frac rp}\leqslant\frac{\|X\|_{L_p}^p}{t^{p-r}} \tag{2.20}$$

for any $1\leqslant r<p$. Therefore,

$$\begin{aligned}\big\|\mathbb{E}\big[\mathbb{E}[X\,|\,\mathcal{A}]\,|\,\mathcal{B}\big]\big\|_{L_r}&\leqslant\big\|\mathbb{E}\big[\mathbb{E}[X-X_t\,|\,\mathcal{A}]\,|\,\mathcal{B}\big]\big\|_{L_r}+\\&\qquad+\big\|\mathbb{E}\big[\mathbb{E}[X_t\,|\,\mathcal{A}]\,|\,\mathcal{B}\big]-\mathbb{E}[X_t]\big\|_{L_r}+\big|\mathbb{E}[X_t]\big|\\&\leqslant\|X-X_t\|_{L_r}+2(4\beta)^{1/r}t+\|X-X_t\|_{L_1}\end{aligned} \tag{2.21}$$

where we have used the contraction property of the conditional expectation, Lemma 2.5 for the random variable X_t, and the fact $\mathbb{E}[X]=0$, respectively. Taking into account (2.20), we conclude that

$$\big\|\mathbb{E}\big[\mathbb{E}[X\,|\,\mathcal{A}]\,|\,\mathcal{B}\big]\big\|_{L_r} \leqslant 2\frac{\|X\|_{L_p}^{p/r}}{t^{\frac{p}{r}-1}} + 8\beta^{1/r}t. \tag{2.22}$$

It remains to optimize the latter with respect to t; the choice $t \coloneqq \beta^{-1/p}\|X\|_{L_p}$ yields the assertion. □

2.5. **Proof of Theorem 2.3.** After normalizing, we may assume that

$$\big\|f(\boldsymbol{X}) - \mathbb{E}[f(\boldsymbol{X})]\big\|_{L_p} = 1. \tag{2.23}$$

Fix an integer k with $d \leqslant k < \lfloor \ell/2 \rfloor$ and $I \in \binom{[n]}{\ell}$, and let $\{\iota_1 < \cdots < \iota_\ell\}$ denote the increasing enumeration of I. Set $m \coloneqq \lfloor \ell/k \rfloor$. Also let $K_1,\dots,K_m \in \binom{[\ell]}{k}$ be successive intervals with $\min(K_1)=1$, and set $J_i \coloneqq \{\iota_\kappa : \kappa \in K_i\}$ for every $i\in[m]$. Thus, the sets $J_1,\dots,J_m$ are successive subsets of I each of cardinality k; also notice that if I is an interval of $[n]$, then the sets $J_1,\dots,J_m$ are intervals too.

Next, denote by $(\Omega,\Sigma,\mathbb{P})$ the underlying probability space on which the random array $\boldsymbol{X}$ is defined, and for every $i\in[m]$ let $\mathcal{F}_{J_i}$ be the σ-algebra generated by the subarray $\boldsymbol{X}_{J_i}$. (See Definition 1.1). We define a filtration $(\mathcal{A}_i)_{i=0}^m$ by setting $\mathcal{A}_0 = \{\emptyset,\Omega\}$ and

$$\mathcal{A}_i \coloneqq \sigma\big(\{\mathcal{F}_{J_1},\dots,\mathcal{F}_{J_i}\}\big) \quad \text{for every } i\in[m]. \tag{2.24}$$

We will use variants of this filtration in Section 9.

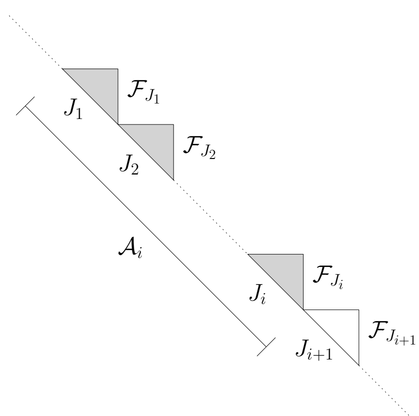

FIGURE 3. The filtration $(\mathcal{A}_i)_{i=0}^m$.

Let $(d_i)_{i=1}^m$ denote the martingale difference sequence of the Doob martingale for $f(\boldsymbol{X})$ with respect to the filtration $(\mathcal{A}_i)_{i=0}^m$, that is, $d_i \coloneqq \mathbb{E}[f(\boldsymbol{X})\,|\,\mathcal{A}_i] - \mathbb{E}[f(\boldsymbol{X})\,|\,\mathcal{A}_{i-1}]$ for every $i \in [m]$. Since $\mathbb{E}[f(\boldsymbol{X}) \,|\, \mathcal{A}_m] - \mathbb{E}[f(\boldsymbol{X})] = \sum_{i=1}^m d_i$, the contractive property of the conditional expectation yields that

$$\Big\| \sum_{i=1}^m d_i \Big\|_{L_p} \leqslant \big\| f(\boldsymbol{X}) - \mathbb{E}[f(\boldsymbol{X})] \big\|_{L_p} \overset{(2.23)}{=} 1. \tag{2.25}$$

Therefore, by Proposition 2.4, there exists an integer $i_0 \in [m]$ so that

$$\|d_{i_0}\|_{L_p} \leqslant \frac{1}{\sqrt{m(p-1)}}. \tag{2.26}$$

We claim that the set $J \coloneqq J_{i_0}$ is as desired.

To this end, fix $1 \leqslant r < p$. First observe that, conditioning further on $\mathcal{F}_{J_{i_0}}$,

$$\big\| \mathbb{E}[f(\boldsymbol{X}) \,|\, \mathcal{F}_{J_{i_0}}] - \mathbb{E}\big[\mathbb{E}[f(\boldsymbol{X}) \,|\, \mathcal{A}_{i_0-1}] \,|\, \mathcal{F}_{J_{i_0}}\big] \big\|_{L_p} = \big\| \mathbb{E}[d_{i_0} \,|\, \mathcal{F}_{J_{i_0}}] \big\|_{L_p} \leqslant \frac{1}{\sqrt{m(p-1)}} \tag{2.27}$$

where we have used the fact that $\mathcal{F}_{J_{i_0}} \subseteq \mathcal{A}_{i_0}$, the contractive property of the conditional expectation once more, and (2.26). By the triangle inequality and taking into account (2.27) and the monotonicity of the L_p-norms, we obtain that

$$\begin{aligned} \big\| \mathbb{E}[f(\boldsymbol{X}) \,|\, \mathcal{F}_{J_{i_0}}] - \mathbb{E}[f(\boldsymbol{X})] \big\|_{L_r} \leqslant & \frac{1}{\sqrt{m(p-1)}} + \\ & + \big\| \mathbb{E}\big[\mathbb{E}[f(\boldsymbol{X}) \,|\, \mathcal{A}_{i_0-1}] \,|\, \mathcal{F}_{J_{i_0}}\big] - \mathbb{E}[f(\boldsymbol{X})] \big\|_{L_r} \end{aligned} \tag{2.28}$$

Finally, by (2.24) and our assumption that the random array $\boldsymbol{X}$ is (β, ℓ)-dissociated, we see that the σ-algebras $\mathcal{F}_{J_{i_0}}$ and $\mathcal{A}_{i_0-1}$ are β-mixing. (See (2.11).) By Proposition 2.7, we conclude that

$$\big\| \mathbb{E}[f(\boldsymbol{X}) \,|\, \mathcal{F}_{J_{i_0}}] - \mathbb{E}[f(\boldsymbol{X})] \big\|_{L_r} \leqslant \frac{1}{\sqrt{m(p-1)}} + 10\beta^{\frac{1}{r} - \frac{1}{p}}. \tag{2.29}$$

and the proof is completed.

2.6. **Necessity of approximate dissociativity.** We close this section with the following proposition which shows that the assumption of approximate dissociativity in Theorem 2.2 is necessary.

Proposition 2.8. *Let n, d, ℓ be positive integers with $n \geqslant \ell \geqslant d$, let $0 < \beta \leqslant 1$, let $\boldsymbol{X}$ be a spreadable, d-dimensional random array on $[n]$ whose entries take vales in a measurable space $\mathcal{X}$, and assume that $\boldsymbol{X}$ is not (β, ℓ)-dissociated. Then there exists a measurable function $f \colon \mathcal{X}^{\binom{[n]}{d}} \to \{0, 1\}$ such that for every $I \in \binom{[n]}{\ell}$ we have*

$$\mathbb{P}\big(\big| \mathbb{E}[f(\boldsymbol{X}) \,|\, \mathcal{F}_I] - \mathbb{E}[f(\boldsymbol{X})] \big| \geqslant \beta/2 \big) \geqslant \beta/2. \tag{2.30}$$

Proof. Since the random array $\boldsymbol{X}$ is spreadable and not (β, ℓ)-dissociated, there exist two integers $j, k \geqslant d$ with $j + k \leqslant \ell$, and two events $A \in \mathcal{F}_{[j]}$ and $B \in \mathcal{F}_K$, where $K \coloneqq \{j+1, \dots, k+j\}$, such that $|\mathbb{P}(A \cap B) - \mathbb{P}(A)\, \mathbb{P}(B)| \geqslant \beta$. We select a measurable subset A' of $\mathcal{X}^{\binom{[j]}{d}}$ such that the events $[\boldsymbol{X}_{[j]} \in A']$ and A agree almost surely, and we set

$\widetilde{A} \coloneqq \pi^{-1}(A')$ where $\pi\colon \mathcal{X}^{\binom{[n]}{d}} \to \mathcal{X}^{\binom{[j]}{d}}$ denotes the natural projection. Finally, we define $f\colon \mathcal{X}^{\binom{[n]}{d}} \to \{0,1\}$ by $f = \mathbf{1}_{\widetilde{A}}$.

We claim that f is as desired. Indeed, let $I \in \binom{[n]}{\ell}$ be arbitrary. We select $L \in \binom{I}{k}$ with $\min(L) > j$. Invoking the spreadability of $\boldsymbol{X}$ and the choice of A and B, we may also select $\Gamma \in \mathcal{F}_L$ such that

$$\big|\mathbb{P}(A \cap \Gamma) - \mathbb{P}(A)\,\mathbb{P}(\Gamma)\big| \geqslant \beta. \tag{2.31}$$

Observing that $\mathbb{P}(A) = \mathbb{E}[f(\boldsymbol{X})]$ and $\mathbb{P}(A\cap\Gamma) = \mathbb{E}[f(\boldsymbol{X})\mathbf{1}_\Gamma]$, and using the fact that $\Gamma \in \mathcal{F}_L \subseteq \mathcal{F}_I$, we obtain that

$$\beta \overset{(2.31)}{\leqslant} \big|\mathbb{E}[\big(f(\boldsymbol{X}) - \mathbb{E}[f(\boldsymbol{X})]\big)\mathbf{1}_\Gamma]\big| = \big|\mathbb{E}[\big(\mathbb{E}[f(\boldsymbol{X})\,|\,\mathcal{F}_I] - \mathbb{E}[f(\boldsymbol{X})]\big)\mathbf{1}_\Gamma]\big| \tag{2.32}$$

which is easily seen to imply (2.30). The proof is completed. □

Remark 2.9. We note that if the random array $\boldsymbol{X}$ in Proposition 2.8 is boolean, then the function f defined above is a polynomial of degree at most $\binom{\ell}{d}$.

3. The box independence condition propagates

3.1. **The main result.** We start by introducing some pieces of notation and some terminology. Let n,d be a positive integers with $n \geqslant 2d$; for every finite sequence $\mathcal{H} = (H_1,\dots,H_d)$ of nonempty finite subsets of $[n]$ with[8] $\max(H_i) < \min(H_{i+1})$ for all $i \in [d-1]$, we set

$$\mathrm{Box}(\mathcal{H}) \coloneqq \Big\{ s \in \binom{[n]}{d} : |s \cap H_i| = 1 \text{ for all } i \in [d] \Big\}. \tag{3.1}$$

If, in addition, we have $|H_i| = 2$ for all $i \in [d]$, then we say that the set $\mathrm{Box}(\mathcal{H})$ is a *d-dimensional box of* $[n]$. By $\mathrm{Box}(d)$ we shall denote the d-dimensional box corresponding to the sequence $(\{1,2\},\dots,\{2d-1,2d\})$, that is,

$$\mathrm{Box}(d) = \Big\{ s \in \binom{[n]}{d} : |s \cap \{2i-1, 2i\}| = 1 \text{ for all } i \in [d] \Big\}. \tag{3.2}$$

We proceed with the following definition. Note that the "$(\vartheta,\mathcal{S})$-box independence" condition introduced below is the one-sided version of (1.6); we will work with this slightly weaker version since it is more amenable to an inductive argument.

Definition 3.1. *Let n,d be positive integers with $n \geqslant 2d$, let $\mathcal{X}$ be a nonempty finite set, and let $\boldsymbol{X} = \langle X_s : s \in \binom{[n]}{d}\rangle$ be an $\mathcal{X}$-valued, d-dimensional random array on $[n]$. Also let $\mathcal{S}$ be a nonempty subset of $\mathcal{X}$.*

(i) (Box independence) *Let $\vartheta > 0$. We say that $\boldsymbol{X}$ is* $(\vartheta,\mathcal{S})$-box independent *if for every d-dimensional box B of $[n]$ and every $a \in \mathcal{S}$ we have*

$$\mathbb{P}\Big(\bigcap_{s\in B}[X_s = a]\Big) \leqslant \prod_{s \in B} \mathbb{P}\big([X_s = a]\big) + \vartheta. \tag{3.3}$$

[8] Note that if $d = 1$, then this condition is superfluous.

(ii) (Approximate independence) *Set* $\ell := \binom{\lfloor n/2 \rfloor}{d}$, *and let* $\boldsymbol{\gamma} = (\gamma_k)_{k=1}^{\ell}$ *be a finite sequence of positive reals. We say that* $\boldsymbol{X}$ *is* $(\boldsymbol{\gamma}, \mathcal{S})$-independent *if for every nonempty subset* $\mathcal{F}$ *of* $\binom{[n]}{d}$ *such that* $\cup\mathcal{F}$ *has cardinality at most* $n/2$, *and every collection* $(a_s)_{s\in\mathcal{F}}$ *of elements of* $\mathcal{S}$ *we have*

$$\Big|\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}}[X_s = a_s]\Big) - \prod_{s\in\mathcal{F}}\mathbb{P}\big([X_s = a_s]\big)\Big| \leqslant \gamma_{|\mathcal{F}|}. \tag{3.4}$$

We are ready to state the main result in this section. It is the higher-dimensional version of Theorem 1.5, and its proof is given in Section 4. (The numerical invariants appearing below are defined in Subsection 4.2.)

Theorem 3.2. *Let* d, n *be positive integers with* $n \geqslant 4d$, *let* $0 < \eta, \vartheta \leqslant 1$, *and set* $\ell := \binom{\lfloor n/2 \rfloor}{d}$. *Then there exists a sequence* $\boldsymbol{\gamma} = (\gamma_k(\eta, \vartheta, d, n))_{k=1}^{\ell}$ *of positive reals such that* $\gamma_k(\eta, \vartheta, d, n)$ *tends to zero as* n *tends to infinity and* η, ϑ *tend to zero, and with the following property.*

Let $\mathcal{X}$ *be a finite set, let* $\mathcal{S}$ *be a nonempty subset of* $\mathcal{X}$, *and let* $\boldsymbol{X}$ *be an* $\mathcal{X}$*-valued,* η*-spreadable,* d*-dimensional random array on* $[n]$. *If* $\boldsymbol{X}$ *is* $(\vartheta, \mathcal{S})$*-box independent, then* $\boldsymbol{X}$ *is also* $(\boldsymbol{\gamma}, \mathcal{S})$*-independent.*

3.2. **Consequences.** The rest of this section is devoted to the proof of two consequences of Theorem 3.2. The first consequence shows that the box independence condition implies approximate dissociativity. Specifically, we have the following corollary.

Corollary 3.3. *For every triple* d, ℓ, m *of positive integers with* $\ell \geqslant 2d$ *and* $m \geqslant 2$, *and every* $\beta > 0$, *there exist an integer* $N \geqslant \ell$ *and two constants* $\eta, \vartheta > 0$ *with the following property.*

Let $n \geqslant N$ *be an integer, let* $\mathcal{X}$ *be a set with* $|\mathcal{X}| = m$, *let* $\mathcal{S}$ *be a subset of* $\mathcal{X}$ *with* $|\mathcal{S}| = |\mathcal{X}| - 1$, *and let* $\boldsymbol{X}$ *be an* $\mathcal{X}$*-valued,* η*-spreadable,* d*-dimensional random array on* $[n]$. *If* $\boldsymbol{X}$ *is* $(\vartheta, \mathcal{S})$*-box independent, then* $\boldsymbol{X}$ *is* (β, ℓ)*-dissociated.* (*See Definition* 2.1.)

The second consequence of Theorem 3.2 shows that the box independence forces all sub-processes indexed by d-dimensional boxes to behave independently. More precisely, we have the following corollary.

Corollary 3.4. *For every pair* d, m *of positive integers with* $m \geqslant 2$, *and every* $\gamma > 0$, *there exist an integer* $N \geqslant 2d$ *and* $\eta, \vartheta > 0$ *with the following property.*

Let $n, \mathcal{X}, \mathcal{S}, \boldsymbol{X} = \langle X_s : s \in \binom{[n]}{d}\rangle$ *be as in Corollary* 3.3. *If* $\boldsymbol{X}$ *is* $(\vartheta, \mathcal{S})$*-box independent, then for every* d*-dimensional box* B *of* $[n]$ *and every collection* $(a_s)_{s\in B}$ *of elements of* $\mathcal{X}$ *we have*

$$\Big|\mathbb{P}\Big(\bigcap_{s\in B}[X_s = a_s]\Big) - \prod_{s\in B}\mathbb{P}\big([X_s = a_s]\big)\Big| \leqslant \gamma. \tag{3.5}$$

Remark 3.5. Although Corollary 3.4 is weaker than Theorem 3.2, a direct proof of the estimate (3.5) is likely to require the whole machinery presented in Section 4.

The deduction of Corollaries 3.3 and 3.4 from Theorem 3.2 is based on the following lemma.

Lemma 3.6. *For every triple d, m, κ of positive integers with $m \geqslant 2$, and every $\gamma > 0$, there exist an integer $N \geqslant 1$ and $\eta, \vartheta > 0$ with the following property.*

Let $n, \mathcal{X}, \mathcal{S}, \boldsymbol{X} = \langle X_s : s \in \binom{[n]}{d} \rangle$ be as in Corollary 3.3. If $\boldsymbol{X}$ is $(\vartheta, \mathcal{S})$-box independent, then for every nonempty subset $\mathcal{F}$ of $\binom{[n]}{d}$ with $|\mathcal{F}| \leqslant \kappa$ and every collection $(a_s)_{s \in \mathcal{F}}$ of elements of $\mathcal{X}$ we have

$$\Big| \mathbb{P}\Big(\bigcap_{s \in \mathcal{F}} [X_s = a_s] \Big) - \prod_{s \in \mathcal{F}} \mathbb{P}\big([X_s = a_s] \big) \Big| \leqslant \gamma. \tag{3.6}$$

We defer the proof of Lemma 3.6 to Subsection 3.3 below. At this point, let us give the proofs of Corollaries 3.3 and 3.4.

Proof of Corollary 3.3. Let d, ℓ, m, β be as in the statement of the corollary. Let N, η, ϑ be as in Lemma 3.6 applied for $\kappa := \binom{\ell}{d}$ and $\gamma := \frac{1}{3} m^{-2\kappa} \beta$. (Clearly, we may assume that $N \geqslant \ell$.) We claim that N, η and ϑ are as desired.

Indeed, fix $n, \mathcal{X}, \mathcal{S}, \boldsymbol{X}$ and recall that we need to show that $\boldsymbol{X}$ is (β, ℓ)-dissociated. To this end, let J, K be subsets of $[n]$ with $|J|, |K| \geqslant d$, $|J| + |K| \leqslant \ell$ and $\max(J) < \min(K)$, and let $A \in \mathcal{F}_J$ and $B \in \mathcal{F}_K$. We will show that $|\mathbb{P}(A \cap B) - \mathbb{P}(A)\,\mathbb{P}(B)| \leqslant \beta$.

Since A belongs to the σ-algebra generated by $\mathbf{X}_J$, there exists a collection $\mathcal{A}$ of maps of the form $\mathbf{a}\colon \binom{J}{d} \to \mathcal{X}$ such that

$$A = \bigcup_{\mathbf{a} \in \mathcal{A}} \bigcap_{s \in \binom{J}{d}} [X_s = \mathbf{a}(s)]. \tag{3.7}$$

Similarly, there exists a collection $\mathcal{B}$ of maps of the form $\mathbf{b}\colon \binom{K}{d} \to \mathcal{X}$ such that

$$B = \bigcup_{\mathbf{b} \in \mathcal{B}} \bigcap_{t \in \binom{K}{d}} [X_t = \mathbf{b}(t)]. \tag{3.8}$$

For every $\mathbf{a} \in \mathcal{A}$ we set $A_{\mathbf{a}} := \bigcap_{s \in \binom{J}{d}} [X_s = \mathbf{a}(s)]$, respectively, for every $\mathbf{b} \in \mathcal{B}$ we set $B_{\mathbf{b}} := \bigcap_{t \in \binom{K}{d}} [X_t = \mathbf{b}(t)]$. By Lemma 3.6, for every $\mathbf{a} \in \mathcal{A}$ and every $\mathbf{b} \in \mathcal{B}$, we have

$$\Big| \mathbb{P}(A_{\mathbf{a}} \cap B_{\mathbf{b}}) - \prod_{s \in \binom{J}{d}} \mathbb{P}\big([X_s = \mathbf{a}(s)] \big) \prod_{t \in \binom{K}{d}} \mathbb{P}\big([X_t = \mathbf{b}(t)] \big) \Big| \leqslant \gamma \tag{3.9}$$

$$\Big| \mathbb{P}(A_{\mathbf{a}}) - \prod_{s \in \binom{J}{d}} \mathbb{P}\big([X_s = \mathbf{a}(s)] \big) \Big| \leqslant \gamma \tag{3.10}$$

$$\Big| \mathbb{P}(B_{\mathbf{b}}) - \prod_{t \in \binom{K}{d}} \mathbb{P}\big([X_t = \mathbf{b}(t)] \big) \Big| \leqslant \gamma; \tag{3.11}$$

consequently, $|\mathbb{P}(A_{\mathbf{a}} \cap B_{\mathbf{b}}) - \mathbb{P}(A_{\mathbf{a}})\,\mathbb{P}(B_{\mathbf{b}})| \leqslant 3\gamma$. On the other hand, by identities (3.7) and (3.8), we see that $A \cap B = \bigcap_{\mathbf{a} \in \mathcal{A}, \mathbf{b} \in \mathcal{B}} A_{\mathbf{a}} \cap B_{\mathbf{b}}$; moreover, the collections $\langle A_{\mathbf{a}} : \mathbf{a} \in \mathcal{A} \rangle$

and $\langle B_{\mathbf{b}} : \mathbf{b} \in \mathcal{B}\rangle$ consist of pairwise disjoint events. Thus, we have

$$\mathbb{P}(A\cap B) = \sum_{\mathbf{a}\in\mathcal{A},\mathbf{b}\in\mathcal{B}} \mathbb{P}(A_{\mathbf{a}}\cap B_{\mathbf{b}}), \ \ \mathbb{P}(A) = \sum_{\mathbf{a}\in\mathcal{A}} \mathbb{P}(A_{\mathbf{a}}) \ \text{ and } \ \mathbb{P}(B) = \sum_{\mathbf{b}\in\mathcal{B}} \mathbb{P}(B_{\mathbf{b}}). \tag{3.12}$$

Therefore, we conclude that

$$\begin{aligned} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\,\mathbb{P}(B)| &\leqslant \sum_{\mathbf{a}\in\mathcal{A},\mathbf{b}\in\mathcal{B}} |\mathbb{P}(A_{\mathbf{a}}\cap B_{\mathbf{b}}) - \mathbb{P}(A_{\mathbf{a}})\,\mathbb{P}(B_{\mathbf{b}})| \\ &\leqslant 3\gamma\,|\mathcal{A}|\,|\mathcal{B}| \leqslant 3\gamma\, m^{2\binom{\ell}{d}} = \beta. \end{aligned} \tag{3.13}$$

$\square$

Proof of Corollary 3.4. It follows from Lemma 3.6 applied for "$\kappa = 2^d$". $\square$

3.3. **Proof of Lemma 3.6.** The result follows from Theorem 3.2 and the inclusion-exclusion formula. Specifically, let d, m, κ, γ be as in the statement of the lemma, and set $\gamma' \coloneqq m^{-\kappa}\gamma$. By Theorem 3.2, there exist an integer $N \geqslant 4d$ and two constants $0<\eta<1$ and $\vartheta>0$ such that

$$\gamma_k(\eta,\vartheta,d,n) < \gamma' \tag{3.14}$$

for every integer $n \geqslant N$ and $k\in[\kappa]$. We claim that N,η and ϑ are as desired.

Indeed, fix $n, \mathcal{X}, \mathcal{S}$ and $\boldsymbol{X} = \langle X_s : s \in \binom{[n]}{d}\rangle$. By (3.14) and Theorem 3.2, for every nonempty $\mathcal{F}^* \subseteq \binom{[n]}{d}$ with $|\mathcal{F}^*| \leqslant \kappa$ and every collection $(a_s)_{s\in\mathcal{F}^*}$ of elements of $\mathcal{S}$, we have

$$\Big|\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}^*}[X_s = a_s]\Big) - \prod_{s\in\mathcal{F}^*}\mathbb{P}\big([X_s=a_s]\big)\Big| \leqslant \gamma'. \tag{3.15}$$

Let $\mathcal{F}$ be a nonempty subset of $\binom{[n]}{d}$ with $|\mathcal{F}|\leqslant\kappa$, and let $(a_s)_{s\in\mathcal{F}}$ be a collection of elements of $\mathcal{X}$. Set $\mathcal{F}' \coloneqq \{s\in\mathcal{F} : a_s\in\mathcal{S}\}$ and $\mathcal{G}\coloneqq\mathcal{F}\setminus\mathcal{F}'$; observe that for every $t\in\mathcal{G}$ the events $\langle [X_t = a] : a\in\mathcal{S}\rangle$ are pairwise disjoint and, moreover,

$$[X_t = a_t] = \Big(\bigcup_{a\in\mathcal{S}}[X_t=a]\Big)^{\complement}. \tag{3.16}$$

(For any event E, by $E^{\complement}$ we denote its complement.) Thus, for every $t\in\mathcal{G}$ we have $\mathbb{P}\big([X_t=a_t]\big) = 1-\sum_{a\in\mathcal{S}}\mathbb{P}\big([X_t=a]\big)$ and, consequently,

$$\begin{aligned} \prod_{s\in\mathcal{F}}\mathbb{P}\big([X_s=a_s]\big) &= \prod_{s\in\mathcal{F}'}\mathbb{P}\big([X_s=a_s]\big)\prod_{t\in\mathcal{G}}\Big(1-\sum_{a\in\mathcal{S}}\mathbb{P}\big([X_t=a]\big)\Big) \\ &= \sum_{\substack{\mathcal{W}\subseteq\mathcal{G}\\ \mathbf{a}\colon \mathcal{W}\to\mathcal{S}}}(-1)^{|\mathcal{W}|}\prod_{t\in\mathcal{W}}\mathbb{P}\big([X_t=\mathbf{a}(t)]\big)\prod_{s\in\mathcal{F}'}\mathbb{P}\big([X_s=a_s]\big) \end{aligned} \tag{3.17}$$

with the convention that the product over an empty index-set is equal to 1. Moreover,

$$\begin{aligned} \mathbb{P}\Big(\bigcap_{s\in\mathcal{F}}[X_s=a_s]\Big) &\overset{(3.16)}{=} \mathbb{P}\Big(\bigcap_{s\in\mathcal{F}'}[X_s=a_s]\cap\bigcap_{t\in\mathcal{G}}\Big(\bigcup_{a\in\mathcal{S}}[X_t=a]\Big)^{\complement}\Big) \\ &= \mathbb{P}\Big(\bigcap_{s\in\mathcal{F}'}[X_s=a_s]\Big) - \mathbb{P}\Big(\bigcap_{s\in\mathcal{F}'}[X_s=a_s]\cap\Big(\bigcup_{t\in\mathcal{G}}\bigcup_{a\in\mathcal{S}}[X_t=a]\Big)\Big). \end{aligned} \tag{3.18}$$

Next observe that for every nonempty subset $\mathcal{W}$ of $\mathcal{G}$ we have

$$\bigcap_{t\in\mathcal{W}}\bigcup_{a\in\mathcal{S}}[X_t=a]=\bigcup_{\mathbf{a}\colon \mathcal{W}\to\mathcal{S}}\Big(\bigcap_{t\in\mathcal{W}}[X_t=\mathbf{a}(t)]\Big) \tag{3.19}$$

and the events $\langle\bigcap_{t\in\mathcal{W}}[X_t=\mathbf{a}(t)] : \mathbf{a}\colon\mathcal{W}\to\mathcal{S}\rangle$ are pairwise disjoint. Hence, by the inclusion-exclusion formula,

$$\begin{aligned}
&\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}'}[X_s=a_s]\cap\Big(\bigcup_{t\in\mathcal{G}}\bigcup_{a\in\mathcal{S}}[X_t=a]\Big)\Big) \\
&\quad=\sum_{\emptyset\neq\mathcal{W}\subseteq\mathcal{G}}(-1)^{|\mathcal{W}|-1}\,\mathbb{P}\Big(\bigcap_{t\in\mathcal{W}}\Big(\bigcap_{s\in\mathcal{F}'}[X_s=a_s]\cap\Big(\bigcup_{a\in\mathcal{S}}[X_t=a]\Big)\Big)\Big) \\
&\quad\overset{(3.19)}{=}\sum_{\emptyset\neq\mathcal{W}\subseteq\mathcal{G}}(-1)^{|\mathcal{W}|-1}\,\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}'}[X_s=a_s]\cap\Big(\bigcup_{\mathbf{a}\colon\mathcal{W}\to\mathcal{S}}\Big(\bigcap_{t\in\mathcal{W}}[X_t=\mathbf{a}(t)]\Big)\Big)\Big) \\
&\quad=\sum_{\emptyset\neq\mathcal{W}\subseteq\mathcal{G}}\ \sum_{\mathbf{a}\colon\mathcal{W}\to\mathcal{S}}(-1)^{|\mathcal{W}|-1}\,\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}'}[X_s=a_s]\cap\bigcap_{t\in\mathcal{W}}[X_t=\mathbf{a}(t)]\Big).
\end{aligned}\tag{3.20}$$

Combining identities (3.18) and (3.20), we see that

$$\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}}[X_s=a_s]\Big)=\sum_{\substack{\mathcal{W}\subseteq\mathcal{G}\\ \mathbf{a}\colon\mathcal{W}\to\mathcal{S}}}(-1)^{|\mathcal{W}|}\,\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}'}[X_s=a_s]\cap\bigcap_{t\in\mathcal{W}}[X_t=\mathbf{a}(t)]\Big) \tag{3.21}$$

with the convention that the intersection over an empty index-set is equal to the whole sample space. Finally, by identities (3.17) and (3.21) and the triangle inequality, we conclude that the quantity $\big|\mathbb{P}\big(\bigcap_{s\in\mathcal{F}}[X_s=a_s]\big)-\prod_{s\in\mathcal{F}}\mathbb{P}\big([X_s=a_s]\big)\big|$ is upper bounded by

$$\begin{aligned}
&\sum_{\substack{\mathcal{W}\subseteq\mathcal{G}\\ \mathbf{a}\colon\mathcal{W}\to\mathcal{S}}}\Big|\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}'}[X_s=a_s]\cap\bigcap_{t\in\mathcal{W}}[X_t=\mathbf{a}(t)]\Big)- \\
&\qquad\qquad-\prod_{s\in\mathcal{F}'}\mathbb{P}\big([X_s=a_s]\big)\prod_{t\in\mathcal{W}}\mathbb{P}\big([X_t=\mathbf{a}(t)]\big)\Big|\overset{(3.15)}{\leqslant}m^{\kappa}\gamma'=\gamma.
\end{aligned}\tag{3.22}$$

The proof of Lemma 3.6 is completed.

4. Proof of Theorem 3.2

This section is devoted to the proof of Theorem 3.2 which proceeds by induction on the dimension d. In a nutshell, the argument is based on repeated averaging and an appropriate version of the weak law of large numbers in order to gradually upgrade the box independence condition. The combinatorial heart of the matter lies in the selection of this averaging.

4.1. Toolbox.

We begin by presenting three lemmas which are needed for the proof of Theorem 3.2, but they are not directly related with the main argument.

Lemma 4.1. *Let m be a positive integer, let $\delta>0$ and let $A_1,\dots,A_m$ be events in a probability space such that for every $i,j\in[m]$ with $i\neq j$ we have*

$$\mathbb{P}(A_i\cap A_j)\leqslant\mathbb{P}(A_i)\,\mathbb{P}(A_j)+\delta. \tag{4.1}$$

Then, setting $Z \coloneqq \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}_{A_i}$, *we have*

$$\mathrm{Var}(Z) \leqslant \frac{1}{m} + \delta. \tag{4.2}$$

Proof. We have

$$\mathrm{Var}(Z) = \mathbb{E}\big[(Z-\mathbb{E}[Z])^2\big] = \frac{1}{m^2}\sum_{i,j\in[m]} \mathbb{E}\big[\big(\mathbf{1}_{A_i}-\mathbb{P}(A_i)\big)\big(\mathbf{1}_{A_j}-\mathbb{P}(A_j)\big)\big]$$

$$= \frac{1}{m^2}\Big[\sum_{i=1}^{m}\big(\mathbb{P}(A_i)-\mathbb{P}(A_i)^2\big) + \sum_{\substack{i,j\in[m]\\ i\neq j}}\big(\mathbb{P}(A_i\cap A_j)-\mathbb{P}(A_i)\,\mathbb{P}(A_j)\big)\Big] \leqslant \frac{1}{m}+\delta. \qquad \square$$

Lemma 4.2. *Let m be a positive integer, let $\eta, \delta > 0$ and let $E, A_1, \dots, A_m$ be events in a probability space such that for every $i, j \in [m]$ with $i \neq j$ we have*

(i) $|\mathbb{P}(A_i) - \mathbb{P}(A_j)| \leqslant \eta$,
(ii) $|\mathbb{P}(E\cap A_i) - \mathbb{P}(E\cap A_j)| \leqslant \eta$, *and*
(iii) $\mathbb{P}(A_i\cap A_j) \leqslant \mathbb{P}(A_i)\,\mathbb{P}(A_j) + \delta$.

Then for every $i \in [m]$ we have

$$\big|\mathbb{P}(E\cap A_i) - \mathbb{P}(E)\,\mathbb{P}(A_i)\big| \leqslant 2\eta + \sqrt{\frac{1}{m}+\delta}. \tag{4.3}$$

Proof. Set $Z \coloneqq \frac{1}{m}\sum_{j=1}^{m}\mathbf{1}_{A_j}$. Let $i\in[m]$. Notice that, by the triangle inequality,

$$\begin{aligned}\big|\mathbb{P}(E\cap A_i) - \mathbb{P}(E)\,\mathbb{P}(A_i)\big| &= \big|\mathbb{E}[\mathbf{1}_E\mathbf{1}_{A_i}] - \mathbb{E}[\mathbf{1}_E\mathbb{P}(A_i)]\big| \\ &\leqslant \big|\mathbb{E}[\mathbf{1}_E\mathbf{1}_{A_i}] - \mathbb{E}[\mathbf{1}_E Z]\big| + \big|\mathbb{E}[\mathbf{1}_E Z] - \mathbb{E}[\mathbf{1}_E\mathbb{E}[Z]]\big| + \\ &\quad + \big|\mathbb{E}[\mathbf{1}_E\mathbb{E}[Z]] - \mathbb{E}[\mathbf{1}_E\mathbb{P}(A_i)]\big|.\end{aligned} \tag{4.4}$$

Invoking the triangle inequality again, we have

$$\big|\mathbb{E}[\mathbf{1}_E\mathbf{1}_{A_i}] - \mathbb{E}[\mathbf{1}_E Z]\big| \leqslant \frac{1}{m}\sum_{j=1}^{m}\big|\mathbb{P}(E\cap A_i) - \mathbb{P}(E\cap A_j)\big| \overset{\text{(ii)}}{\leqslant} \eta \tag{4.5}$$

$$\big|\mathbb{E}[\mathbf{1}_E\mathbb{E}[Z]] - \mathbb{E}[\mathbf{1}_E\mathbb{P}(A_i)]\big| \leqslant \mathbb{P}(E)\frac{1}{m}\sum_{j=1}^{m}\big|\mathbb{P}(A_j) - \mathbb{P}(A_i)\big| \overset{\text{(i)}}{\leqslant} \eta. \tag{4.6}$$

Finally, by the Cauchy–Schwarz inequality, hypothesis (iii) and Lemma 4.1,

$$\big|\mathbb{E}[\mathbf{1}_E Z] - \mathbb{E}[\mathbf{1}_E\mathbb{E}[Z]]\big| \leqslant \sqrt{\mathbb{P}(E)}\;\|Z-\mathbb{E}[Z]\|_{L_2} \leqslant \sqrt{\frac{1}{m}+\delta}. \tag{4.7}$$

The estimate (4.3) follows from (4.4)–(4.7). $\square$

Lemma 4.3. *Let $m \geqslant 1$ be an integer, let $\eta > 0$, and let $(A_i)_{i=1}^{m}$ be an η-spreadable sequence*[9] *of events in a probability space. Then for every $i, j \in [m]$ with $i \neq j$,*

$$\mathbb{P}(A_i\cap A_j) \geqslant \mathbb{P}(A_i)\,\mathbb{P}(A_j) - \frac{1}{m} - 3\eta. \tag{4.8}$$

[9] That is, the random vector $(\mathbf{1}_{A_1}, \dots, \mathbf{1}_{A_m})$ is η-spreadable according to Definition 1.2.

Proof. Set $Z := \frac{1}{m}\sum_{k=1}^{m}\mathbf{1}_{A_k}$. Fix $i,j\in[m]$ with $i\neq j$. Then we have

$$\big|\mathbb{P}(A_i\cap A_j)-\mathbb{E}[Z^2]\big| = \Big|\mathbb{P}(A_i\cap A_j)-\frac{1}{m^2}\sum_{k=1}^{m}\mathbb{P}(A_k)-\frac{1}{m^2}\sum_{\substack{k,\ell\in[m]\\ k\neq\ell}}\mathbb{P}(A_k\cap A_\ell)\Big| \leqslant \frac{1}{m}+\eta.$$

Also notice that

$$\big|\mathbb{P}(A_i)\,\mathbb{P}(A_j)-\mathbb{E}[Z]^2\big| \leqslant \mathbb{E}[Z]\big|\mathbb{P}(A_j)-\mathbb{E}[Z]\big| + \mathbb{P}(A_j)\big|\mathbb{P}(A_i)-\mathbb{E}[Z]\big| \leqslant 2\eta.$$

Since $\mathbb{E}[Z]^2\leqslant\mathbb{E}[Z^2]$, inequality (4.8) follows from the previous two estimates. □

4.2. **Initializing various numerical parameters.** Our goal in this subsection is to define, by recursion on d, the numbers $\gamma_k(\eta,\vartheta,d,n)$ as well as some other numerical invariants which are needed for the proof of Theorem 3.2. (The reader is advised to skip this subsection at first reading.)

We start by setting

$$\gamma_k(\eta,\vartheta,1,n) := (3k-1)\eta+(k-1)\sqrt{\frac{1}{\lfloor n/2\rfloor}+\vartheta} \tag{4.9}$$

for every $0<\eta\leqslant 1$, every $\vartheta>0$ and every pair of positive integers k,n with $n\geqslant 2$ and $k\leqslant n/2$.

Let $d\geqslant 2$ be an integer, and assume that the numbers $\gamma_k(\eta,\vartheta,d-1,n)$ have been defined for every choice of admissible parameters. Fix $0<\eta\leqslant 1$ and $\vartheta>0$, and let n be an integer with $n\geqslant 4d$. We set

$$\vartheta_1(\eta,\vartheta,d,n) := (n-2d+2)^{-1/2}+(2^d+5)\sqrt{\eta}+\sqrt{\vartheta} \tag{4.10}$$

$$\vartheta_2(\eta,\vartheta,d,n) := \frac{2^{d-1}}{n-d+1}+2^d3\eta+\vartheta \tag{4.11}$$

$$\vartheta_3(\eta,\vartheta,d,n) := \frac{d-1}{(n-2d+2)^{1/2^{d-1}}}+(2^d+5)\eta^{1/2^{d-1}}+\vartheta^{1/2^{d-1}}+3\eta. \tag{4.12}$$

Next, for every positive integer k with $k\leqslant\binom{\lfloor (n-1)/2\rfloor}{d-1}$ we set

$$\gamma^{(1)}_k(\eta,\vartheta,d,n) := \gamma_k\big(\eta,\vartheta_1(\eta,\vartheta,d,n),d,n),d-1,n-1\big)+(k+1)\eta \tag{4.13}$$

$$\gamma^{(2)}_k(\eta,\vartheta,d,n) := \gamma_k\big(\eta,\vartheta_2(\eta,\vartheta,d,n),d-1,n-2\big) \tag{4.14}$$

$$\gamma^{(3)}_k(\eta,\vartheta,d,n) := 2\gamma^{(1)}_k(\eta,\vartheta,d,n)+\gamma^{(2)}_k(\eta,\vartheta,d,n)+k\,\vartheta_3(\eta,\vartheta,d,n) \tag{4.15}$$

$$\gamma^{(4)}_k(\eta,\vartheta,d,n) := \big(\gamma^{(3)}_k(\eta,\vartheta,d,n)+\lfloor n/2\rfloor^{-1}+(2k+1)\eta\big)^{1/2}+2\eta. \tag{4.16}$$

Moreover, for every positive integer u with $u\leqslant n/2$ and every choice $k_1,\dots,k_u$ of positive integers with $k_1,\dots,k_u\leqslant\binom{\lfloor (n-2)/2\rfloor}{d-1}$ set

$$\gamma^{(5)}(\eta,\vartheta,d,n,(k_i)_{i=1}^u) := \gamma^{(1)}_{k_1}(\eta,\vartheta,d,n)+\sum_{i=2}^{u}\big(\gamma^{(1)}_{k_i}(\eta,\vartheta,d,n)+\gamma^{(4)}_{k_i}(\eta,\vartheta,d,n)\big) \tag{4.17}$$

with the convention that the sum in (4.17) is equal to 0 if $u = 1$. Finally, for every positive integer k with $k \leqslant \binom{\lfloor n/2 \rfloor}{d}$ we define

$$\gamma_k(\eta, \vartheta, d, n) \coloneqq (k+1)\,\eta + \max\{\gamma^{(5)}(\eta, \vartheta, d, n, (k_i)_{i=1}^u)\} \tag{4.18}$$

where the above maximum is taken over all choices of positive integers $u, k_1, \dots, k_u$ satisfying $u \leqslant n/2 - d$, $k_1, \dots, k_u \leqslant \binom{\lfloor (n-2)/2 \rfloor}{d-1}$ and $k_1 + \dots + k_u = k$.

4.3. **The inductive hypothesis.** For every positive integer d by P(d) we shall denote the following statement.

Let $n \geqslant 2d$ be an integer, let $0 < \eta < 1$, let $\vartheta > 0$, let $\mathcal{X}$ be a nonempty finite set, and let $\mathcal{S}$ be a nonempty subset of $\mathcal{X}$. Set $\ell \coloneqq \binom{\lfloor n/2 \rfloor}{d}$ and let $\boldsymbol{\gamma} = (\gamma_k(\eta, \vartheta, d, n))_{k=1}^{\ell}$ be as in Subsection 4.2. Let $\boldsymbol{X}$ be an $\mathcal{X}$-valued, η-spreadable, d-dimensional random array $\boldsymbol{X}$ on $[n]$. If $\boldsymbol{X}$ is $(\vartheta, \mathcal{S})$-box independent, then $\boldsymbol{X}$ is $(\boldsymbol{\gamma}, \mathcal{S})$-independent.

It is clear that Theorem 3.2 is equivalent to the validity of P(d) for every integer $d \geqslant 1$.

4.4. **The base case "$d = 1$".** The initial step of the induction follows from the following lemma.

Lemma 4.4. *Let $n, \eta, \vartheta, \mathcal{X}$ and $\mathcal{S}$ be as in the statement of P(1), and assume that $\boldsymbol{X} = (X_1, \dots, X_n)$ is an $\mathcal{X}$-valued, η-spreadable, random vector. Assume, moreover, that for every $i, j \in [n]$ with $i \neq j$ and every $a \in \mathcal{S}$ we have*

$$\mathbb{P}\big([X_i = a] \cap [X_j = a]\big) \leqslant \mathbb{P}\big([X_i = a]\big)\,\mathbb{P}\big([X_j = a]\big) + \vartheta. \tag{4.19}$$

Then for every nonempty $\mathcal{F} \subseteq [n]$ with $|\mathcal{F}| \leqslant n/2$ and every collection $(a_i)_{i \in \mathcal{F}}$ of elements of $\mathcal{S}$, we have

$$\Big|\mathbb{P}\Big(\bigcap_{i \in \mathcal{F}} [X_i = a_i]\Big) - \prod_{i \in \mathcal{F}} \mathbb{P}\big([X_i = a_i]\big)\Big| \leqslant \gamma_{|\mathcal{F}|}(\eta, \vartheta, 1, n) \tag{4.20}$$

where $(\gamma_k(\eta, \vartheta, 1, n))_{k=1}^{\lfloor n/2 \rfloor}$ is as in (4.9). In particular, P(1) holds true.

Proof. Observe that, by the η-spreadability of $\boldsymbol{X}$, it is enough to show that for every $k \in \{1, \dots, \lfloor n/2 \rfloor\}$ and every $a_1, \dots, a_k \in \mathcal{S}$ we have

$$\Big|\mathbb{P}\Big(\bigcap_{i=1}^{k} [X_i = a_i]\Big) - \prod_{i=1}^{k} \mathbb{P}\big([X_i = a_i]\big)\Big| \leqslant (k-1)\Big(2\eta + \sqrt{\frac{1}{\lfloor n/2 \rfloor} + \vartheta}\Big). \tag{4.21}$$

To this end, we proceed by induction of k. The case "$k = 1$" is straightforward. Let k be a positive integer with $k < \lfloor n/2 \rfloor$, and assume that (4.21) has been verified up to k. Fix $a_1, \dots, a_{k+1} \in \mathcal{S}$. Set $m \coloneqq \lfloor n/2 \rfloor$ and $E \coloneqq \bigcap_{i=1}^{k} [X_i = a_i]$. Also set $A_j \coloneqq [X_{k+j} = a_{k+1}]$ for every $j \in [m]$. Using the η-spreadability of $\boldsymbol{X}$, for every $j, j' \in [m]$ with $j \neq j'$ we have

(i) $|\mathbb{P}(A_j) - \mathbb{P}(A_{j'})| \leqslant \eta$, and
(ii) $|\mathbb{P}(E \cap A_j) - \mathbb{P}(E \cap A_{j'})| \leqslant \eta$.

Moreover, since $a_{k+1} \in \mathcal{S}$, we have

(iii) $\mathbb{P}(A_j \cap A_{j'}) \leqslant \mathbb{P}(A_j)\,\mathbb{P}(A_{j'}) + \vartheta$.

Applying Lemma 4.2 for "$\delta = \vartheta$" and using the definition of A_1, we see that

$$\big|\mathbb{P}\big(E \cap [X_{k+1} = a_{k+1}]\big) - \mathbb{P}(E)\,\mathbb{P}\big([X_{k+1} = a_{k+1}]\big)\big| \leqslant 2\eta + \sqrt{\frac{1}{m} + \vartheta}. \tag{4.22}$$

On the other hand, by our inductive assumptions, we have

$$\Big|\mathbb{P}(E) - \prod_{j=1}^{k} \mathbb{P}\big([X_j = a_j]\big)\Big| \leqslant (k-1)\Big(2\eta + \sqrt{\frac{1}{m} + \vartheta}\Big). \tag{4.23}$$

Combining (4.22) and (4.23), we see that (4.21) is satisfied, as desired. □

4.5. **The general inductive step.** We now enter into the main part of the proof of Theorem 3.2. *Specifically, fix an integer $d \geqslant 2$. Throughout this subsection, we will assume that $\mathrm{P}(d-1)$ has been proved.*

4.5.1. *Step 1: preparatory lemmas.* Our goal in this step is to prove two probabilistic lemmas which will be used in the third and the fourth step of the proof respectively. Strictly speaking, these lemmas are not part of the proof of $\mathrm{P}(d)$ since in their proofs we do not use the inductive assumptions. (In particular, this subsection can be read independently.)

The first lemma essentially shows that the reverse inequality in (3.3) always holds true.

Lemma 4.5. *Let n be an integer with $n \geqslant 2d$, let $0 < \eta < 1$, let $\mathcal{X}$ be a nonempty finite set, and let $\boldsymbol{X} = \langle X_s : s \in \binom{[n]}{d}\rangle$ be an $\mathcal{X}$-valued, η-spreadable, d-dimensional random array on $[n]$. Then for every $t \in \binom{[n-2]}{d-1}$ and every $a \in \mathcal{X}$ we have*

$$\mathbb{P}\big([X_{t\cup\{n-1\}} = a]\big)\,\mathbb{P}\big([X_{t\cup\{n\}} = a]\big) \leqslant \mathbb{P}\big([X_{t\cup\{n-1\}} = a] \cap [X_{t\cup\{n\}} = a]\big) + \frac{1}{n-d+1} + 6\eta. \tag{4.24}$$

Proof. Fix $t \in \binom{[n-2]}{d-1}$ and $a \in \mathcal{X}$. Set $t_0 \coloneqq [d-1]$, and $A_i \coloneqq [X_{t_0\cup\{d-1+i\}} = a]$ for every $i \in [n-d+1]$. Observe that the sequence $(A_1 \dots, A_{n-d+1})$ is η-spreadable[10]. By Lemma 4.3, we obtain that

$$\mathbb{P}(A_1)\,\mathbb{P}(A_2) \leqslant \mathbb{P}(A_1 \cap A_2) + \frac{1}{n-d+1} + 3\eta. \tag{4.25}$$

By (4.25) and the η-spreadability of $\boldsymbol{X}$, the estimate (4.24) follows. □

The second lemma shows that the box independence condition (3.3) is inherited by the two-dimensional faces of d-dimensional boxes.

Lemma 4.6. *Let n be an integer with $n \geqslant 2d$, let $0 < \eta < 1$, let $\vartheta > 0$, let $\mathcal{X}$ be a nonempty finite set, let $\mathcal{S}$ be a nonempty subset of $\mathcal{X}$, and let $\boldsymbol{X} = \langle X_s : s \in \binom{[n]}{d}\rangle$*

[10]Recall that this means that the random vector $(\mathbf{1}_{A_1}, \dots, \mathbf{1}_{A_{n-d+1}})$ is η-spreadable.

be an $\mathcal{X}$-valued, η-spreadable, d-dimensional random array on $[n]$ which is $(\vartheta,\mathcal{S})$-box independent. Then for every $t\in\binom{[n-2]}{d-1}$ and every $a\in\mathcal{S}$ we have

$$\mathbb{P}\big([X_{t\cup\{n-1\}}=a]\cap[X_{t\cup\{n\}}=a]\big)\leqslant\mathbb{P}\big([X_{t\cup\{n-1\}}=a]\big)\,\mathbb{P}\big([X_{t\cup\{n\}}=a]\big)+ + \vartheta_3(\eta,\vartheta,d,n) \tag{4.26}$$

where $\vartheta_3(\eta,\vartheta,d,n)$ is as defined in (4.12).

Proof. Fix $a\in\mathcal{S}$. For every $i\in[d]$ set $t_{i-1}\coloneqq[i-1]$ (where, by convention, $[0]=\emptyset$) and $H_i\coloneqq\{n-2d+2i-1,n-2d+2i\}$. We define, recursively, a finite sequence $(\vartheta_i)_{i=0}^{d-1}$ by setting $\vartheta_0=\vartheta$ and

$$\vartheta_{r+1}=\Big(\frac{1}{n-2d+r+2}+(2^{d-r}+5)\eta+\vartheta_r\Big)^{1/2}. \tag{4.27}$$

By induction on $r\in\{0,\dots,d-1\}$, we will show that

$$\mathbb{P}\Big(\bigcap_{v\in B_r}[X_{t_r\cup v}=a]\Big)\leqslant\prod_{v\in B_r}\mathbb{P}\big([X_{t_r\cup v}=a]\big)+\vartheta_r \tag{4.28}$$

where $B_r\coloneqq\mathrm{Box}\big((H_{r+1},\dots,H_d)\big)$ is the $(d-r)$-dimensional box determined by the sequence $(H_{r+1},\dots,H_d)$. (See (3.1).) The case "$r=0$" follows from the fact that the random array $\boldsymbol{X}$ is $(\vartheta,\mathcal{S})$-box independent. Next, let $r\in\{0,\dots,d-2\}$ and assume that (4.28) has been proved up to r. For every $j\in[n-2d+r+2]$ set

$$A_j\coloneqq\bigcap_{v\in B_{r+1}}[X_{t_r\cup\{r+j\}\cup v}=a]. \tag{4.29}$$

Since $\boldsymbol{X}$ is η-spreadable, the sequence $(A_1,\dots,A_{n-2d+r+2})$ is η-spreadable. Using this observation and the inductive assumptions, we see that

$$\begin{aligned}\mathbb{P}(A_1\cap A_2)&\leqslant\mathbb{P}(A_{n-2d+r+1}\cap A_{n-2d+r+2})+\eta\\&=\mathbb{P}\Big(\bigcap_{v\in B_r}[X_{t_r\cup v}=a]\Big)+\eta\leqslant\prod_{v\in B_r}\mathbb{P}\big([X_{t_r\cup v}=a]\big)+\eta+\vartheta_r.\end{aligned} \tag{4.30}$$

On the other hand, since $\boldsymbol{X}$ is η-spreadable, we have

$$\prod_{v\in B_r}\mathbb{P}\big([X_{t_r\cup v}=a]\big)\leqslant\Big(\prod_{v\in B_{r+1}}\mathbb{P}\big([X_{t_{r+1}\cup v}=a]\big)\Big)^2+2^{d-r}\eta. \tag{4.31}$$

Moreover, by Lemma 4.3 applied to the η-spreadable sequence $(A_j)_{j=1}^{n-2d+r+2}$,

$$\mathbb{P}(A_1\cap A_2)\geqslant\mathbb{P}(A_1)\,\mathbb{P}(A_2)-\frac{1}{n-2d+r+2}-3\eta. \tag{4.32}$$

By (4.30)–(4.32) and using the η-spreadability of the sequence $(A_j)_{j=1}^{n-2d+r+2}$ once again, we obtain that

$$\mathbb{P}\Big(\bigcap_{v\in B_{r+1}}[X_{t_{r+1}\cup v}=a]\Big)^2=\mathbb{P}(A_1)^2\leqslant \mathbb{P}(A_1)\,\mathbb{P}(A_2)+\eta \leqslant \Big(\prod_{v\in B_{r+1}}\mathbb{P}\big([X_{t_{r+1}\cup v}=a]\big)\Big)^2+\vartheta_{r+1}^2. \tag{4.33}$$

Taking square-roots, this estimate completes the inductive proof of (4.28).

Now notice that

$$\begin{aligned}\vartheta_{d-1}&\leqslant \vartheta^{\frac{1}{2^{d-1}}}+\sum_{j=1}^{d-1}(n-2d+2)^{-1/2^j}+((2^d+5)\eta)^{1/2^j}\\ &\leqslant \frac{d-1}{(n-2d+2)^{1/2^{d-1}}}+(d-1)(2^d+5)^{1/2}\eta^{1/2^{d-1}}+\vartheta^{1/2^{d-1}}\\ &\leqslant \frac{d-1}{(n-2d+2)^{1/2^{d-1}}}+(2^d+5)\eta^{1/2^{d-1}}+\vartheta^{1/2^{d-1}}.\end{aligned} \tag{4.34}$$

Setting $s_1:=[d-1]\cup\{n-1\}$ and $s_2:=[d-1]\cup\{n\}$, by (4.28) and (4.34), we have

$$\begin{aligned}\mathbb{P}\big([X_{s_1}=a]\cap[X_{s_2}=a]\big)&\leqslant \mathbb{P}\big([X_{s_1}=a]\big)\,\mathbb{P}\big([X_{s_2}=a]\big)+\\ &\quad+\frac{d-1}{(n-2d+2)^{1/2^{d-1}}}+(2^d+5)\eta^{1/2^{d-1}}+\vartheta^{1/2^{d-1}}.\end{aligned} \tag{4.35}$$

Taking into account the η-spreadability of $\boldsymbol{X}$ and the definition of $\vartheta_3(\eta,\vartheta,d,n)$, the estimate (4.26) follows from (4.35). □

4.5.2. *Step 2: rewriting the inductive assumptions.* We proceed with the following lemma which will enable us to use P$(d-1)$ in a more convenient form.

Lemma 4.7. *Let $n,\eta,\vartheta,\mathcal{X},\mathcal{S}$ be as in the statement of P(d), and let $\boldsymbol{X}=\langle X_s : s\in\binom{[n]}{d}\rangle$ be an $\mathcal{X}$-valued, η-spreadable, d-dimensional random array on $[n]$ which is $(\vartheta,\mathcal{S})$-box independent. We define $\widetilde{\boldsymbol{X}}=\langle \widetilde{X}_t : t\in\binom{[n-1]}{d-1}\rangle$ by setting*

$$\widetilde{X}_t:=X_{t\cup\{n\}}. \tag{4.36}$$

Then the random array $\widetilde{\boldsymbol{X}}$ is $\mathcal{X}$-valued, η-spreadable and $(\vartheta_1(\eta,\vartheta,d,n),\mathcal{S})$-box independent, where $\vartheta_1(\eta,\vartheta,d,n)$ is as (4.10).

Proof. Since $\boldsymbol{X}$ is $\mathcal{X}$-valued and η-spreadable, by (4.36), we see that these properties are inherited to $\widetilde{\boldsymbol{X}}$. Thus, we only need to check that $\widetilde{\boldsymbol{X}}$ is $(\vartheta_1(\eta,\vartheta,d,n),\mathcal{S})$-box independent.

To this end, fix $a\in\mathcal{S}$ and a finite sequence $\mathcal{H}=(H_1,\dots,H_{d-1})$ of 2-element subsets of $[n-1]$ with $\max(H_i)<\min(H_{i+1})$ for all $i\in[d-2]$; let $B:=\mathrm{Box}(\mathcal{H})$ denote the $(d-1)$-dimensional box determined by the sequence $\mathcal{H}$. Moreover, set

$$B_0:=\mathrm{Box}\big((\{1,2\},\dots,\{2d-3,2d-2\})\big)$$

and $A_r \coloneqq \bigcap_{t\in B_0}[X_{t\cup\{2d-2+r\}} = a]$ for every $r \in [n-2d+2]$. Notice that the sequence $(A_1,\dots,A_{n-2d+2})$ is η-spreadable. Therefore, by Lemma 4.3,

$$\mathbb{P}(A_1)^2 \leqslant \mathbb{P}(A_1)\,\mathbb{P}(A_2) + \eta \leqslant \mathbb{P}(A_1\cap A_2) + \frac{1}{n-2d+2} + 4\eta. \tag{4.37}$$

Next, set $B' \coloneqq \mathrm{Box}\big((\{1,2\},\dots,\{2d-3,2d-2\},\{2d-1,2d\})\big)$, and observe that B' is a d-dimensional box and $A_1\cap A_2 = \bigcap_{s\in B'}[X_s = a]$. Since $\boldsymbol{X}$ is $(\vartheta,\mathcal{S})$-box independent and $a\in\mathcal{S}$, we see that

$$\begin{aligned}\mathbb{P}(A_1\cap A_2) &\leqslant \prod_{s\in B'}\mathbb{P}\big([X_s=a]\big) + \vartheta \\ &= \Big(\prod_{t\in B_0}\mathbb{P}\big([X_{t\cup\{2d-1\}}=a]\big)\Big)\Big(\prod_{t\in B_0}\mathbb{P}\big([X_{t\cup\{2d\}}=a]\big)\Big) + \vartheta \\ &\leqslant \Big(\prod_{t\in B_0}\mathbb{P}\big([X_{t\cup\{2d-1\}}=a]\big)\Big)^2 + 2^{d-1}\eta + \vartheta\end{aligned} \tag{4.38}$$

where the last inequality follows from the η-spreadability of $\boldsymbol{X}$. By (4.37), (4.38) and the definition of A_1, we obtain

$$\begin{aligned}\mathbb{P}\Big(\bigcap_{t\in B_0}[X_{t\cup\{2d-1\}}=a]\Big) &\leqslant \prod_{t\in B_0}\mathbb{P}\big([X_{t\cup\{2d-1\}}=a]\big) + \\ &\qquad + \Big(\frac{1}{n-2d+2} + (2^{d-1}+4)\eta + \vartheta\Big)^{1/2} \\ &\leqslant \prod_{t\in B_0}\mathbb{P}\big([X_{t\cup\{2d-1\}}=a]\big) + \\ &\qquad + (n-2d+2)^{-1/2} + (2^{d-1}+4)\sqrt{\eta} + \sqrt{\vartheta}.\end{aligned} \tag{4.39}$$

On the other hand, using the η-spreadability of $\boldsymbol{X}$, we have

$$\Big|\mathbb{P}\Big(\bigcap_{t\in B_0}[X_{t\cup\{2d-1\}}=a]\Big) - \mathbb{P}\Big(\bigcap_{t\in B}[\widetilde{X}_t=a]\Big)\Big| \leqslant \eta \leqslant \sqrt{\eta} \tag{4.40}$$

$$\Big|\prod_{t\in B_0}\mathbb{P}\big([X_{t\cup\{2d-1\}}=a]\big) - \prod_{t\in B}\mathbb{P}\big([\widetilde{X}_t=a]\big)\Big| \leqslant 2^{d-1}\eta \leqslant 2^{d-1}\sqrt{\eta}. \tag{4.41}$$

Combining (4.39)–(4.41) and invoking the definition of $\vartheta_1(\eta,\vartheta,d,n)$ in (4.10), we conclude that

$$\mathbb{P}\Big(\bigcap_{t\in B}[\tilde{X}_t=\alpha]\Big) \leqslant \prod_{t\in B}\mathbb{P}\big([\tilde{X}_t=a]\big) + \vartheta_1(\eta,\vartheta,d,n). \tag{4.42}$$

Since a and B were arbitrary, the result follows. □

By Lemma 4.7 and P$(d-1)$, we have the following corollary.

Corollary 4.8. *Let $n,\eta,\vartheta,\mathcal{X},\mathcal{S},\boldsymbol{X}$ be as in Lemma 4.7. Then for every nonempty subset $\mathcal{G}$ of $\binom{[n-1]}{d-1}$ with $|\cup\mathcal{G}|\leqslant(n-1)/2$, every collection $(a_t)_{t\in\mathcal{G}}$ of elements of $\mathcal{S}$, and every*

$r \in [n]$ with $r > \max(\cup\mathcal{G})$ we have

$$\Big|\mathbb{P}\Big(\bigcap_{t\in\mathcal{G}}[X_{t\cup\{r\}} = a_t]\Big) - \prod_{t\in\mathcal{G}}\mathbb{P}\big([X_{t\cup\{r\}} = a_t]\big)\Big| \leqslant \gamma^{(1)}_{|\mathcal{G}|}(\eta,\vartheta,d,n) \tag{4.43}$$

where $\gamma^{(1)}_{|\mathcal{G}|}(\eta,\vartheta,d,n)$ is as in (4.13).

4.5.3. *Step 3: doubling.* The following lemma complements Lemma 4.7. It is also based on the inductive hypothesis P($d-1$), but it will enable to use it in a rather different form.

Lemma 4.9 (Doubling)**.** *Let $n,\eta,\vartheta,\mathcal{X},\mathcal{S}$ be as in the statement of P(d), and assume that $\boldsymbol{X} = \langle X_s : s \in \binom{[n]}{d}\rangle$ is an $\mathcal{X}$-valued, η-spreadable, d-dimensional random array on $[n]$ which is $(\vartheta,\mathcal{S})$-box independent. We define a $(d-1)$-dimensional random array $\widetilde{\boldsymbol{X}}' = \langle \widetilde{X}'_t : t \in \binom{[n-2]}{d-1}\rangle$ by setting*

$$\widetilde{X}'_t \coloneqq (X_{t\cup\{n-1\}}, X_{t\cup\{n\}}). \tag{4.44}$$

Then $\widetilde{\boldsymbol{X}}'$ is $(\mathcal{X}\times\mathcal{X})$-valued, η-spreadable and $(\vartheta_2(\eta,\vartheta,d,n),\{(a,a) : a \in \mathcal{S}\})$-box independent, where $\vartheta_2(\eta,\vartheta,d,n)$ is as defined in (4.11).

Proof. It is clear that $\widetilde{\boldsymbol{X}}'$ is $(\mathcal{X}\times\mathcal{X})$-valued and η-spreadable. So, we only need to show that $\widetilde{\boldsymbol{X}}'$ is $(\vartheta_2(\eta,\vartheta,d,n),\{(a,a) : a \in \mathcal{S}\})$-box independent.

Let $H_1,\dots,H_{d-1}$ be 2-element subsets of $[n-2]$ with $\max(H_i) < \min(H_{i+1})$ for all $i \in [d-2]$. Also let $a \in \mathcal{S}$. Set $\widetilde{B} \coloneqq \mathrm{Box}\big((H_1,\dots,H_{d-1})\big)$; also set $H_d \coloneqq \{n-1,n\}$ and $B \coloneqq \mathrm{Box}\big((H_1,\dots,H_{d-1},H_d)\big)$. Since $\boldsymbol{X}$ is $(\vartheta,\mathcal{S})$-box independent, we see that

$$\begin{aligned}\mathbb{P}\Big(\bigcap_{t\in\widetilde{B}}[\widetilde{X}'_t = (a,a)]\Big) &= \mathbb{P}\Big(\bigcap_{t\in\widetilde{B}}\big([X_{t\cup\{n-1\}} = a]\cap[X_{t\cup\{n\}} = a]\big)\Big) \\ &= \mathbb{P}\Big(\bigcap_{s\in B}[X_s = a]\Big) \leqslant \prod_{s\in B}\mathbb{P}\big([X_s = a]\big) + \vartheta.\end{aligned} \tag{4.45}$$

By Lemma 4.5, we have

$$\begin{aligned}\prod_{s\in B}\mathbb{P}\big([X_s = a]\big) &= \prod_{t\in\widetilde{B}}\mathbb{P}\big([X_{t\cup\{n-1\}} = a]\big)\,\mathbb{P}\big([X_{t\cup\{n\}} = a]\big) \\ &\leqslant \prod_{t\in\widetilde{B}}\mathbb{P}\big([X_{t\cup\{n-1\}} = a]\cap[X_{t\cup\{n\}} = a]\big) + \frac{2^{d-1}}{n-d+1} + 2^{d-1}6\eta \\ &= \prod_{t\in\widetilde{B}}\mathbb{P}\big([\widetilde{X}'_t = (a,a)]\big) + \frac{2^{d-1}}{n-d+1} + 2^{d}3\eta.\end{aligned} \tag{4.46}$$

By (4.45) and (4.46) and the definition of $\vartheta_2(\eta,\vartheta,d,n)$, the result follows. □

The following corollary—which is an immediate consequence of Lemma 4.9 and the inductive assumption P($d-1$)—is the analogue of Corollary 4.8.

Corollary 4.10. *Let $n,\eta,\vartheta,\mathcal{X},\mathcal{S},\boldsymbol{X},\widetilde{\boldsymbol{X}}'$ be as in Lemma 4.9. Then the random array $\widetilde{\boldsymbol{X}}'$ is $\big((\gamma^{(2)}_k(\eta,\vartheta,d,n))_{k=1}^{\ell},\{(a,a) : a \in \mathcal{S}\}\big)$-independent, where $\ell = \binom{\lfloor (n-2)/2\rfloor}{d-1}$ and $(\gamma^{(2)}_k(\eta,\vartheta,d,n))_{k=1}^{\ell}$ is as in (4.14).*

4.5.4. *Step 4: gluing.* This is the main step of the proof. Specifically, our goal is to prove the following proposition.

Proposition 4.11 (Gluing). *Let $n \geqslant 2d+2$ be an integer, let $\eta,\vartheta,\mathcal{X},\mathcal{S}$ be as in the statement of P(d), and assume that $\boldsymbol{X}=\langle X_s : s\in\binom{[n]}{d}\rangle$ is an $\mathcal{X}$-valued, η-spreadable, d-dimensional random array on $[n]$ which is $(\vartheta,\mathcal{S})$-box independent. Finally, let r be an integer with $d<r\leqslant n/2$, let $\mathcal{G}$ be a nonempty subset of $\binom{[r-1]}{d-1}$, let $(a_t)_{t\in\mathcal{G}}$ be a collection of elements of $\mathcal{S}$, let $\mathcal{F}$ be a nonempty subset of $\binom{[r-1]}{d}$, and let $(b_s)_{s\in\mathcal{F}}$ be a collection of elements of $\mathcal{S}$. Then we have*

$$\Big|\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}}[X_s=b_s]\cap\bigcap_{t\in\mathcal{G}}[X_{t\cup\{r\}}=a_t]\Big)- \\ -\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}}[X_s=b_s]\Big)\,\mathbb{P}\Big(\bigcap_{t\in\mathcal{G}}[X_{t\cup\{r\}}=a_t]\Big)\Big|\leqslant\gamma^{(4)}_{|G|}(\eta,\vartheta,d,n) \tag{4.47}$$

where $\gamma^{(4)}_{|\mathcal{G}|}(\eta,\vartheta,d,n)$ is as in (4.16).

Proposition 4.11 follows by carefully selecting a sequence of events, and then applying the averaging argument presented in Lemma 4.2. In order to do so, we need to control the variances of the corresponding averages. This is, essentially, the content of the following lemma.

Lemma 4.12 (Variance estimate). *Let $n,\eta,\vartheta,\mathcal{X},\mathcal{S}$ be as in the statement of P(d), and assume that $\boldsymbol{X}=\langle X_s : s\in\binom{[n]}{d}\rangle$ is an $\mathcal{X}$-valued, η-spreadable, d-dimensional random array on $[n]$ which is $(\vartheta,\mathcal{S})$-box independent. Then for every nonempty subset $\mathcal{G}$ of $\binom{[n-2]}{d-1}$ with $|\cup\mathcal{G}|\leqslant(n-2)/2$, and every collection $(a_t)_{t\in\mathcal{G}}$ of elements of $\mathcal{S}$ we have*

$$\mathbb{P}\Big(\bigcap_{t\in\mathcal{G}}[X_{t\cup\{n-1\}}=a_t]\cap\bigcap_{t\in\mathcal{G}}[X_{t\cup\{n\}}=a_t]\Big) \\ \leqslant\mathbb{P}\Big(\bigcap_{t\in\mathcal{G}}[X_{t\cup\{n-1\}}=a_t]\Big)\,\mathbb{P}\Big(\bigcap_{t\in\mathcal{G}}[X_{t\cup\{n\}}=a_t]\Big)+\gamma^{(3)}_{|\mathcal{G}|}(\eta,\vartheta,d,n) \tag{4.48}$$

where $\gamma^{(3)}_{|\mathcal{G}|}(\eta,\vartheta,d,n)$ is as in (4.15).

Proof. Let $\mathcal{G}$ be a subset of $\binom{[n-2]}{d-1}$ with $|\cup\mathcal{G}|\leqslant(n-2)/2$, and let $(a_t)_{t\in\mathcal{G}}$ be a collection of elements of $\mathcal{S}$. By Corollary 4.10, we have

$$\mathbb{P}\Big(\bigcap_{t\in\mathcal{G}}[X_{t\cup\{n-1\}}=a_t]\cap\bigcap_{t\in\mathcal{G}}[X_{t\cup\{n\}}=a_t]\Big) \\ \leqslant\prod_{t\in\mathcal{G}}\mathbb{P}\big([X_{t\cup\{n-1\}}=a_t]\cap[X_{t\cup\{n\}}=a_t]\big)+\gamma^{(2)}_{|\mathcal{G}|}(\eta,\vartheta,d,n). \tag{4.49}$$

Moreover, by Lemma 4.6,

$$\prod_{t\in\mathcal{G}}\mathbb{P}\big([X_{t\cup\{n-1\}}=a_t]\cap[X_{t\cup\{n\}}=a_t]\big) \\ \leqslant\prod_{t\in\mathcal{G}}\mathbb{P}\big([X_{t\cup\{n-1\}}=a_t]\big)\,\mathbb{P}\big([X_{t\cup\{n\}}=a_t]\big)+\,|\mathcal{G}|\,\vartheta_3(\eta,\vartheta,d,n). \tag{4.50}$$

Finally, by Corollary 4.8, we see that

$$\prod_{t\in\mathcal{G}} \mathbb{P}\big([X_{t\cup\{n-1\}} = a_t]\big) \leqslant \mathbb{P}\Big(\bigcap_{t\in\mathcal{G}} [X_{t\cup\{n-1\}} = a_t]\Big) + \gamma^{(1)}_{|\mathcal{G}|}(\eta,\vartheta,d,n) \tag{4.51}$$

$$\prod_{t\in\mathcal{G}} \mathbb{P}\big([X_{t\cup\{n\}} = a_t]\big) \leqslant \mathbb{P}\Big(\bigcap_{t\in\mathcal{G}} [X_{t\cup\{n\}} = a_t]\Big) + \gamma^{(1)}_{|\mathcal{G}|}(\eta,\vartheta,d,n). \tag{4.52}$$

The estimate (4.48) follows by combining (4.49)–(4.52) and invoking the definition of the constant $(\gamma^{(3)}_{|\mathcal{G}|}(\eta,\vartheta,d,n)$ in (4.15). □

We are now ready to give the proof of Proposition 4.11.

Proof of Proposition 4.11. Set $E \coloneqq \bigcap_{s\in\mathcal{F}}[X_s = b_s]$ and $A_i \coloneqq \bigcap_{t\in\mathcal{G}}[X_{t\cup\{r-1+i\}}]$ for every $i \in \{1,\dots,\lfloor n/2\rfloor\}$. Since $\boldsymbol{X}$ is η-spreadable, for every $i,j \in \{1,\dots,\lfloor n/2\rfloor\}$ with $i \neq j$ we have

(i) $|\mathbb{P}(A_i) - \mathbb{P}(A_j)| \leqslant \eta$, and
(ii) $|\mathbb{P}(E\cap A_i) - \mathbb{P}(E\cap A_j)| \leqslant \eta$.

Moreover, applying Lemma 4.12 and using the η-spreadability of $\boldsymbol{X}$ again, for every $i,j \in \{1,\dots,\lfloor n/2\rfloor\}$ with $i\neq j$ we have

(iii) $\mathbb{P}(A_i\cap A_j) \leqslant \mathbb{P}(A_i)\,\mathbb{P}(A_j) + \gamma^{(3)}_{|\mathcal{G}|}(\eta,\vartheta,d,n) + (2|\mathcal{G}|+1)\eta$.

By Lemma 4.2 applied for "$\delta = \gamma^{(3)}_{|\mathcal{G}|}(\eta,\vartheta,d,n) + (2|\mathcal{G}|+1)\eta$" and taking into account the definition of the constant $\gamma^{(4)}_{|\mathcal{G}|}(\eta,\vartheta,d,n)$, we conclude that (4.47) is satisfied. □

4.5.5. *Step 5: completion of the proof.* This is the last step of the proof. Recall that we need to prove that the statement P(d) holds true, or equivalently, that the estimate (3.4) is satisfied for the sequence $\boldsymbol{\gamma} = (\gamma_k(\eta,\vartheta,d,n))_{k=1}^{\ell}$ defined in Subsection 4.2. As expected, the verification of this estimate will be reduced to Proposition 4.11. To this end, we will decompose an arbitrary nonempty subset $\mathcal{F}$ of $\binom{[n]}{d}$ into several components which are easier to handle. The details of this decomposition are presented in the following definition.

Definition 4.13 (Slicing profile)**.** *Let n,d be positive integers with $n\geqslant d$ and let $\mathcal{F}$ be a nonempty subset of $\binom{[n]}{d}$. There exist, unique,*

- *$u\in[n]$,*
- *$r_1,\dots,r_u\in[n]$ with $d\leqslant r_1<\dots<r_u$, and*
- *for every $i\in[u]$ a nonempty subset $\mathcal{G}_i$ of $\binom{[r_i-1]}{d-1}$,*

such that

$$\mathcal{F} = \bigcup_{i=1}^{u}\big\{t\cup\{r_i\} : t\in\mathcal{G}_i\big\}. \tag{4.53}$$

We refer to the triple $(u,(r_i)_{i=1}^u,(\mathcal{G}_i)_{i=1}^u)$ as the slicing *of $\mathcal{F}$, and to the sequence $(|\mathcal{G}_i|)_{i=1}^u$ as the* slicing profile *of $\mathcal{F}$. Finally, we denote by $\mathrm{SP}(n)$ the set of all nonempty finite sequences $(k_i)_{i=1}^u$ which are the slicing profile of some nonempty subset $\mathcal{F}$ of $\binom{[n]}{d}$.*

Example 4.14. Let $d=2$, $n=6$, and let $\mathcal{F}$ be the subset of $\binom{[6]}{2}$ defined by

$$\mathcal{F} \coloneqq \big\{\{1,3\},\{1,6\},\{2,3\},\{2,5\},\{4,5\}\big\}.$$

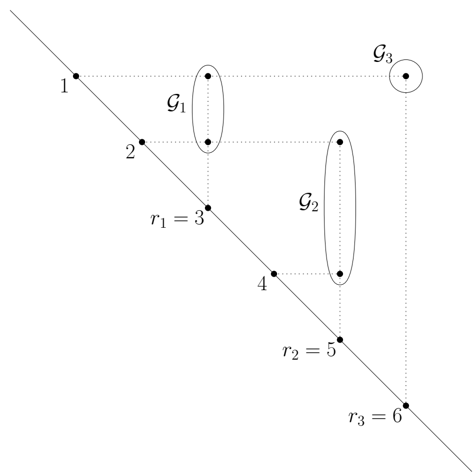

FIGURE 4. The slicing profile of $\mathcal{F}$.

Then the slicing of $\mathcal{F}$ is the triple $\big(3,(r_1,r_2,r_3),(\mathcal{G}_1,\mathcal{G}_2,\mathcal{G}_3)\big)$ where $r_1=3$, $r_2=5$, $r_3=6$, $\mathcal{G}_1=\{1,2\}$, $\mathcal{G}_2=\{2,4\}$ and $\mathcal{G}_3=\{1\}$; in particular, the slicing profile of $\mathcal{F}$ is the sequence $(2,2,1)$.

We have the following lemma.

Lemma 4.15. *Let $n,\eta,\vartheta,\mathcal{X},\mathcal{S}$ be as in the statement of P(d). Let $\boldsymbol{X}=\langle X_s : s\in\binom{[n]}{d}\rangle$ be an $\mathcal{X}$-valued, η-spreadable, d-dimensional random array on $[n]$ which is $(\vartheta,\mathcal{S})$-box independent. Also let $u\leqslant (n/2)-d+1$ be a positive integer, and let $(k_i)_{i=1}^u\in \mathrm{SP}(\lfloor n/2\rfloor)$. If $\mathcal{F}$ is a nonempty subset of $\binom{[\lfloor n/2\rfloor]}{d}$ with slicing profile $(k_i)_{i=1}^u$, then for every collection $(a_s)_{s\in\mathcal{F}}$ of elements of $\mathcal{S}$, we have*

$$\Big|\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}}[X_s=a_s]\Big)-\prod_{s\in\mathcal{F}}\mathbb{P}\big([X_s=a_s]\big)\Big|\leqslant \gamma^{(5)}(\eta,\vartheta,d,n,(k_i)_{i=1}^u) \tag{4.54}$$

where $\gamma^{(5)}(\eta,\vartheta,d,n,(k_i)_{i=1}^u)$ is as in (4.17).

Proof. We proceed by induction on u. The case "$u=1$" follows from Corollary 4.8. Let $u<(n/2)-d$ be a positive integer, and assume that (4.54) has been proved up to u. Let $(k_i)_{i=1}^{u+1}\in \mathrm{SP}(\lfloor n/2\rfloor)$, let $\mathcal{F}$ be a subset of $\binom{[\lfloor n/2\rfloor]}{d}$ with slicing profile $(k_i)_{i=1}^{u+1}$, and let $(a_s)_{s\in\mathcal{F}}$ be a collection of elements of $\mathcal{S}$.

First observe that $n\geqslant 2d+2$ since there exists a nonempty subset of $\binom{[\lfloor n/2\rfloor]}{d}$ with slicing profile of length at least 2; in particular, in what follows, Proposition 4.11 can be applied. Let $(u,(r_i)_{i=1}^{u+1},(\mathcal{G}_i)_{i=1}^{u+1})$ denote the slicing of $\mathcal{F}$, and decompose $\mathcal{F}$ as $\mathcal{F}_1\cup\mathcal{F}_2$ where

$$\mathcal{F}_1 \coloneqq \big\{t\cup\{r_i\} : t\in\mathcal{G}_i, i\in[u]\big\} \ \text{ and } \ \mathcal{F}_2 \coloneqq \big\{t\cup\{r_{u+1}\} : t\in\mathcal{G}_{u+1}\big\}. \tag{4.55}$$

Notice that $d\leqslant r_1<r_{u+1}\leqslant n/2$, $\mathcal{G}_{u+1}\subseteq\binom{[r_{u+1}-1]}{d-1}$ and $|\mathcal{G}_{u+1}|=k_{u+1}$. By Proposition 4.11 applied for "$r=r_{u+1}$", "$\mathcal{G}=\mathcal{G}_{u+1}$", "$(a_t)_{t\in\mathcal{G}}=(a_{t\cup\{r_{u+1}\}})_{t\in\mathcal{G}_{u+1}}$", "$\mathcal{F}=\mathcal{F}_1$ and "$(b_s)_{s\in\mathcal{F}}=(a_s)_{s\in\mathcal{F}_1}$", we have

$$\begin{aligned}\Big|\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}_1}[X_s=a_s]\cap\bigcap_{s\in\mathcal{F}_2}[X_s=a_s]\Big)- \\ -\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}_1}[X_s=a_s]\Big)\,\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}_2}[X_s=a_s]\Big)\Big|\leqslant\gamma^{(4)}_{k_{u+1}}(\eta,\vartheta,d,n).\end{aligned} \tag{4.56}$$

On the other hand, by our inductive assumptions, we obtain that

$$\Big|\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}_1}[X_s=a_s]\Big)-\prod_{s\in\mathcal{F}_1}\mathbb{P}\big([X_s=a_s]\big)\Big|\leqslant\gamma^{(5)}(\eta,\vartheta,d,n,(k_i)_{i=1}^{u}). \tag{4.57}$$

Moreover, since $|\mathcal{G}_{u+1}|=k_{u+1}$, by Corollary 4.8,

$$\Big|\mathbb{P}\Big(\bigcap_{s\in\mathcal{F}_2}[X_s=a_s]\Big)-\prod_{s\in\mathcal{F}_2}\mathbb{P}\big([X_s=a_s]\big)\Big|\leqslant\gamma^{(1)}_{k_{u+1}}(\eta,\vartheta,d,n). \tag{4.58}$$

The inductive step is completed by combining (4.57) and (4.58) and using the definition of the constant $\gamma^{(5)}(\eta,\vartheta,d,n,(k_i)_{i=1}^{u+1})$ in (4.17). □

It is clear that Lemma 4.15 implies that P(d) holds true. This completes the proof of the general inductive step, and so the entire proof of Theorem 3.2 is completed.

5. Proof of Theorem 1.5

It is not hard to verify that the recursive selection in Subsection 4.2 implies that for every pair of positive integers n,d with $n\geqslant 4d$, every $0<\eta,\vartheta\leqslant 1$, and every positive integer k with $k\leqslant\binom{\lfloor n/2\rfloor}{d}$ we have

$$\gamma_k(\eta,\vartheta,d,n)\leqslant 100k\,2^d\,\big(\sqrt[4^d]{1/n}+\sqrt[4^d]{\eta}+\sqrt[4^d]{\vartheta}\big). \tag{5.1}$$

By Theorem 3.2, this estimate clearly yields Theorem 1.5.

6. Proof of Theorem 1.4 and its higher-dimensional version

The following theorem is the higher-dimensional version of Theorem 1.4. (Also note that the case "$d=1$" corresponds to random vectors.)

Theorem 6.1. *Let d,m be two positive integers with $m\geqslant 2$, let $1<p\leqslant 2$, let $0<\varepsilon\leqslant 1$, let $k\geqslant d$ be an integer, and set*

$$C=C(d,m,p,\varepsilon,k)\coloneqq\exp\Big(\frac{24d\ln m}{\varepsilon^{4d}(p-1)^d}\,k^d\Big). \tag{6.1}$$

Also let $n \geqslant C$ be an integer, let $\mathcal{X}$ be a set with $|\mathcal{X}| = m$, and let $\boldsymbol{X} = \langle X_s : s \in \binom{[n]}{d} \rangle$ be an $\mathcal{X}$-valued, $(1/C)$-spreadable, d-dimensional random array on $[n]$. Assume that there exists $\mathcal{S} \subseteq \mathcal{X}$ with $|\mathcal{S}| = |\mathcal{X}| - 1$ such that for every $a \in \mathcal{S}$ we have

$$\Big|\mathbb{P}\Big(\bigcap_{s\in \mathrm{Box}(d)} [X_s = a]\Big) - \prod_{s\in \mathrm{Box}(d)} \mathbb{P}\big([X_s = a]\big)\Big| \leqslant \frac{1}{C} \tag{6.2}$$

where $\mathrm{Box}(d)$ denotes the d-dimensional box defined in (3.2). Then for every function $f\colon \mathcal{X}^{\binom{[n]}{d}} \to \mathbb{R}$ with $\mathbb{E}[f(\boldsymbol{X})] = 0$ and $\|f(\boldsymbol{X})\|_{L_p} = 1$ there exists an interval I of $[n]$ with $|I| = k$ such that for every $J \subseteq I$ with $|J| \geqslant d$ we have

$$\mathbb{P}\big(\big|\mathbb{E}[f(\boldsymbol{X}) \,|\, \mathcal{F}_J]\big| \leqslant \varepsilon\big) \geqslant 1 - \varepsilon. \tag{6.3}$$

Theorems 1.4 and 6.1 are immediate consequences of Theorem 2.2 and Corollary 3.3. The estimates in (1.3) and (6.1) follow by combining (5.1) with

- the choice of β and ℓ in (2.2) and (2.3) respectively, and
- the choice of the constants in the proof of Lemma 3.6.

7. Extensions/Refinements

7.1. **Dissociated random arrays.** The following theorem is the analogue of Theorem 6.1 for the case of dissociated random arrays. (The proof follows arguing as in Section 2, and it is left to the interested reader.)

Theorem 7.1. *Let $1 < p \leqslant 2$, let $0 < \varepsilon \leqslant 1$, and set*

$$c = c(\varepsilon, p) \coloneqq \frac{1}{4}\, \varepsilon^{\frac{2(p+1)}{p}} (p-1). \tag{7.1}$$

Also let n, d be positive integers with $n \geqslant 2d/c$, and let $\boldsymbol{X}$ be a dissociated, d-dimensional random array on $[n]$ whose entries take values in a measurable space $\mathcal{X}$. Then for every measurable function $f\colon \mathcal{X}^{\binom{[n]}{d}} \to \mathbb{R}$ with $\mathbb{E}[f(\boldsymbol{X})] = 0$ and $\|f(\boldsymbol{X})\|_{L_p} = 1$ there exists an interval I of $[n]$ with $|I| \geqslant cn$ such that for every $J \subseteq I$ with $|J| \geqslant d$ we have

$$\mathbb{P}\big(\big|\mathbb{E}[f(\boldsymbol{X}) \,|\, \mathcal{F}_J]\big| \leqslant \varepsilon\big) \geqslant 1 - \varepsilon. \tag{7.2}$$

Note that Theorem 7.1 improves upon Theorem 6.1 in two ways. Firstly, observe that in Theorem 7.1 no restriction is imposed on the distributions of the entries of $\boldsymbol{X}$. Secondly, note that the random variable $f(\boldsymbol{X})$ becomes concentrated by conditioning it on a subarray whose size is proportional to n.

An important—especially, from the point of view of applications—class of random arrays for which Theorem 7.1 is applicable consists of those random arrays whose entries are of the form (1.12) where $(\xi_1, \dots, \xi_n)$ is a random vector with independent (but not necessarily identically distributed) entries.

Remark 7.2. Observe that the size of the set I obtained by Theorem 7.1 depends polynomially on the parameter ε and, in particular, it becomes smaller as ε gets smaller. We note that this sort of dependence is actually necessary. This can be seen by considering (appropriately normalized) linear functions of i.i.d. Bernoulli random variables and invoking the Berry–Esseen theorem.

7.2. **Vector-valued functions of random arrays.** Recall that a Banach space E is called *uniformly convex* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every $x, y \in E$ with $\|x\|_E = \|y\|_E = 1$ and $\|x - y\|_E \geqslant \varepsilon$ we have that $\|(x+y)/2\|_E \leqslant 1 - \delta$. It is a classical fact (see [Ja72, GG71]) that for every uniformly convex Banach space E and every $p > 1$ there exist an exponent $q \geqslant 2$ and a constant $C > 0$ such that for every E-valued martingale difference sequence $(d_i)_{i=1}^m$ we have

$$\Big(\sum_{i=1}^m \|d_i\|_{L_p(E)}^q \Big)^{1/q} \leqslant C \Big\| \sum_{i=1}^m d_i \Big\|_{L_p(E)}. \tag{7.3}$$

(See, also, [Pi11, Pi16] for a proof and a detailed presentation of related material.) Using (7.3) instead of Proposition 2.4 and arguing precisely as in Section 2, we obtain the following vector-valued version of Theorem 6.1.

Theorem 7.3. *For every uniformly convex Banach space E, every pair d, m of positive integers with $m \geqslant 2$, every $p > 1$, every $0 < \varepsilon \leqslant 1$ and every integer $k \geqslant d$, there exists a constant $C > 0$ with the following property.*

Let $n \geqslant C$ be an integer, let $\mathcal{X}$ be a set with $|\mathcal{X}| = m$ and let $\boldsymbol{X} = \langle X_s : s \in \binom{[n]}{d} \rangle$ be an $\mathcal{X}$-valued, $(1/C)$-spreadable, d-dimensional random array on $[n]$. Assume that there exists $\mathcal{S} \subseteq \mathcal{X}$ with $|\mathcal{S}| = |\mathcal{X}| - 1$ such that for every $a \in \mathcal{S}$ we have

$$\Big| \mathbb{P}\Big(\bigcap_{s \in \mathrm{Box}(d)} [X_s = a] \Big) - \prod_{s \in \mathrm{Box}(d)} \mathbb{P}\big([X_s = a]\big) \Big| \leqslant \frac{1}{C} \tag{7.4}$$

where $\mathrm{Box}(d)$ denotes the d-dimensional box defined in (3.2). Then for every function $f \colon \mathcal{X}^{\binom{[n]}{d}} \to E$ with $\mathbb{E}[f(\boldsymbol{X})] = 0$ and $\|f(\boldsymbol{X})\|_{L_p(E)} = 1$ there exists an interval I of $[n]$ with $|I| = k$ such that for every $J \subseteq I$ with $|J| \geqslant d$ we have

$$\mathbb{P}\big(\big\| \mathbb{E}[f(\boldsymbol{X}) \,|\, \mathcal{F}_J] \big\|_E \leqslant \varepsilon \big) \geqslant 1 - \varepsilon. \tag{7.5}$$

7.3. **Simultaneous conditional concentration.** Our last result in this section can be loosely described as "simultaneous conditional concentration"; it asserts that we can achieve concentration by conditioning on the same subarray for almost all members of a given family of approximate spreadable random arrays with the box independence condition.

Theorem 7.4. *Let d, m, p, ε, k be as in Theorem 6.1, set $C' \coloneqq C(d, m, p, \varepsilon^3/4, k)$ where $C(d, m, p, \varepsilon^3/4, k)$ is as in (6.1), and let $n \geqslant C'$ be an integer. Also let $(\mathcal{V}, \lambda)$ be a finite probability space, and for every $v \in \mathcal{V}$ let $\boldsymbol{X}_v = \langle X_s^v : s \in \binom{[n]}{d} \rangle$ be an $(1/C')$-spreadable, d-dimensional random array on $[n]$ which takes values in a set $\mathcal{X}_v$ with $|\mathcal{X}_v| = m$; assume* *that there exists* $\mathcal{S}_v \subseteq \mathcal{X}_v$ *with* $|\mathcal{S}_v| = |\mathcal{X}_v| - 1$ *such that for every* $a \in \mathcal{S}_v$ *we have*

$$\Big|\mathbb{P}\Big(\bigcap_{s\in\mathrm{Box}(d)} [X^v_s = a]\Big) - \prod_{s\in\mathrm{Box}(d)} \mathbb{P}\big([X^v_s = a]\big)\Big| \leqslant \frac{1}{C'} \tag{7.6}$$

where $\mathrm{Box}(d)$ *is as in* (3.2)*. Finally, for every* $v \in \mathcal{V}$ *let* $f_v \colon \mathcal{X}_v^{\binom{[n]}{d}} \to \mathbb{R}$ *be a function such that* $\mathbb{E}[f_v(\boldsymbol{X}_v)] = 0$ *and* $\|f_v(\boldsymbol{X}_v)\|_{L_p} = 1$*. Then there exist* $G \subseteq \mathcal{V}$ *with* $\lambda(G) \geqslant 1-\varepsilon$ *and an interval* I *of* $[n]$ *with* $|I| = k$ *such that for every* $v \in G$ *and every* $J \subseteq I$ *with* $|J| \geqslant d$ *we have*

$$\mathbb{P}\big(\big|\mathbb{E}[f_v(\boldsymbol{X}_v)\,|\,\mathcal{F}_J]\big| \leqslant \varepsilon\big) \geqslant 1-\varepsilon. \tag{7.7}$$

Proof. For every $v \in \mathcal{V}$ let $(\mathcal{A}^v_i)_{i=0}^m$ be the filtration defined in (2.24) for the random array $\boldsymbol{X}_v$ and let $(d^v_i)_{i=1}^m$ denote the martingale difference sequence of the Doob martingale for $f_v(\boldsymbol{X}_v)$. By Proposition 2.4 and our assumptions, we see that

$$\sum_{i=1}^m \Big(\mathop{\mathbb{E}}_{v\sim\lambda} \|d^v_i\|^2_{L_p}\Big) = \mathop{\mathbb{E}}_{v\sim\lambda} \Big(\sum_{i=1}^m \|d^v_i\|^2_{L_p}\Big) \overset{(2.9)}{\leqslant} \frac{1}{p-1} \tag{7.8}$$

and so, there exists $i_0 \in [m]$ such that

$$\mathop{\mathbb{E}}_{v\sim\lambda} \|d^v_{i_0}\|_{L_p} \leqslant \frac{1}{\sqrt{m(p-1)}}. \tag{7.9}$$

By Markov's inequality, there exists $G \subseteq \mathcal{V}$ with $\lambda(G) \geqslant 1-(m(p-1))^{-1/4}$ such that for every $v \in G$ we have

$$\|d^v_{i_0}\|_{L_p} \leqslant \frac{1}{\sqrt[4]{m(p-1)}}. \tag{7.10}$$

Using these observations, the result follows arguing precisely as in the proof of Theorem 6.1 in Section 6. □

Remark 7.5. We note that there is also an extension of Theorem 7.1 in the spirit of Theorem 7.4. More precisely, if we assume in Theorem 7.4 that for every $v \in \mathcal{V}$ the random array $\boldsymbol{X}_v$ is dissociated (not necessarily finite-valued), then the interval I can be selected so as $|I| \geqslant c'n$ where $c' \coloneqq \frac{1}{4}\,\varepsilon^{\frac{2(2p+1)}{p}}(p-1)$.

Part 2. Connection with combinatorics

8. Random arrays arising from combinatorial structures

In this section we present examples of boolean, spreadable, high-dimensional random arrays which arise from combinatorial structures and they satisfy the box independence condition and/or are approximately dissociated.

8.1. **From graphs and hypergraphs to spreadable random arrays.** Let $d \geqslant 2$ be an integer, and let V be a finite set with $|V| \geqslant d$. With every subset A of V^d we associate a boolean, spreadable, d-dimensional random array $\boldsymbol{X}_A = \langle X^A_s : s \in \binom{\mathbb{N}}{d} \rangle$ on $\mathbb{N}$ defined by setting for every $s = \{i_1 < \cdots < i_d\} \in \binom{\mathbb{N}}{d}$,

$$X^A_s \coloneqq \mathbf{1}_A(\xi_{i_1}, \dots, \xi_{i_d}). \tag{8.1}$$

where (ξ_i) is a sequence of i.i.d. random variables uniformly distributed on V. (Notice that if A is a nonempty proper subset of V^d, then the entries of $\boldsymbol{X}_A$ are not independent.)

A special case of this construction, which is relevant in the ensuing discussion, is obtained by considering a d-uniform hypergraph on V. Specifically, given a d-uniform hypergraph G on V, we identify G with a subset $\mathcal{G}$ of V^d via the rule

$$(v_1, \dots, v_d) \in \mathcal{G} \Leftrightarrow \{v_1, \dots, v_d\} \in G, \tag{8.2}$$

and we define $\boldsymbol{X}_G = \langle X^G_s : s \in \binom{\mathbb{N}}{d} \rangle$ to be the random array in (8.1) which corresponds to the set $\mathcal{G}$. Note that this definition is canonical, in the sense that various combinatorial parameters of G can be expressed as functions of the finite subarrays of $\boldsymbol{X}_G$. For instance, let $n \geqslant d$ be an integer, and let F be a d-uniform hypergraph on $[n]$; then, denoting by $t(F, G)$ the homomorphism density of F in G (see [Lov12, Chapter 5]), we have

$$t(F, G) = \mathbb{E}[f_F(\boldsymbol{X}_{G,n})] \tag{8.3}$$

where $f_F \colon \mathbb{R}^{\binom{[n]}{d}} \to \mathbb{R}$ is defined by setting for every $\boldsymbol{x} = (x_t)_{t \in \binom{[n]}{d}} \in \mathbb{R}^{\binom{[n]}{d}}$

$$f_F(\boldsymbol{x}) \coloneqq \prod_{s \in F} x_s \tag{8.4}$$

and $\boldsymbol{X}_{G,n}$ denotes the subarray of $\boldsymbol{X}_G$ determined by $[n]$. (See Definition 1.1.) Of course, similar identities are valid for weighted uniform hypergraphs.

As we shall see shortly in Proposition 8.2 below, in this general framework the box independence condition of the random array $\boldsymbol{X}_G$ is in fact equivalent to a well-known combinatorial property of G, namely its *quasirandomness*.

8.1.1. *Quasirandom graphs and hypergraphs.* Quasirandom objects are deterministic discrete structures which behave like random ones for most practical purposes. The phenomenon was first discovered in the context of graphs by Chung, Graham and Wilson [CGW88, CGW89] who build upon previous work of Thomason [Tho87]. The last twenty years the theory was also extended to hypergraphs, and it has found numerous significant applications in number theory and theoretical computer science (see, *e.g.*, [Rő15]).

8.1.1.1. Much of the modern theory of quasirandomness is developed using the box norms introduced by Gowers [Go07]. Specifically, let $d \geqslant 2$ be an integer, let (Ω, Σ, μ) be a probability space, and let Ω^d be equipped with the product measure. For every integrable random variable $f\colon \Omega^d \to \mathbb{R}$ we define its *box norm* $\|f\|_\square$ by the rule

$$\|f\|_\square \coloneqq \Big(\int \prod_{\boldsymbol{\epsilon}\in\{0,1\}^d} f(\boldsymbol{\omega}_{\boldsymbol{\epsilon}})\, d\boldsymbol{\mu}(\boldsymbol{\omega})\Big)^{1/2^d} \tag{8.5}$$

where $\boldsymbol{\mu}$ denotes the product measure on Ω^{2d} and, for every $\boldsymbol{\omega} = (\omega_1^0, \omega_1^1, \dots, \omega_d^0, \omega_d^1) \in \Omega^{2d}$ and every $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_d) \in \{0,1\}^d$ we have $\boldsymbol{\omega}_{\boldsymbol{\epsilon}} \coloneqq (\omega_1^{\epsilon_1}, \dots, \omega_d^{\epsilon_d}) \in \Omega^d$; by convention, we set $\|f\|_\square \coloneqq +\infty$ if the integral in (8.5) does not exist. The quantity $\|\cdot\|_\square$ is a norm on the vector space $\{f \in L_1 : \|f\|_\square < +\infty\}$, and it satisfies the following inequality, known as the *Gowers–Cauchy–Schwarz inequality*: for every collection $\langle f_{\boldsymbol{\epsilon}} : \epsilon \in \{0,1\}^d\rangle$ of integrable random variables on Ω^d we have

$$\Big| \int \prod_{\boldsymbol{\epsilon}\in\{0,1\}^d} f_{\boldsymbol{\epsilon}}(\boldsymbol{\omega}_{\boldsymbol{\epsilon}})\, d\boldsymbol{\mu}(\boldsymbol{\omega})\Big| \leqslant \prod_{\boldsymbol{\epsilon}\in\{0,1\}^d} \|f_{\boldsymbol{\epsilon}}\|_\square. \tag{8.6}$$

For proofs of these basic facts, as well as for a more complete presentation of related material, we refer to [GT10, Appendix B].

8.1.1.2. The link between the box norms and quasirandomness is given in the following definition.

Definition 8.1 (Box uniformity)**.** *Let $d \geqslant 2$, let V be a finite set with $|V| \geqslant d$, and let $\varrho > 0$. We say that a d-uniform hypergraph G on V is ϱ-*box uniform *(or* box uniform *if ϱ is understood) provided that*

$$\big\|\mathbf{1}_{\mathcal{G}} - \mathbb{E}[\mathbf{1}_{\mathcal{G}}]\big\|_\square \leqslant \varrho \tag{8.7}$$

where $\mathcal{G}$ is as in (8.2). (Here, we view V as a discrete probability space equipped with the uniform probability measure.)

Of course, Definition 8.1 is interesting when the parameter ϱ is much smaller than $\mathbb{E}[\mathbf{1}_{\mathcal{G}}]$. We also note that although box uniformity is defined analytically, it has a number of equivalent combinatorial formulations. For instance, it is easy to see that a graph G is box uniform if and only if it has roughly the expected number of 4-cycles; see [CGW88, CGW89, Rő15] for more information.

8.1.2. *The box independence condition via quasirandomness.* We have the following proposition. (See part (i) of Definition 3.1 for the definition of box independence.)

Proposition 8.2. *Let $d \geqslant 2$ be an integer, and let V be a finite set with $|V| \geqslant d$. Also let G be a d-uniform hypergraph on V, let $\boldsymbol{X}_G = \langle X_s^G : s \in \binom{\mathbb{N}}{d}\rangle$ be the random array associated with G via (8.1), and for every integer $n \geqslant d$ let $\boldsymbol{X}_{G,n}$ denote the subarray of $\boldsymbol{X}_G$ determined by $[n]$. Finally, let $\varrho, \vartheta > 0$. Then the following hold.*

(i) *If G is ϱ-box uniform, then for every integer $n \geqslant d$ the random array $\boldsymbol{X}_{G,n}$ is $(2^d\varrho, \{1\})$-box independent.*

(ii) *Conversely, if $\boldsymbol{X}_{G,n}$ is $(\vartheta, \{1\})$-box independent for some (equivalently, every) integer $n \geqslant d$, then G is $(12\,\vartheta^{1/8^d})$-box uniform.*

Proof. We start with the following observation which follows readily from (8.1). Let $\mathrm{Box}(d)$ be the d-dimensional box defined in (3.2), and let F be a nonempty subset of $\mathrm{Box}(d)$. Then there exists a subset[11] H of $\{0,1\}^d$ with $|F| = |H|$ and such that

$$\mathbb{E}\Big[\prod_{s\in F} X_s^G\Big] = \int \prod_{\boldsymbol{\epsilon}\in H} \mathbf{1}_{\mathcal{G}}(\boldsymbol{\omega}_{\boldsymbol{\epsilon}})\, d\boldsymbol{\mu}(\boldsymbol{\omega}). \tag{8.8}$$

(Here, by $\boldsymbol{\mu}$ we denote the uniform probability measure on V^{2d}, and we follow the conventions described right after (8.5).)

We proceed to the proof of part (i). Notice that $\|\mathbf{1}_{\mathcal{G}}\|_{\square} \leqslant \|\mathbf{1}_{\mathcal{G}}\|_{L_\infty} \leqslant 1$ and, moreover, $\mathbb{E}[X_s^G] = \mathbb{E}[\mathbf{1}_{\mathcal{G}}]$ for every $s \in \binom{\mathbb{N}}{d}$. Taking into account these observations and using our assumption, identity (8.8), a telescopic argument and the Gowers–Cauchy–Schwarz inequality (8.6), we obtain that

$$\Big|\mathbb{E}\Big[\prod_{s\in \mathrm{Box}(d)} X_s^G\Big] - \prod_{s\in \mathrm{Box}(d)} \mathbb{E}[X_s^G]\Big| \leqslant 2^d \varrho. \tag{8.9}$$

Since the random array $\boldsymbol{X}_G$ is spreadable, by Definition (3.1) and (8.9), we see $\boldsymbol{X}_G$ is $(2^d\varrho, \{1\})$-box independent.

For the proof of part (ii) we will need the following fact which follows from Theorem 3.2, the numerical invariants introduced in Subsection 4.2—see also (5.1)—and the fact that the random array $\boldsymbol{X}_G$ is spreadable.

Fact 8.3. *Let the notation and assumptions be as in part* (ii) *of Proposition* 8.2*. Then for every nonempty subset F of* $\mathrm{Box}(d)$ *we have*

$$\Big|\mathbb{E}\Big[\prod_{s\in F} X_s^G\Big] - \prod_{s\in F} \mathbb{E}[X_s^G]\Big| \leqslant 300\, 2^{2d}\, \vartheta^{1/4^d}. \tag{8.10}$$

Using Fact 8.3, we shall estimate the quantity

$$\big\|\mathbf{1}_{\mathcal{G}} - \mathbb{E}[\mathbf{1}_{\mathcal{G}}]\big\|_{\square}^{2^d} \overset{(8.5)}{=} \sum_{H\subseteq\{0,1\}^d} (-1)^{2^d-|H|}\mathbb{E}[\mathbf{1}_{\mathcal{G}}]^{2^d-|H|} \int \prod_{\boldsymbol{\epsilon}\in H} \mathbf{1}_{\mathcal{G}}(\boldsymbol{\omega}_{\boldsymbol{\epsilon}})\, d\boldsymbol{\mu}(\boldsymbol{\omega}). \tag{8.11}$$

(Here, as in the proof of Lemma 3.6, we use the convention that the product over an empty index-set is equal to 1.) Fix a nonempty subset H of $\{0,1\}^d$, and let F be the subset of $\mathrm{Box}(d)$ with $|F| = |H|$ and such that (8.8) is satisfied; since $\mathbb{E}[X_s^G] = \mathbb{E}[\mathbf{1}_{\mathcal{G}}]$ for every $s \in \binom{\mathbb{N}}{d}$, by Fact 8.3, we have

$$\Big|\int \prod_{\boldsymbol{\epsilon}\in H} \mathbf{1}_{\mathcal{G}}(\boldsymbol{\omega}_{\boldsymbol{\epsilon}})\, d\boldsymbol{\mu}(\boldsymbol{\omega}) - \mathbb{E}[\mathbf{1}_{\mathcal{G}}]^{|H|}\Big| \leqslant 300\, 2^{2d}\, \vartheta^{1/4^d}. \tag{8.12}$$

By (8.11) and (8.12), we see that G is $(12\,\vartheta^{1/8^d})$-box uniform, as desired. □

[11] Note that this subset is essentially unique.

8.2. **Mixtures.** An important property of the class of boolean, spreadable random arrays is that it is closed under mixtures. More precisely, let n, d, J be positive integers with $n \geqslant d \geqslant 2$ and let $\boldsymbol{X}_1 = \langle X^1_s : s \in \binom{[n]}{d}\rangle, \dots, \boldsymbol{X}_J = \langle X^J_s : s \in \binom{[n]}{d}\rangle$ be boolean, spreadable, d-dimensional random arrays on $[n]$. Then, for any choice $\lambda_1, \dots, \lambda_J$ of convex coefficients, there exists a boolean, spreadable, d-dimensional random array $\boldsymbol{X} = \langle X_s : s \in \binom{[n]}{d}\rangle$ on $[n]$ which satisfies

$$\mathbb{E}\Big[\prod_{s\in\mathcal{F}} X_s\Big] = \sum_{j=1}^{J} \lambda_j \mathbb{E}\Big[\prod_{s\in\mathcal{F}} X^j_s\Big] \tag{8.13}$$

for every nonempty finite subset $\mathcal{F}$ of $\binom{[n]}{d}$.

It turns out that the class of boolean, spreadable random arrays which satisfy the box independence condition is also closed under mixtures. In particular, we have the following proposition. (Its proof follows from a direct computation.)

Proposition 8.4. *Let n, d, J be positive integers with $n \geqslant d \geqslant 2$, and let $\delta, \vartheta > 0$. For every $j \in [J]$ let $\boldsymbol{X}_j = \langle X^j_s : s \in \binom{[n]}{d}\rangle$ be a boolean, spreadable, d-dimensional random array on $[n]$ which is $(\vartheta, \{1\})$-box independent and satisfies $|\mathbb{E}[X^j_s] - \delta| \leqslant \vartheta$ for all $s \in \binom{[n]}{d}$. If $\boldsymbol{X}$ is any mixture of $\boldsymbol{X}_1, \dots, \boldsymbol{X}_J$, then $\boldsymbol{X}$ is $(2^{d+2}\vartheta, \{1\})$-box independent.*

Observe that, by Propositions 8.2 and 8.4, if $G_1, \dots, G_J$ are quasirandom, d-uniform hypergraphs with the same edge density, then any mixture of the finite subarrays of $\boldsymbol{X}_{G_1}, \dots, \boldsymbol{X}_{G_J}$ satisfies the box independent condition. We note that this fact essentially characterizes the box independence condition. Specifically, it follows from [DTV21, Propositions 8.3 and 3.1] that for every boolean, spreadable, d-dimensional random array $\boldsymbol{X}$ which satisfies the box independence condition, there exist quasirandom, d-uniform hypergraphs $G_1, \dots, G_J$ with the same edge density, such that the law of $\boldsymbol{X}$ is close, in the total variation distance, to the law of a mixture of the finite subarrays of $\boldsymbol{X}_{G_1}, \dots, \boldsymbol{X}_{G_J}$.

8.3. **Further combinatorial structures.** Let n, k, d be positive integers with $n \geqslant d \geqslant 2$ and $k \leqslant \binom{n}{d}$, and let $\boldsymbol{\Xi} = \langle \xi_e : e \in \binom{[n]}{d}\rangle$ be a d-dimensional random array with boolean entries which are uniformly distributed on the set of all $\boldsymbol{x} \in \{0,1\}^{\binom{[n]}{d}}$ which have exactly k one's. (In particular, $\boldsymbol{\Xi}$ is exchangeable.) The random array $\boldsymbol{\Xi}$ generates the classical fixed size Erdős–Rényi random graph/hypergraph, and it is clear that it satisfies the box independence condition. By taking products[12] of the entries of $\boldsymbol{\Xi}$ as in (1.12), one also obtains exchangeable random arrays which are approximately dissociated.

Spreadable random arrays—and, in particular, spreadable random arrays which satisfy the box independence condition—are also closely related to a class of stochastic processes introduced by Furstenberg and Katznelson [FK91] in their proof of the density Hales–Jewett theorem (see also [Au11, DT21]). Unfortunately, this relation is not so transparent

[12]These products have a natural combinatorial interpretation; *e.g.*, they can be used to count subgraphs of random graphs.

as in case of graphs and hypergraphs, and we shall refrain from discussing it further since it requires several probabilistic and Ramsey-theoretic tools in order to be properly exposed.

9. Quasirandom families of graphs: proof of Theorem 1.8

9.1. **Isomorphic invariant families of graphs.** Let $n \geqslant 2$ be an integer, and recall that by μ we denote the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}}$. Also recall that a family of graphs $\mathcal{A} \subseteq \{0,1\}^{\binom{[n]}{2}}$ is called *isomorphic invariant* if for every $G \subseteq \binom{[n]}{2}$ we have that G belongs to $\mathcal{A}$ if every isomorphic copy of G belongs to $\mathcal{A}$. (See also (1.9).) This subsection is devoted to the proof of the following proposition.

Proposition 9.1. *Let $n \geqslant 2^{32}$ be an integer, and let $\mathcal{A} \subseteq \{0,1\}^{\binom{[n]}{2}}$ be isomorphic invariant. Then there exists a nonnegative parameter $\gamma(\mathcal{A})$ with*

$$\gamma(\mathcal{A}) \geqslant \mu(\mathcal{A})^4 - \frac{21\sqrt{2}}{\sqrt{\log_2 n}} \tag{9.1}$$

such that for every $U = \{i<j<k<\ell\} \in \binom{[n]}{4}$, denoting by $\mathbf{P}$ the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}\setminus\binom{U}{2}}$, we have

$$\mathbf{P}\big(W : W \cup \{i,k\}, W \cup \{i,\ell\}, W \cup \{j,k\}, W \cup \{j,\ell\} \in \mathcal{A}\big) = \gamma(\mathcal{A}). \tag{9.2}$$

Before we proceed to the proof of Proposition 9.1 we need to introduce some pieces of notation that will be used throughout this section. Specifically, if $n \geqslant 2$ is an integer and I is a nonempty subset of $\binom{[n]}{2}$, then for every $z \in \{0,1\}^{\binom{[n]}{2}}$ by $z \upharpoonright I \in \{0,1\}^I$ we shall denote the restriction of z on I. We will also view the set $\{0,1\}^I$ as a discrete probability space equipped with the uniform probability measure; we shall recall this particular convention whenever necessary in order to facilitate the reader.

Proof of Proposition 9.1. Since $\mathcal{A}$ is isomorphic invariant, there exists a unique nonnegative parameter $\gamma(\mathcal{A})$ which satisfies (9.2). Thus, we only need to show (9.1).

We start by setting

$$k \coloneqq \Big\lfloor \sqrt{\log_2(\sqrt{n})} \Big\rfloor, \quad \varepsilon \coloneqq \frac{1}{\sqrt[4]{n}}, \quad \delta \coloneqq \varepsilon\, 2^{-\frac{k^2}{2}}, \quad \ell \coloneqq \lfloor \delta^{-1} \rfloor \tag{9.3}$$

and we observe that $\ell k \leqslant n$ and $k \geqslant 4$. Let μ_1 denote the uniform probability measure on $\{0,1\}^{\binom{[k]}{2}}$, and for every $x \in \{0,1\}^{\binom{[k]}{2}}$ let $\mathcal{A}_x \coloneqq \big\{y \in \{0,1\}^{\binom{[n]}{2}\setminus\binom{[k]}{2}} : x \cup y \in \mathcal{A}\big\}$ denote the section of $\mathcal{A}$ at x. Also let μ_2 denote the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}\setminus\binom{[k]}{2}}$. Arguing as in the proof of Theorem 2.3 in Subsection 2.5 and using the fact that $\mathcal{A}$ is isomorphic invariant, we see that

$$\big\|\mathbb{E}[\mathbf{1}_{\mathcal{A}} \,|\, \mathcal{F}] - \mu(\mathcal{A})\big\|_{L_2}^2 \leqslant \delta \tag{9.4}$$

where $\mathcal{F}$ denotes the σ-algebra of $\{0,1\}^{\binom{[n]}{2}}$ generated by the partition

$$\Big\{\big\{z \in \{0,1\}^{\binom{[n]}{2}} : z \upharpoonright \binom{[k]}{2} = x\big\} : x \in \{0,1\}^{\binom{[k]}{2}}\Big\}. \tag{9.5}$$

(Note that here we view $\{0,1\}^{\binom{[n]}{2}}$ as a discrete probability space equipped with μ.) By (9.3), (9.4) and Chebyshev's inequality, we obtain that

$$\mu_1\big(x \in \{0,1\}^{\binom{[k]}{2}} : |\mu_2(\mathcal{A}_x) - \mu(\mathcal{A})| \geqslant \varepsilon\big) \leqslant 2^{-\frac{k^2}{2}} < 2^{-\binom{[k]}{2}}; \tag{9.6}$$

consequently, for every $x \in \{0,1\}^{\binom{[k]}{2}}$ we have

$$\mu_2(\mathcal{A}_x) \geqslant \mu(\mathcal{A}) - \varepsilon. \tag{9.7}$$

Now for every $i \in [k-1]$ let $x_i \in \{0,1\}^{\binom{[k]}{2}}$ be such that $x_i^{-1}(\{1\}) = \big\{\{1, i+1\}\big\}$; namely, x_i is the graph on $[k]$ with the single edge $\{1, i+1\}$. Setting $\delta_1 \coloneqq \mu_2(\mathcal{A}_{x_1})$ and $\delta_2 \coloneqq \mu_2(\mathcal{A}_{x_1} \cap \mathcal{A}_{x_2})$ and using again the fact that $\mathcal{A}$ is isomorphic invariant, for every $i, j \in [k-1]$ with $i \neq j$ we have

$$\mu_2(\mathcal{A}_{x_i}) = \delta_1 \quad \text{and} \quad \mu_2(\mathcal{A}_{x_i} \cap \mathcal{A}_{x_j}) = \delta_2. \tag{9.8}$$

Therefore, by the Cauchy–Schwarz inequality,

$$\begin{aligned}\big(\mu(\mathcal{A}) - \varepsilon\big)^2 &\overset{(9.7)}{\leqslant} \mathbb{E}\Big[\frac{1}{k-1}\sum_{i=1}^{k-1} \mathbf{1}_{\mathcal{A}_{x_i}}\Big]^2 \leqslant \frac{1}{(k-1)^2}\mathbb{E}\Big[\Big(\sum_{i=1}^{k-1} \mathbf{1}_{\mathcal{A}_{x_i}}\Big)^2\Big] \\ &= \frac{\delta_1}{k-1} + \frac{k-2}{k-1}\delta_2 \leqslant \frac{1}{k-2} + \delta_2.\end{aligned} \tag{9.9}$$

Next, for every $i \in [k-2]$ let $x_{1,i}, x_{2,i} \in \{0,1\}^{\binom{[k]}{2}}$ be defined by $x_{1,i}^{-1}(\{1\}) = \big\{\{i, k-1\}\big\}$ and $x_{2,i}^{-1}(\{1\}) = \big\{\{i, k\}\big\}$. Using once again the isomorphic invariance of $\mathcal{A}$ and setting $\delta_4 \coloneqq \mu_2(\mathcal{A}_{x_{1,1}} \cap \mathcal{A}_{x_{2,1}} \cap \mathcal{A}_{x_{1,2}} \cap \mathcal{A}_{x_{2,2}})$, we see that

$$\mu_2(\mathcal{A}_{x_{1,i}} \cap \mathcal{A}_{x_{2,i}}) = \delta_2 \quad \text{and} \quad \mu_2(\mathcal{A}_{x_{1,i}} \cap \mathcal{A}_{x_{2,i}} \cap \mathcal{A}_{x_{1,j}} \cap \mathcal{A}_{x_{2,j}}) = \delta_4 \tag{9.10}$$

for every $i, j \in [k-1]$ with $i \neq j$. Hence, by the Cauchy–Schwarz inequality,

$$\begin{aligned}\Big((\mu(\mathcal{A}) - \varepsilon)^2 - \frac{1}{k-2}\Big)^2 &\overset{(9.9)}{\leqslant} \delta_2^2 = \mathbb{E}\Big[\frac{1}{k-2}\sum_{i=1}^{k-2} \mathbf{1}_{\mathcal{A}_{x_{1,i}} \cap \mathcal{A}_{x_{2,i}}}\Big]^2 \\ &\leqslant \frac{1}{(k-2)^2}\mathbb{E}\Big[\Big(\sum_{i=1}^{k-2} \mathbf{1}_{\mathcal{A}_{x_{1,i}} \cap \mathcal{A}_{x_{2,i}}}\Big)^2\Big] = \frac{\delta_2}{k-2} + \frac{k-3}{k-2}\delta_4 \leqslant \frac{1}{k-3} + \delta_4\end{aligned} \tag{9.11}$$

which yields that

$$\delta_4 \geqslant \mu(\mathcal{A})^4 - 8\varepsilon - \frac{3}{k-3}. \tag{9.12}$$

We will show that the parameter $\gamma(\mathcal{A})$ is roughly equal to δ_4. Clearly, this is enough to compete the proof.

To this end we will argue as in (9.4). Precisely, define $z_1, z_2, z_3, z_4 \in \{0,1\}^{\binom{[4]}{2}}$ by $z_1^{-1}(\{1\}) = \big\{\{1,3\}\big\}$, $z_2^{-1}(\{1\}) = \big\{\{1,4\}\big\}$, $z_3^{-1}(\{1\}) = \big\{\{2,3\}\big\}$ and $z_4^{-1}(\{1\}) = \big\{\{2,4\}\big\}$, and for every $i \in [4]$ let $\mathcal{A}_{z_i} \coloneqq \big\{y \in \{0,1\}^{\binom{[n]}{2} \setminus \binom{[4]}{2}} : z_i \cup y \in \mathcal{A}\big\}$ denote the section of $\mathcal{A}$ at z_i. Also set $S \coloneqq \mathcal{A}_{z_1} \cap \mathcal{A}_{z_2} \cap \mathcal{A}_{z_3} \cap \mathcal{A}_{z_4}$ and for every $u \in \binom{[k]}{2} \setminus \binom{[4]}{2}$ let $S_u \coloneqq \big\{y \in \{0,1\}^{\binom{[n]}{2} \setminus \binom{[k]}{2}} : u \cup y \in S\big\}$ denote the section of S at u. Notice that, by the choice of $\gamma(\mathcal{A})$ in (9.2), we have $\mathbf{P}(S)=\gamma(\mathcal{A})$ where $\mathbf{P}$ is the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}\setminus\binom{[4]}{2}}$. Taking into account this observation, the isomorphic invariance of $\mathcal{A}$ and proceeding[13] as in the proof of Theorem 2.3, we obtain that

$$\big\|\mathbb{E}[\mathbf{1}_S\,|\,\mathcal{G}]-\gamma(\mathcal{A})\big\|_{L_2}^2\leqslant\delta \tag{9.13}$$

where $\mathcal{G}$ stands for the σ-algebra of $\{0,1\}^{\binom{[n]}{2}\setminus\binom{[4]}{2}}$ generated by the partition

$$\Big\{\big\{z\in\{0,1\}^{\binom{[n]}{2}\setminus\binom{[4]}{2}}: z\upharpoonright\binom{[k]}{2}\Big\backslash\binom{[4]}{2}=u\big\}: u\in\{0,1\}^{\binom{[k]}{2}\setminus\binom{[4]}{2}}\Big\}. \tag{9.14}$$

By (9.3), (9.13) and Chebyshev's inequality, we have

$$\mu_3\big(u\in\{0,1\}^{\binom{[k]}{2}\setminus\binom{[4]}{2}}:|\mu_2(S_u)-\gamma(\mathcal{A})|\geqslant\varepsilon\big)\leqslant 2^{-\frac{k^2}{2}}<2^{-\binom{[k]}{2}+\binom{[4]}{2}} \tag{9.15}$$

where μ_3 denotes the uniform probability measure on $\{0,1\}^{\binom{[k]}{2}\setminus\binom{[4]}{2}}$ and, as above, μ_2 is the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}\setminus\binom{[k]}{2}}$; thus, for every $u\in\{0,1\}^{\binom{[k]}{2}\setminus\binom{[4]}{2}}$,

$$|\mu_2(S_u)-\gamma(\mathcal{A})|<\varepsilon. \tag{9.16}$$

Let $u_0\in\{0,1\}^{\binom{[k]}{2}\setminus\binom{[4]}{2}}$ be such that $u_0^{-1}(\{1\})=\emptyset$. Then, by the definitions of S and δ_4 and the isomorphic invariance of $\mathcal{A}$, we see that $\mu_2(S_{u_0})=\delta_4$ and consequently, by (9.12) and (9.16), we conclude that

$$\gamma(\mathcal{A})\geqslant\mu(\mathcal{A})^4-9\varepsilon-\frac{3}{k-3}\geqslant\mu(\mathcal{A})^4-\frac{21}{k}\overset{(9.3)}{\geqslant}\mu(\mathcal{A})^4-\frac{21}{\sqrt{\log_2(\sqrt{n})}}. \qquad\square \tag{9.17}$$

9.2. **Proof of Theorem 1.8.** The main goal of the proof is to extract out of the quasirandom family $\mathcal{A}$ a boolean two-dimensional approximately spreadable random array $\boldsymbol{X}$ which satisfies the box independence condition; once this is done, the proof will be completed by an application of Theorem 1.5.

9.2.1. *Preliminary tools.* We start with a more precise, quantitative, version of Proposition 1.3 for boolean two-dimensional random arrays. Specifically, let $\ell,m,r\geqslant 2$ be integers with $\ell\leqslant m$, and recall that the multicolor hypergraph Ramsey number $R_\ell(m,r)$ is the least integer $N\geqslant m$ such that for every set X with $|X|\geqslant N$ and every coloring $c\colon\binom{X}{\ell}\to[r]$ there exists $Y\in\binom{X}{m}$ such that c is constant on $\binom{Y}{\ell}$. It is a classical result due to Erdős and Rado [ER52] that the numbers $R_\ell(m,r)$ have (at most) a tower-type dependence with respect to the parameters ℓ,m,r. The following fact is the promised quantitative version of Proposition 1.3; the proof follows by invoking the relevant definitions, and it is left to the reader.

[13] Again we point out that here we view $\{0,1\}^{\binom{[n]}{2}\setminus\binom{[4]}{2}}$ as a discrete probability space equipped with $\mathbf{P}$. We also note that in this case the filtration is slightly different since we are working with the index-set $\binom{[n]}{2}\setminus\binom{[4]}{2}$ instead of $\binom{[n]}{2}$. We will use a similar variation in the next subsection.

Fact 9.2. *Let $0<\eta\leqslant 1$, let $\ell\geqslant 2$ be an integer, and let N be an integer such that*

$$N \geqslant R_\ell\Big(2\ell, \big(2^{\binom{\ell}{2}}\lceil \eta^{-1}\rceil\big)^{\binom{\ell}{2}}\Big). \tag{9.18}$$

Then for every boolean two-dimensional random array $\boldsymbol{X}$ on $[N]$ there exists $L\in\binom{[N]}{\ell}$ such that the subarray $\boldsymbol{X}_L$ is η-spreadable. (*See Definition 1.1.*)

We proceed by introducing some terminology and some pieces of notation. Let $m\geqslant \ell$ be positive integers and let $F\in\binom{[m]}{\ell}$; given two subsets $L\subseteq M$ of $\mathbb{N}$ with $|L|=\ell$ and $|M|=m$, we say that the *relative position of L inside M is F* if, denoting by $\{i_1<\cdots<i_m\}$ the increasing enumeration of M, we have that $L=\{i_j: j\in F\}$.

Moreover, for every finite subset M of $\mathbb{N}$ with $|M|\geqslant 2$ every $e\in\binom{M}{2}$ we shall denote by $x(e,M)\in\{0,1\}^{\binom{M}{2}}$ the unique element satisfying $x(e,M)^{-1}(\{1\})=\{e\}$.

It is also convenient to introduce the following definition. (Recall that for every integer $n\geqslant 2$ by μ we denote the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}}$.)

Definition 9.3 (Admissibility)**.** *Let $0<\eta\leqslant 1$, let $n\geqslant m\geqslant \ell\geqslant 2$ be integers, let $\mathcal{A}\subseteq\{0,1\}^{\binom{[n]}{2}}$ and let $F\in\binom{[m]}{\ell}$. Given $P\subseteq[n]$ with $|P|\geqslant m$, we say that P is* $(\mathcal{A},\eta,F)$-admissible *if for every $M\in\binom{P}{m}$, denoting by ν the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}\setminus\binom{M}{2}}$, the following hold.*

($\mathcal{P}1$) *For every $x\in\{0,1\}^{\binom{M}{2}}$ we have $|\nu(\mathcal{A}_x)-\mu(\mathcal{A})|\leqslant\eta$ where $\mathcal{A}_x$ is the section of $\mathcal{A}$ at x.*

($\mathcal{P}2$) *If $L\in\binom{M}{\ell}$ is the unique subset of M whose relative position inside M is F, then the two-dimensional random array $\langle \mathbf{1}_{\mathcal{A}_{x(e,M)}}: e\in\binom{L}{2}\rangle$ is η-spreadable.* (*Here, we view $\{0,1\}^{\binom{[n]}{2}\setminus\binom{M}{2}}$ as a discrete probability space equipped with ν and we denote by $\mathcal{A}_{x(e,M)}$ the section of $\mathcal{A}$ at $x(e,M)$.*)

We have the following lemma.

Lemma 9.4. *Let $0<\eta\leqslant 1$, let $\ell\geqslant 2$ be an integer, and set*

$$m=m(\eta,\ell)\coloneqq R_\ell\Big(2\ell, \big(2^{\binom{\ell}{2}}\lceil \eta^{-1}\rceil\big)^{\binom{\ell}{2}}\Big). \tag{9.19}$$

Then for every integer $p\geqslant m$ there exists an integer $q\geqslant p$ with the following property. If $n\geqslant q$ is an integer and $\mathcal{A}\subseteq\{0,1\}^{\binom{[n]}{2}}$ is a family of graphs, then for every $Q\in\binom{[n]}{q}$ there exist $F\in\binom{[m]}{\ell}$ and $P\in\binom{Q}{p}$ such that P is $(\mathcal{A},\eta,F)$-admissible.

Proof. Fix $p\geqslant m$ and set

$$q_1\coloneqq R_\ell\Big(p,\binom{m}{\ell}\Big) \quad\text{and}\quad q\coloneqq \max\big\{\lceil\eta^{-2}\rceil, 2^{q_1^2}\big\}\cdot q_1 \tag{9.20}$$

We claim that q is as desired. Indeed, let $n,\mathcal{A}$ be as in the statement of the lemma, and let $Q\in\binom{[n]}{q}$ be arbitrary. Setting $\delta\coloneqq\max\big\{\eta^2,2^{-q_1^2}\big\}$ and arguing[14] as in Theorem 2.3,

[14] Note that, although here we proceed exactly as in (9.4), we cannot guarantee that the set Q_1 is an initial interval of Q since the family $\mathcal{A}$ is not necessarily isomorphic invariant.

we select $Q_1 \in \binom{Q}{q_1}$ such that

$$\big\|\mathbb{E}[\mathbf{1}_{\mathcal{A}}\,|\,\mathcal{F}] - \mu(\mathcal{A})\big\|_{L_2}^2 \leqslant \delta \tag{9.21}$$

where $\mathcal{F}$ denotes the σ-algebra of $\{0,1\}^{\binom{[n]}{2}}$ generated by the partition

$$\Big\{\big\{z \in \{0,1\}^{\binom{[n]}{2}} : z \upharpoonright \binom{Q_1}{2} = x\big\} : x \in \{0,1\}^{\binom{Q_1}{2}}\Big\}. \tag{9.22}$$

(Recall that we view $\{0,1\}^{\binom{[n]}{2}}$ as a discrete probability space equipped with μ.) Since $\sqrt{\delta} < 2^{-\binom{q_1}{2}}$, by (9.21) and Chebyshev's inequality, for every $x \in \{0,1\}^{\binom{Q_1}{2}}$ we have

$$|\mu_1(\mathcal{A}_x) - \mu(\mathcal{A})| \leqslant \eta \tag{9.23}$$

where μ_1 denotes the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}\setminus\binom{Q_1}{2}}$ and, as usual, $\mathcal{A}_x$ is the section of $\mathcal{A}$ at x. By double averaging, this yields that for every $M \in \binom{Q_1}{m}$ and every $x \in \{0,1\}^{\binom{M}{2}}$ we have $|\nu(\mathcal{A}_x) - \mu(\mathcal{A})| \leqslant \eta$ where ν is the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}\setminus\binom{M}{2}}$. In other words, property ($\mathcal{P}1$) in Definition 9.3 will be satisfied as long as the desired set P is contained in Q_1.

For property ($\mathcal{P}2$) we argue as follows. Let $M \in \binom{Q_1}{m}$ be arbitrary; by the choice of the constant m in (9.19) and Fact 9.2 applied to the boolean, two-dimensional random array $\langle \mathbf{1}_{\mathcal{A}_{x(e,M)}} : e \in \binom{M}{2}\rangle$, there exists $F_M \in \binom{[m]}{\ell}$ such that if $L \in \binom{M}{\ell}$ is the unique subset of M whose relative position inside M is F_M, then the random array $\langle \mathbf{1}_{\mathcal{A}_{x(e,M)}} : e \in \binom{L}{2}\rangle$ is η-spreadable. By the choice of q_1 in (9.20) and another application of Ramsey's theorem, there exist $P \in \binom{Q_1}{p}$ and $F \in \binom{[m]}{\ell}$ such that $F_M = F$ for every $M \in \binom{P}{m}$. That is, property ($\mathcal{P}2$) is satisfied for P, as desired. □

9.2.2. *Numerical parameters.* Our next step is to introduce some numerical parameters. We fix $0 < \delta \leqslant 1$ and an integer $k \geqslant 2$, and we begin by selecting $0 < \eta, \theta_0 \leqslant 1$ and an integer $\ell \geqslant 4k$ such that

$$\eta + \Big(\lfloor \ell/k \rfloor^{-1} + 2k^2\eta + 800k^2\big(\ell^{-1/16} + \eta^{1/16} + (10\eta + \theta_0)^{1/16}\big)\Big)^{1/2} < \delta^{\binom{k}{2}+1}. \tag{9.24}$$

Next we set

$$m \coloneqq m(\eta,\ell) \overset{(9.19)}{=} R_\ell\Big(2\ell, \big(2^{\binom{\ell}{2}}\lceil \eta^{-1}\rceil\big)^{\binom{\ell}{2}}\Big) \tag{9.25}$$

$$\varrho \coloneqq \eta^2 2^{-m^2} \tag{9.26}$$

$$p \coloneqq 5m\lceil \varrho^{-1}\rceil + 4 \tag{9.27}$$

$$q_1 \coloneqq \mathrm{R}_m\Big(p, \binom{m}{\ell}\Big). \tag{9.28}$$

Finally, we define

$$q_0 \coloneqq \max\big\{\lceil \eta^2\rceil, 2^{q_1^2}\big\}\cdot q_1 \quad \text{and} \quad \theta \coloneqq \frac{1}{2}\cdot \min\Big\{\theta_0, \binom{q_0}{4}^{-1}\Big\}. \tag{9.29}$$

Notice, in particular, that with this choice we also have that

$$\eta + \Big(\lfloor \ell/k \rfloor^{-1} + 2k^2\eta + 800k^2\big(\ell^{-1/16} + \eta^{1/16} + (10\eta+\theta)^{1/16}\big)\Big)^{1/2} < \delta^{\binom{k}{2}+1}. \tag{9.30}$$

9.2.3. *Completion of the proof.* We are ready for the main part of the argument. As above, let $0 < \delta \leqslant 1$ and $k \geqslant 2$. Also let $n \geqslant q_0$ be an integer and let $\mathcal{A} \subseteq \{0,1\}^{\binom{[n]}{2}}$ be a θ-quasirandom family of graphs with $\mu(\mathcal{A}) \geqslant \delta$, where q_0, θ are as in (9.29).

By Lemma 9.4, for every $Q \in \binom{[n]}{q_0}$ we fix $P_Q \in \binom{Q}{p}$ and $F_Q \in \binom{[m]}{\ell}$ such that P_Q is $(\mathcal{A}, \eta, F_Q)$-admissible in the sense of Definition 9.3. Moreover, denoting by $\{r_1 < \dots < r_p\}$ the increasing enumeration of P_Q, we set $U_Q \coloneqq \big\{ r_{jm\lceil 1/\varrho\rceil + j} : j \in \{1,2,3,4\}\big\} \in \binom{Q}{4}$. Then observe that

$$\Big|\Big\{ U_Q : Q \in \binom{[n]}{q_0}\Big\}\Big| \geqslant \binom{n}{4}\binom{q_0}{4}^{-1}. \tag{9.31}$$

Since the family $\mathcal{A}$ is θ-quasirandom, by Definition 1.7 and the choice of θ in (9.29), there exists $Q_0 \in \binom{[n]}{q_0}$ such that, writing $U_{Q_0} = \{u_1 < u_2 < u_3 < u_4\}$ and setting[15] $B \coloneqq \mathcal{A}_{x(\{u_1,u_3\},U_{Q_0})} \cap \mathcal{A}_{x(\{u_1,u_4\},U_{Q_0})} \cap \mathcal{A}_{x(\{u_2,u_3\},U_{Q_0})} \cap \mathcal{A}_{x(\{u_2,u_4\},U_{Q_0})}$, we have

$$\mathbf{P}(B) \leqslant \mu(\mathcal{A})^4 + \theta \tag{9.32}$$

where $\mathbf{P}$ denotes the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}\setminus\binom{U_{Q_0}}{2}}$.

As above, let $\{r_1 < \dots < r_p\}$ denote the increasing enumeration of the set P_{Q_0}. For every $i \in \big\{1, \dots, \lceil \varrho^{-1}\rceil\big\}$ set

$$R_i \coloneqq \{ r_{j+(t-1)(m\lceil 1/\varrho\rceil+1)} : t \in [5] \text{ and } j \in [im]\} \tag{9.33}$$

and notice that $R_i \cap U_{Q_0} = \emptyset$; also let $\mathcal{F}_i$ be the σ-algebra of $\{0,1\}^{\binom{[n]}{2}\setminus\binom{U_{Q_0}}{2}}$ generated by the partition

$$\Big\{ \big\{ y \in \{0,1\}^{\binom{[n]}{2}\setminus\binom{U_{Q_0}}{2}} : y \upharpoonright \binom{R_i \cup U_{Q_0}}{2} \setminus \binom{U_{Q_0}}{2} = z \big\} : z \in \{0,1\}^{\binom{R_i\cup U_{Q_0}}{2}\setminus\binom{U_{Q_0}}{2}} \Big\}.$$

Finally, set

$$R_0 \coloneqq \emptyset \quad \text{and} \quad \mathcal{F}_0 \coloneqq \Big\{ \emptyset, \{0,1\}^{\binom{[n]}{2}\setminus\binom{U_{Q_0}}{2}} \Big\}. \tag{9.34}$$

Consider the Doob martingale for $\mathbf{1}_B$ with respect to the filtration $\mathcal{F}_0, \mathcal{F}_1, \dots, \mathcal{F}_{\lceil \varrho^{-1}\rceil}$; using the elementary fact that martingale difference sequences are orthogonal in L_2, we may select $i_0 \in \big\{1, \dots, \lceil \varrho^{-1}\rceil\big\}$ such that

$$\big\| \mathbb{E}[\mathbf{1}_B \,|\, \mathcal{F}_{i_0}] - \mathbb{E}[\mathbf{1}_B \,|\, \mathcal{F}_{i_0-1}] \big\|_{L_2}^2 \leqslant \varrho. \tag{9.35}$$

(Here, we view $\{0,1\}^{\binom{[n]}{2}\setminus\binom{U_{Q_0}}{2}}$ as a discrete probability space equipped with $\mathbf{P}$.) Next, we select $M \in \binom{U_{Q_0}\cup(R_{i_0}\setminus R_{i_0-1})}{m} \subseteq \binom{P_{Q_0}}{m}$ such that, denoting by L the unique element of $\binom{M}{\ell}$ whose relative position inside M is F_{Q_0}, we have that $U_{Q_0} \subseteq L$; this selection is possible

---

[15] Recall that by $x(\{u_1,u_3\}, U_{Q_0})$ we denote the unique element of $\{0,1\}^{\binom{Q_0}{2}}$ such that $x(\{u_1,u_3\}, U_{Q_0})^{-1}(\{1\}) = \{u_1,u_3\}$, and similarly for $x(\{u_1,u_4\}, U_{Q_0})$, $x(\{u_2,u_3\}, U_{Q_0})$, $x(\{u_2,u_4\}, U_{Q_0})$.

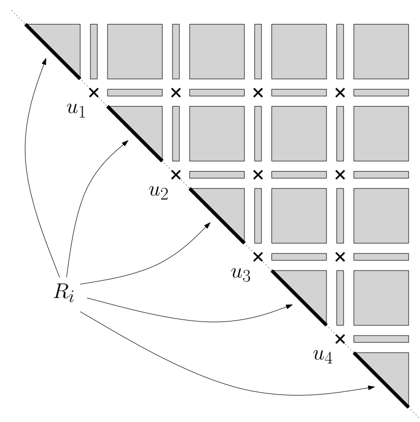

FIGURE 5. The σ-algebra $\mathcal{F}_i$ associated with the set R_i.

by the choice of R_{i_0} in (9.33). Finally, denote by $\mathcal{F}$ the σ-algebra of $\{0,1\}^{\binom{[n]}{2}\setminus\binom{U_{Q_0}}{2}}$ generated by the partition

$$\Big\{\big\{y\in\{0,1\}^{\binom{[n]}{2}\setminus\binom{U_{Q_0}}{2}} : y\upharpoonright\binom{M}{2}\setminus\binom{U_{Q_0}}{2}=z\big\} : z\in\{0,1\}^{\binom{M}{2}\setminus\binom{U_{Q_0}}{2}}\Big\}$$

and observe that $\mathcal{F}$ is a sub-σ-algebra of $\mathcal{F}_{i_0}$, while the σ-algebras $\mathcal{F}$ and $\mathcal{F}_{i_0-1}$ are independent. Therefore, by (9.35) and the fact that the conditional expectation is a linear contraction on L_2, we see that

$$\big\|\mathbb{E}[\mathbf{1}_B\,|\,\mathcal{F}]-\mathbf{P}(B)\big\|^2_{L_2}\leqslant\varrho. \tag{9.36}$$

As we have already noted, this estimate together with Chebyshev's inequality and the choice of ϱ in (9.26) yield that for every $z\in\{0,1\}^{\binom{M}{2}\setminus\binom{U_{Q_0}}{2}}$ we have

$$|\nu(B_z)-\mathbf{P}(B)|\leqslant\eta \tag{9.37}$$

where ν is the uniform probability measure on $\{0,1\}^{\binom{[n]}{2}\setminus\binom{M}{2}}$ and B_z is the section of B at z.

Now let $z_0\in\{0,1\}^{\binom{M}{2}\setminus\binom{U_{Q_0}}{2}}$ be such that $z_0^{-1}(\{1\})=\emptyset$. By (9.32) and (9.37), we have

$$\nu(B_{z_0})\leqslant\mu(\mathcal{A})^4+\eta+\theta. \tag{9.38}$$

Also notice that $B_{z_0}=\mathcal{A}_{x(\{u_1,u_3\},M)}\cap\mathcal{A}_{x(\{u_1,u_4\},M)}\cap\mathcal{A}_{x(\{u_2,u_3\},M)}\cap\mathcal{A}_{x(\{u_2,u_4\},M)}$. On the other hand, recall that P_{Q_0} is $(\mathcal{A},\eta,F_{Q_0})$-admissible and that L is the unique subset of $M\in\binom{P_{Q_0}}{m}$ whose relative position inside M is F_{Q_0}. Taking into account these observations and using properties ($\mathcal{P}1$) and ($\mathcal{P}2$) in Definition 9.3, we see that the boolean random array $\langle\mathbf{1}_{\mathcal{A}_{x(e,M)}}:e\in\binom{L}{2}\rangle$ is η-stationary and it satisfies (1.7) with $\vartheta=10\eta+\theta$. Let $\{s_1<\dots<s_m\}$ denote the increasing enumeration of M, and for every

$j \in \{1, \dots, \lfloor \ell/k \rfloor\}$ set

$$K_j \coloneqq \{s_i : i \in [jk] \setminus [(j-1)k]\} \quad \text{and} \quad \Gamma_j \coloneqq \bigcap_{e \in \binom{K_j}{2}} \mathcal{A}_{x(e,M)}. \tag{9.39}$$

By Theorem 1.5, property ($\mathcal{P}1$) in Definition 9.3 and the previous discussion, for every $j, l \in \{1, \dots, \lfloor \ell/k \rfloor\}$ with $j \neq l$ we have

$$\big|\nu(\Gamma_j) - \mu(\mathcal{A})^{\binom{k}{2}}\big| \leqslant \eta \binom{k}{2} + 400 \binom{k}{2} \big(\ell^{-1/16} + \eta^{1/16} + (10\eta + \theta)^{1/16}\big) \tag{9.40}$$

and

$$\big|\nu(\Gamma_j \cap \Gamma_l) - \mu(\mathcal{A})^{2\binom{k}{2}}\big| \leqslant \eta\, 2\binom{k}{2} + 400 \cdot 2\binom{k}{2} \big(\ell^{-1/16} + \eta^{1/16} + (10\eta + \theta)^{1/16}\big). \tag{9.41}$$

Introduce the random variable

$$X \coloneqq \frac{1}{\lfloor \ell/k \rfloor} \sum_{j=1}^{\lfloor \ell/k \rfloor} \mathbf{1}_{\Gamma_j} \tag{9.42}$$

and observe that, by (9.40) and (9.41), we have

$$\|X - \mu(\mathcal{A})^{\binom{k}{2}}\|_{L_2} \leqslant \Big(\lfloor \ell/k \rfloor^{-1} + 2k^2\eta + 800k^2\big(\ell^{-1/16} + \eta^{1/16} + (30\eta + 6\theta)^{1/16}\big)\Big)^{1/2}. \tag{9.43}$$

Let $x_0 \in \{0,1\}^{\binom{M}{2}}$ be the unique element satisfying $x_0^{-1}(\{1\}) = \emptyset$ and note that

$$\begin{aligned} \mathbb{E}[X\mathbf{1}_{\mathcal{A}_{x_0}}] &\geqslant \mu(\mathcal{A})^{\binom{k}{2}}\nu(\mathcal{A}_{x_0}) - \sqrt{\nu(\mathcal{A}_{x_0})}\,\|X - \mu(\mathcal{A})^{\binom{k}{2}}\|_{L_2} \\ &\overset{(\mathcal{P}1),(9.43)}{\geqslant} \delta^{\binom{k}{2}+1} - \eta - \Big(\lfloor \ell/k \rfloor^{-1} + 2k^2\eta + 800k^2\big(\ell^{-1/16} + \eta^{1/16} + (10\eta + \theta)^{1/16}\big)\Big)^{1/2}. \end{aligned} \tag{9.44}$$

Therefore, by (9.30), we have $\mathbb{E}[X\mathbf{1}_{\mathcal{A}_{x_0}}] > 0$ which in turn implies that there exists $j_0 \in \{1, \dots, \lfloor \ell/k \rfloor\}$ such that $\mathcal{A}_{x_0} \cap \Gamma_{j_0} \neq \emptyset$. By the choices of Γ_{j_0} in (9.39) and x_0, it is clear that the set $K_{j_0} \in \binom{[n]}{k}$ is as desired. The proof of Theorem 1.8 is completed.

Remark 9.5 (Analysis of the bounds). Using the Erdős–Rado theorem [ER52], it is not hard to see that the proof of Theorem 1.8 yields a tower-type dependence of θ and ℓ_0 with respect to the parameters δ and k. More precisely, there exists a primitive recursive $\psi\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ function belonging to the class $\mathcal{E}^4$ of Grzegorczyk's hierarchy[16] such that $\theta^{-1}, \ell_0 \leqslant \psi(\lceil \delta^{-1} \rceil, k)$ for every $0 < \delta \leqslant 1$ and every integer $k \geqslant 2$.

Remark 9.6 (Extensions to families of uniform hypergraphs). Theorem 1.8 can be extended to families of d-uniform hypergraphs $\mathcal{A} \subseteq \{0,1\}^{\binom{[n]}{d}}$ for any integer $d \geqslant 2$; this can be done by using Theorem 3.2 instead Theorem 1.5 and appropriately modifying the notion of quasirandomness in Definition 1.7. We leave the (fairly straightforward) formulations of these extensions to the interested reader.

[16]See [DK16, Appendix A] for an introduction to Grzegorczyk's hierarchy and a discussion of its role in analyzing bounds in Ramsey theory.

Remark 9.7. Let $\mathcal{A} \subseteq \{0,1\}^{\binom{[n]}{2}}$ be a family of graphs on $[n]$, let $K \subseteq [n]$ with $|K| \geqslant 2$, and let $\mathcal{S} \subseteq \{0,1\}^{\binom{K}{2}}$ be a family of graphs on K. We say that $\mathcal{A}$ *smashes* $\mathcal{S}$ if there exists $W \subseteq \binom{[n]}{2} \setminus \binom{K}{2}$ such that $W \cup H \in \mathcal{A}$ for every $H \in \mathcal{S}$. With this terminology, Theorem 1.8 is equivalent to saying that if the family $\mathcal{A}$ is dense and quasirandom—in the sense of Definition 1.7—then it smashes all graphs with at most one edge on some $K \in \binom{[n]}{k}$. It would be interesting to find quasirandomness conditions which ensure that the family $\mathcal{A}$ smashes richer families of small graphs. In this direction, the following problem is the most intriguing.

Problem 9.8. *Find natural quasirandomness conditions on a family of graphs $\mathcal{A}$ which ensure that $\mathcal{A}$ smashes* all *graphs on some $K \in \binom{[n]}{k}$.*

Appendix A. Examples

Our goal in this appendix is to present examples which show that the box independence condition in Theorems 1.4 and 6.1 is essentially optimal. We focus on boolean random arrays as this case already covers all underlying phenomena.

A.1. **Boxes and faces.** We start by introducing some terminology which will be used throughout this section. Let $d \geqslant 2$ be an integer; we say that a subset B of $\binom{\mathbb{N}}{d}$ is a *d-dimensional box of* $\mathbb{N}$ if it is a d-dimensional box of $[n]$ for some integer $n \geqslant 2d$. (See Subsection 3.1.) Moreover, we say that a subset F of $\binom{\mathbb{N}}{d}$ is a $(d-1)$-*face of* $\mathbb{N}$ if it is of the form $\mathrm{Box}(\mathcal{H})$ where $\mathcal{H} = (H_1, \dots, H_d)$ is a finite sequence of nonempty subsets of $\mathbb{N}$ of cardinality at most 2 with $\max(H_i) < \min(H_{i+1})$ for all $i \in [d-1]$, and such that $\sum_{i=1}^{d} |H_i| = 2d - 1$. (Thus, $|H_i| = 2$ for all but at one $i \in [d]$.)

A.2. **The two-dimensional case.** We have the following proposition.

Proposition A.1. *There exists a boolean, exchangeable, two-dimensional random array $\boldsymbol{X} = \langle X_s : s \in \binom{\mathbb{N}}{2} \rangle$ on $\mathbb{N}$ with the following properties.*

- ($\mathcal{P}1$) *For every $s \in \binom{\mathbb{N}}{2}$ we have $\mathbb{E}[X_s] = \frac{1}{2}$.*
- ($\mathcal{P}2$) *For every distinct $s, t \in \binom{\mathbb{N}}{2}$ we have $\mathbb{E}[X_s X_t] = \frac{1}{4}$.*
- ($\mathcal{P}3$) *For every 2-dimensional box B of $\mathbb{N}$ and every nonempty subset G of B with $G \neq B$ we have $\mathbb{E}\big[\prod_{s \in G} X_s\big] = (\frac{1}{2})^{|G|}$.*
- ($\mathcal{P}4$) *For every 2-dimensional box B of $\mathbb{N}$ we have $\mathbb{E}\big[\prod_{s \in B} X_s\big] = \frac{3}{2}(\frac{1}{2})^4$.*
- ($\mathcal{P}5$) *Let $n \geqslant 8$ be an integer, and let $\boldsymbol{X}_n$ denote the subarray of $\boldsymbol{X}$ determined by $[n]$. (See Definition 1.1.) Then there exists a translated multilinear polynomial $f \colon \mathbb{R}^{\binom{[n]}{2}} \to \mathbb{R}$ of degree 4 with $\mathbb{E}[f(\boldsymbol{X}_n)] = 0$ and $\|f(\boldsymbol{X}_n)\|_{L_\infty} \leqslant 1$, such that for every subset I of $[n]$ with $|I| \geqslant 8$ we have $\mathbb{P}\big(\big|\mathbb{E}[f(\boldsymbol{X}_n) \,|\, \mathcal{F}_I]\big| \geqslant 2^{-11}\big) \geqslant 2^{-11}$.*

Proof. We will define the random array $\boldsymbol{X}$ by providing an integral representation of its distribution. (Of course, this maneuver is expected by the Aldous–Hoover representation theorem [Ald81, Hoo79].) Specifically, set $V \coloneqq \{0,1\}$ and $A \coloneqq \{(0,0),(1,1)\} \subseteq V^2$; we view V as a discrete probability space equipped with the uniform probability measure. We also set $\Omega \coloneqq \{0,1\}^{\binom{\mathbb{N}}{2}}$ and we equip Ω with the product σ-algebra which we denote by Σ. Let $\mathbb{P}$ be the unique probability measure on (Ω,Σ) which satisfies, for every nonempty finite subset $\mathcal{F}$ of $\binom{\mathbb{N}}{2}$, that

$$\mathbb{P}\Big(\big\{(x_t)_{t\in\binom{\mathbb{N}}{2}}\in\Omega : x_s = 1 \text{ for all } s\in\mathcal{F}\big\}\Big) = \frac{1}{2}\Big(\frac{1}{2}\Big)^{|\mathcal{F}|} + \frac{1}{2}\int \prod_{s\in\mathcal{F}} \mathbf{1}_A(\boldsymbol{v}_s)\, d\boldsymbol{\mu}(\boldsymbol{v}) \tag{A.1}$$

where: (i) $\boldsymbol{\mu}$ denotes the product measure on $V^{\mathbb{N}}$ obtained by equipping each factor with the uniform probability measure on V, and (ii) for every $\boldsymbol{v}=(v_i)\in V^{\mathbb{N}}$ and every $s=\{i_1<i_2\}\in\binom{\mathbb{N}}{2}$ by $\boldsymbol{v}_s=(v_{i_1},v_{i_2})\in V^2$ we denote the restriction of $\boldsymbol{v}$ on the coordinates determined by s. Next, for every $s\in\binom{\mathbb{N}}{2}$ let $X_s\colon \Omega\to\{0,1\}$ denote the projection on the s-th coordinate, that is, $X_s\big((x_t)_{t\in\binom{\mathbb{N}}{2}}\big)=x_s$ for every $(x_t)_{t\in\binom{\mathbb{N}}{2}}\in\Omega$. The fact that the set A is symmetric implies that the random array $\boldsymbol{X}=\langle X_s : s\in\binom{\mathbb{N}}{2}\rangle$ is exchangeable; moreover, for every nonempty finite subset $\mathcal{F}$ of $\binom{\mathbb{N}}{2}$ we have

$$\mathbb{E}\Big[\prod_{s\in\mathcal{F}} X_s\Big] = \frac{1}{2}\Big(\frac{1}{2}\Big)^{|\mathcal{F}|} + \frac{1}{2}\int \prod_{s\in\mathcal{F}} \mathbf{1}_A(\boldsymbol{v}_s)\, d\boldsymbol{\mu}(\boldsymbol{v}). \tag{A.2}$$

Using (A.2), properties ($\mathcal{P}1$)–($\mathcal{P}4$) follow from a direct computation.

In order to verify property ($\mathcal{P}5$) we argue as in the proof of Proposition 2.8. Fix an integer $n\geqslant 8$. Let $\mathrm{Box}(2)$ be the 2-dimensional box of $\mathbb{N}$ defined in (3.2). We define $f\colon \mathbb{R}^{\binom{[n]}{2}}\to\mathbb{R}$ by setting for every $\boldsymbol{x}=(x_t)_{t\in\binom{[n]}{2}}\in\mathbb{R}^{\binom{[n]}{2}}$

$$\begin{aligned} f(\boldsymbol{x}) &\coloneqq \prod_{s\in\mathrm{Box}(2)} x_s - \mathbb{E}\Big[\prod_{s\in\mathrm{Box}(2)} X_s\Big] \\ &\overset{(3.2)}{=} x_{\{1,3\}}x_{\{1,4\}}x_{\{2,3\}}x_{\{2,4\}} - \mathbb{E}[X_{\{1,3\}}X_{\{1,4\}}X_{\{2,3\}}X_{\{2,4\}}]. \end{aligned} \tag{A.3}$$

It is clear that f is a translated multilinear polynomial of degree 4 with $\mathbb{E}[f(\boldsymbol{X}_n)]=0$ and $\|f(\boldsymbol{X}_n)\|_{L_\infty}\leqslant 1$. (Recall that $\boldsymbol{X}_n$ denotes the subarray of $\boldsymbol{X}$ determined by $[n]$.) Let I be an arbitrary subset of $[n]$ with $|I|\geqslant 8$. Since $|I|\geqslant 8$, there exists a 2-dimensional box B of $\mathbb{N}$ with $B\subseteq\binom{I}{2}$ and such that $\min(s)\geqslant 5$ for every $s\in B$. Set $C\coloneqq\bigcap_{s\in B}[X_s=1]$ and observe that $C\in\mathcal{F}_I$. Hence, by the exchangeability of $\boldsymbol{X}$, we have

$$\begin{aligned} \mathbb{E}\big[\mathbb{E}[f(\boldsymbol{X}_n)\,|\,\mathcal{F}_I]\,\mathbf{1}_C\big] &= \mathbb{E}[f(\boldsymbol{X}_n)\mathbf{1}_C] \\ &= \mathbb{E}\Big[\prod_{s\in\mathrm{Box}(2)\cup B} X_s\Big] - \mathbb{E}\Big[\bigcap_{s\in\mathrm{Box}(2)} X_s\Big]^2 \overset{(A.2)}{=} \frac{1}{2^{10}} \end{aligned} \tag{A.4}$$

which implies that $\mathbb{P}\big(\big|\mathbb{E}[f(\boldsymbol{X}_n)\,|\,\mathcal{F}_I]\big|\geqslant 2^{-11}\big)\geqslant 2^{-11}$. The proof is completed. □

A.3. **The higher-dimensional case.** The following result is the higher-dimensional analogue of Proposition A.1.

Proposition A.2. *Let $d \geqslant 3$ be an integer. Also let $\delta > 0$. Then there exists a boolean, exchangeable, d-dimensional random array $\boldsymbol{X} = \langle X_s : s \in \binom{\mathbb{N}}{d}\rangle$ on $\mathbb{N}$ with the following properties.*

($\mathcal{P}1$) *For every $s \in \binom{\mathbb{N}}{d}$ we have $\big|\mathbb{E}[X_s] - \frac{1}{2}\big| \leqslant \delta$.*
($\mathcal{P}2$) *For every distinct $s, t \in \binom{\mathbb{N}}{d}$ we have $\big|\mathbb{E}[X_s X_t] - \frac{1}{4}\big| \leqslant \delta$.*
($\mathcal{P}3$) *For every $(d-1)$-face F of $\mathbb{N}$ we have $\big|\mathbb{E}\big[\prod_{s\in F} X_s\big] - (\frac{1}{2})^{|F|}\big| \leqslant \delta$.*
($\mathcal{P}4$) *For every d-dimensional box B of $\mathbb{N}$ we have $\big|\mathbb{E}\big[\prod_{s\in B} X_s\big] - \frac{3}{2}(\frac{1}{2})^{|B|}\big| \leqslant \delta$.*
($\mathcal{P}5$) *Set $\vartheta \coloneqq 16^{-1}2^{-2^{d+1}}$. (Note that ϑ does not dependent on δ.) Let $n \geqslant 4d$ be an integer, and let $\boldsymbol{X}_n$ denote the subarray of $\boldsymbol{X}$ determined by $[n]$. Then there exists a translated multilinear polynomial $f\colon \mathbb{R}^{\binom{[n]}{d}} \to \mathbb{R}$ of degree 2^d with $\mathbb{E}[f(\boldsymbol{X}_n)] = 0$ and $\|f(\boldsymbol{X}_n)\|_{L_\infty} \leqslant 1$, such that for every subset I of $[n]$ with $|I| \geqslant 4d$ we have $\mathbb{P}\big(\big|\mathbb{E}[f(\boldsymbol{X}_n)\,|\,\mathcal{F}_I]\big| \geqslant \vartheta\big) \geqslant \vartheta$.*

Remark A.3. We point out that property ($\mathcal{P}3$) is rather strong. Indeed, arguing as in the proof of Theorem 3.2, it is not hard to show that if $\boldsymbol{X} = \langle X_s : s \in \binom{\mathbb{N}}{d}\rangle$ is any boolean, spreadable, d-dimensional random array on $\mathbb{N}$ which satisfies properties ($\mathcal{P}1$) and ($\mathcal{P}3$) of Proposition A.2, then for every d-dimensional box B of $\mathbb{N}$ and every nonempty subset G of B with $G \neq B$ we have

$$\Big|\mathbb{E}\Big[\prod_{s\in G} X_s\Big] - \Big(\frac{1}{2}\Big)^{|G|}\Big| = o_{\delta\to 0;d}(1). \tag{A.5}$$

Note that (A.5) barely misses to imply that $\boldsymbol{X}$ satisfies the box independence condition.

The examples provided by Proposition A.2 can be roughly described as semi-random, in the sense that they are part random and part deterministic. The following lemma provides us with the random component.

Lemma A.4. *Let $d \geqslant 3$ be an integer, and let $\varepsilon > 0$. Then there exist a nonempty finite set V and a symmetric[17] subset A of V^{d-1} such that, denoting by $A^{\complement}$ the complement of A, for every pair F, G of disjoint (possibly empty) subsets of $\binom{[2d]}{d-1}$ we have*

$$\Big|\int \Big(\prod_{s\in F} \mathbf{1}_A(\boldsymbol{v}_s)\Big)\cdot\Big(\prod_{s\in G} \mathbf{1}_{A^{\complement}}(\boldsymbol{v}_s)\Big)\, d\boldsymbol{\mu}(\boldsymbol{v}) - \Big(\frac{1}{2}\Big)^{|F|+|G|}\Big| \leqslant \varepsilon \tag{A.6}$$

where: (i) *$\boldsymbol{\mu}$ denotes the product measure on $V^{\mathbb{N}}$ obtained by equipping each factor with the uniform probability measure on V,* (ii) *for every $\boldsymbol{v} = (v_i) \in V^{\mathbb{N}}$ and every $s = \{i_1 < \cdots < i_{d-1}\} \in \binom{\mathbb{N}}{d}$ we have $\boldsymbol{v}_s = (v_{i_1}, \dots, v_{i_{d-1}}) \in V^{d-1}$, and* (iii) *in (A.6) we use the convention that the product of an empty family of functions is equal to the constant function 1.*

[17]That is, for every $(v_1, \dots, v_{d-1}) \in V^{d-1}$ and every permutation π of $[d-1]$ we have that $(v_1, \dots, v_{d-1}) \in A$ if and only if $(v_{\pi(1)}, \dots, v_{\pi(d-1)}) \in A$.

Lemma A.4 follows from a standard random selection and the Azuma–Hoeffding inequality; see, *e.g.*, [DTV21, Fact 3.3 and Lemma 3.4] for a proof.

We are ready to proceed to the proof of Proposition A.2.

Proof of Proposition A.2. Let V and A be the sets obtained by Lemma A.4 applied for

$$\varepsilon \coloneqq \min\big\{\delta 2^{-d2^d}, 8^{-1}2^{-(d+2)2^d}\big\} \tag{A.7}$$

and observe that V can be selected so that its cardinality is an even positive integer. We also note that in the rest of the proof we follow the notational conventions in Lemma A.4.

First, for every $i \in [d]$ we define $h_i^0, h_i^1 \colon V^d \to \{0,1\}$ by setting for every $\boldsymbol{v} \in V^d$

$$h_i^0(\boldsymbol{v}) \coloneqq \mathbf{1}_A(\boldsymbol{v}_{[d]\setminus\{i\}}) \ \text{ and } \ h_i^1(\boldsymbol{v}) \coloneqq \mathbf{1}_{A^\complement}(\boldsymbol{v}_{[d]\setminus\{i\}}). \tag{A.8}$$

Next, for every $\boldsymbol{x} \in \{0,1\}^d$ define $h_{\boldsymbol{x}} \colon V^d \to \{0,1\}$ by

$$h_{\boldsymbol{x}} \coloneqq \prod_{i=1}^{d} h_i^{\boldsymbol{x}(i)}. \tag{A.9}$$

Finally, set

$$\mathbb{A} \coloneqq \big\{\boldsymbol{x} \in \{0,1\}^d : \boldsymbol{x}(1) + \cdots + \boldsymbol{x}(d) \text{ is even}\big\} \tag{A.10}$$

and define $H \colon V^d \to \{0,1\}$ by

$$H \coloneqq \sum_{\boldsymbol{x}\in\mathbb{A}} h_{\boldsymbol{x}}. \tag{A.11}$$

For instance, if $d=3$, then

$$\begin{aligned} H(v_1,v_2,v_3) = {} & \mathbf{1}_A(v_1,v_2)\mathbf{1}_A(v_2,v_3)\mathbf{1}_A(v_1,v_3) + \mathbf{1}_{A^\complement}(v_1,v_2)\mathbf{1}_{A^\complement}(v_2,v_3)\mathbf{1}_A(v_1,v_3) + \\ & + \mathbf{1}_{A^\complement}(v_1,v_2)\mathbf{1}_A(v_2,v_3)\mathbf{1}_{A^\complement}(v_1,v_3) + \mathbf{1}_A(v_1,v_2)\mathbf{1}_{A^\complement}(v_2,v_3)\mathbf{1}_{A^\complement}(v_1,v_3). \end{aligned}$$

Note that the function H is symmetric[18]. In the following series of claims we isolate several properties of H which will be used in the proofs of properties ($\mathcal{P}1$)–($\mathcal{P}5$).

Claim A.5. *For every $k \in [d+1]$ set $t_k \coloneqq \{k,\dots,k+d-1\} \in \binom{\mathbb{N}}{d}$. Then we have*

$$\Big|\int H(\boldsymbol{v}_{t_1})\, d\boldsymbol{\mu}(\boldsymbol{v}) - \frac{1}{2}\Big| \leqslant 2^{d-1}\varepsilon. \tag{A.12}$$

Moreover, for every $k \in \{2,\dots,d+1\}$ we have

$$\Big|\int H(\boldsymbol{v}_{t_1})H(\boldsymbol{v}_{t_k})\, d\boldsymbol{\mu}(\boldsymbol{v}) - \frac{1}{4}\Big| \leqslant 2^{2d-3}\varepsilon. \tag{A.13}$$

[18] That is, we have $H(v_1,\dots,v_d) = H(v_{\pi(1)},\dots,v_{\pi(d)})$ for every $(v_1,\dots,v_d) \in V^d$ and every permutation π of $[d]$.

Proof of Claim A.5. First observe that (A.12) follows from (A.6), the fact that $|\mathbb{A}| = 2^{d-1}$ and the definition of H. Next, fix $k \in \{2, \dots, d+1\}$. Then, for every $\boldsymbol{v} \in V^{\mathbb{N}}$ we have

$$H(\boldsymbol{v}_{t_1})H(\boldsymbol{v}_{t_k}) = \sum_{\boldsymbol{x},\boldsymbol{y}\in\mathbb{A}} \Big(\prod_{i=1}^{d} h_i^{\boldsymbol{x}(i)}(\boldsymbol{v}_{t_1})\Big)\Big(\prod_{j=1}^{d} h_j^{\boldsymbol{y}(j)}(\boldsymbol{v}_{t_k})\Big). \tag{A.14}$$

Therefore, if $k > 2$, then (A.13) also follows from (A.6) and the fact that $|\mathbb{A}| = 2^{d-1}$. So assume that $k = 2$. By (A.14), for every $\boldsymbol{v} \in V^{\mathbb{N}}$ we have

$$H(\boldsymbol{v}_{t_1})H(\boldsymbol{v}_{t_2}) = \sum_{\boldsymbol{x},\boldsymbol{y}\in\mathbb{A}} \Big(\prod_{i=2}^{d} h_i^{\boldsymbol{x}(i)}(\boldsymbol{v}_{t_1})\Big)\Big(\prod_{j=1}^{d-1} h_j^{\boldsymbol{y}(j)}(\boldsymbol{v}_{t_2})\Big) h_1^{\boldsymbol{x}(1)}(\boldsymbol{v}_{t_1}) h_d^{\boldsymbol{y}(d)}(\boldsymbol{v}_{t_2}). \tag{A.15}$$

Notice that for every $\boldsymbol{v} \in V^{\mathbb{N}}$ we have $h^0_{\boldsymbol{v}_{t_1}} = h^0_{\boldsymbol{v}_{t_2}}$ and $h^1_{\boldsymbol{v}_{t_1}} = h^1_{\boldsymbol{v}_{t_2}}$. Thus, setting $\mathcal{W} \coloneqq \{(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{A}\times\mathbb{A} : \boldsymbol{x}(1) = \boldsymbol{y}(d)\}$, we see that $h_1^{\boldsymbol{x}(1)}(\boldsymbol{v}_{t_1}) h_d^{\boldsymbol{y}(d)}(\boldsymbol{v}_{t_2}) = 0$ for every $(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{A}\times\mathbb{A}\setminus\mathcal{W}$. Combining this information with (A.15), we obtain that

$$H(\boldsymbol{v}_{t_1})H(\boldsymbol{v}_{t_2}) = \sum_{(\boldsymbol{x},\boldsymbol{y})\in\mathcal{W}} \Big(\prod_{i=2}^{d} h_i^{\boldsymbol{x}(i)}(\boldsymbol{v}_{t_1})\Big)\Big(\prod_{j=1}^{d-1} h_j^{\boldsymbol{y}(j)}(\boldsymbol{v}_{t_2})\Big) h_1^{\boldsymbol{x}(1)}(\boldsymbol{v}_{t_1}) h_d^{\boldsymbol{y}(d)}(\boldsymbol{v}_{t_2}) \tag{A.16}$$

for every $\boldsymbol{v} \in V^{\mathbb{N}}$. On the other hand, by (A.6), for every $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{W}$ we have

$$\Big|\int \Big(\prod_{i=2}^{d} h_i^{\boldsymbol{x}(i)}(\boldsymbol{v}_{t_1})\Big)\Big(\prod_{j=1}^{d-1} h_j^{\boldsymbol{y}(j)}(\boldsymbol{v}_{t_2})\Big) h_1^{\boldsymbol{x}(1)}(\boldsymbol{v}_{t_1}) h_d^{\boldsymbol{y}(d)}(\boldsymbol{v}_{t_2})\, d\boldsymbol{\mu}(\boldsymbol{v}) - \Big(\frac{1}{2}\Big)^{2d-1}\Big| \leqslant \varepsilon. \tag{A.17}$$

Since $|\mathcal{W}| = 2^{2d-3}$, we conclude that (A.13) for $k = 2$ follows from (A.16) and (A.17). The proof of Claim A.5 is completed. □

Claim A.6. *Set* $C \coloneqq \big\{u \cup \{2d-1\} : u \in \mathrm{Box}(d-1)\big\}$ *where* $\mathrm{Box}(d-1)$ *is as in* (3.2). (*Notice that* $C \subseteq \binom{\mathbb{N}}{d}$.) *Then we have*

$$\Big|\int \prod_{s\in C} H(\boldsymbol{v}_s)\, d\boldsymbol{\mu}(\boldsymbol{v}) - \Big(\frac{1}{2}\Big)^{|C|}\Big| \leqslant (d+1)2^{d-2+(d-1)2^{d-2}}\varepsilon. \tag{A.18}$$

Proof of Claim A.6. We start by setting $j_i^0 \coloneqq 2i-1$ and $j_i^1 \coloneqq 2i$ for every $i \in [d-1]$. Next, for every $\boldsymbol{\epsilon} = (\epsilon_i)_{i=1}^{d-1} \in \{0,1\}^{d-1}$ set $s(\boldsymbol{\epsilon}) \coloneqq \big\{j_i^{\epsilon_i} : i \in [d-1]\big\} \cup \{2d-1\}$, and notice that $C = \big\{s(\boldsymbol{\epsilon}) : \boldsymbol{\epsilon} \in \{0,1\}^{d-1}\big\}$. Moreover, by (A.9) and (A.11), we have

$$\begin{aligned}\int \prod_{s\in C} H(\boldsymbol{v}_s)\, d\boldsymbol{\mu}(\boldsymbol{v}) &= \int \prod_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}} H(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})\, d\boldsymbol{\mu}(\boldsymbol{v}) = \\ &= \sum_{(\boldsymbol{x}_{\boldsymbol{\epsilon}})_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}}\in\mathbb{A}^{\{0,1\}^{d-1}}} \int \prod_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}} h_{\boldsymbol{x}_{\boldsymbol{\epsilon}}}(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})\, d\boldsymbol{\mu}(\boldsymbol{v}).\end{aligned} \tag{A.19}$$

We define a subset $\mathcal{R}$ of $\mathbb{A}^{\{0,1\}^{d-1}}$ by the rule

$$(\boldsymbol{x}_{\boldsymbol{\epsilon}})_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}} \in \mathcal{R} \Leftrightarrow \begin{array}{l} \text{for every } j\in[d-1], \text{ every } \boldsymbol{\epsilon}=(\epsilon_i)_{i=1}^{d-1}\in\{0,1\}^{d-1} \\ \text{and every } \boldsymbol{\epsilon}'=(\epsilon'_i)_{i=1}^{d-1}\in\{0,1\}^{d-1} \\ \text{with } \epsilon_i=\epsilon'_i \text{ for all } i\in[d-1]\setminus\{j\} \\ \text{we have } \boldsymbol{x}_{\boldsymbol{\epsilon}}(j)=\boldsymbol{x}_{\boldsymbol{\epsilon}'}(j). \end{array} \tag{A.20}$$

Observe that for every $j\in[d-1]$ and every $\boldsymbol{\epsilon}=(\epsilon_i)_{i=1}^{d-1}, \boldsymbol{\epsilon}'=(\epsilon'_i)_{i=1}^{d-1}\in\{0,1\}^{d-1}$ with

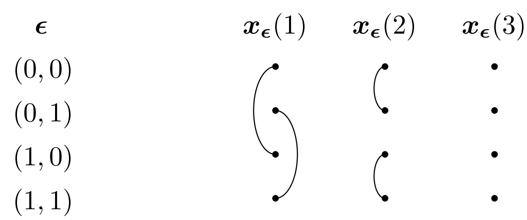

FIGURE 6. The structure of the set $\mathcal{R}$ for $d=3$. Connected dots imply equality of the corresponding coordinates.

$\epsilon_i=\epsilon'_i$ for all $i\in[d-1]\setminus\{j\}$ we have $h^0_j(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})=h^0_j(\boldsymbol{v}_{s(\boldsymbol{\epsilon}')})$ and $h^1_j(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})=h^1_j(\boldsymbol{v}_{s(\boldsymbol{\epsilon}')})$ for every $\boldsymbol{v}\in V^{\mathbb{N}}$ which, in turn, implies that $h^0_j(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})h^1_j(\boldsymbol{v}_{s(\boldsymbol{\epsilon}')})=0$. Consequently, $\prod_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}} h_{\boldsymbol{x}_{\boldsymbol{\epsilon}}}(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})=0$ for every $(\boldsymbol{x}_{\boldsymbol{\epsilon}})_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}}\in\mathbb{A}^{\{0,1\}^{d-1}}\setminus\mathcal{R}$ and every $\boldsymbol{v}\in V^{\mathbb{N}}$. Therefore, by (A.19), we obtain that

$$\int\prod_{s\in C} H(\boldsymbol{v}_s)\,d\boldsymbol{\mu}(\boldsymbol{v}) = \sum_{(\boldsymbol{x}_{\boldsymbol{\epsilon}})_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}}\in\mathcal{R}} \int \prod_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}} h_{\boldsymbol{x}_{\boldsymbol{\epsilon}}}(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})\,d\boldsymbol{\mu}(\boldsymbol{v}). \tag{A.21}$$

On the other hand, by (A.6), for every $(\boldsymbol{x}_{\boldsymbol{\epsilon}})_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}}\in\mathcal{R}$ we have

$$\begin{aligned} &\Big|\int \prod_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}} h_{\boldsymbol{x}_{\boldsymbol{\epsilon}}}(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})\,d\boldsymbol{\mu}(\boldsymbol{v}) - \Big(\frac{1}{2}\Big)^{(d+1)2^{d-2}}\Big| \\ &= \Big|\int \prod_{i=1}^{d}\prod_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}} h_i^{\boldsymbol{x}_{\boldsymbol{\epsilon}}(i)}(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})\,d\boldsymbol{\mu}(\boldsymbol{v}) - \Big(\frac{1}{2}\Big)^{(d+1)2^{d-2}}\Big| \\ &= \Big|\int \Big(\prod_{i=1}^{d-1}\prod_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}} h_i^{\boldsymbol{x}_{\boldsymbol{\epsilon}}(i)}(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})\Big)\Big(\prod_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}} h_d^{\boldsymbol{x}_{\boldsymbol{\epsilon}}(d)}(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})\Big)\,d\boldsymbol{\mu}(\boldsymbol{v}) - \Big(\frac{1}{2}\Big)^{(d+1)2^{d-2}}\Big| \\ &\leqslant (d{+}1)2^{d-2}\varepsilon. \end{aligned} \tag{A.22}$$

The estimate (A.18) follows from (A.21), (A.22), and the fact that $|\mathcal{R}|=2^{(d-1)2^{d-2}}$ and $|C|=2^{d-1}$. The proof of Claim A.6 is completed. □

Claim A.7. *Let* Box(d) *be as in* (3.2). *Then we have*

$$\Big|\int\prod_{s\in\mathrm{Box}(d)} H(\boldsymbol{v}_s)d\boldsymbol{\mu}(\boldsymbol{v}) - 2\Big(\frac{1}{2}\Big)^{|\mathrm{Box}(d)|}\Big| \leqslant d2^{d+(d-2)2^{d-1}}\varepsilon. \tag{A.23}$$

Proof of Claim A.7. As in the proof of Claim A.6, for every $i \in [d]$ set $j_i^0 \coloneqq 2i-1$ and $j_i^1 \coloneqq 2i$. Moreover, for every $\boldsymbol{\epsilon} = (\epsilon_i)_{i=1}^d \in \{0,1\}^d$ set $s(\boldsymbol{\epsilon}) \coloneqq \big\{ j_i^{\epsilon_i} : i \in [d]\big\}$, and observe that $\mathrm{Box}(d) = \big\{ s(\boldsymbol{\epsilon}) : \boldsymbol{\epsilon} \in \{0,1\}^d\big\}$. We define a subset $\mathcal{Q}$ of $\mathbb{A}^{\{0,1\}^d}$ by setting

$$(\text{A.24}) \qquad (\boldsymbol{x}_{\boldsymbol{\epsilon}})_{\boldsymbol{\epsilon}\in\{0,1\}^d} \in \mathcal{Q} \Leftrightarrow \begin{array}{l} \text{for every } j \in [d], \text{ every } \boldsymbol{\epsilon} = (\epsilon_i)_{i=1}^d \in \{0,1\}^d \\ \text{and every } \boldsymbol{\epsilon}' = (\epsilon'_i)_{i=1}^d \in \{0,1\}^d \\ \text{with } \epsilon_i = \epsilon'_i \text{ for all } i \in [d]\setminus\{j\} \\ \text{we have } \boldsymbol{x}_{\boldsymbol{\epsilon}}(j) = \boldsymbol{x}_{\boldsymbol{\epsilon}'}(j). \end{array}$$

By (A.9), (A.11), the definition of $\mathcal{Q}$ and arguing as in the proof of Claim A.6,

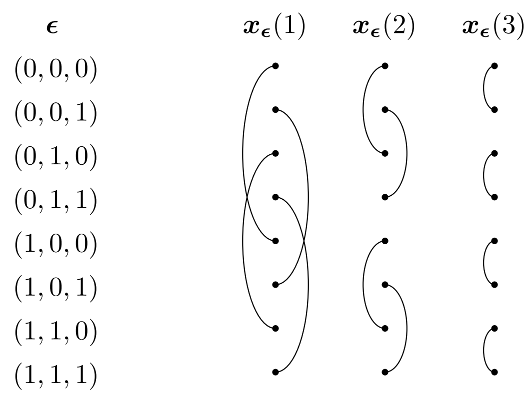

FIGURE 7. The structure of the set $\mathcal{Q}$ for $d=3$. As in Figure 6, connected dots imply equality of the corresponding coordinates.

$$(\text{A.25}) \qquad \begin{aligned} \int \prod_{s\in\mathrm{Box}(d)} H(\boldsymbol{v}_s)\, d\boldsymbol{\mu}(\boldsymbol{v}) &= \int \prod_{\boldsymbol{\epsilon}\in\{0,1\}^d} H(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})\, d\boldsymbol{\mu}(\boldsymbol{v}) \\ &= \sum_{(\boldsymbol{x}_{\boldsymbol{\epsilon}})_{\boldsymbol{\epsilon}\in\{0,1\}^d}\in\mathbb{A}^{\{0,1\}^d}} \int \prod_{\boldsymbol{\epsilon}\in\{0,1\}^d} h_{\boldsymbol{x}_{\boldsymbol{\epsilon}}}(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})\, d\boldsymbol{\mu}(\boldsymbol{v}) \\ &= \sum_{(\boldsymbol{x}_{\boldsymbol{\epsilon}})_{\boldsymbol{\epsilon}\in\{0,1\}^d}\in\mathcal{Q}} \int \prod_{\boldsymbol{\epsilon}\in\{0,1\}^d} h_{\boldsymbol{x}_{\boldsymbol{\epsilon}}}(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})\, d\boldsymbol{\mu}(\boldsymbol{v}). \end{aligned}$$

By (A.6), for every $(\boldsymbol{x}_{\boldsymbol{\epsilon}})_{\boldsymbol{\epsilon}\in\{0,1\}^{d-1}} \in \mathcal{Q}$ we have

$$(\text{A.26}) \qquad \Big| \int \prod_{\boldsymbol{\epsilon}\in\{0,1\}^d} h_{\boldsymbol{x}_{\boldsymbol{\epsilon}}}(\boldsymbol{v}_{s(\boldsymbol{\epsilon})})\, d\boldsymbol{\mu}(\boldsymbol{v}) - \Big(\frac{1}{2}\Big)^{d2^{d-1}} \Big| \leqslant d2^{d-1}\varepsilon.$$

Finally, note that $|Q| = 2^{(d-2)2^{d-1}+1}$. Using this information, (A.23) follows from (A.25), (A.26) and the fact that $|\mathrm{Box}(d)| = 2^d$. The proof of Claim A.7 is completed. $\square$

Claim A.8. *Let B be a d-dimensional box of $\mathbb{N}$ such that* $\min(s) \geqslant 2d+1$ *for every* $s \in B$. *Then we have*

$$\Big| \int \prod_{s\in \mathrm{Box}(d)\cup B} H(\boldsymbol{v}_s)\, d\boldsymbol{\mu}(\boldsymbol{v}) - 4\Big(\frac{1}{2}\Big)^{2|\mathrm{Box}(d)|}\Big| \leqslant d2^{d+1+(d-2)2^{d-1}}\varepsilon. \tag{A.27}$$

Proof of Claim A.8. It follows immediately by Claim A.7. □

After this preliminary discussion, we now enter into the main part of the proof. Let $\boldsymbol{X} = \langle X_s : s \in \binom{\mathbb{N}}{d}\rangle$ be a boolean, exchangeable, d-dimensional random array on $\mathbb{N}$ whose distribution satisfies

$$\mathbb{E}\Big[\prod_{s\in\mathcal{F}} X_s\Big] = \frac{1}{2}\Big(\frac{1}{2}\Big)^{|\mathcal{F}|} + \frac{1}{2}\int \prod_{s\in\mathcal{F}} H(\boldsymbol{v}_s)\, d\boldsymbol{\mu}(\boldsymbol{v}) \tag{A.28}$$

for every nonempty finite subset $\mathcal{F}$ of $\binom{\mathbb{N}}{d}$. (The existence of such a random array follows arguing precisely as in the proof of Proposition A.1.)

First, we will show that $\boldsymbol{X}$ satisfies properties ($\mathcal{P}1$) up to ($\mathcal{P}4$). For property ($\mathcal{P}1$), let $s \in \binom{\mathbb{N}}{d}$ be arbitrary and notice that, by the exchangeability of $\boldsymbol{X}$ and (A.28),

$$\mathbb{E}[X_s] = \frac{1}{4} + \frac{1}{2}\int H(\boldsymbol{v}_{t_1})\, d\boldsymbol{\mu}(\boldsymbol{v}) \tag{A.29}$$

where, as in Claim A.5, we have $t_1 = \{1,\dots,d\}$. By (A.12) and the choice of ε in (A.7), we obtain that $\big|\mathbb{E}[X_s] - \frac{1}{2}\big| \leqslant 2^{d-2}\varepsilon \leqslant \delta$. For property ($\mathcal{P}2$), let $s,t \in \binom{\mathbb{N}}{d}$ be distinct, and set $k := d - |s\cap t| + 1$. Since $\boldsymbol{X}$ is exchangeable, by (A.28), we have

$$\mathbb{E}[X_sX_t] = \frac{1}{8} + \frac{1}{2}\int H(\boldsymbol{v}_{t_1})H(\boldsymbol{v}_{t_k})\, d\boldsymbol{\mu}(\boldsymbol{v}) \tag{A.30}$$

where t_1 and t_k are as in Claim A.5. By (A.13), (A.30) and invoking again (A.7), we see that $\big|\mathbb{E}[X_sX_t] - \frac{1}{4}\big| \leqslant 2^{2d-4}\varepsilon \leqslant \delta$. For property ($\mathcal{P}3$), let F be a $(d-1)$-face of $\mathbb{N}$. By the exchangeability of $\boldsymbol{X}$, (A.28) and the choice of the set C in Claim A.6,

$$\mathbb{E}\Big[\prod_{s\in F} X_s\Big] = \frac{1}{2}\Big(\frac{1}{2}\Big)^{|F|} + \frac{1}{2}\int \prod_{s\in C} H(\boldsymbol{v}_s)\, d\boldsymbol{\mu}(\boldsymbol{v}) \tag{A.31}$$

which implies, by (A.18), that

$$\Big|\mathbb{E}\Big[\prod_{s\in F} X_s\Big] - \Big(\frac{1}{2}\Big)^{|F|}\Big| \leqslant (d+1)2^{d-3+(d-1)2^{d-2}}\varepsilon \overset{(A.7)}{\leqslant} \delta. \tag{A.32}$$

Lastly, for property ($\mathcal{P}4$), let B be a d-dimensional box of $\mathbb{N}$. Using once again the exchangeability of $\boldsymbol{X}$ and (A.28), we see that

$$\mathbb{E}\Big[\prod_{s\in B} X_s\Big] = \frac{1}{2}\Big(\frac{1}{2}\Big)^{|B|} + \frac{1}{2}\int \prod_{s\in \mathrm{Box}(d)} H(\boldsymbol{v}_s)\, d\boldsymbol{\mu}(\boldsymbol{v}) \tag{A.33}$$

and so, by (A.23),

$$\Big|\mathbb{E}\Big[\prod_{s\in B} X_s\Big] - \frac{3}{2}\Big(\frac{1}{2}\Big)^{|B|}\Big| \leqslant d2^{d-1+(d-2)2^{d-1}}\varepsilon \overset{(A.7)}{\leqslant} \delta. \tag{A.34}$$

Thus, it remains to verify property ($\mathcal{P}5$). As expected, we will argue as in Proposition 2.8. Specifically, fix an integer $n \geqslant 4d$, and define $f\colon \mathbb{R}^{\binom{[n]}{d}} \to \mathbb{R}$ by setting for every $\boldsymbol{x} = (x_t)_{t\in\binom{[n]}{d}} \in \mathbb{R}^{\binom{[n]}{d}}$

$$f(\boldsymbol{x}) \coloneqq \prod_{s\in \mathrm{Box}(d)} x_s - \mathbb{E}\Big[\prod_{s\in\mathrm{Box}(d)} X_s\Big]. \tag{A.35}$$

Clearly, f is a translated multilinear polynomial of degree 2^d and satisfies $\mathbb{E}[f(\boldsymbol{X}_n)] = 0$ and $\|f(\boldsymbol{X}_n)\|_{L_\infty} \leqslant 1$. On the other hand, if B is a d-dimensional box on $\mathbb{N}$ such that $\min(s) \geqslant 2d+1$ for every $s\in B$, then

$$\mathbb{E}\Big[\prod_{s\in\mathrm{Box}(d)\cup B} X_s\Big] \overset{\text{(A.28)}}{=} \frac{1}{2}\Big(\frac{1}{2}\Big)^{2^{d+1}} + \frac{1}{2}\int \prod_{s\in\mathrm{Box}(d)\cup B} H(\boldsymbol{v}_s)\, d\boldsymbol{\mu}(\boldsymbol{v}) \tag{A.36}$$

and so, by (A.27),

$$\Big|\mathbb{E}\Big[\prod_{s\in\mathrm{Box}(d)\cup B} X_s\Big] - \frac{5}{2}\Big(\frac{1}{2}\Big)^{2^{d+1}}\Big| \leqslant d2^{d+(d-2)2^{d-1}}\varepsilon. \tag{A.37}$$

Using this estimate, property ($\mathcal{P}4$) and arguing as in the proof of Proposition A.1, it is easy to verify that the function f satisfies property ($\mathcal{P}5$). The proof of Proposition A.2 is completed. □

References

[AdWo15] R. Adamczak and P. Wolff, *Concentration inequalities for non-Lipschitz functions with bounded derivatives of higher order*, Probab. Theory Related Fields 162 (2015), 531–586.

[Ald81] D. J. Aldous, *Representations for partially exchangeable arrays of random variables*, J. Multivariate Anal. 11 (1981), 581–597.

[Au08] T. Austin, *On exchangeable random variables and the statistics of large graphs and hypergraphs*, Probability Surveys 5 (2008), 80–145.

[Au11] T. Austin, *Deducing the density Hales–Jewett theorem from an infinitary removal lemma*, J. Theor. Probab. 24 (2011), 615–633.

[Au13] T. Austin, *Exchangeable random arrays*, preprint (2013), available at https://www.math.ucla.edu/~tim/ExchnotesforIISc.pdf.

[Ber96] V. Bergelson, *Ergodic Ramsey theory—an update*, in "Ergodic Theory of $\mathbb{Z}^d$-Actions", London Mathematical Society Lecture Note Series, Vol. 228, Cambridge University Press, 1996, 1–61.

[Bob04] S. Bobkov, *Concentration of normalized sums and a central limit theorem for noncorrelated random variables*, Ann. Probab. 32 (2004), 2884–2908.

[BLM13] S. Boucheron, G. Lugosi and P. Massart, *Concentration Inequalities. A Nonasymptotic Theory of Independence*, Oxford University Press, 2013.

[Ch06] S. Chatterjee, *A generalization of the Lindeberg principle*, Ann. Probab. 34 (2006), 2061–2076.

[CGW88] F. R. K. Chung, R. L. Graham and R. M. Wilson, *Quasi-random graphs*, Proc. Natl. Acad. Sci. USA 85 (1988), 969–970.

[CGW89] F. R. K. Chung, R. L. Graham and R. M. Wilson, *Quasi-random graphs*, Combinatorica 9 (1989), 345–362.

[DF80] P. Diaconis and D. Freedman, *Finite exchangeable sequences*, Ann. Probab. 8 (19180), 745–764.

[DK16] P. Dodos and V. Kanellopoulos, *Ramsey Theory for Product Spaces*, Mathematical Surveys and Monographs, Vol. 212, American Mathematical Society, 2016.

[DKK16] P. Dodos, V. Kanellopoulos and Th. Karageorgos, *Szemerédi's regularity lemma via martingales*, Electron. J. Combin. 23 (2016), Research Paper P3.11, 1–24.

[DKT16] P. Dodos, V. Kanellopoulos and K. Tyros, *A concentration inequality for product spaces*, J. Funct. Anal. 270 (2016), 609–620.

[DT21] P. Dodos and K. Tyros, *A structure theorem for stochastic processes indexed by the discrete hypercube*, Forum Math. Sigma 9 (2021), Paper No. e8, 1–30.

[DTV21] P. Dodos, K. Tyros and P. Valettas, *Decompositions of finite high-dimensional random arrays*, preprint (2021), available at https://arxiv.org/abs/2102.11102.

[ER52] P. Erdős and R. Rado, *Combinatorial theorems on classifications of subsets of a given set*, Proc. London Math. Soc. 2 (1952), 417–439.

[FT85] D. H. Fremlin and M. Talagrand, *Subgraphs of random graphs*, Trans. Amer. Math. Soc. 291 (1985), 551–582.

[FK91] H. Furstenberg and Y. Katznelson, *A density version of the Hales–Jewett theorem*, J. Anal. Math. 57 (1991), 64–117.

[GSS19] F. Götze, H. Sambale and A. Sinulis, *Concentration inequalities for polynomials in α-subexponential random variables*, preprint (2019), available at https://arxiv.org/abs/1903.05964.

[Go07] W. T. Gowers, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Ann. Math. 166 (2007), 897–946.

[Go09] W. T. Gowers, *The first unknown case of polynomial DHJ*, blog post (2009), available at https://gowers.wordpress.com/2009/11/14/the-first-unknown-case-of-polynomial-dhj/.

[GT10] B. Green and T. Tao, *Linear equations in primes*, Ann. Math. 171 (2010), 1753–1850.

[GG71] V. I. Gurarii and N. I. Gurarii, *On bases in uniformly convex and uniformly smooth Banach spaces*, Izv. Akad. Nauk SSSR Ser. Mat. 35 (1971), 210–215.

[Hoo79] D. N. Hoover, *Relations on probability spaces and arrays of random variables*, preprint (1979), available at https://www.stat.berkeley.edu/~aldous/Research/hoover.pdf.

[Ja72] R. C. James, *Super-reflexive spaces with bases*, Pac. J. Math. 41 (1972), 409–417.

[Kal92] O. Kallenberg, *Symmetries on random arrays and set-indexed processes*, J. Theor. Probab. 5 (1992), 727–765.

[Kal05] O. Kallenberg, *Probabilistic Symmetries and Invariance Principles*, Probability and its Applications (New York), Springer, 2005.

[La06] R. Latala, *Estimates of moments and tails of Gaussian chaoses*, Ann. Probab. 34 (2006), 2315–2331.

[Le01] M. Ledoux, *The Concentration of Measure Phenomenon*, Mathematical Surveys and Monographs, Vol. 89, American Mathematical Society, 2001.

[Lov12] L. Lovász, *Large networks and graph limits*, Colloquium Publications, Vol. 60, American Mathematical Society, 2012.

[MS75] W. G. Moginley and R. Sibson, *Dissociated random variables*, Math. Proc. Cambridge Philos. Soc. 77 (1975), 185–187.

[Pi11] G. Pisier, *Martingales in Banach Spaces* (*in connection with Type and Cotype*), preprint (2011), available at http://www.math.jussieu.fr/~pisier/ihp-pisier.pdf.

[Pi16] G. Pisier, *Martingales in Banach Spaces*, Cambridge Studies in Advanced Mathematics, Vol. 155, Cambridge University Press, 2016.

[Ra30] F. P. Ramsey, *On a problem of formal logic*, Proc. London Math. Soc. 30 (1930), 264–286.

[RX16] E. Ricard and Q. Xu, *A noncommutative martingale convexity inequality*, Ann. Probab. 44 (2016), 867–882.

[Rő15] V. Rődl, *Quasi-randomness and the regularity method in hypergraphs*, in "Proceedings of the International Congress of Mathematicians" Vol. I, 571–599, 2015.

[Tao06] T. Tao, *Szemerédi's regularity lemma revisited*, Contrib. Discrete Math. 1 (2006), 8–27.

[Tao08] T. Tao, *Structure and Randomness: Pages from Year One of a Mathematical Blog*, American Mathematical Society, 2007.

[Tho87] A. Thomason, *Pseudo-random graphs*, Ann. Discrete Math. 33 (1987), 307–331.

[V19] R. Vershynin, *Concentration inequalities for random tensors*, Bernoulli 26 (2020), 3139–3162.

Department of Mathematics, University of Athens, Panepistimiopolis 157 84, Athens, Greece
Email address: pdodos@math.uoa.gr

Department of Mathematics, University of Athens, Panepistimiopolis 157 84, Athens, Greece
Email address: ktyros@math.uoa.gr

Mathematics Department, University of Missouri, Columbia, MO, 65211
Email address: valettasp@missouri.edu