L_p **REGULAR SPARSE HYPERGRAPHS**

PANDELIS DODOS, VASSILIS KANELLOPOULOS AND THODORIS KARAGEORGOS

ABSTRACT. We study sparse hypergraphs which satisfy a mild pseudorandomness condition known as L_p *regularity*. We prove appropriate regularity and counting lemmas, and we extend the relative removal lemma of Tao in this setting. This answers a question from [3].

CONTENTS

1. INTRODUCTION

1.1. In this paper we study a family of sparse weighted uniform hypergraphs which satisfy a mild pseudorandomness condition. Our primary objective is to prove appropriate regularity and counting lemmas, and to extend the relative removal lemma for this class of hypergraphs. Beside their intrinsic interest, these types of results have found significant applications in number theory and Ramsey theory. To a large extend, the present paper was motivated by these applications.

1.2. To put our discussion in a proper context it is useful to recall one of the most well-known pseudorandomness conditions for graphs, introduced in [18, 19]. Specifically, let $G = (V, E)$ be a finite graph and let $p := |E|/\binom{|V|}{2}$ denote the edge density of G; the reader should have in mind that we are interested in the case where G is sparse, that is, in the regime $p = o(|V|^2)$. Also let $D \geqslant 1$ and $0 < \gamma \leqslant 1$, and recall that the graph G is said to be (D, γ)-*bounded* provided that

for every pair X, Y of disjoint subsets of V with $|X| \geqslant \gamma|V|$ and $|Y| \geqslant \gamma|V|$, we have $|E \cap (X \times Y)| \leqslant Dp|X| |Y|$. This natural condition expresses the fact that the graph G has "no large dense spots", and is satisfied by several models of sparse random graphs (see, e.g., [2]).

1.3. The notion of pseudorandomness which we will consider in this paper is an extension of the aforementioned boundedness hypothesis. It is relevant also in the continuous setting—e.g., in the context of graph limits [20]—and as such we will be working with the following measure theoretic structures. As usual, for every positive integer n we set $[n] := \{1, \ldots, n\}$.

Definition 1.1 ([33]). *A hypergraph system is a triple*

$$(1.1) \qquad \mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$$

where n is a positive integer, $\langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle$ is a finite sequence of probability spaces and $\mathcal{H}$ is a hypergraph on $[n]$. If $\mathcal{H}$ is r-uniform, then $\mathscr{H}$ will be called an r-uniform hypergraph system. On the other hand, if $\mu_i(A) \leqslant \eta$ for every $i \in [n]$ and every atom[1] A of (X_i, Σ_i, μ_i), then $\mathscr{H}$ will be called η-nonatomic.

Observe that a hypergraph system which consists of discrete probability spaces is just a vertex-weighted n-partite hypergraph. Also note that if μ is the uniform probability measure on a nonempty finite set X, then every atom of (X, μ) has probability $|X|^{-1}$. Therefore, η-nonatomicity is just an asymptotic condition.

At this point we need to recall some basic definitions. Given a hypergraph system $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$ by $(\boldsymbol{X}, \boldsymbol{\Sigma}, \boldsymbol{\mu})$ we denote the product of the spaces $\langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle$. More generally, for every nonempty $e \subseteq [n]$ let $(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ be the product of the spaces $\langle (X_i, \Sigma_i, \mu_i) : i \in e \rangle$ and observe that the σ-algebra $\boldsymbol{\Sigma}_e$ can be "lifted" to $\boldsymbol{X}$ by setting

$$(1.2) \qquad \mathcal{B}_e = \left\{ \pi_e^{-1}(\mathbf{A}) : \mathbf{A} \in \boldsymbol{\Sigma}_e \right\}$$

where $\pi_e \colon \boldsymbol{X} \to \boldsymbol{X}_e$ is the natural projection. If, in addition, $|e| \geqslant 2$, then let $\partial e = \{e' \subseteq e : |e'| = |e| - 1\}$ and define

$$(1.3) \qquad \mathcal{S}_{\partial e} := \left\{ \bigcap_{e' \in \partial e} A_{e'} : A_{e'} \in \mathcal{B}_{e'} \text{ for every } e' \in \partial e \right\} \subseteq \mathcal{B}_e.$$

The family $\mathcal{S}_{\partial e}$ is very easy to grasp if $n = 2$ and $e = \{1, 2\}$. It consists of all measurable rectangles of $X_1 \times X_2$, that is,

$$\mathcal{S}_{\partial e} = \{A \times B : A \in \Sigma_1 \text{ and } B \in \Sigma_2\}.$$

On the other hand, if $n = 3$ and $e = \{1, 2, 3\}$, then observe that

$$\mathcal{S}_{\partial e} = \{A \cap B \cap C : A \in \mathcal{B}_{\{1,2\}}, B \in \mathcal{B}_{\{2,3\}} \text{ and } C \in \mathcal{B}_{\{1,3\}}\}$$

---

[1] Recall that an *atom* of a probability space is a measurable event with positive probability which contains no measurable event with smaller positive probability.

where $\mathcal{B}_{\{1,2\}} = \{\mathbf{A} \times X_3 : \mathbf{A} \subseteq X_1 \times X_2 \text{ is measurable}\}$, and similarly for $\mathcal{B}_{\{2,3\}}$ and $\mathcal{B}_{\{1,3\}}$. In general, notice that every member of $\mathcal{S}_{\partial e}$ is the intersection of events which depend on fewer coordinates, and so it is useful to view the elements of $\mathcal{S}_{\partial e}$ as "lower-complexity" events. Much of the interest in these families of sets stems from the fact that they are "characteristic" for a number of structural properties of hypergraphs (see, e.g., [8, 36]).

1.4. We are now in a position to introduce the class of weighted hypergraphs which we will consider in this paper. It was recently defined[2] in [3].

Definition 1.2. *Let $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$ be a hypergraph system. Also let $C, \eta > 0$ and $1 \leqslant p \leqslant \infty$, and let $e \in \mathcal{H}$ with $|e| \geqslant 2$. A random variable $f \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ is said to be (C, η, p)-regular (or simply L_p regular if C and η are understood) provided that for every partition $\mathcal{P}$ of $\boldsymbol{X}$ with $\mathcal{P} \subseteq \mathcal{S}_{\partial e}$ and $\boldsymbol{\mu}(A) \geqslant \eta$ for every $A \in \mathcal{P}$ we have*

$$(1.4) \qquad \qquad \|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} \leqslant C$$

where $\mathcal{A}_{\mathcal{P}}$ is the σ-algebra on $\boldsymbol{X}$ generated by $\mathcal{P}$ and $\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})$ is the conditional expectation of f with respect to $\mathcal{A}_{\mathcal{P}}$.

The main point in Definition 1.2 is that, even though we make no assumption on the existence of moments, an L_p regular random variable *behaves* like a function in L_p as long as we project it on sufficiently "nice" σ-algebras of $\boldsymbol{X}$.

Notice that L_p regularity becomes weaker as p becomes smaller. In particular, the case "$p = 1$" is essentially of no interest since every integrable random variable is L_1 regular. On the other hand, in the context of graphs, L_∞ regularity reduces to the boundedness hypothesis that we mentioned above. Indeed, it is not hard to see that a bipartite graph $G = (V_1, V_2, E)$ with edge density p is (D, γ)-bounded for some D, γ if and only if the random variable $\mathbf{1}_E/p$ is L_∞ regular. (Here, we view V_1 and V_2 as discrete probability spaces equipped with the uniform probability measures.) For weighted hypergraphs, however, L_∞ regularity is a more subtle property. It is implied by the pseudorandomness conditions appearing in the work of Green and Tao [15, 16], though closer to the spirit of this paper is the work of Tao in [34]. We will comment further on these connections in Subsection 2.2.

Between the above extremes there is a large class of sparse weighted hypergraphs (namely, those which are L_p regular for some $1 < p < \infty$) which are, as we shall see, particularly well-behaved.

2. STATEMENT OF THE MAIN RESULTS

2.1. **Regularity lemma.** We first recall some pieces of notation and terminology. We say that a function $F \colon \mathbb{N} \to \mathbb{R}$ is a *growth function* provided that: (i) F is

---

[2]Actually, in [3] only L_p regular graphs were considered, but it is straightforward to extend the definition to hypergraphs.

increasing, and (ii) $F(n) \geqslant n + 1$ for every $n \in \mathbb{N}$. Moreover, for every hypergraph system $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$, every $e \in \mathcal{H}$ with $|e| \geqslant 2$ and every $f \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ let

$$(2.1) \qquad \|f\|_{\mathcal{S}_{\partial e}} = \sup \left\{ \Big| \int_A f \, d\boldsymbol{\mu} \Big| : A \in \mathcal{S}_{\partial e} \right\}$$

where $\mathcal{S}_{\partial e}$ is as in (1.3). The quantity $\|f\|_{\mathcal{S}_{\partial e}}$ is a measure of "uniformity" of f and it has been extensively used in extremal combinatorics (see, e.g., [12, 13, 20, 33]). It is sometimes referred to in the literature as the *cut norm* of f.

The following decomposition of L_p regular random variables is the first main result of this paper. As in Definition 1.2, for every partition $\mathcal{P}$ of a nonempty set X by $\mathcal{A}_{\mathcal{P}}$ we denote the σ-algebra on X generated by $\mathcal{P}$.

Theorem 2.1. *Let $n, r \in \mathbb{N}$ with $n \geqslant r \geqslant 2$, and let $C > 0$ and $1 < p \leqslant \infty$. Also let $F \colon \mathbb{N} \to \mathbb{R}$ be a growth function and $0 < \sigma \leqslant 1$. Then there exists a positive integer $\mathrm{Reg} = \mathrm{Reg}(n, r, C, p, F, \sigma)$ such that, setting $\eta = 1/\mathrm{Reg}$, the following holds. Let $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$ be an η-nonatomic, r-uniform hypergraph system. For every $e \in \mathcal{H}$ let $f_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ be nonnegative and (C, η, p)-regular. Then there exist*

- (a) *a positive integer M with $M \leqslant \mathrm{Reg}$,*
- (b) *for every $e \in \mathcal{H}$ a partition $\mathcal{P}_e$ of $\boldsymbol{X}$ with $\mathcal{P}_e \subseteq \mathcal{S}_{\partial e}$ and $\boldsymbol{\mu}(A) \geqslant 1/M$ for every $A \in \mathcal{P}_e$, and*
- (c) *for every $e \in \mathcal{H}$ a refinement $\mathcal{Q}_e$ of $\mathcal{P}_e$ with $\mathcal{Q}_e \subseteq \mathcal{S}_{\partial e}$ and $\boldsymbol{\mu}(A) \geqslant \eta$ for every $A \in \mathcal{Q}_e$,*

such that for every $e \in \mathcal{H}$, writing $f_e = f_{\mathrm{str}}^e + f_{\mathrm{err}}^e + f_{\mathrm{unf}}^e$ with

$$(2.2) \quad f_{\mathrm{str}}^e = \mathbb{E}(f_e \,|\, \mathcal{A}_{\mathcal{P}_e}), \;\; f_{\mathrm{err}}^e = \mathbb{E}(f_e \,|\, \mathcal{A}_{\mathcal{Q}_e}) - \mathbb{E}(f_e \,|\, \mathcal{A}_{\mathcal{P}_e}), \;\; f_{\mathrm{unf}}^e = f_e - \mathbb{E}(f_e \,|\, \mathcal{A}_{\mathcal{Q}_e}),$$

we have the estimates

$$(2.3) \qquad \|f_{\mathrm{str}}^e\|_{L_p} \leqslant C, \;\; \|f_{\mathrm{err}}^e\|_{L_{p^\dagger}} \leqslant \sigma \;\; and \;\; \|f_{\mathrm{unf}}^e\|_{\mathcal{S}_{\partial e}} \leqslant \frac{1}{F(M)}$$

where $p^\dagger = \min\{2, p\}$.

Note that, unless $p = \infty$, the structured part of the above decomposition (namely, the function f_{str}^e) is not uniformly bounded. This is an intrinsic feature of L_p regular hypergraphs, and it is an important difference between Theorem 2.1 and several related results (see, e.g., [2, 5, 6, 14, 15, 18, 23, 35, 37]). Observe, however, that, by part (b) and (2.3), one has a very good control on the correlation between f_{str}^e and $f_{\mathrm{unf}}^{e'}$ for every $e, e' \in \mathcal{H}$. Hence, by appropriately selecting the growth function F, we can force the function f_{str}^e to behave like a bounded function for many practical purposes.

The proof of Theorem 2.1 is given in Section 5 and relies in an essential way on a uniform convexity estimate for martingale difference sequences in L_p spaces (Proposition 3.1 in the main text), and on a Hölder-type inequality for L_p regular

random variables which is presented in Section 4. The argument is technically demanding mainly because, on the one hand, we aim to the strong norm estimates in (2.3) and, on the other hand, we need to ensure that the cells of each partition $\mathcal{P}_e$ are fairly large. The reader is also referred to [9] where a similar method is used in a simpler (but closely related) context, as well as to [4] where the algorithmic aspects of the proof of Theorem 2.1 are analysed and applications are given for sparse instances of NP-hard problems.

2.2. **Relative counting and removal lemmas.** The second part of this paper deals with the existence of counting and removal lemmas for L_p regular hypergraphs. Firstly, it is necessary to note that L_p regularity alone is not strong enough to yield results of this type, even if we restrict our attention to L_∞ regular graphs. It is in fact a rather delicate question to find natural conditions which imply sparse versions of the counting and removal lemmas[3]. In this direction, a powerful transference method was introduced by Green and Tao [15, 16] which can be applied to relatively dense subsets of several sufficiently pseudorandom combinatorial structures. For hypergraph systems, the method was developed by Tao in [34] who obtained a "relative removal lemma" which essentially covers the case of L_∞ regular hypergraphs. Our second main result extends the relative removal lemma in the whole range of admissible p's.

Theorem 2.2. *Let $n, r \in \mathbb{N}$ with $n \geqslant r \geqslant 2$, and let $C \geqslant 1$ and $1 < p \leqslant \infty$. Then for every $0 < \varepsilon \leqslant 1$ there exist two strictly positive constants $\eta = \eta(n, r, C, p, \varepsilon)$ and $\delta = \delta(n, r, C, p, \varepsilon)$, and a positive integer $k = k(n, r, C, p, \varepsilon)$ with the following property. Let $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$ be an η-nonatomic, r-uniform hypergraph system and let $\langle \nu_e : e \in \mathcal{H} \rangle$ be a (C, η, p)-pseudorandom family. For every $e \in \mathcal{H}$ let $f_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ with $0 \leqslant f_e \leqslant \nu_e$ such that*

$$(2.4) \qquad \int \prod_{e \in \mathcal{H}} f_e \, d\boldsymbol{\mu} \leqslant \delta.$$

Then for every $e \in \mathcal{H}$ there exists $F_e \in \mathcal{B}_e$ with

$$(2.5) \qquad \int_{\boldsymbol{X} \setminus F_e} f_e \, d\boldsymbol{\mu} \leqslant \varepsilon \quad \text{and} \quad \bigcap_{e \in \mathcal{H}} F_e = \emptyset.$$

Moreover, there exists a collection $\langle \mathcal{P}_{e'} : e' \subseteq e \text{ for some } e \in \mathcal{H} \rangle$ of partitions of $\boldsymbol{X}$ such that: (i) $\mathcal{P}_{e'} \subseteq \mathcal{B}_{e'}$ and $|\mathcal{P}_{e'}| \leqslant k$ for every $e' \subseteq e \in \mathcal{H}$, and (ii) for every $e \in \mathcal{H}$ the set F_e belongs to the algebra generated by the family $\bigcup_{e' \subsetneq e} \mathcal{P}_{e'}$.

Here, a (C, η, p)-*pseudorandom family* is a collection $\langle \nu_e : e \in \mathcal{H} \rangle$ of nonnegative random variables each of which effectively resembles a function $\psi_e \in L_p(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\Sigma})$ with $\|\psi_e\|_{L_p} \leqslant C$ (in particular, each ν_e is L_p regular). The precise definition is given in Section 6, along with some basic properties. If $\psi_e = 1$ for every $e \in \mathcal{H}$,

---

[3]Recently, there was progress in this direction for sparse random graphs; see [28] and the references therein.

then this definition coincides[4] with the "linear forms condition" introduced in [34], and so Theorem 2.2 yields the stronger version of the relative removal lemma obtained recently in [6]. More important, in the general case we do not demand that the functions $\langle \psi_e : e \in \mathcal{H} \rangle$ are independent and, in fact, there are natural examples of pseudorandom families consisting of mutually correlated random variables (see [10]).

As we have already indicated, the proof of Theorem 2.2 follows the general philosophy of the transfer method of Green and Tao. It is based on Theorem 2.1 as well as on a relative counting lemma for pseudorandom families (Theorem 7.1 in the main text). This relative counting lemma in turn relies on a new decomposition method which we shall discuss in detail in Subsection 7.1. The reader is also referred to [7] for further applications of this method in an arithmetic context. We also remark that the problem of obtaining a counting lemma for L_p regular hypergraphs was asked in [3, Section 8].

2.3. **Consequences.** We conclude our discussion in this section by noting that Theorem 2.2 yields a new Szemerédi-type theorem for sparse pseudorandom subsets of finite additive groups. We present this result in Subsection 8.3. Readers mainly interested in the integers will also find in Subsection 8.3 a purely arithmetic version of Theorem 2.2 (Theorem 8.3 in the main text).

3. Background material

Our general notation and terminology is standard. By $\mathbb{N} = \{0, 1, \dots\}$ we denote the set of all natural numbers.

3.1. **Martingale difference sequences.** Let (X, Σ, μ) be a probability space and recall that a finite sequence $(d_i)_{i=0}^n$ of integrable real-valued random variables on (X, Σ, μ) is said to be a *martingale difference sequence* if there exists a martingale $(f_i)_{i=0}^n$ such that $d_0 = f_0$ and $d_i = f_i - f_{i-1}$ if $n \geqslant 1$ and $i \in [n]$.

It is clear that every square-integrable martingale difference sequence $(d_i)_{i=0}^n$ is orthogonal in L_2 and, therefore,

$$(3.1) \qquad \Big(\sum_{i=0}^n \|d_i\|_{L_2}^2 \Big)^{1/2} = \Big\| \sum_{i=0}^n d_i \Big\|_{L_2}.$$

We will need the following extension of this basic fact.

Proposition 3.1. *Let (X, Σ, μ) be a probability space and $1 < p \leqslant 2$. Then for every martingale difference sequence $(d_i)_{i=0}^n$ in $L_p(X, \Sigma, \mu)$ we have*

$$(3.2) \qquad \Big(\sum_{i=0}^n \|d_i\|_{L_p}^2 \Big)^{1/2} \leqslant \Big(\frac{1}{p-1} \Big)^{1/2} \Big\| \sum_{i=0}^n d_i \Big\|_{L_p}.$$

---

[4]Actually, it is weaker than the linear forms condition, and formally coincides with the conditions which are discussed in [6].

It is a remarkable fact that the constant $(p-1)^{-1/2}$ appearing in the right-hand side of (3.2) is best possible. This sharp estimate was recently proved by Ricard and Xu [24] who deduced it from the following uniform convexity inequality for L_p spaces (see [1] or [22, Lemma 4.32]).

Proposition 3.2. *Let (X, Σ, μ) be an arbitrary measure space and $1 < p \leqslant 2$. Then for every $x, y \in L_p(X, \Sigma, \mu)$ we have*

$$(3.3) \qquad \|x\|_{L_p}^2 + (p-1)\|y\|_{L_p}^2 \leqslant \frac{\|x+y\|_{L_p}^2 + \|x-y\|_{L_p}^2}{2}.$$

3.2. **Probability spaces with small atoms.** Recall that a classical theorem of Sierpiński asserts that for every nonatomic finite measure space (X, Σ, μ) and every $0 \leqslant c \leqslant \mu(X)$ there exists $C \in \Sigma$ with $\mu(C) = c$. We will need the following approximate version of this result.

Proposition 3.3. *Let $0 < \eta < 1$ and let (X, Σ, μ) be a probability space such that $\mu(A) \leqslant \eta$ for every atom A of (X, Σ, μ). Also let $B \in \Sigma$ with $\mu(B) > \eta$, and let $\eta \leqslant c < \mu(B)$. Then there exists $C \in \Sigma$ with $C \subseteq B$ and $c \leqslant \mu(C) < c + \eta$.*

Proposition 3.3 is straightforward for discrete probability spaces. The general case follows from the aforementioned result of Sierpiński and a transfinite exhaustion argument. We give a proof for the convenience of the reader.

Proof of Proposition 3.3. Assume not, that is,

(H) for every $C \in \Sigma$ with $C \subseteq B$ either $\mu(C) < c$ or $\mu(C) \geqslant c + \eta$.

We will use hypothesis (H) to construct an increasing family $\langle Z_\alpha : \alpha < \omega_1 \rangle$ of measurable events of (X, Σ, μ) such that $Z_\alpha \subseteq B$, $\mu(Z_\alpha) < c$ and $\mu(Z_{\alpha+1} \setminus Z_\alpha) > 0$ for every $\alpha < \omega_1$. Clearly, this leads to a contradiction.

We begin by setting $Z_0 = \emptyset$. If α is a limit ordinal, then we set $Z_\alpha = \bigcup_{\beta < \alpha} Z_\beta$; notice that $\mu(Z_\alpha) \leqslant c$ and so, by hypothesis (H), we have $\mu(Z_\alpha) < c$. Finally, let $\alpha = \beta + 1$ be a successor ordinal. By Sierpiński's result and hypothesis (H), the set $B \setminus Z_\beta$ must contain an atom A of (X, Σ, μ). We set $Z_\alpha = Z_\beta \cup A$ and we observe that $\mu(Z_{\beta+1} \setminus Z_\beta) = \mu(A) > 0$. Also notice that $\mu(Z_\alpha) < c + \eta$. Thus, invoking hypothesis (H) once again, we conclude that $\mu(Z_\alpha) < c$ and the proof of Proposition 3.3 is completed. $\qquad\square$

3.3. **Removal lemma for hypergraph systems.** We will need the following version of the removal lemma for hypergraph systems which is due to Tao [33] (see also [8] for an exposition). Closely related discrete analogues were obtained earlier by Gowers [13] and, independently, by Nagle, Rödl, Schacht and Skokan [21, 26].

Theorem 3.4. *For every $n, r \in \mathbb{N}$ with $n \geqslant r \geqslant 2$ and every $0 < \varepsilon \leqslant 1$ there exist a strictly positive constant $\Delta(n, r, \varepsilon)$ and a positive integer $K(n, r, \varepsilon)$ with*

the following property. Let $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$ be an r-uniform hypergraph system and for every $e \in \mathcal{H}$ let $E_e \in \mathcal{B}_e$ such that

$$(3.4) \qquad \boldsymbol{\mu}\Big(\bigcap_{e \in \mathcal{H}} E_e \Big) \leqslant \Delta(n, r, \varepsilon).$$

Then for every $e \in \mathcal{H}$ there exists $F_e \in \mathcal{B}_e$ with

$$(3.5) \qquad \boldsymbol{\mu}(E_e \setminus F_e) \leqslant \varepsilon \quad and \quad \bigcap_{e \in \mathcal{H}} F_e = \emptyset.$$

Moreover, there exists a collection $\langle \mathcal{P}_{e'} : e' \subseteq e$ for some $e \in \mathcal{H} \rangle$ of partitions of $\boldsymbol{X}$ such that: (i) $\mathcal{P}_{e'} \subseteq \mathcal{B}_{e'}$ and $|\mathcal{P}_{e'}| \leqslant K(n, r, \varepsilon)$ for every $e' \subseteq e \in \mathcal{H}$, and (ii) for every $e \in \mathcal{H}$ the set F_e belongs to the algebra generated by the family $\bigcup_{e' \subsetneq e} \mathcal{P}_{e'}$.

4. A HÖLDER-TYPE INEQUALITY

In this section, our goal is to prove a Hölder-type inequality for L_p regular random variables. We begin with a brief motivating discussion. Let (X, Σ, μ) be a probability space and let f be a nonnegative real-valued random variable which belongs to $L_p(X, \Sigma, \mu)$ for some $1 < p \leqslant \infty$. Let q denote the conjugate exponent of p and notice that, by Höder's inequality, for every $A \in \Sigma$ we have

$$(4.1) \qquad \Big(\int_A f \, d\mu \Big)^q \leqslant \|f\|_{L_p}^q \cdot \mu(A).$$

Of course, this uniform estimate does not hold true if f is not in $L_p(X, \Sigma, \mu)$. Nevertheless, we do have a version of (4.1) if we merely assume that f is L_p regular. This is the content of the following proposition.

Proposition 4.1. Let $n, r \in \mathbb{N}$ with $n \geqslant r \geqslant 2$ and $0 < \eta \leqslant (r+1)^{-1}$. Also let $C > 0$ and $1 < p \leqslant \infty$, and let q denote the conjugate exponent of p. Finally, let $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$ be an η-nonatomic hypergraph system, $e \in \mathcal{H}$ with $|e| = r$, and let $f \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ be nonnegative. Then the following hold.

(a) If f is (C, η, p)-regular, then for every $A \in \mathcal{S}_{\partial_e}$ we have

$$(4.2) \qquad \Big(\int_A f \, d\boldsymbol{\mu} \Big)^q \leqslant C^q \big(\boldsymbol{\mu}(A) + (r+3)\eta \big).$$

(b) On the other hand, if (4.2) is satisfied for every $A \in \mathcal{S}_{\partial_e}$, then the random variable f is (K, η, p)-regular where $K = C(r+4)^{1/q}\eta^{-1/p}$. In particular, if $p = \infty$, then f is $(C(r+4), \eta, \infty)$-regular.

Proposition 4.1 is based on the simple (but quite useful) observation that for every $A \in \mathcal{S}_{\partial e}$ with $\boldsymbol{\mu}(A) \geqslant \eta$ we can find a partition of $\boldsymbol{X}$ which almost contains the set A, and whose members are contained in $\mathcal{S}_{\partial e}$ and are not too small. We present this fact in a slightly more general form in Subsection 4.1 (this form is needed later on in Section 5). The proof of Proposition 4.1 is completed in Subsection 4.2.

4.1. **A partition lemma.** For every probability space (X, Σ, μ) and every finite partition $\mathcal{P}$ of X with $\mathcal{P} \subseteq \Sigma$ we set

$$(4.3) \qquad \iota(\mathcal{P}) = \min\{\mu(P) : P \in \mathcal{P}\}.$$

We have the following lemma.

Lemma 4.2. *Let r be a positive integer and $0 < \theta < 1$. Also let (X, Σ, μ) be a probability space, $(\mathcal{B}_i)_{i=1}^r$ a finite sequence of sub-σ-algebras of Σ, and set*

$$\mathcal{S} = \Big\{ \bigcap_{i=1}^r A_i : A_i \in \mathcal{B}_i \text{ for every } i \in [r] \Big\}.$$

Then for every $A \in \mathcal{S}$ with $\mu(A) \geqslant \theta$ there exist: (i) a partition $\mathcal{Q}$ of X with $\mathcal{Q} \subseteq \mathcal{S}$ and $\iota(\mathcal{Q}) \geqslant \theta$, (ii) a set $B \in \mathcal{Q}$ with $A \subseteq B$, and (iii) pairwise disjoint sets $B_1, \ldots, B_r \in \mathcal{S}$ with $\mu(B_i) < \theta$ for every $i \in [r]$, such that $B \setminus A = \bigcup_{i=1}^r B_i$.

Proof. Fix $A \in \mathcal{S}$ with $\mu(A) \geqslant \theta$ and write $A = \bigcap_{i=1}^r A_i$ where $A_i \in \mathcal{B}_i$ for every $i \in [r]$. For every nonempty $I \subseteq [r]$ and every $i \in I$ let

$$C_{I,i} = \Big(\bigcap_{j \in \{\ell \in I : \ell < i\}} A_j \Big) \cap (X \setminus A_i)$$

with the convention that $C_{I,i} = X \setminus A_i$ if $i = \min(I)$. It is clear that $C_{I,i} \in \mathcal{S}$ for every $i \in I$. Moreover, notice that the family $\{C_{I,i} : i \in I\}$ is a partition of $X \setminus \bigcap_{i \in I} A_i$. We set $G = \big\{ i \in [r] : \mu(C_{[r],i}) \geqslant \theta \big\}$ and we observe that if $G = \emptyset$, then the trivial partition $\mathcal{Q} = \{X\}$ and the sets $C_{[r],1}, \ldots, C_{[r],r}$ satisfy the requirements of the lemma. So, assume that G is nonempty and let

$$B = \bigcap_{i \in G} A_i \quad \text{and} \quad \mathcal{Q} = \{B\} \cup \big\{ C_{G,i} : i \in G \big\}.$$

Also let $B_i = B \cap C_{[r]\setminus G,i}$ if $i \notin G$, and $B_i = \emptyset$ if $i \in G$. We will show that $\mathcal{Q}, B$ and $B_1, \ldots, B_r$ are as desired.

Indeed, notice first that $\mathcal{Q}$ is a partition of X with $\mathcal{Q} \subseteq \mathcal{S}$, $B \in \mathcal{Q}$ and $A \subseteq B$. Next, let $Q \in \mathcal{Q}$ be arbitrary. If $Q = B$, then $\mu(Q) = \mu(B) \geqslant \mu(A) \geqslant \theta$. Otherwise, there exists $i \in G$ such that $Q = C_{G,i}$. Since $C_{[r],i} \subseteq C_{G,i}$ and $i \in G$, we see that $\mu(Q) = \mu(C_{G,i}) \geqslant \mu(C_{[r],i}) \geqslant \theta$. Thus, we have $\iota(\mathcal{Q}) \geqslant \theta$. Finally, observe that $B_1, \ldots, B_r \in \mathcal{S}$ are pairwise disjoint and

$$B \setminus A = \bigcup_{i=1}^r (B \cap C_{[r],i}) = \bigcup_{i \notin G} (B \cap C_{[r]\setminus G,i}) = \bigcup_{i=1}^r B_i.$$

Moreover, for every $i \notin G$ we have

$$B_i = B \cap C_{[r]\setminus G,i} = \Big(\bigcap_{j \in G} A_j \Big) \cap C_{[r]\setminus G,i} \subseteq C_{[r],i}$$

and so $\mu(B_i) \leqslant \mu(C_{[r],i}) < \theta$. The proof of Lemma 4.2 is completed. $\qquad\square$

4.2. **Proof of Proposition 4.1.** We will need the following lemma. It is a version of Proposition 3.3 in the context of hypergraph systems.

Lemma 4.3. *Let $n, r \in \mathbb{N}$ with $n \geqslant r \geqslant 2$ and $0 < \alpha, \eta < 1$ with $r\eta \leqslant 1 - a$. Also let $\mathscr{H} = (n, \langle(X_i, \Sigma_i, \mu_i) : i \in [n]\rangle, \mathcal{H})$ be an η-nonatomic hypergraph system, and let $e \in \mathcal{H}$ with $|e| = r$. Then for every $A \in \mathcal{S}_{\partial e}$ with $\boldsymbol{\mu}(A) < a$ there exists $B \in \mathcal{S}_{\partial e}$ with $A \subseteq B$ and $a \leqslant \boldsymbol{\mu}(B) < a + 2\eta$.*

Proof. We argue as in the proof of Proposition 3.3. Specifically, fix $A \in \mathcal{S}_{\partial e}$ with $\boldsymbol{\mu}(A) < a$ and assume, towards a contradiction, that

(H) for every $B \in \mathcal{S}_{\partial e}$ with $A \subseteq B$ either $\boldsymbol{\mu}(B) < a$ or $\boldsymbol{\mu}(B) \geqslant a + 2\eta$.

For every $e' \in \partial e$ we select $A_{e'} \in \mathcal{B}_{e'}$ such that $A = \bigcap_{e' \in \partial e} A_{e'}$ and we observe that

$$(4.4) \qquad \sum_{e' \in \partial e} \boldsymbol{\mu}(\boldsymbol{X} \setminus A_{e'}) \geqslant \boldsymbol{\mu}\Big(\boldsymbol{X} \setminus \bigcap_{e' \in \partial e} A_{e'}\Big) > 1 - a \geqslant r\eta.$$

Therefore, there exists $e_1' \in \partial e$ such that $\boldsymbol{\mu}(\boldsymbol{X} \setminus A_{e_1'}) > \eta$. Since $\mathscr{H}$ is η-nonatomic, we see that $\boldsymbol{\mu}(Z) \leqslant \eta^{r-1} \leqslant \eta$ for every atom Z of $(\boldsymbol{X}, \mathcal{B}_{e_1'}, \boldsymbol{\mu})$. Hence, by Proposition 3.3 applied for "$B = \boldsymbol{X} \setminus A_{e_1'}$" and "$c = \eta$", there exists $B_{e_1'} \in \mathcal{B}_{e_1'}$ with $B_{e_1'} \subseteq \boldsymbol{X} \setminus A_{e_1'}$ and $\eta \leqslant \boldsymbol{\mu}(B_{e_1'}) < 2\eta$. We set $A_{e_1'}^1 = A_{e_1'} \cup B_{e_1'}$ and $A_{e'}^1 = A_{e'}$ if $e' \in \partial e \setminus \{e_1'\}$. Notice that: (i) $\boldsymbol{\mu}(A_{e_1'}^1) \geqslant \boldsymbol{\mu}(A_{e_1'}) + \eta$, (ii) $\bigcap_{e' \in \partial e} A_{e'}^1 \in \mathcal{S}_{\partial e}$, and (iii) $A \subseteq \bigcap_{e' \in \partial e} A_{e'}^1$. Moreover, $\boldsymbol{\mu}\big(\bigcap_{e' \in \partial e} A_{e'}^1\big) \leqslant \boldsymbol{\mu}(A) + 2\eta < a + 2\eta$ and so, by hypothesis (H), we obtain that $\boldsymbol{\mu}\big(\bigcap_{e' \in \partial e} A_{e'}^1\big) < a$. It follows, in particular, that the estimate in (4.4) is satisfied for the family $\langle A_{e'}^1 : e' \in \partial e\rangle$.

Thus, setting $M = \lceil 2r/\eta \rceil$, we select recursively: (a) a finite sequence $(e_m')_{m=1}^M$ in ∂e, and (b) for every $e' \in \partial e$ a finite sequence $(A_{e'}^m)_{m=0}^M$ in $\mathcal{B}_{e'}$ with $A_{e'}^0 = A_{e'}$, such that for every $m \in [M]$ the following hold.

(C1) For every $e' \in \partial e$ we have $A_{e'}^{m-1} \subseteq A_{e'}^m$. Moreover, $\boldsymbol{\mu}(A_{e_m'}^m) \geqslant \boldsymbol{\mu}(A_{e_m'}^{m-1}) + \eta$.
(C2) We have $\boldsymbol{\mu}\big(\bigcap_{e' \in \partial e} A_{e'}^m\big) < a$.

By the classical pigeonhole principle, there exist $L \subseteq [M]$ with $|L| \geqslant M/r$ and $g \in \partial e$ such that $e_m' = g$ for every $m \in L$. If $\ell = \max(L)$, then by (C1) we conclude that $\boldsymbol{\mu}(A_g^\ell) \geqslant 2$ which is clearly a contradiction. $\qquad\square$

We are ready to give the proof of Proposition 4.1.

Proof of Proposition 4.1. (a) Fix $A \in \mathcal{S}_{\partial e}$. If $\eta \leqslant \boldsymbol{\mu}(A)$, then we claim that

$$(4.5) \qquad \Big(\int_A f \, d\boldsymbol{\mu}\Big)^q \leqslant C^q\big(\boldsymbol{\mu}(A) + r\eta\big).$$

Indeed, by Lemma 4.2, there exist a partition $\mathcal{Q}$ of $\boldsymbol{X}$ with $\mathcal{Q} \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{Q}) \geqslant \eta$, and $B \in \mathcal{Q}$ with $A \subseteq B$ and $\boldsymbol{\mu}(B \setminus A) < r\eta$. Since f is (C, η, p)-regular we see that

$$\frac{\int_B f \, d\boldsymbol{\mu}}{\boldsymbol{\mu}(B)} \, \boldsymbol{\mu}(B)^{1/p} \leqslant \|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}})\|_{L_p} \leqslant C.$$

(Here, we have $\boldsymbol{\mu}(B)^{1/p} = 1$ if $p = \infty$.) Hence,

$$\Big(\int_A f \, d\boldsymbol{\mu}\Big)^q \leqslant \Big(\int_B f \, d\boldsymbol{\mu}\Big)^q \leqslant C^q \boldsymbol{\mu}(B) \leqslant C^q\big(\boldsymbol{\mu}(A) + r\eta\big).$$

Next, assume that $0 \leqslant \boldsymbol{\mu}(A) < \eta$. Our hypothesis that $0 < \eta \leqslant (r+1)^{-1}$ yields that $r\eta \leqslant 1 - \eta$ and so, by Lemma 4.3, there exists $B \in \mathcal{S}_{\partial e}$ with $A \subseteq B$ and $\eta \leqslant \boldsymbol{\mu}(B) < 3\eta$. Therefore,

$$(4.6) \quad \Big(\int_A f \, d\boldsymbol{\mu} \Big)^q \leqslant \Big(\int_B f \, d\boldsymbol{\mu} \Big)^q \overset{(4.5)}{\leqslant} C^q \big(\boldsymbol{\mu}(B) + r\eta \big) \leqslant C^q \big(\boldsymbol{\mu}(A) + (r+3)\eta \big)$$

and the proof of part (a) is completed.

(b) Let $\mathcal{P}$ be an arbitrary partition of $\boldsymbol{X}$ with $\mathcal{P} \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{P}) \geqslant \eta$. By (4.2), for every $P \in \mathcal{P}$ we have $\int_P f \, d\boldsymbol{\mu} \leqslant C(r+4)^{1/q} \boldsymbol{\mu}(P)^{1/q}$. Therefore, if $1 < p < \infty$,

$$\begin{aligned}
\|\mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}})\|_{L_p}^p &= \sum_{P \in \mathcal{P}} \Big(\frac{\int_P f \, d\boldsymbol{\mu}}{\boldsymbol{\mu}(P)} \Big)^p \boldsymbol{\mu}(P) \leqslant C^p (r+4)^{p/q} \sum_{P \in \mathcal{P}} \boldsymbol{\mu}(P)^{p/q+1-p} \\
&= C^p (r+4)^{p/q} |\mathcal{P}| \leqslant C^p (r+4)^{p/q} \eta^{-1}.
\end{aligned}$$

On the other hand, if $p = \infty$,

$$\|\mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}})\|_{L_\infty} = \max \Big\{ \frac{\int_P f \, d\boldsymbol{\mu}}{\boldsymbol{\mu}(P)} : P \in \mathcal{P} \Big\} \leqslant \frac{C \big(\boldsymbol{\mu}(P) + (r+3)\eta \big)}{\boldsymbol{\mu}(P)} \leqslant C(r+4)$$

as desired. □

5. Regularity lemma for L_p regular hypergraphs

In this section we give the proof of Theorem 2.1. Specifically, let $n, r \in \mathbb{N}$ with $n \geqslant r \geqslant 2$, and let $C > 0$ and $1 < p \leqslant \infty$. Let q denote the conjugate exponent of p and set $p^\dagger = \min\{2, p\}$. Also let $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$ be an r-uniform hypergraph system. *These data will be fixed throughout this section.*

5.1. **Preliminary tools.** The following result is a refinement of the partition lemma presented in Subsection 4.1. Recall that for every probability space (X, Σ, μ) and every partition $\mathcal{P}$ of X with $\mathcal{P} \subseteq \Sigma$ we set $\iota(\mathcal{P}) = \min\{\mu(P) : P \in \mathcal{P}\}$.

Lemma 5.1. *Let $0 < \vartheta, \eta < 1$ and $e \in \mathcal{H}$. Let $f \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ be nonnegative and (C, η, p)-regular, and $\mathcal{P}$ a finite partition of $\boldsymbol{X}$ with $\mathcal{P} \subseteq \mathcal{S}_{\partial e}$. Assume that*

$$(5.1) \qquad\qquad \eta \leqslant \big(\vartheta \cdot \iota(\mathcal{P}) \big)^q$$

and that $\mathscr{H}$ is η-nonatomic. Then for every $A \in \mathcal{S}_{\partial e}$ there exist: (i) *a refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q} \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{Q}) \geqslant \big(\vartheta \cdot \iota(\mathcal{P}) \big)^q$, and* (ii) *a set $B \in \mathcal{A}_{\mathcal{Q}}$, such that*

$$(5.2) \qquad \int_{A \triangle B} \mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}}) \, d\boldsymbol{\mu} \leqslant Cr\vartheta \quad and \quad \int_{A \triangle B} f \, d\boldsymbol{\mu} \leqslant 5Cr^2\vartheta.$$

Proof. It is based on the following greedy algorithm. Let P be a cell of the partition $\mathcal{P}$ and assume that the trace $A \cap P$ of the given set A on P is a relatively large subset of P. In this case, using Lemma 4.2 applied for the conditional probability measure with respect to P, we may select a partition $\mathcal{Q}_P$ of P and an element B_P of $\mathcal{Q}_P$ which is a rather strong approximation of $A \cap P$. On the other hand, if the measure of $A \cap P$ is negligible when compared with the measure of P, then we do

not split P. After scanning all the cells of the partition $\mathcal{P}$, this process will eventually produce a refinement $\mathcal{Q}$ of $\mathcal{P}$ and a fairly accurate global approximation B of A which belongs to the σ-algebra generated by $\mathcal{Q}$. Once this is done, the estimates in (5.2) will follow from Hölder's inequality and Proposition 4.1 respectively.

We proceed to the details. We fix $A \in \mathcal{S}_{\partial e}$ and we set

$$(5.3) \qquad \theta = \vartheta^q \cdot \iota(\mathcal{P})^{q-1}.$$

First, for every $P \in \mathcal{P}$ we select a partition $\mathcal{Q}_P$ of P with $\mathcal{Q}_P \subseteq \mathcal{S}_{\partial e}$ and a set $B_P \in \mathcal{S}_{\partial e}$ as follows. Let $P \in \mathcal{P}$ be arbitrary. If $\boldsymbol{\mu}(A \cap P) < \theta \boldsymbol{\mu}(P)$, then we set $\mathcal{Q}_P = \{P\}$ and $B_P = \emptyset$. Otherwise, let $(P, \boldsymbol{\Sigma}_P, \boldsymbol{\mu}_P)$ be the probability space where $\boldsymbol{\Sigma}_P = \{C \cap P : C \in \boldsymbol{\Sigma}\}$ and $\boldsymbol{\mu}_P$ is the conditional probability measure of $\boldsymbol{\mu}$ with respect to P, that is, $\boldsymbol{\mu}_P(C) = \boldsymbol{\mu}(C \cap P)/\boldsymbol{\mu}(P)$ for every $C \in \boldsymbol{\Sigma}$. Write $\partial e = \{e'_1, \ldots, e'_r\}$ and for every $i \in [r]$ let $\mathcal{B}_i = \{B \cap P : B \in \mathcal{B}_{e'_i}\}$; observe that $\mathcal{B}_i$ is a sub-σ-algebra of $\boldsymbol{\Sigma}_P$. Also let $\mathcal{S} = \big\{ \bigcap_{i=1}^r B_i : B_i \in \mathcal{B}_i \text{ for every } i \in [r] \big\} \subseteq \mathcal{S}_{\partial e}$. By Lemma 4.2 applied to the probability space $(P, \boldsymbol{\Sigma}_P, \boldsymbol{\mu}_P)$ and the set $A \cap P \in \mathcal{S}$, we obtain: (i) a partition $\mathcal{Q}_P$ of P with $\mathcal{Q}_P \subseteq \mathcal{S}$ and $\iota(\mathcal{Q}_P) \geqslant \theta$, (ii) a set $B_P \in \mathcal{Q}_P$ with $A \cap P \subseteq B_P$, and (iii) pairwise disjoint sets $B_1^P, \ldots, B_r^P \in \mathcal{S}$ with $\boldsymbol{\mu}_P(B_i^P) < \theta$ for every $i \in [r]$, such that $B_P \setminus (A \cap P) = \bigcup_{i=1}^r B_i^P$.

Next, we define

$$(5.4) \qquad \mathcal{Q} = \bigcup_{P \in \mathcal{P}} \mathcal{Q}_P \ \text{ and } \ B = \bigcup_{P \in \mathcal{P}} B_P.$$

Observe that $\mathcal{Q}$ is a refinement of $\mathcal{P}$ with $\mathcal{Q} \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{Q}) \geqslant \theta \cdot \iota(\mathcal{P}) = \big(\vartheta \cdot \iota(\mathcal{P})\big)^q$. Also note that $B \in \mathcal{A}_{\mathcal{Q}}$ and, setting $\mathcal{P}^* = \{P \in \mathcal{P} : \boldsymbol{\mu}(A \cap P) \geqslant \theta \boldsymbol{\mu}(P)\}$, we have

$$(5.5) \qquad A \triangle B = \Big(\bigcup_{P \in \mathcal{P} \setminus \mathcal{P}^*} (A \cap P) \Big) \cup \Big(\bigcup_{P \in \mathcal{P}^*} \big(\bigcup_{i=1}^r B_i^P \big) \Big)$$

where for every $P \in \mathcal{P}^*$ the sets $B_1^P, \ldots, B_r^P$ are as in (iii) above. In particular, noticing that $\boldsymbol{\mu}(A \cap P) < \theta \boldsymbol{\mu}(P)$ for every $P \notin \mathcal{P}^*$ and $\boldsymbol{\mu}(B_i^P) < \theta \boldsymbol{\mu}(P)$ for every $P \in \mathcal{P}^*$ and every $i \in [r]$, we see that

$$(5.6) \qquad \boldsymbol{\mu}(A \triangle B) \leqslant r\theta.$$

On the other hand, by (5.1), we have $\iota(\mathcal{P}) \geqslant \eta$, and so $\|\mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}})\|_{L_p} \leqslant C$ since f is (C, η, p)-regular. Hence, by Hölder's inequality, we obtain that

$$\int_{A \triangle B} \mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}}) \, d\boldsymbol{\mu} \leqslant \|\mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}})\|_{L_p} \cdot \boldsymbol{\mu}(A \triangle B)^{1/q} \overset{(5.6)}{\leqslant} Cr^{1/q}\theta^{1/q} \overset{(5.3)}{\leqslant} Cr\vartheta.$$

We proceed to show that $\int_{A \triangle B} f \, d\boldsymbol{\mu} \leqslant 5Cr^2\vartheta$. To this end notice first that

$$(5.7) \qquad \int_{A \triangle B} f \, d\boldsymbol{\mu} = \sum_{P \in \mathcal{P} \setminus \mathcal{P}^*} \int_{A \cap P} f \, d\boldsymbol{\mu} + \sum_{P \in \mathcal{P}^*} \sum_{i=1}^r \int_{B_i^P} f \, d\boldsymbol{\mu}.$$

By (5.1) and (5.3), we have $\eta \leqslant \theta \boldsymbol{\mu}(P)$ for every $P \in \mathcal{P}$. Hence, if $P \in \mathcal{P} \setminus \mathcal{P}^*$, then, by Proposition 4.1,

$$
\begin{aligned}
\Big(\int_{A \cap P} f \, d\boldsymbol{\mu} \Big)^q &\leqslant C^q \big(\boldsymbol{\mu}(A \cap P) + (r+3)\eta \big) \\
&\leqslant C^q \big(\theta \boldsymbol{\mu}(P) + (r+3)\theta \boldsymbol{\mu}(P) \big) \leqslant 5 C^q r \theta \boldsymbol{\mu}(P)
\end{aligned}
$$

and so

$$
(5.8) \qquad \sum_{P \in \mathcal{P} \setminus \mathcal{P}^*} \int_{A \cap P} f \, d\boldsymbol{\mu} \leqslant 5 C r \theta^{1/q} \sum_{P \in \mathcal{P} \setminus \mathcal{P}^*} \boldsymbol{\mu}(P)^{1/q}.
$$

Respectively, for every $P \in \mathcal{P}^*$ and every $i \in [r]$ we have

$$
\Big(\int_{B_i^P} f \, d\boldsymbol{\mu} \Big)^q \leqslant C^q \big(\boldsymbol{\mu}(B_i^P) + (r+3)\eta \big) \leqslant 5 C^q r \theta \boldsymbol{\mu}(P)
$$

which yields that

$$
(5.9) \qquad \sum_{P \in \mathcal{P}^*} \sum_{i=1}^{r} \int_{B_i^P} f \, d\boldsymbol{\mu} \leqslant 5 C r^2 \theta^{1/q} \sum_{P \in \mathcal{P}^*} \boldsymbol{\mu}(P)^{1/q}.
$$

Finally, notice that the function $x^{1/q}$ is concave on $\mathbb{R}^+$ since $q \geqslant 1$. Therefore,

$$
(5.10) \qquad \sum_{P \in \mathcal{P}} \boldsymbol{\mu}(P)^{1/q} \leqslant |\mathcal{P}|^{1/p} \leqslant \iota(\mathcal{P})^{-1/p}.
$$

Combining (5.7)–(5.10) we conclude that

$$
\int_{A \triangle B} f \, d\boldsymbol{\mu} \leqslant 5 C r^2 \theta^{1/q} \sum_{P \in \mathcal{P}} \boldsymbol{\mu}(P)^{1/q} \leqslant 5 C r^2 \theta^{1/q} \cdot \iota(\mathcal{P})^{-1/p} \stackrel{(5.3)}{=} 5 C r^2 \vartheta
$$

and the proof of Lemma 5.1 is completed. $\qquad \square$

Lemma 5.1 is used in the proof of the following result which asserts (roughly speaking) that if a given approximation of an L_p regular random variable f is not sufficiently close to f in the cut norm, then we can find a much finer approximation.

Lemma 5.2. *Let $0 < \delta, \eta < 1$ and set $\vartheta = \delta(12 C r^2)^{-1}$. Also let $e \in \mathcal{H}$ and let $f \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ be nonnegative and (C, η, p)-regular. Finally, let $\mathcal{P}$ be a finite partition of $\boldsymbol{X}$ with $\mathcal{P} \subseteq \mathcal{S}_{\partial e}$ such that $\| f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}}) \|_{\mathcal{S}_{\partial e}} > \delta$. Assume that*

$$
(5.11) \qquad \eta \leqslant \big(\vartheta \cdot \iota(\mathcal{P}) \big)^q
$$

and that $\mathcal{H}$ is η-nonatomic. Then there exists a refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q} \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{Q}) \geqslant \big(\vartheta \cdot \iota(\mathcal{P}) \big)^q$, such that $\| \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}}) \|_{L_{p^\dagger}} \geqslant \delta/2$.

Proof. We select $A \in \mathcal{S}_{\partial e}$ such that

$$
(5.12) \qquad \Big| \int_A \big(f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}}) \big) \, d\boldsymbol{\mu} \Big| > \delta.
$$

Next, we apply Lemma 5.1 and we obtain a refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q} \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{Q}) \geqslant \big(\vartheta \cdot \iota(\mathcal{P})\big)^q$, and a set $B \in \mathcal{A}_{\mathcal{Q}}$ such that $\int_{A \triangle B} \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}}) \, d\boldsymbol{\mu} \leqslant Cr\vartheta$ and $\int_{A \triangle B} f \, d\boldsymbol{\mu} \leqslant 5Cr^2\vartheta$. Then, by the choice of ϑ, we have

$$\big| \int_A \big(f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\big) \, d\boldsymbol{\mu} - \int_B \big(f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\big) \, d\boldsymbol{\mu} \big| \leqslant$$

$$\leqslant \int_{A \triangle B} f \, d\boldsymbol{\mu} + \int_{A \triangle B} \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}}) \, d\boldsymbol{\mu} \leqslant 5Cr^2\vartheta + Cr\vartheta \leqslant 6Cr^2\vartheta = \delta/2$$

and so, by (5.12),

$$(5.13) \qquad \big| \int_B \big(f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\big) \, d\boldsymbol{\mu} \big| \geqslant \delta/2.$$

On the other hand, the fact that $B \in \mathcal{A}_{\mathcal{Q}}$ yields that

$$(5.14) \qquad \int_B \big(f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\big) \, d\boldsymbol{\mu} = \int_B \big(\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\big) \, d\boldsymbol{\mu}.$$

Therefore, by the monotonicity of the L_p norms, we conclude that

$$\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_{p^\dagger}} \geqslant \|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_1} \geqslant$$

$$\big| \int_B \big(\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\big) \, d\boldsymbol{\mu} \big| \stackrel{(5.14)}{=} \big| \int_B \big(f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\big) \, d\boldsymbol{\mu} \big| \stackrel{(5.13)}{\geqslant} \delta/2$$

as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

5.2. **Proof of Theorem 2.1.** We now enter into the main part of the proof. We begin with the following lemma.

Lemma 5.3. *Let $0 < \delta, \eta < 1$ and $0 < \sigma \leqslant 1$, and set $\vartheta = \delta(12Cr^2)^{-1}$ and $N = \lceil 4(p^\dagger - 1)^{-1} \sigma^2 \delta^{-2} \rceil$. Also let $e \in \mathcal{H}$, let $f \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ be nonnegative and (C, η, p)-regular, and let $\mathcal{P}$ be a finite partition of $\boldsymbol{X}$ with $\mathcal{P} \subseteq \mathcal{S}_{\partial e}$. Assume that*

$$(5.15) \qquad\qquad \eta \leqslant \big(\vartheta^N \cdot \iota(\mathcal{P})\big)^{q^N}$$

and that $\mathcal{H}$ is η-nonatomic. Then there exists a refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q} \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{Q}) \geqslant \big(\vartheta^N \cdot \iota(\mathcal{P})\big)^{q^N}$, such that either

(a) $\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_{p^\dagger}} > \sigma$, *or*

(b) $\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_{p^\dagger}} \leqslant \sigma$ *and* $\|f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}})\|_{\mathcal{S}_{\partial e}} \leqslant \delta$.

The case "$p \geqslant 2$" in Lemma 5.3 can be proved with an "energy increment strategy", a method which originates from the work of Szemerédi [31, 32] and is essentially based upon the fact that martingale difference sequences are orthogonal in L_2 (see, e.g., [8, 33]). The general case follows a method introduced recently in [9] and relies on Proposition 3.1 (see also [11] for another application).

Proof of Lemma 5.3. Assume that part (a) is not satisfied, that is,

(H1) for every refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q} \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{Q}) \geqslant \big(\vartheta^N \cdot \iota(\mathcal{P})\big)^{q^N}$ we have $\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_{p^\dagger}} \leqslant \sigma$.

We claim that there exists a refinement $\mathcal{Q}$ of $\mathcal{P}$ which satisfies the second part of the lemma. Indeed, if not, then, by (H1) and Lemma 5.2, we see that

(H2) for every refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q} \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{Q}) \geqslant \left(\vartheta^N \cdot \iota(\mathcal{P})\right)^{q^N}$ there exists a refinement $\mathcal{R}$ of $\mathcal{Q}$ with $\mathcal{R} \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{R}) \geqslant \left(\vartheta \cdot \iota(\mathcal{Q})\right)^q$ such that $\|(f \mid \mathcal{A}_{\mathcal{R}}) - \mathbb{E}(f \mid \mathcal{A}_{\mathcal{Q}})\|_{L_{p^\dagger}} > \delta/2$.

Recursively and using (H2), we select partitions $\mathcal{P}_0, \dots, \mathcal{P}_N$ of $\boldsymbol{X}$ with $\mathcal{P}_0 = \mathcal{P}$ such that for every $i \in [N]$ we have: (P1) $\mathcal{P}_i$ is a refinement of $\mathcal{P}_{i-1}$ with $\mathcal{P}_i \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{P}_i) \geqslant \left(\vartheta \cdot \iota(\mathcal{P}_{i-1})\right)^q$, and (P2) $\|\mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}_i}) - \mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}_{i-1}})\|_{L_{p^\dagger}} > \delta/2$.

Next, set $g = f - \mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}})$ and let $(d_i)_{i=0}^N$ be the difference sequence associated with the finite martingale $\mathbb{E}(g \mid \mathcal{A}_{\mathcal{P}_0}), \dots, \mathbb{E}(g \mid \mathcal{A}_{\mathcal{P}_N})$. Notice that for every $i \in [N]$ we have $d_i = \mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}_i}) - \mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}_{i-1}})$ which implies, by (P2), that $\|d_i\|_{L_{p^\dagger}} > \delta/2$. Therefore, by the choice of N and Proposition 3.1,

$$\sigma \leqslant (p^\dagger - 1)^{1/2} \frac{\delta}{2} N^{1/2} \quad < \quad (p^\dagger - 1)^{1/2} \Big(\sum_{i=0}^N \|d_i\|_{L_{p^\dagger}}^2 \Big)^{1/2}$$

$$\leqslant \quad \Big\| \sum_{i=0}^N d_i \Big\|_{L_{p^\dagger}} = \|\mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}_N}) - \mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}})\|_{L_{p^\dagger}}.$$

On the other hand, by (P1), we see that $\mathcal{P}_N$ is a refinement of $\mathcal{P}$ with $\mathcal{P}_N \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{Q}) \geqslant \left(\vartheta^N \cdot \iota(\mathcal{P})\right)^{q^N}$ and so, by (H1), $\|\mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}_N}) - \mathbb{E}(f \mid \mathcal{A}_{\mathcal{P}})\|_{L_{p^\dagger}} \leqslant \sigma$ which contradicts, of course, the above estimate. The proof is thus completed. $\square$

We introduce some numerical invariants. For every growth function $F \colon \mathbb{N} \to \mathbb{R}$ and every $0 < \sigma \leqslant 1$ we define, recursively, a sequence (N_m) in $\mathbb{N}$ and two sequences (η_m) and (ϑ_m) in $(0, 1]$ by setting $N_0 = 0$, $\eta_0 = 1$, $\theta_0 = \left(12Cr^2 F(1)\right)^{-1}$ and

(5.16)
$$\begin{cases} N_{m+1} = \lceil 4(p^\dagger - 1)^{-1} \sigma^2 F(\lceil \eta_m^{-1} \rceil)^2 \rceil, \\ \eta_{m+1} = (\vartheta_m^{N_{m+1}} \cdot \eta_m)^{q^{N_{m+1}}}, \\ \vartheta_{m+1} = \left(12Cr^2 F(\lceil \eta_{m+1}^{-1} \rceil)\right)^{-1}. \end{cases}$$

The following lemma is the last step of the proof of Theorem 2.1.

Lemma 5.4. *Let* $0 < \sigma \leqslant 1$ *and* $F \colon \mathbb{N} \to \mathbb{R}$ *a growth function. Set*

(5.17)
$$L = \lceil C^2 (p^\dagger - 1)^{-1} \sigma^{-2} n^r \rceil$$

and let (η_m) *be as in (5.16). Let* $0 < \eta \leqslant \eta_L$ *and assume that* $\mathscr{H}$ *is* η*-nonatomic. For every* $e \in \mathcal{H}$ *let* $f_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ *be nonnegative and* (C, η, p)*-regular. Then there exist: (i) a positive integer* $m \in \{0, \dots, L-1\}$, *(ii) for every* $e \in \mathcal{H}$ *a partition* $\mathcal{P}_e$ *of* $\boldsymbol{X}$ *with* $\mathcal{P}_e \subseteq \mathcal{S}_{\partial e}$ *and* $\iota(\mathcal{P}_e) \geqslant \eta_m$, *and (iii) for every* $e \in \mathcal{H}$ *a refinement* $\mathcal{Q}_e$ *of* $\mathcal{P}_e$ *with* $\mathcal{Q}_e \subseteq \mathcal{S}_{\partial e}$ *and* $\iota(\mathcal{Q}_e) \geqslant \eta_{m+1}$, *such that for every* $e \in \mathcal{H}$ *we have* $\|\mathbb{E}(f_e \mid \mathcal{A}_{\mathcal{Q}_e}) - \mathbb{E}(f_e \mid \mathcal{A}_{\mathcal{P}_e})\|_{L_{p^\dagger}} \leqslant \sigma$ *and* $\|f_e - \mathbb{E}(f_e \mid \mathcal{A}_{\mathcal{Q}_e})\|_{\mathcal{S}_{\partial e}} \leqslant 1/F(\lceil \eta_m^{-1} \rceil)$.

Proof. It is similar to the proof of Lemma 5.3 and so we will briefly sketch the argument. If the lemma is false, then using Lemma 5.3 we select for every $e \in \mathcal{H}$

partitions $\mathcal{P}_0^e, \ldots, \mathcal{P}_L^e$ of $\boldsymbol{X}$ with $\mathcal{P}_0^e = \{\boldsymbol{X}\}$ as well as $e_1, \ldots, e_L \in \mathcal{H}$ such that
for every $m \in [L]$ we have: (P1) $\mathcal{P}_m^e$ is a refinement of $\mathcal{P}_{m-1}^e$ with $\mathcal{P}_m^e \subseteq \mathcal{S}_{\partial e}$ and
$\iota(\mathcal{P}_m^e) \geqslant \eta_m$ for every $e \in \mathcal{H}$, and (P2) $\|\mathbb{E}(f_{e_m} \,|\, \mathcal{A}_{\mathcal{P}_m^{e_m}}) - \mathbb{E}(f_{e_m} \,|\, \mathcal{A}_{\mathcal{P}_{m-1}^{e_m}})\|_{L_{p^\dagger}} > \sigma$.
By the pigeonhole principle, there exist $\boldsymbol{e} \in \mathcal{H}$ and $I \subseteq [L]$ with $|I| \geqslant L/n^r$, such
that $\boldsymbol{e} = e_m$ for every $m \in I$. Let $(d_m)_{m=0}^L$ be the difference sequence associated
with the finite martingale $\mathbb{E}(f_{\boldsymbol{e}} \,|\, \mathcal{A}_{\mathcal{P}_0^e}), \ldots, \mathbb{E}(f_{\boldsymbol{e}} \,|\, \mathcal{A}_{\mathcal{P}_L^e})$ and notice that $\|d_m\|_{L_{p^\dagger}} > \sigma$
for every $m \in I$. Moreover, since $f_{\boldsymbol{e}}$ is (C, η, p)-regular and $\iota(\mathcal{P}_m^{\boldsymbol{e}}) \geqslant \eta_m \geqslant \eta_L \geqslant \eta$
we see that $\|\mathbb{E}(f_{\boldsymbol{e}} \,|\, \mathcal{A}_{\mathcal{P}_m^e})\|_{L_{p^\dagger}} \leqslant \|\mathbb{E}(f_{\boldsymbol{e}} \,|\, \mathcal{A}_{\mathcal{P}_m^e})\|_{L_p} \leqslant C$ for every $m \in [L]$. Hence, by
the choice of L in (5.17) and Proposition 3.1, we conclude that

$$C < (p^\dagger - 1)^{1/2} \Big(\sum_{m=0}^L \|d_m\|_{L_{p^\dagger}}^2 \Big)^{1/2} \leqslant \big\| \sum_{m=0}^L d_m \big\|_{L_{p^\dagger}} = \|\mathbb{E}(f_{\boldsymbol{e}} \,|\, \mathcal{A}_{\mathcal{P}_L^e})\|_{L_{p^\dagger}} \leqslant C$$

which is clearly a contradiction. □

We are ready to complete the proof of Theorem 2.1.

Proof of Theorem 2.1. Let $F \colon \mathbb{N} \to \mathbb{R}$ be a growth function and $0 < \sigma \leqslant 1$, and let
L and η_L be as in (5.17) and (5.16) respectively. We set $\mathrm{Reg} = \lceil \eta_L^{-1} \rceil$ and we claim
that with this choice the result follows. Indeed, set $\eta := 1/\mathrm{Reg} \leqslant \eta_L$ and assume
that $\mathscr{H}$ is η-nonatomic. For every $e \in \mathcal{H}$ let $f_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ be nonnegative and
(C, η, p)-regular. Let $m \in \{0, \ldots, L-1\}$, $\langle \mathcal{P}_e : e \in \mathcal{H} \rangle$ and $\langle \mathcal{Q}_e : e \in \mathcal{H} \rangle$ be as in
Lemma 5.4 and define $M = \lceil \eta_m^{-1} \rceil$. It is clear that M, $\langle \mathcal{P}_e : e \in \mathcal{H} \rangle$ and $\langle \mathcal{Q}_e : e \in \mathcal{H} \rangle$
are as desired. □

6. Pseudorandom families

Throughout this section let n, r be two positive integers with $n \geqslant r \geqslant 2$ and let
$\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$ be an r-uniform hypergraph system.

6.1. **Definitions and basic properties.** We begin by introducing some pieces
of notation. Let $e \subseteq [n]$ be nonempty and recall that $\pi_e \colon \boldsymbol{X} \to \boldsymbol{X}_e$ stands for the
natural projection. Observe that for every $f \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ there exists a unique
random variable $\mathbf{f} \in L_1(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ such that

$$(6.1) \qquad\qquad\qquad f = \mathbf{f} \circ \pi_e$$

and note that the map $L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu}) \ni f \mapsto \mathbf{f} \in L_1(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ is a linear isometry.
If, in addition, the set $[n] \setminus e$ is nonempty, then for every $\mathbf{x}_e \in \boldsymbol{X}_e$ and every
$\mathbf{x}_{[n] \setminus e} \in \boldsymbol{X}_{[n] \setminus e}$ by $(\mathbf{x}_e, \mathbf{x}_{[n] \setminus e})$ we denote the unique element $\mathbf{x}$ of $\boldsymbol{X}$ such that
$\mathbf{x}_e = \pi_e(\mathbf{x})$ and $\mathbf{x}_{[n] \setminus e} = \pi_{[n] \setminus e}(\mathbf{x})$. Finally, for every $f \colon \boldsymbol{X} \to \mathbb{R}$ and every $\mathbf{x}_e \in \boldsymbol{X}_e$
let $f_{\mathbf{x}_e} \colon \boldsymbol{X}_{[n] \setminus e} \to \mathbb{R}$ be the section of f at $\mathbf{x}_e$, that is, $f_{\mathbf{x}_e}(\mathbf{x}_{[n] \setminus e}) = f(\mathbf{x}_e, \mathbf{x}_{[n] \setminus e})$.

Definition 6.1. *Let $C \geqslant 1$ and $0 < \eta < 1$. Also let $1 < p \leqslant \infty$ and let q denote the
conjugate exponent of p. For every $e \in \mathcal{H}$ let $\nu_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ be a nonnegative
random variable. We say that the family $\langle \nu_e : e \in \mathcal{H} \rangle$ is (C, η, p)-pseudorandom
provided that the following conditions are satisfied.*

(C1) (Copies of sub-hypergraphs of $\mathcal{H}$) *For every nonempty $\mathcal{G} \subseteq \mathcal{H}$ we have*

$$(6.2) \qquad \int \prod_{e \in \mathcal{G}} \nu_e \, d\boldsymbol{\mu} \geqslant 1 - \eta.$$

(C2) *For every $e \in \mathcal{H}$ there exists $\psi_e \in L_p(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ with $\|\psi_e\|_{L_p} \leqslant C$ and satisfying the following properties.*

(a) (The cut norm of $\nu_e - \psi_e$ is negligible) *We have $\|\nu_e - \psi_e\|_{\mathcal{S}_{\partial e}} \leqslant \eta$.*

(b) (Local linear forms) *For every $j \in \{0,1\}$ and every $e' \in \mathcal{H} \setminus \{e\}$ let $g_{e'}^{(j)} \in L_1(\boldsymbol{X}, \mathcal{B}_{e'}, \boldsymbol{\mu})$ such that either $0 \leqslant g_{e'}^{(j)} \leqslant \nu_{e'}$ or $0 \leqslant g_{e'}^{(j)} \leqslant 1$. If $\boldsymbol{\nu}_e$ and $\boldsymbol{\psi}_e$ are as in (6.1) for ν_e and ψ_e respectively, then*

$$(6.3) \qquad \Big| \int (\boldsymbol{\nu}_e - \boldsymbol{\psi}_e)(\mathbf{x}_e) \prod_{j \in \{0,1\}} \Big(\int \prod_{e' \in \mathcal{H} \setminus \{e\}} (g_{e'}^{(j)})_{\mathbf{x}_e} \, d\boldsymbol{\mu}_{[n] \setminus e} \Big) d\boldsymbol{\mu}_e \Big| \leqslant \eta.$$

(C3) (Integrability of marginals) *Let $e \in \mathcal{H}$ and let $\mathcal{G} \subseteq \mathcal{H} \setminus \{e\}$ be nonempty, and define $\boldsymbol{\nu}_{e,\mathcal{G}} \colon \boldsymbol{X}_e \to \mathbb{R}$ by $\boldsymbol{\nu}_{e,\mathcal{G}}(\mathbf{x}_e) = \int \prod_{e' \in \mathcal{G}} (\nu_{e'})_{\mathbf{x}_e} \, d\boldsymbol{\mu}_{[n] \setminus e}$. Then, setting*

$$(6.4) \qquad \ell := \min \Big\{ 2n : n \in \mathbb{N} \text{ and } 2n \geqslant 2q + \Big(1 - \frac{1}{C} \Big) + \frac{1}{p} \Big\}$$

(where $1/p = 0$ if $p = \infty$), we have

$$(6.5) \qquad \int \boldsymbol{\nu}_{e,\mathcal{G}}^{\ell} \, d\boldsymbol{\mu}_e \leqslant C + \eta.$$

As we have noted, examples of pseudorandom families are presented in [10]. Here, as an illustration, we will describe a class of these examples which are quite easy to grasp. (The verification of their properties, although not particularly difficult, is somewhat lengthy and is given in [10].) They are of the form $\langle \nu_e + \varphi_e : e \in \mathcal{H} \rangle$ where: (i) $\mathcal{H}$ is the r-simplex, that is, the complete r-uniform hypergraph on $[r+1]$, (ii) each $\nu_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ is nonnegative and the family $\langle \nu_e : e \in \mathcal{H} \rangle$ satisfies a slight modification of the "linear forms condition" introduced in [34, Definition 2.8], and (iii) each $\varphi_e \in L_p(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ is also nonnegative and satisfies $\|\varphi_e^p\|_{\Box_\ell(\boldsymbol{X}_e)} \leqslant C^p$ where ℓ is as in (6.4) and $\| \cdot \|_{\Box_\ell(\boldsymbol{X}_e)}$ is a variant of the Gowers box norm [13] which was first considered by Hatami [17]. In other words, the family $\langle \nu_e + \varphi_e : e \in \mathcal{H} \rangle$ is a perturbation of a system of measures which appears in [34], the main point being that only integrability conditions are imposed on each "noise" φ_e.

We now proceed to discuss some properties of pseudorandom families.

6.1.1. First observe that condition (C1) expresses a natural combinatorial requirement, namely that the weighted hypergraph $\langle \nu_e : e \in \mathcal{H} \rangle$ contains many copies of every sub-hypergraph of $\mathcal{H}$.

6.1.2. Condition (C2.a) is also rather mild and implies that each ν_e is, to some extend, well-behaved. Specifically, we have the following lemma.

Lemma 6.2. *If the family $\langle \nu_e : e \in \mathcal{H} \rangle$ satisfies condition (C2.a), then for every $e \in \mathcal{H}$ the random variable ν_e is $(C + 1, \eta, p)$-regular.*

Proof. Let $e \in \mathcal{H}$ and let $\mathcal{P}$ be a partition of $\boldsymbol{X}$ with $\mathcal{P} \subseteq \mathcal{S}_{\partial e}$ and $\boldsymbol{\mu}(P) \geqslant \eta$ for every $P \in \mathcal{P}$. By condition (C2.a), for every $P \in \mathcal{P}$ we have

$$\frac{|\int_P (\nu_e - \psi_e)\,d\boldsymbol{\mu}|}{\boldsymbol{\mu}(P)} \leqslant 1$$

and, consequently, $\|\mathbb{E}(\nu_e - \psi_e \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_\infty} \leqslant 1$. Therefore, by the triangle inequality and the monotonicity of the L_p norms, we conclude that

$$\|\mathbb{E}(\nu_e \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} \leqslant \|\mathbb{E}(\psi_e \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} + \|\mathbb{E}(\nu_e - \psi_e \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} \leqslant C + 1$$

and the proof is completed. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

Condition (C2.b), the local linear forms condition, is the strongest (and as such, the most restrictive) condition of all. In the case where $\psi_e = 1$ for every $e \in \mathcal{H}$ it was explicitly isolated[5] in [6, Lemma 6.3], though closely related variants had appeared earlier in the work of Green and Tao [15]. One of the signs of the strength of the local linear forms condition is that it implies condition (C2.a) as long as the hypergraph $\mathcal{H}$ is not too sparse. More precisely, assume that for every $e \in \mathcal{H}$ we have $\partial e \subseteq \{e' \cap e : e' \in \mathcal{H}\}$ (this is the case, for instance, if $\mathcal{H}$ is the r-simplex). Fix $e \in \mathcal{H}$ and for every $f \in \partial e$ let $A_f \in \mathcal{B}_f$. We set $g_{e'}^{(0)} = \mathbf{1}_{A_f}$ if $e' \cap e = f$; otherwise, let $g_{e'}^{(j)} = 1$. By (6.3), we see that $|\int (\nu_e - \psi_e) \prod_{f \in \partial e} \mathbf{1}_{A_f}\,d\boldsymbol{\mu}| \leqslant \eta$ which implies, of course, that $\|\nu_e - \psi_e\|_{\mathcal{S}_{\partial e}} \leqslant \eta$.

6.1.3. Condition (C3) can be seen as an instance of the general fact that by taking averages we improve integrability. In a nutshell, we demand that the marginal $\boldsymbol{\nu}_{e,\mathcal{G}}$ belongs to $L_\ell(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ where ℓ is an even integer strictly bigger than $2q$, though in the special case $C = 1$ and $p = \infty$ we can simply take $\ell = 2$. The exact choice of ℓ in (6.4) expresses these requirements in a formal way, but the reader may actually think of[6] the parameter ℓ as being equal to $2(q + 1)$.

The following lemma reduces condition (C3) to certain moment estimates. We point out that these estimates are needed for the analysis of pseudorandom families.

Lemma 6.3. *If the family $\langle \nu_e : e \in \mathcal{H} \rangle$ satisfies condition* (C3), *then for every $e \in \mathcal{H}$ and every nonempty $\mathcal{G} \subseteq \mathcal{H} \setminus \{e\}$ the following hold.*

(a) *If either $C > 1$ or $1 < p < \infty$, then $\ell > 2q$ and for every $\mathbf{A} \in \boldsymbol{\Sigma}_e$ we have*

$$(6.6) \qquad\qquad \int_{\mathbf{A}} \boldsymbol{\nu}_{e,\mathcal{G}}^{2q}\,d\boldsymbol{\mu}_e \leqslant (C + 1)\,\boldsymbol{\mu}_e(\mathbf{A})^{e(C,p)}$$

where $e(C, p) = (4pq)^{-1}$ if $1 < p < \infty$, and $e(C, \infty) = 1/2$ if $C > 1$.

---

[5]We remark that in [6] condition (C2.b) is referred to as the "strong linear forms" condition.

[6]Note that this simplification is not far from the truth. Indeed, if q is an integer and $C > 1$, then, by (6.4), we have that $\ell = 2(q + 1)$. It is also useful to point out that in the proofs of Theorems 2.2 and 7.1 (which are our main results concerning pseudorandom families), we will not use the fact that the number ℓ in (6.4) is an even integer. However, this property is needed when one is working with the Gowers box norms and their variants (see, e.g., [10, Proposition 2.1]), and so it is reasonable to add this harmless condition in Definition 6.1.

(b) *Assume that the family $\langle \nu_e : e \in \mathcal{H} \rangle$ also satisfies condition* (C1), *and that $C = 1$ and $p = \infty$. Then $\ell = 2$ and $\|\boldsymbol{\nu}_{e,\mathcal{G}} - 1\|_{L_2}^2 \leqslant 4\eta^{1/2}$. In particular, for every $\mathbf{A} \in \boldsymbol{\Sigma}_e$ we have*

$$(6.7) \qquad \int_{\mathbf{A}} \boldsymbol{\nu}_{e,\mathcal{G}}^2 \, d\boldsymbol{\mu}_e \leqslant 2\boldsymbol{\mu}_e(\mathbf{A}) + 8\eta^{1/2}.$$

Proof. (a) The fact that $\ell > 2q$ follows immediately by (6.4). Next, fix $\mathbf{A} \in \boldsymbol{\Sigma}_e$. By Hölder's inequality, we have

$$(6.8) \qquad \int_{\mathbf{A}} \boldsymbol{\nu}_{e,\mathcal{G}}^{2q} \, d\boldsymbol{\mu}_e \leqslant \|\boldsymbol{\nu}_{e,\mathcal{G}}\|_{L_\ell}^{2q} \cdot \boldsymbol{\mu}_e(\mathbf{A})^{1 - \frac{2q}{\ell}} \overset{(6.5)}{\leqslant} (C + 1) \, \boldsymbol{\mu}_e(\mathbf{A})^{1 - \frac{2q}{\ell}}.$$

On the other hand, by (6.4) and the choice of $e(C, p)$, we see that $1 - \frac{2q}{\ell} \geqslant e(C, p)$. By (6.8), the proof of part (a) is completed.

(b) First observe that $\ell = 2$. Moreover, by Fubini's theorem and Jensen's inequality,

$$1 - \eta \overset{(C1)}{\leqslant} \int \prod_{e' \in \mathcal{G}} \nu_{e'} \, d\boldsymbol{\mu} = \int \boldsymbol{\nu}_{e,\mathcal{G}} \, d\boldsymbol{\mu}_e \leqslant \Big(\int \boldsymbol{\nu}_{e,\mathcal{G}}^2 \, d\boldsymbol{\mu}_e \Big)^{1/2} \overset{(C3)}{\leqslant} (1 + \eta)^{1/2}$$

and, consequently, $|\int (\boldsymbol{\nu}_{e,\mathcal{G}}^2 - 1) \, d\boldsymbol{\mu}_e| \leqslant 2\eta$ and $|\int (\boldsymbol{\nu}_{e,\mathcal{G}} - 1) \, d\boldsymbol{\mu}_e| \leqslant \eta^{1/2}$. Therefore,

$$
\begin{aligned}
(6.9) \quad \|\boldsymbol{\nu}_{e,\mathcal{G}} - 1\|_{L_2}^2 \;&=\; \int (\boldsymbol{\nu}_{e,\mathcal{G}}^2 - 2\boldsymbol{\nu}_{e,\mathcal{G}} + 1) \, d\boldsymbol{\mu}_e \\
&\leqslant\; \Big| \int (\boldsymbol{\nu}_{e,\mathcal{G}}^2 - 1) \, d\boldsymbol{\mu}_e \Big| + 2 \Big| \int (\boldsymbol{\nu}_{e,\mathcal{G}} - 1) \, d\boldsymbol{\mu}_e \Big| \leqslant 4\eta^{1/2}.
\end{aligned}
$$

Now let $\mathbf{A} \in \boldsymbol{\Sigma}_e$ and note that $\|\boldsymbol{\nu}_{e,\mathcal{G}} \cdot \mathbf{1_A} - \mathbf{1_A}\|_{L_2} \leqslant \|\boldsymbol{\nu}_{e,\mathcal{G}} - 1\|_{L_2}$. Hence, by (6.9) and the triangle inequality, we have $\|\boldsymbol{\nu}_{e,\mathcal{G}} \cdot \mathbf{1_A}\|_{L_2} \leqslant \|\mathbf{1_A}\|_{L_2} + (4\eta^{1/2})^{1/2}$ and so

$$\int_{\mathbf{A}} \boldsymbol{\nu}_{e,\mathcal{G}}^2 \, d\boldsymbol{\mu}_e \leqslant \big(\boldsymbol{\mu}_e(\mathbf{A})^{1/2} + (4\eta^{1/2})^{1/2} \big)^2 \leqslant 2\boldsymbol{\mu}_e(\mathbf{A}) + 8\eta^{1/2}$$

as desired. $\qquad\square$

6.2. **Conditions on the majorants.** Although for the analysis of pseudorandom families we need precisely conditions (C1)–(C3), in practice some of these conditions are not so easily checked. This is the case, for instance, with the local linear forms condition, since it requires verifying the estimate in (6.3) not only for the "majorants" $\langle \nu_e : e \in \mathcal{H} \rangle$ but also for all nonnegative functions which are pointwise bounded by them. However, this problem can be effectively resolved by imposing some slightly stronger conditions on $\langle \nu_e : e \in \mathcal{H} \rangle$, and then reducing (6.3) to these conditions by repeated applications of the Cauchy–Schwarz inequality. This method was developed extensively by Green and Tao [15, 16] and has become standard in the field. As such, our presentation in this subsection will be rather brief.

First we need to introduce some notation. Recall that by n, r we denote two positive integers with $n \geqslant r \geqslant 2$ and $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$ stands for an r-uniform hypergraph system. For every $m \in \mathbb{N}$ with $m \geqslant 2$ and every nonempty $e \subseteq [n]$ let $(\boldsymbol{X}_e^m, \boldsymbol{\Sigma}_e^m, \boldsymbol{\mu}_e^m)$ be the product of m copies of the probability space $(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$. (If $e = [n]$, then this product will be denoted by $(\boldsymbol{X}^m, \boldsymbol{\Sigma}^m, \boldsymbol{\mu}^m)$.)

Next, let $h \subseteq [n]$ with $e \subseteq h$. For every $\mathbf{x}_h^{(0)} = (x_i^{(0)})_{i \in h}, \ldots, \mathbf{x}_h^{(m-1)} = (x_i^{(m-1)})_{i \in h}$ in $\boldsymbol{X}_h$ and every $\omega = (\omega_i)_{i \in e} \in \{0, \ldots, m-1\}^e$ we set

$$\mathbf{x}_e^{(\omega)} = (x_i^{(\omega_i)})_{i \in e} \in \boldsymbol{X}_e.$$

Note that if $\omega = j^e$ for some $j \in \{0, \ldots, m-1\}$ (that is, $\omega = (\omega_i)_{i \in e}$ with $\omega_i = j$ for every $i \in e$), then $\mathbf{x}_e^{(\omega)} = \pi_{h,e}(\mathbf{x}_h^{(j)})$ where $\pi_{h,e} \colon \boldsymbol{X}_h \to \boldsymbol{X}_e$ is the natural projection.

Proposition 6.4. *Let $C \geqslant 1$ and $0 < \eta < 1$. Also let $1 < p \leqslant \infty$ and let q denote the conjugate exponent of p. For every $e \in \mathcal{H}$ let $\nu_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ be a nonnegative random variable and let $\boldsymbol{\nu}_e$ be as in (6.1) for ν_e. Assume that the following properties are satisfied.*

(P1) *If ℓ is as in (6.4), then*

$$1 - \eta \leqslant \int \prod_{e \in \mathcal{H}} \prod_{\omega \in \{0,\ldots,\ell-1\}^e} \boldsymbol{\nu}_e^{n_{e,\omega}}(\mathbf{x}_e^{(\omega)}) \, d\boldsymbol{\mu}^{\ell} \leqslant C + \eta$$

for any choice of $n_{e,\omega} \in \{0, 1\}$.

(P2) *For every $e \in \mathcal{H}$ there exists $\psi_e \in L_p(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ with $\|\psi_e\|_{L_p} \leqslant C$ such that*

$$\Big| \int \prod_{\omega \in \{0,1\}^e} (\boldsymbol{\nu}_e - \boldsymbol{\psi}_e)(\mathbf{x}_e^{(\omega)}) \prod_{e' \in \mathcal{H} \setminus \{e\}} \prod_{\omega \in \{0,1\}^{e'}} \boldsymbol{\nu}_{e'}^{n_{e',\omega}}(\mathbf{x}_{e'}^{(\omega)}) \, d\boldsymbol{\mu}^2 \Big| \leqslant \eta$$

for any choice of $n_{e',\omega} \in \{0, 1\}$.

Then $\langle \nu_e : e \in \mathcal{H} \rangle$ is a (C, η', p)-pseudorandom family where $\eta' = (C+1) \eta^{1/2^r}$.

For comparison, it is useful at this point to recall that a family $\langle \nu_e : e \in \mathcal{H} \rangle$ satisfies the "linear forms condition" in the sense of [34, Definition 2.8] if we have

(6.10)
$$\int \prod_{e \in \mathcal{H}} \prod_{\omega \in \{0,1\}^e} \boldsymbol{\nu}_e^{n_{e,\omega}}(\mathbf{x}_e^{(\omega)}) \, d\boldsymbol{\mu}^2 = 1 + o(1)$$

for any choice of $n_{e,\omega} \in \{0, 1\}$. (Here, the quantity $o(1)$ is an error term that can be made arbitrarily small.) Thus, we see that if $C = 1$, $p = \infty$ and $\psi_e = 1$ for every $e \in \mathcal{H}$, then properties (P1) and (P2) in Proposition 6.4 are equivalent to (6.10).

For the proof of Proposition 6.4 we need the following lemmas. The first one is a straightforward consequence of the Gowers–Cauchy–Schwarz inequality (see, e.g., [16, Lemma B.2]).

Lemma 6.5. *Let $e \in \mathcal{H}$ and $f \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$. If $\mathbf{f}$ is as in (6.1) for f, then*

$$\|f\|_{\mathcal{S}_{\partial e}} \leqslant \Big(\int \prod_{\omega \in \{0,1\}^e} \mathbf{f}(\mathbf{x}_e^{(\omega)}) \, d\boldsymbol{\mu}_e^2 \Big)^{1/2^r}.$$

The second lemma follows arguing precisely as in the proof of [6, Lemma 6.3] or [34, Proposition 5.1]. We leave the details to the interested reader.

Lemma 6.6. *Let $C \geqslant 1$, $0 < \eta < 1$ and $1 < p \leqslant \infty$. For every $e \in \mathcal{H}$ let $\nu_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ be a nonnegative random variable and assume that the family $\langle \nu_e : e \in \mathcal{H} \rangle$ satisfies properties* (P1) *and* (P2) *in Proposition* 6.4. *Fix $e \in \mathcal{H}$, and for every $j \in \{0,1\}$ and every $e' \in \mathcal{H} \setminus \{e\}$ let $g_{e'}^{(j)} \in L_1(\boldsymbol{X}, \mathcal{B}_{e'}, \boldsymbol{\mu})$ such that either $0 \leqslant g_{e'}^{(j)} \leqslant \nu_{e'}$ or $0 \leqslant g_{e'}^{(j)} \leqslant 1$. Then we have*

$$\Big| \int (\boldsymbol{\nu}_e - \boldsymbol{\psi}_e)(\mathbf{x}_e) \prod_{j \in \{0,1\}} \Big(\int \prod_{e' \in \mathcal{H} \setminus \{e\}} (g_{e'}^{(j)})_{\mathbf{x}_e} \, d\boldsymbol{\mu}_{[n] \setminus e} \Big) \, d\boldsymbol{\mu}_e \Big| \leqslant (C+1) \, \eta^{1/2^r}.$$

That is, condition (C2.b) *is satisfied for the parameter $(C+1) \, \eta^{1/2^r}$.*

We are ready to give the proof of Proposition 6.4.

Proof of Proposition 6.4. Let $\mathcal{G} \subseteq \mathcal{H}$ be nonempty and observe that

$$\int \prod_{e \in \mathcal{G}} \nu_e \, d\boldsymbol{\mu} = \int \prod_{e \in \mathcal{H}} \prod_{\omega \in \{0,\dots,\ell-1\}^e} \boldsymbol{\nu}_e^{n_{e,\omega}}(\mathbf{x}_e^{(\omega)}) \, d\boldsymbol{\mu}^\ell$$

where $n_{e,\omega} \in \{0,1\}$ and $n_{e,\omega} = 1$ if and only if $e \in \mathcal{G}$ and $\omega = 0^e$. Hence, by property (P1), condition (C1) is satisfied.

Next, let $e \in \mathcal{H}$ by arbitrary. By property (P2), we have

$$\int \prod_{\omega \in \{0,1\}^e} (\boldsymbol{\nu}_e - \boldsymbol{\psi}_e)(\mathbf{x}_e^{(\omega)}) \, d\boldsymbol{\mu}_e^2 \leqslant \eta$$

and so, by Lemma 6.5, we see that condition (C2.a) is satisfied. Condition (C2.b) follows, of course, from Lemma 6.6.

It remains to verify condition (C3). To this end, fix $e \in \mathcal{H}$ and let $\mathcal{G}$ be a nonempty subset of $\mathcal{H} \setminus \{e\}$. Then,

$$\begin{aligned} \int \boldsymbol{\nu}_{e,\mathcal{G}}^\ell \, d\boldsymbol{\mu}_e &= \int \Big(\int \prod_{e' \in \mathcal{G}} \boldsymbol{\nu}_{e'}(\mathbf{x}_{e' \cap e}, \mathbf{x}_{e' \setminus e}) \, d\boldsymbol{\mu}_{[n] \setminus e} \Big)^\ell \, d\boldsymbol{\mu}_e \\ &= \int \Big(\int \prod_{e' \in \mathcal{G}} \prod_{j=0}^{\ell-1} \boldsymbol{\nu}_{e'}(\mathbf{x}_{e' \cap e}, \mathbf{x}_{e' \setminus e}^{(j^{e' \setminus e})}) \, d\boldsymbol{\mu}_{[n] \setminus e}^\ell \Big) \, d\boldsymbol{\mu}_e \\ &= \int \prod_{e' \in \mathcal{G}} \prod_{\omega \in \{0,\dots,\ell-1\}^{e'}} \boldsymbol{\nu}_{e'}^{n_{e',\omega}}(\mathbf{x}_{e'}^{(\omega)}) \, d\boldsymbol{\mu}^\ell \end{aligned}$$

where $n_{e',\omega} \in \{0,1\}$ and $n_{e',\omega} = 1$ if and only if $\omega \in \{(0^{e' \cap e}, j^{e' \setminus e}) : 0 \leqslant j \leqslant \ell - 1\}$. By property (P1), we conclude that condition (C3) is satisfied, and the proof of Proposition 6.4 is completed. $\qquad \square$

7. Relative counting lemma for pseudorandom families

The following theorem is the main result in this section.

Theorem 7.1. *Let $n, r \in \mathbb{N}$ with $n \geqslant r \geqslant 2$, and let $C \geqslant 1$ and $1 < p \leqslant \infty$. Also let $\zeta \geqslant 1$ and $0 < \gamma \leqslant 1$. Then there exist two strictly positive constants $\eta = \eta(n, r, C, p, \zeta, \gamma)$ and $\alpha = \alpha(n, r, C, p, \zeta, \gamma)$ with the following property. Let*

$\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) \colon i \in [n] \rangle, \mathcal{H})$ *be an r-uniform hypergraph system, and let*
$\langle \nu_e : e \in \mathcal{H} \rangle$ *be a (C, η, p)-pseudorandom family. Moreover, for every $e \in \mathcal{H}$ let*
$g_e, h_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ *such that $0 \leqslant g_e \leqslant \nu_e$, $0 \leqslant h_e \leqslant \zeta$ and $\|g_e - h_e\|_{\mathcal{S}_{\partial e}} \leqslant \alpha$.*
Then we have

$$(7.1) \qquad \Big| \int \prod_{e \in \mathcal{H}} g_e \, d\boldsymbol{\mu} - \int \prod_{e \in \mathcal{H}} h_e \, d\boldsymbol{\mu} \Big| \leqslant \gamma.$$

The hypotheses of Theorem 7.1 might appear rather strong: on the one hand the function g_e is dominated by ν_e (and so, by Lemma 6.2, it is L_p regular), but on the other hand it is approximated in the cut norm by a nonnegative function h_e with $\|h_e\|_{L_\infty} \leqslant \zeta$. It turns out, however, that for *every* $0 \leqslant f_e \leqslant \nu_e$ we can indeed satisfy these requirements by slightly truncating f_e. More precisely, we have the following proposition.

Proposition 7.2. *Let n, r, C, p and $\mathscr{H}$ be as in Theorem 7.1, and let M be a positive integer, $0 < \alpha \leqslant 1$ and $e \in \mathcal{H}$. Also let $\mathcal{P}_e$ be a partition of $\boldsymbol{X}$ with $\mathcal{P}_e \subseteq \mathcal{S}_{\partial e}$ and $\boldsymbol{\mu}(P) \geqslant 1/M$ for every $P \in \mathcal{P}_e$, and let $\mathcal{Q}_e$ be a finite refinement of $\mathcal{P}_e$ with $\mathcal{Q}_e \subseteq \mathcal{S}_{\partial e}$. Finally, let $f_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ be nonnegative and write $f_e = f_{\mathrm{str}}^e + f_{\mathrm{err}}^e + f_{\mathrm{unf}}^e$ where $f_{\mathrm{str}}^e, f_{\mathrm{err}}^e$ and f_{unf}^e are as in (2.2). Assume that the estimates in (2.3) are satisfied for $\sigma = \alpha/2$ and a growth function $F \colon \mathbb{N} \to \mathbb{R}$ with $F(m) \geqslant 2\alpha^{-1}m$ for every $m \in \mathbb{N}$. Then the following hold.*

(a) *For every $A \in \mathcal{A}_{\mathcal{P}_e}$ we have $\|f_e \cdot \boldsymbol{1}_A - f_{\mathrm{str}}^e \cdot \boldsymbol{1}_A\|_{\mathcal{S}_{\partial e}} \leqslant \alpha$.*

(b) *Assume that $1 < p < \infty$. Let $\zeta \geqslant 1$ and set $A = [f_{\mathrm{str}}^e \leqslant \zeta]$. Then we have $A \in \mathcal{A}_{\mathcal{P}_e}$ and $\boldsymbol{\mu}(\boldsymbol{X} \setminus A) \leqslant (C/\zeta)^p$. Moreover,*

$$(7.2) \qquad \int_{\boldsymbol{X} \setminus A} f_e \, d\boldsymbol{\mu} \leqslant C^p \zeta^{1-p} + \alpha \quad and \quad \int_{\boldsymbol{X} \setminus A} f_{\mathrm{str}}^e \, d\boldsymbol{\mu} \leqslant C^p \zeta^{1-p}.$$

The rest of this section is devoted to the proof of Theorem 7.1. As we have mentioned in the introduction, it is based on a decomposition method which is discussed in Subsection 7.1. We also note that Proposition 7.2 is not used in the argument, though it is needed for the proof of Theorem 2.2. As such, we defer its proof to Subsection 8.1.

7.1. **Proof of Theorem 7.1.** First we need to do some preparatory work. Let $n, r \in \mathbb{N}$ with $n \geqslant r \geqslant 2$, and let $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) \colon i \in [n] \rangle, \mathcal{H})$ be an r-uniform hypergraph system. Also let $C \geqslant 1$ and $1 < p \leqslant \infty$, and denote by q the conjugate exponent of p. *These data will be fixed throughout the proof.*

Next, observe that it suffices to prove Theorem 7.1 only for the case "$\zeta = 1$". Indeed, if the numbers $\eta(n, r, C, p, 1, \gamma)$ and $\alpha(n, r, C, p, 1, \gamma)$ have been determined, then it is easy to see that for every $\zeta \geqslant 1$ Theorem 7.1 holds true for the parameters $\eta(n, r, C, p, 1, \gamma\zeta^{-n^r})$ and $\zeta \cdot \alpha(n, r, C, p, 1, \gamma\zeta^{-n^r})$. Thus, in what follows we will assume that $\zeta = 1$. To avoid trivialities, we will also assume that $|\mathcal{H}| \geqslant 2$.

We proceed to introduce some numerical invariants. For every $0 < \gamma \leqslant 1$ we set

$$(7.3) \qquad \beta(\gamma) = \left(10(C+1)^2 \gamma^{-1}\right)^{2q/x(C,p)} \text{ and } \theta(\gamma) = \left(20(C+1)\beta(\gamma)\right)^{-2q} \gamma^{2q}$$

where $x(C,p) = (4pq)^{-1}$ if $1 < p < \infty$, $x(C,\infty) = 1/2$ if $C > 1$, and $x(1,\infty) = 1$. Moreover, for every $m \in \{0, \ldots, n^r\}$ and every $0 < \gamma \leqslant 1$ we define $\alpha_m(\gamma)$ and $\eta_m(\gamma)$ in $(0,1]$ recursively by the rule

$$(7.4) \qquad\qquad \alpha_0(\gamma) = \gamma/5 \text{ and } \alpha_{m+1}(\gamma) = \alpha_m\left(\theta(\gamma)\right)$$

and

$$(7.5) \qquad\qquad \eta_0(\gamma) = \left(30(C+1)\right)^{-4q} \gamma^{4q} \text{ and } \eta_{m+1}(\gamma) = \eta_m\left(\theta(\gamma)\right).$$

Notice that $\alpha_{m+1}(\gamma) \leqslant \alpha_m(\gamma)$ and $\eta_{m+1}(\gamma) \leqslant \eta_m(\gamma)$ for every $0 < \gamma \leqslant 1$.

After this preliminary discussion we are ready to enter into the main part of the proof which proceeds by induction. Specifically, let $\langle \nu_e : e \in \mathcal{H} \rangle$ be a family of nonnegative random variables such that $\nu_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ for every $e \in \mathcal{H}$. By induction on $m \in \{0, \ldots, |\mathcal{H}|\}$ we will show that for every $0 < \gamma \leqslant 1$ if the family $\langle \nu_e : e \in \mathcal{H} \rangle$ is $(C, \eta_m(\gamma), p)$-pseudorandom where $\eta_m(\gamma)$ is as in (7.5), then the estimate (7.1) is satisfied for any collection $\langle g_e, h_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu}) : e \in \mathcal{H} \rangle$ with the following properties: (P1) for every $e \in \mathcal{H}$ we have that either $0 \leqslant g_e \leqslant \nu_e$ or $g_e = h_e$, (P2) for every $e \in \mathcal{H}$ we have $0 \leqslant h_e \leqslant 1$ and $\|g_e - h_e\|_{\mathcal{S}_{\partial e}} \leqslant \alpha_m(\gamma)$ where $\alpha_m(\gamma)$ is as in (7.4), and (P3) $|\{e \in \mathcal{H} : g_e \neq h_e\}| \leqslant m$.

The initial case "$m = 0$" is straightforward, and so let $m \in \{1, \ldots, |\mathcal{H}|\}$ and assume that the induction has been carried out up to $m - 1$. Fix $0 < \gamma \leqslant 1$ and let $\langle g_e, h_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu}) : e \in \mathcal{H} \rangle$ be a collection satisfying properties (P1)–(P3). Set

$$\Delta := \int \prod_{e \in \mathcal{H}} g_e \, d\boldsymbol{\mu} - \int \prod_{e \in \mathcal{H}} h_e \, d\boldsymbol{\mu}$$

and recall that we need to show that $|\Delta| \leqslant \gamma$. To this end, we may assume that $|\{e \in \mathcal{H} : g_e \neq h_e\}| = m$ (otherwise, the desired estimate follows immediately from the inductive assumptions). Thus, we may select $e_0 \in \mathcal{H}$ with $g_{e_0} \neq h_{e_0}$; note that, by property (P1), we have $0 \leqslant g_{e_0} \leqslant \nu_{e_0}$. We set $\mathcal{G} = \{e \in \mathcal{H} \setminus \{e_0\} : g_e \neq h_e\}$ and we define $G, H : \boldsymbol{X}_{e_0} \to \mathbb{R}$ by the rule

$$G(\mathbf{x}_{e_0}) = \int \prod_{e \in \mathcal{H} \setminus \{e_0\}} (g_e)_{\mathbf{x}_{e_0}} \, d\boldsymbol{\mu}_{[n] \setminus e_0} \text{ and } H(\mathbf{x}_{e_0}) = \int \prod_{e \in \mathcal{H} \setminus \{e_0\}} (h_e)_{\mathbf{x}_{e_0}} \, d\boldsymbol{\mu}_{[n] \setminus e_0}.$$

Observe that $0 \leqslant H \leqslant 1$. Moreover, if $\mathcal{G}$ is nonempty, then we have $0 \leqslant G \leqslant \boldsymbol{\nu}_{e_0, \mathcal{G}}$ where $\boldsymbol{\nu}_{e_0, \mathcal{G}}$ is as in Definition 6.1. On the other hand, notice that $G = H$ if $\mathcal{G} = \emptyset$. In the following claim we obtain a first estimate for $|\Delta|$.

Claim 7.3. *We have*

$$(7.6) \qquad |\Delta| \leqslant 2(C+1)\left(\|G - H\|_{L_{2q}} + \eta_m(\gamma)^{1/2}\right) + \|g_{e_0} - h_{e_0}\|_{\mathcal{S}_{\partial e_0}}.$$

The proof of Claim 7.3 is given in Subsection 7.2. It is important to note that it does not use the inductive assumptions and relies, instead, on the local linear forms condition (condition (C2.b) in Definition 6.1) and Hölder's inequality. Closely related estimates appear in [6, 34].

The next claim is the second step of the proof.

Claim 7.4. *If $\beta(\gamma)$ and $\theta(\gamma)$ are as in (7.3), then we have*

$$(7.7) \qquad \int (G - H)^{2q} \, d\boldsymbol{\mu}_{e_0} \leqslant 2\beta(\gamma)^{2q}\theta(\gamma) + (C + 1)^2 \beta(\gamma)^{-x(C,p)} + 8\eta_m(\gamma)^{1/2}.$$

Granting Claims 7.3 and 7.4, the proof of the inductive step (and, consequently, of Theorem 7.1) is completed as follows. First observe that, by (7.5), we have $\eta_m(\gamma) \leqslant \big(30(C+1)\big)^{-4q}\gamma^{4q}$; in particular, $8\eta_m(\gamma)^{1/2} \leqslant \big(10(C+1)\big)^{-2q}\gamma^{2q}$. On the other hand, by Claim 7.4 and the choice of $\beta(\gamma)$ and $\theta(\gamma)$ in (7.3), it is easy to see that $\|G - H\|_{L_{2q}} \leqslant 3\big(10(C+1)\big)^{-1}\gamma$. Therefore, by Claim 7.3 and property (P2),

$$\begin{aligned}
|\Delta| &\leqslant 2(C+1)\big(\|G-H\|_{L_{2q}} + \eta_m(\gamma)^{1/2}\big) + \|g_{e_0} - h_{e_0}\|_{\mathcal{S}_{\partial e_0}} \\
&\leqslant 4\gamma/5 + \alpha_m(\gamma) \leqslant 4\gamma/5 + \alpha_0(\gamma) \leqslant 4\gamma/5 + \gamma/5 = \gamma.
\end{aligned}$$

We close this subsection by commenting on the proof of Claim 7.4 (the details are given in Subsection 7.3). Estimates of this form are usually obtained for stronger norms than the cut norm, and as such, they depend on stronger pseudorandomness conditions. The first general method available in this context was developed recently in [6]. It is known as "densification" and consists of taking successive marginals in order to arrive at an expression which involves only bounded functions (see also [29]). Another approach (which ultimately relies on Hölder's inequality) was introduced by Tao and Ziegler [38, 39].

We propose a new method to deal with these types of problems. It appears to be quite versatile—see, e.g., [7]—and is based on a simple decomposition scheme. The method is best seen in action: we first observe the pointwise bound

$$(G - H)^{2q} \leqslant (G - H)^{2q} \, \mathbf{1}_{[G \geqslant H]} + (H - G) \, H^{2q-1}\mathbf{1}_{[G < H]}.$$

Since $0 \leqslant H^{2q-1}\mathbf{1}_{[G<H]} \leqslant 1$ the expectation of the second term of the above decomposition can be estimated using our inductive hypotheses. For the first term we select a cut-off parameter $\beta \geqslant 1$ and we decompose further as

$$(G - H)^{2q} \, \mathbf{1}_{[G \geqslant H]} \leqslant G^{2q} \, \mathbf{1}_{[G \geqslant H]}\mathbf{1}_{[G > \beta]} + (G - H) \, G^{2q-1}\mathbf{1}_{[G \geqslant H]}\mathbf{1}_{[G \leqslant \beta]}.$$

If β is large enough, then we can effectively bound the expectation of the first term of the new decomposition using Lemma 6.3 and Markov's inequality. On the other hand, we have $0 \leqslant G^{2q-1}\mathbf{1}_{[G \geqslant H]}\mathbf{1}_{[G \leqslant \beta]} \leqslant \beta^{2q-1}$ and so the second term can also be handled by our inductive assumptions. By optimizing the parameter β, we obtain the estimate in (7.7) thus completing the proof of Claim 7.4.

7.2. Proof of Claim 7.3. Let $\mathbf{g}_{e_0}$ be as in (6.1) for g_{e_0}. Set

$$I_1 = \int \mathbf{g}_{e_0}(G - H) \, d\boldsymbol{\mu}_{e_0} \ \text{ and } \ I_2 = \int (g_{e_0} - h_{e_0}) \prod_{e \in \mathcal{H} \setminus \{e_0\}} h_e \, d\boldsymbol{\mu}$$

and notice that $|\Delta| \leqslant |I_1| + |I_2|$. Next, observe that

$$(7.8) \qquad\qquad |I_2| \leqslant \|g_{e_0} - h_{e_0}\|_{\mathcal{S}_{\partial e_0}}.$$

This follows by Fubini's theorem and the following well-known fact (see, e.g., [13]). We recall the proof for the convenience of the reader.

Fact 7.5. *Let* $e \in \mathcal{H}$ *with* $|e| \geqslant 2$ *and* $g_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$. *For every* $f \in \partial e$ *let* $u_f \in L_\infty(\boldsymbol{X}, \mathcal{B}_f, \boldsymbol{\mu})$ *with* $0 \leqslant u_f \leqslant 1$. *Then we have* $|\int g_e \prod_{f \in \partial e} u_f \, d\boldsymbol{\mu}| \leqslant \|g_e\|_{\mathcal{S}_{\partial e}}$.

Proof. Set $k = |e|$ and let $\{f_1, \ldots, f_k\}$ be an enumeration of ∂e. We define $Z \colon [0,1]^k \to \mathbb{R}$ by the rule $Z(t_1, \ldots, t_k) = \int g_e \prod_{i=1}^k \mathbf{1}_{[u_{f_i} > t_i]} \, d\boldsymbol{\mu}$. Notice that $\bigcap_{i=1}^k [u_{f_i} > t_i] \in \mathcal{S}_{\partial e}$ for every $(t_1, \ldots, t_k) \in [0,1]^k$ and so $\|Z\|_{L_\infty} \leqslant \|g_e\|_{\mathcal{S}_{\partial e}}$. On the other hand, denoting by $\boldsymbol{\lambda}$ the Lebesgue measure on $[0,1]^k$, by Fubini's theorem we have $\int g_e \prod_{f \in \partial e} u_f \, d\boldsymbol{\mu} = \int Z \, d\boldsymbol{\lambda}$ and the result follows. $\qquad\square$

We proceed to estimate $|I_1|$. First, by the Cauchy–Schwarz inequality and the fact that $0 \leqslant g_{e_0} \leqslant \nu_{e_0}$, we obtain

$$|I_1|^2 \leqslant \int \mathbf{g}_{e_0} \, d\boldsymbol{\mu}_{e_0} \cdot \int \mathbf{g}_{e_0}(G - H)^2 \, d\boldsymbol{\mu}_{e_0} \leqslant \int \boldsymbol{\nu}_{e_0} \, d\boldsymbol{\mu}_{e_0} \cdot \int \boldsymbol{\nu}_{e_0}(G - H)^2 \, d\boldsymbol{\mu}_{e_0}.$$

Let $\psi_{e_0} \in L_p(\boldsymbol{X}, \mathcal{B}_{e_0}, \boldsymbol{\mu})$ with $\|\psi_{e_0}\|_{L_p} \leqslant C$ be as in Definition 6.1 and notice that by condition (C2.a) we have $|\int (\nu_{e_0} - \psi_{e_0}) \, d\boldsymbol{\mu}| \leqslant \eta_m(\gamma)$. This is easily seen to imply that $\int \nu_{e_0} \, d\boldsymbol{\mu} \leqslant C + 1$ and so, by the previous estimate, we have

$$|I_1|^2 \leqslant (C+1) \cdot \left(\Big| \int \boldsymbol{\psi}_{e_0}(G - H)^2 \, d\boldsymbol{\mu}_{e_0} \Big| + \Big| \int (\boldsymbol{\nu}_{e_0} - \boldsymbol{\psi}_{e_0})(G - H)^2 \, d\boldsymbol{\mu}_{e_0} \Big| \right)$$

where $\boldsymbol{\psi}_{e_0}$ is as in (6.1) for ψ_{e_0}. Next, writing $(G - H)^2 = G^2 - 2GH + H^2$ and applying (6.3), we see that $|\int (\boldsymbol{\nu}_{e_0} - \boldsymbol{\psi}_{e_0})(G - H)^2 \, d\boldsymbol{\mu}_{e_0}| \leqslant 4\eta_m(\gamma)$. On the other hand, by Hölder's inequality, $|\int \boldsymbol{\psi}_{e_0}(G - H)^2 \, d\boldsymbol{\mu}_{e_0}| \leqslant C\|G - H\|_{L_{2q}}^2$. Therefore,

$$(7.9) \qquad\qquad |I_1| \leqslant 2(C+1)\big(\|G - H\|_{L_{2q}} + \eta_m(\gamma)^{1/2}\big).$$

Combining (7.8) and (7.9) we conclude that the estimate in (7.6) is satisfied, as desired.

7.3. Proof of Claim 7.4. Recall that $\mathcal{G}$ stands for the set $\{e \in \mathcal{H} \setminus \{e_0\} : g_e \neq h_e\}$. We may assume, of course, that $\mathcal{G}$ is nonempty and, consequently, that $G \neq H$. Set $\mathbf{A} = [G < H]$, $\mathbf{B} = [G \geqslant H] \cap [G \leqslant \beta(\gamma)]$ and $\mathbf{C} = [G \geqslant H] \cap [G > \beta(\gamma)]$, and notice that $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \boldsymbol{\Sigma}_{e_0}$. Next, define

$$I_1 = \int (H - G) H^{2q-1} \mathbf{1}_{\mathbf{A}} \, d\boldsymbol{\mu}_{e_0}, \ I_2 = \int (G - H) G^{2q-1} \mathbf{1}_{\mathbf{B}} \, d\boldsymbol{\mu}_{e_0}, \ I_3 = \int_{\mathbf{C}} G^{2q} \, d\boldsymbol{\mu}_{e_0}$$

and observe that $I_1, I_2, I_3 \geqslant 0$ and $\int (G - H)^{2q} \, d\boldsymbol{\mu}_{e_0} \leqslant I_1 + I_2 + I_3$. Thus, it suffices to estimate I_1, I_2 and I_3.

First we argue for I_1. Let $h'_{e_0} = (H^{2q-1}\mathbf{1_A}) \circ \pi_{e_0} \in L_1(\boldsymbol{X}, \mathcal{B}_{e_0}, \boldsymbol{\mu})$ and notice that $0 \leqslant h'_{e_0} \leqslant 1$. By the definition of G and H, we see that

$$I_1 = \big| \int \prod_{e \in \mathcal{H}\setminus\{e_0\}} g_e \cdot h'_{e_0} \, d\boldsymbol{\mu} - \int \prod_{e \in \mathcal{H}\setminus\{e_0\}} h_e \cdot h'_{e_0} \, d\boldsymbol{\mu} \big|.$$

Moreover, by (7.4) and property (P2), we have $\|g_e - h_e\|_{\mathcal{S}_{\partial e}} \leqslant \alpha_{m-1}\big(\theta(\gamma)\big)$ for every $e \in \mathcal{H} \setminus \{e_0\}$. Since $\eta_m(\gamma) = \eta_{m-1}\big(\theta(\gamma)\big)$, by our inductive assumptions, we obtain

$$(7.10) \qquad\qquad\qquad I_1 \leqslant \theta(\gamma).$$

The estimation of I_2 is similar. Indeed, observe that

$$I_2 = \beta(\gamma)^{2q-1} \int (G - H)\big(G/\beta(\gamma)\big)^{2q-1}\mathbf{1_B} \, d\boldsymbol{\mu}_{e_0}$$

and $0 \leqslant \big(G/\beta(\gamma)\big)^{2q-1}\mathbf{1_B} \leqslant 1$. Therefore,

$$(7.11) \qquad\qquad\qquad I_2 \leqslant \beta(\gamma)^{2q-1}\theta(\gamma).$$

We proceed to estimate I_3. Let $\boldsymbol{\nu}_{e_0,\mathcal{G}}$ and ℓ be as in Definition 6.1, and recall that $0 \leqslant G \leqslant \boldsymbol{\nu}_{e_0,\mathcal{G}}$. By Markov's inequality and the monotonicity of the L_p norms,

$$\boldsymbol{\mu}_{e_0}(\mathbf{C}) \leqslant \boldsymbol{\mu}_{e_0}\big([\boldsymbol{\nu}_{e_0,\mathcal{G}} \geqslant \beta(\gamma)]\big) \leqslant \frac{\int \boldsymbol{\nu}_{e_0,\mathcal{G}} \, d\boldsymbol{\mu}_{e_0}}{\beta(\gamma)} \leqslant \frac{\|\boldsymbol{\nu}_{e_0,\mathcal{G}}\|_{L_\ell}}{\beta(\gamma)} \overset{(6.5)}{\leqslant} \frac{C+1}{\beta(\gamma)}.$$

Thus, by Lemma 6.3 and the choice of $x(C, p)$, we have

$$(7.12) \qquad I_3 \;\leqslant\; \int_{\mathbf{C}} \boldsymbol{\nu}_{e_0,\mathcal{G}}^{2q} \, d\boldsymbol{\mu}_{e_0} \leqslant (C+1)\,\boldsymbol{\mu}_{e_0}(\mathbf{C})^{x(C,p)} + 8\eta_m(\gamma)^{1/2}$$

$$\leqslant\; (C+1)^2 \beta(\gamma)^{-x(C,p)} + 8\eta_m(\gamma)^{1/2}.$$

Combining (7.10)–(7.12) we conclude that the estimate in (7.7) is satisfied. The proof of Claim 7.4 is completed.

8. Relative removal lemma for pseudorandom families

This section is devoted to the proof of Theorem 2.2 and its consequences. As we have mentioned, in Subsection 8.1 we give the proof of Proposition 7.2. The proof of Theorem 2.2 is then completed in Subsection 8.2. Finally, in Subsection 8.3 we present arithmetic versions of Theorem 2.2.

8.1. **Proof of Proposition 7.2.** For part (a), fix $A \in \mathcal{A}_{\mathcal{P}_e}$ and let $\mathcal{P}' \subseteq \mathcal{P}_e$ such that $A = \bigcup \mathcal{P}'$. Notice that $|\mathcal{P}'| \leqslant |\mathcal{P}_e| \leqslant M$ and

$$f_e \cdot \mathbf{1}_A - f_{\text{str}}^e \cdot \mathbf{1}_A = f_{\text{err}}^e \cdot \mathbf{1}_A + \sum_{P \in \mathcal{P}'} f_{\text{unf}}^e \cdot \mathbf{1}_P.$$

Therefore, for any $B \in \mathcal{S}_{\partial_e}$ we have

$$\big| \int_B (f_e \cdot \mathbf{1}_A - f_{\text{str}}^e \cdot \mathbf{1}_A) \, d\boldsymbol{\mu} \big| \;\leqslant\; \big| \int_{B \cap A} f_{\text{err}}^e \, d\boldsymbol{\mu} \big| + \sum_{P \in \mathcal{P}'} \big| \int_{B \cap P} f_{\text{unf}}^e \, d\boldsymbol{\mu} \big|$$

$$\leqslant\; \|f_{\text{err}}^e\|_{L_{p^\dagger}} + |\mathcal{P}'| \cdot \|f_{\text{unf}}^e\|_{\mathcal{S}_{\partial_e}} \leqslant \sigma + \frac{M}{F(M)} \leqslant \alpha$$

which implies, of course, that $\|f_e \cdot \mathbf{1}_A - f_{\mathrm{str}}^e \cdot \mathbf{1}_A\|_{\mathcal{S}_{\partial e}} \leqslant \alpha$.

For part (b), let $\zeta \geqslant 1$ be arbitrary and set $A = [f_{\mathrm{str}}^e \leqslant \zeta]$. First observe that $A \in \mathcal{A}_{\mathcal{P}_e}$ since $f_{\mathrm{str}}^e = \mathbb{E}(f_e \,|\, \mathcal{A}_{\mathcal{P}_e})$. Next, by Markov's inequality, we have

$$\boldsymbol{\mu}(\boldsymbol{X} \setminus A) \leqslant \frac{\int (f_{\mathrm{str}}^e)^p \, d\boldsymbol{\mu}}{\zeta^p} \leqslant C^p \zeta^{-p}$$

and so, by Hölder's inequality,

$$\int_{\boldsymbol{X} \setminus A} f_{\mathrm{str}}^e \, d\boldsymbol{\mu} \leqslant \|f_{\mathrm{str}}^e\|_{L_p} \cdot \boldsymbol{\mu}(\boldsymbol{X} \setminus A)^{1/q} \leqslant C^p \zeta^{1-p}.$$

Finally, by part (a) and the fact that $\boldsymbol{X} \setminus A \in \mathcal{A}_{\mathcal{P}_e}$, we conclude that

$$
\begin{aligned}
\int_{\boldsymbol{X} \setminus A} f_e \, d\boldsymbol{\mu} \;&\leqslant\; \int_{\boldsymbol{X} \setminus A} f_{\mathrm{str}}^e \, d\boldsymbol{\mu} + \Big| \int_{\boldsymbol{X} \setminus A} (f_e - f_{\mathrm{str}}^e) \, d\boldsymbol{\mu} \Big| \\
&\leqslant\; C^p \zeta^{1-p} + \|f_e \cdot \mathbf{1}_{\boldsymbol{X} \setminus A} - f_{\mathrm{str}}^e \cdot \mathbf{1}_{\boldsymbol{X} \setminus A}\|_{\mathcal{S}_{\partial e}} \leqslant C^p \zeta^{1-p} + \alpha
\end{aligned}
$$

and the proof of Proposition 7.2 is completed.

8.2. **Proof of Theorem 2.2.** Let n, r, C, p and ε be as in the statement of the theorem and let q denote the conjugate exponent of p. We begin by introducing some numerical invariants. First, we set

$$\zeta = \zeta(C, p, \varepsilon) = (C+1)^q (\varepsilon/6)^{1-q}.$$

Also let $\Delta(n, r, \frac{\varepsilon}{6\zeta})$ and $K(n, r, \frac{\varepsilon}{6\zeta})$ be as in Theorem 3.4 and note that we may assume that $\Delta(n, r, \frac{\varepsilon}{6\zeta}) \leqslant \frac{\varepsilon}{6\zeta}$. We define

$$(8.1) \quad \delta = \delta(n, r, C, p, \varepsilon) = \frac{\Delta(n, r, \frac{\varepsilon}{6\zeta})^{n^r}}{2} \quad \text{and} \quad k = k(n, r, C, p, \varepsilon) = K\Big(n, r, \frac{\varepsilon}{6\zeta}\Big).$$

Next, let $\alpha(n, r, C, p, \zeta, \delta)$ and $\eta(n, r, C, p, \zeta, \delta)$ be as in Theorem 7.1 and set

$$\alpha = \min\{k^{-2^r}(\varepsilon/3), \alpha(n, r, C, p, \zeta, \delta)\} \quad \text{and} \quad \mathrm{Reg} = \mathrm{Reg}(n, r, C+1, p, F, \alpha/2)$$

where $F \colon \mathbb{N} \to \mathbb{R}$ is the growth function defined by the rule $F(m) = 2\alpha^{-1}(m+1)$ and $\mathrm{Reg}(n, r, C+1, p, F, \alpha/2)$ is as in Theorem 2.1. Finally, we define

$$(8.2) \qquad \eta = \eta(n, r, C, p, \varepsilon) = \min\{1/\mathrm{Reg}, \eta(n, r, C, p, \zeta, \delta)\}.$$

We will show that the parameters η, δ and k are as desired.

Indeed, let $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$ be an η-nonatomic, r-uniform hypergraph system and let $\langle \nu_e : e \in \mathcal{H} \rangle$ be a (C, η, p)-pseudorandom family. For every $e \in \mathcal{H}$ let $f_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ with $0 \leqslant f_e \leqslant \nu_e$ and assume that

$$(8.3) \qquad \int \prod_{e \in \mathcal{H}} f_e \, d\boldsymbol{\mu} \leqslant \delta.$$

By Lemma 6.2, for every $e \in \mathcal{H}$ the random variable ν_e is $(C+1, \eta, p)$-regular and, consequently, so is f_e. Therefore, by (8.2), we may apply Theorem 2.1 and we obtain: (a) a positive integer M with $M \leqslant \mathrm{Reg}$, (b) for every $e \in \mathcal{H}$ a partition $\mathcal{P}_e$ of $\boldsymbol{X}$ with $\mathcal{P}_e \subseteq \mathcal{S}_{\partial e}$ and $\boldsymbol{\mu}(A) \geqslant 1/M$ for every $A \in \mathcal{P}_e$, and (c) for every $e \in \mathcal{H}$ a

finite refinement $\mathcal{Q}_e$ of $\mathcal{P}_e$, such that for every $e \in \mathcal{H}$, writing $f_e = f^e_{\text{str}} + f^e_{\text{err}} + f^e_{\text{unf}}$ where f^e_{str}, f^e_{err} and f^e_{unf} are as in (2.2), we have the estimates

$$(8.4) \qquad \|f^e_{\text{str}}\|_{L_p} \leqslant C + 1, \quad \|f^e_{\text{err}}\|_{L_{p^{\dagger}}} \leqslant \alpha/2 \quad \text{and} \quad \|f^e_{\text{unf}}\|_{\mathcal{S}_{\partial e}} \leqslant \frac{1}{F(M)}$$

where $p^{\dagger} = \min\{2, p\}$. For every $e \in \mathcal{H}$ let

$$(8.5) \qquad A_e = [f^e_{\text{str}} \leqslant \zeta], \quad g_e = f_e \cdot \mathbf{1}_{A_e} \quad \text{and} \quad h_e = f^e_{\text{str}} \cdot \mathbf{1}_{A_e}$$

and notice that $0 \leqslant g_e \leqslant \nu_e$ and $0 \leqslant h_e \leqslant \zeta$. Moreover, by Proposition 7.2, we see that $\|g_e - h_e\|_{\mathcal{S}_{\partial e}} \leqslant \alpha$.

Claim 8.1. *We have* $\int \prod_{e \in \mathcal{H}} h_e \, d\boldsymbol{\mu} \leqslant \Delta(n, r, \frac{\varepsilon}{6\zeta})^{n^r}$.

Proof. First observe that, by the choice of α and Theorem 7.1,

$$(8.6) \qquad \Big| \int \prod_{e \in \mathcal{H}} g_e \, d\boldsymbol{\mu} - \int \prod_{e \in \mathcal{H}} h_e \, d\boldsymbol{\mu} \Big| \leqslant \delta.$$

On the other hand, we have $0 \leqslant g_e \leqslant f_e$ for every $e \in \mathcal{H}$. Hence, by (8.3) and (8.6),

$$\int \prod_{e \in \mathcal{H}} h_e \, d\boldsymbol{\mu} \leqslant \int \prod_{e \in \mathcal{H}} f_e \, d\boldsymbol{\mu} + \Big| \int \prod_{e \in \mathcal{H}} h_e \, d\boldsymbol{\mu} - \int \prod_{e \in \mathcal{H}} g_e \, d\boldsymbol{\mu} \Big| \leqslant 2\delta.$$

Finally, by (8.1), we have $2\delta \leqslant \Delta(n, r, \frac{\varepsilon}{6\zeta})^{n^r}$ and the proof is completed. $\qquad \square$

Now for every $e \in \mathcal{H}$ set $E_e = [h_e \geqslant \Delta(n, r, \frac{\varepsilon}{6\zeta})]$. Since $|\mathcal{H}| \leqslant \binom{n}{r} \leqslant n^r - 1$ and $\Delta(n, r, \frac{\varepsilon}{6\zeta}) \leqslant 1$, by Claim 8.1 and Markov's inequality, we have

$$\boldsymbol{\mu}\Big(\bigcap_{e \in \mathcal{H}} E_e \Big) \leqslant \boldsymbol{\mu}\Big(\Big\{ \mathbf{x} \in \boldsymbol{X} : \prod_{e \in \mathcal{H}} h_e(\mathbf{x}) \geqslant \Delta\Big(n, r, \frac{\varepsilon}{6\zeta}\Big)^{|\mathcal{H}|} \Big\} \Big) \leqslant \Delta\Big(n, r, \frac{\varepsilon}{6\zeta}\Big).$$

Thus, by Theorem 3.4, for every $e \in \mathcal{H}$ there exists $F_e \in \mathcal{B}_e$ with

$$(8.7) \qquad \boldsymbol{\mu}(E_e \setminus F_e) \leqslant \frac{\varepsilon}{6\zeta} \quad \text{and} \quad \bigcap_{e \in \mathcal{H}} F_e = \emptyset.$$

Moreover, by (8.1), there exists a collection $\langle \mathcal{P}_{e'} : e' \subseteq e \text{ for some } e \in \mathcal{H} \rangle$ of partitions of $\boldsymbol{X}$ such that: (i) $\mathcal{P}_{e'} \subseteq \mathcal{B}_{e'}$ and $|\mathcal{P}_{e'}| \leqslant k$ for every $e' \subseteq e \in \mathcal{H}$, and (ii) for every $e \in \mathcal{H}$ the set F_e belongs to the algebra generated by the family $\bigcup_{e' \subsetneq e} \mathcal{P}_{e'}$. Therefore, the proof of the theorem will be completed once we show that

$$(8.8) \qquad \int_{\boldsymbol{X} \setminus F_e} f_e \, d\boldsymbol{\mu} \leqslant \varepsilon$$

for every $e \in \mathcal{H}$. To this end, fix $e \in \mathcal{H}$ and notice that

$$(8.9) \quad \int_{\boldsymbol{X} \setminus F_e} f_e \, d\boldsymbol{\mu} \leqslant \int_{\boldsymbol{X} \setminus F_e} h_e \, d\boldsymbol{\mu} + \Big| \int_{\boldsymbol{X} \setminus F_e} (g_e - h_e) \, d\boldsymbol{\mu} \Big| + \Big| \int_{\boldsymbol{X} \setminus F_e} (f_e - g_e) \, d\boldsymbol{\mu} \Big|.$$

Next observe that, by the definition of E_e and the fact that $0 \leqslant h_e \leqslant \zeta$, we have

$$(8.10) \qquad \int_{\boldsymbol{X} \setminus F_e} h_e \, d\boldsymbol{\mu} \quad \leqslant \quad \int_{\boldsymbol{X} \setminus E_e} h_e \, d\boldsymbol{\mu} + \int_{E_e \setminus F_e} h_e \, d\boldsymbol{\mu}$$

$$\leqslant \quad \Delta\Big(n, r, \frac{\varepsilon}{6\zeta}\Big) + \zeta \, \boldsymbol{\mu}(E_e \setminus F_e) \overset{(8.7)}{\leqslant} \varepsilon/3.$$

To estimate the second term in the right-hand side of (8.9), let $\mathcal{A}$ denote the algebra on $\boldsymbol{X}$ generated by the family $\bigcup_{e' \subsetneq e} \mathcal{P}_{e'}$ and note that every atom of $\mathcal{A}$ is of the form $\bigcap_{e' \subsetneq e} A_{e'}$ where $A_{e'} \in \mathcal{P}_{e'}$ for every $e' \subsetneq e$. It follows that the number of atoms of $\mathcal{A}$ is less than k^{2^r} and, moreover, every atom of $\mathcal{A}$ belongs to $\mathcal{S}_{\partial e}$. In particular, there exists a family $\mathcal{F} \subseteq \mathcal{S}_{\partial e}$ consisting of pairwise disjoint sets with $|\mathcal{F}| \leqslant k^{2^r}$ and such that $\boldsymbol{X} \setminus F_e = \bigcup \mathcal{F}$. Therefore, by the fact that $\|g_e - h_e\|_{\mathcal{S}_{\partial e}} \leqslant \alpha$ and the choice of α, we have

$$(8.11) \qquad \Big| \int_{\boldsymbol{X} \setminus F_e} (g_e - h_e) \, d\boldsymbol{\mu} \Big| \leqslant \sum_{A \in \mathcal{F}} \Big| \int_A (g_e - h_e) \, d\boldsymbol{\mu} \Big| \leqslant |\mathcal{F}| \, \alpha \leqslant k^{2^r} \alpha \leqslant \varepsilon/3.$$

Finally, to estimate the last term in the right-hand side of (8.9), notice that if $p = \infty$, then this term is equal to zero. (Indeed, in this case we have $\zeta = C + 1$ and $A_e = \boldsymbol{X}$.) On the other hand, if $1 < p < \infty$, then, by Proposition 7.2 and the choice of ζ and α, we obtain that

$$(8.12) \qquad \Big| \int_{\boldsymbol{X} \setminus F_e} (f_e - g_e) \, d\boldsymbol{\mu} \Big| \;=\; \int_{\boldsymbol{X} \setminus F_e} f_e \cdot \mathbf{1}_{\boldsymbol{X} \setminus A_e} \, d\boldsymbol{\mu} \leqslant \int_{\boldsymbol{X} \setminus A_e} f_e \, d\boldsymbol{\mu}$$
$$\leqslant \; (C+1)^p \zeta^{1-p} + \alpha \leqslant \varepsilon/3.$$

Combining (8.9)–(8.12) we conclude that (8.8) is satisfied, and so the entire proof of Theorem 2.2 is completed.

8.3. **Consequences.** As we have noted, Theorem 2.2 yields a Szemerédi-type result for sparse pseudorandom subsets of finite additive groups. (Recall that an *additive group* is just an abelian group written additively.) The argument for deducing this result is well-known—see, e.g., [13, 25, 30, 34]—and originates from the work of Ruzsa and Szemerédi [27].

First we introduce some notation. If $A_1, \ldots, A_d$ are nonempty finite sets and $f \colon A_1 \times \cdots \times A_d \to \mathbb{R}$ is a function, then by $\mathbb{E}[f(x_1, \ldots, x_d) \,|\, x_1 \in A_1, \ldots, x_d \in A_d]$ we shall denote the average of f, that is,

$$\mathbb{E}[f(x_1, \ldots, x_d) \,|\, x_1 \in A_1, \ldots, x_d \in A_d] \coloneqq \frac{1}{|\prod_{i=1}^d A_i|} \sum_{(x_1, \ldots, x_d) \in \prod_{i=1}^d A_i} f(x_1, \ldots, x_d).$$

The following relative multidimensional Szemerédi theorem is the main result in this subsection. It follows from Theorem 2.2 arguing precisely as in the proof of [34, Theorem 2.18].

Theorem 8.2. *For every integer $k \geqslant 3$, every $C \geqslant 1$, every $1 < p \leqslant \infty$ and every $0 < \delta \leqslant 1$ there exist a positive integer $N = N(k, C, p, \delta)$ and a strictly positive constant $c = c(k, C, p, \delta)$ with the following property. Let Z, Z' be finite additive groups and let $\langle \phi_i : i \in [k] \rangle$ be a collection of group homomorphisms from Z into Z' such that the set $\{\phi_i(d) - \phi_j(d) : i, j \in [k] \text{ and } d \in Z\}$ generates Z'. Consider the $(k-1)$-uniform hypergraph system $\mathscr{H} = (k, \langle (X_i, \mu_i) : i \in [k] \rangle, \mathcal{H})$ where: (a) $\mathcal{H} = \binom{k}{k-1}$, and (b) (X_i, μ_i) is the discrete probability space with $X_i = Z$ and*

μ_i the uniform probability measure on Z for every $i \in [k]$. Also let $\nu \colon Z' \to \mathbb{R}$ be a nonnegative function and for every $j \in [k]$ define $\nu_{[k]\setminus\{j\}} \colon \boldsymbol{X} \to \mathbb{R}$ by the rule

$$(8.13) \qquad \nu_{[k]\setminus\{j\}}\big((x_i)_{i\in[k]}\big) = \nu\Big(\sum_{i\in[k]} \big(\phi_i(x_i) - \phi_j(x_i)\big)\Big).$$

(Here, we have $\boldsymbol{X} = X_1 \times \cdots \times X_k$). Assume that the family $\langle \nu_{[k]\setminus\{j\}} : j \in [k] \rangle$ is (C, N^{-1}, p)-pseudorandom and that $|Z| \geqslant N$. Then for every $f \colon Z' \to \mathbb{R}$ with $0 \leqslant f \leqslant \nu$ and $\mathbb{E}[f(x) \,|\, x \in Z'] \geqslant \delta$ we have

$$(8.14) \qquad \mathbb{E}\Big[\prod_{j\in[k]} f\big(a + \phi_j(d)\big) \,\Big|\, a \in Z', d \in Z \Big] \geqslant c.$$

We close this subsection by discussing in more detail the pseudorandomness hypotheses appearing in Theorem 8.2 for the important special case of the cyclic group $\mathbb{Z}_n \coloneqq \mathbb{Z}/n\mathbb{Z}$. Specifically, let k, C and p be as in Theorem 8.2 and let $0 < \eta \leqslant 1$. Also let n be a positive integer. We say that a function $\nu \colon \mathbb{Z}_n \to \mathbb{R}$ is (k, C, η, p)-pseudorandom if it is nonnegative and satisfies the following conditions.

(C1) If ℓ is as in (6.4), then we have

$$1 - \eta \leqslant \mathbb{E}\Big[\prod_{j=1}^{k} \prod_{\omega \in \{0, \ldots, \ell-1\}^{[k]\setminus\{j\}}} \nu\big(\sum_{i=1}^{k}(i - j)\, x_i^{(\omega_i)}\big)^{n_{j,\omega}}$$

$$\Big|\, x_1^{(0)}, \ldots, x_1^{(\ell-1)}, \ldots, x_k^{(0)}, \ldots, x_k^{(\ell-1)} \in \mathbb{Z}_n \Big] \leqslant C + \eta$$

for any choice of $n_{j,\omega} \in \{0, 1\}$.

(C2) For every $j \in [k]$ there exists $\boldsymbol{\psi}_j \colon \mathbb{Z}_n^{[k]\setminus\{j\}} \to \mathbb{R}$ with $\mathbb{E}\big[|\boldsymbol{\psi}_j|^p\big] \leqslant C^p$ and

$$\Big| \mathbb{E}\Big[\prod_{\omega \in \{0,1\}^{[k]\setminus\{j\}}} \Big(\nu\big(\sum_{i=1}^{k}(i-j)\, x_i^{(\omega_i)}\big) - \boldsymbol{\psi}_j\big((x_i^{(\omega_i)})_{i\in[k]\setminus\{j\}}\big) \Big) \times$$

$$\times \prod_{j'\in[k]\setminus\{j\}} \prod_{\omega\in\{0,1\}^{[k]\setminus\{j'\}}} \nu\big(\sum_{i=1}^{k}(i-j')\, x_i^{(\omega_i)}\big)^{n_{j',\omega}} \,\Big|\, x_1^{(0)}, x_1^{(1)}, \ldots, x_k^{(0)}, x_k^{(1)} \in \mathbb{Z}_n \Big] \Big| \leqslant \eta$$

for any choice of $n_{j',\omega} \in \{0, 1\}$.

By Proposition 6.4 and Theorem 8.2, we have the following theorem.

Theorem 8.3. *Let $k \in \mathbb{N}$ with $k \geqslant 3$, $C \geqslant 1$ and $1 < p \leqslant \infty$, and let $0 < \delta \leqslant 1$. Also let $N = N(k, C, p, \delta)$ and $c = c(k, C, p, \delta)$ be as in Theorem 8.2 and set*

$$\eta = \big((C+1)N\big)^{-2^{k-1}}.$$

If $n \in \mathbb{N}$ with $n \geqslant N$ and $\nu \colon \mathbb{Z}_n \to \mathbb{R}$ is (k, C, η, p)-pseudorandom, then for every $f \colon \mathbb{Z}_n \to \mathbb{R}$ with $0 \leqslant f \leqslant \nu$ and $\mathbb{E}[f(x) \,|\, x \in \mathbb{Z}_n] \geqslant \delta$ we have

$$(8.15) \qquad \mathbb{E}\Big[\prod_{j\in[k]} f\big(a + jd\big) \,\Big|\, a, d \in \mathbb{Z}_n \Big] \geqslant c.$$

References

[1] K. Ball, E. A. Carlen and E. H. Lieb, *Sharp uniform convexity and smoothness inequalities for trace norms*, Invent. Math. 115 (1994), 463–482.

[2] B. Bollobás and O. Riordan, *Metrics for sparse graphs*, in "Surveys in Combinatorics 2009", London Mathematical Society Lecture Note Series, Vol. 365, Cambridge University Press, 2009, 211–287.

[3] C. Borgs, J. T. Chayes, H. Cohn and Y. Zhao, *An L^p theory of sparse graph convergence I: limits, sparse random graph models, and power law distributions*, preprint (2014), available at `arXiv:1401.2906`.

[4] S. Brazitikos and Th. Karageorgos, *An algorithmic regularity lemma for L_p regular sparse matrices*, SIAM J. Discrete Math. 31 (2017), 2301–2313.

[5] A. Coja-Oghlan, C. Cooper and A. Frieze, *An efficient sparse regularity concept*, SIAM J. Discrete Math. 23 (2010), 2000–2034.

[6] D. Conlon, J. Fox and Y. Zhao, *A relative Szemerédi theorem*, Geom. Funct. Anal. 25 (2015), 733–762.

[7] P. Dodos and V. Kanellopoulos, *Uniformity norms, their weaker versions, and applications*, preprint (2016), available at `arXiv:1610.00487`.

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

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

[10] P. Dodos, V. Kanellopoulos and Th. Karageorgos, *L_p regular sparse hypergraphs: box norms*, Fund. Math. 248 (2020), 49–77.

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

[12] A. Frieze and R. Kannan, *Quick approximation to matrices and applications*, Combinatorica 19 (1999), 175–220.

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

[14] W. T. Gowers, *Decompositions, approximate structure, transference, and the Hahn–Banach theorem*, Bull. Lond. Math. Soc. 42 (2010), 573–606.

[15] B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. Math. 167 (2008), 481–547.

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

[17] H. Hatami, *On Generalizations of Gowers Norms*, Thesis, University of Toronto, 2009.

[18] Y. Kohayakawa, *Szemerédi's regularity lemma for sparse graphs*, in "Foundations of Computational Mathematics", Springer, 1997, 216–230.

[19] Y. Kohayakawa and V. Rödl, *Szemerédi's regularity lemma and quasi-randomness*, in "Recent Advances in Algorithms and Combinatorics", CMS Books in Mathematics, Vol. 11, Springer, 2003, 289–351.

[20] L. Lovász, *Large Networks and Graph Limits*, American Mathematical Society Colloquium Publications, Vol. 60, American Mathematical Society, Providence, RI, 2012.

[21] B. Nagle, V. Rödl and M. Schacht, *The counting lemma for regular k-uniform hypergraphs*, Random Struct. Algorithms 28 (2006), 113–179.

[22] 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`.

[23] O. Reingold, L. Trevisan, M. Tulsiani and S. Vadhan, *Dense subsets of pseudorandom sets*, in "49th Annual IEEE Symposium on Foundations of Computer Science", IEEE Computer Society, 2008, 76–85.

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

[25] V. Rödl, M. Schacht, E. Tengan and N. Tokushige, *Density theorems and extremal hypergraph problems*, Israel J. Math. 152 (2006), 371–380.

[26] V. Rödl and J. Skokan, *Regularity lemma for k-uniform hypergraphs*, Random Struct. Algorithms 25 (2004), 1–42.

[27] I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, in "Combinatorics" (Keszthely, 1976), Colloquia Mathematics Societies János Bolyai, Vol. 18, 1978, 939–945.

[28] D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math. 201 (2015), 925–992.

[29] X. Shao, *Narrow arithmetic progressions in the primes*, preprint (2015), available at `arXiv:1509.04955`.

[30] J. Solymosi, *A note on a question of Erdős and Graham*, Combin. Probab. Comput. 13 (2004), 263–267.

[31] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27 (1975), 199–245.

[32] E. Szemerédi, *Regular partitions of graphs*, in "Problèmes combinatoires et théorie des graphes" (Colloq. Internat. du CNRS, Univ. Orsay, Orsay, 1976), Colloques Internationaux du CNRS, Vol. 260, CNRS, 1978, 399–401.

[33] T. Tao, *A variant of the hypergraph removal lemma*, J. Comb. Theory, Ser. A 113 (2006), 1257–1280.

[34] T. Tao, *The Gaussian primes contain arbitrarily shaped constellations*, J. Anal. Math. 99 (2006), 109–176.

[35] T. Tao, *Structure and randomness in combinatorics*, in "48th Annual IEEE Symposium on Foundations of Computer Science", IEEE Computer Society, 2007, 3–15.

[36] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge Studies in Advanced Mathematics, Vol. 105, Cambridge University Press, 2006.

[37] T. Tao and T. Ziegler, *The primes contain arbitrarily long polynomial progressions*, Acta Math. 201 (2008), 213–305.

[38] T. Tao and T. Ziegler, *Narrow progressions in the primes*, in "Analytic Number Theory, in honor of Helmut Maier's 60th Birthday" (to appear), available at `arXiv:1409.1327`.

[39] T. Tao and T. Ziegler, *Polynomial patterns in the primes*, preprint (2016), available at `arXiv:1603.07817`.

Department of Mathematics, University of Athens, Panepistimiopolis 157 84, Athens, Greece

E-mail address: `pdodos@math.uoa.gr`

National Technical University of Athens, Faculty of Applied Sciences, Department of Mathematics, Zografou Campus, 157 80, Athens, Greece

E-mail address: `bkanel@math.ntua.gr`

Department of Mathematics, University of Athens, Panepistimiopolis 157 84, Athens, Greece

E-mail address: `tkarageo@math.uoa.gr`