National and Kapodistrian University of Athens
Department of Mathematics

A dissertation submitted for the degree of Doctor of Philosophy

Analytic properties of sparse graphs and hypergraphs

Author
Karageorgos Theodoros

Supervisor
Dodos Pandelis

The present Ph.D. dissertation was funded by the State Scholarships Foundation (I.K.Y.), by virtue of the Act "Second-cycle post-graduate studies scholarship programme (Ph.D.)". The Act is implemented through resources from the "Operational Programme Human Resources Development, Education and Lifelong Learning", for the period 2014-2020, which is funded by the European Social Fund (ESF) and the Greek State.

2014–2018

Ph.D. Dissertation Committee:

Dodos Pandelis, Assistant Professor, NKUA[1] (Supervisor)

Gatzouras Dimitrios, Professor, NKUA

Giannopoulos Apostolos, Professor, NKUA

Thilikos Dimitrios, Professor, NKUA

Kanellopoulos Vasileios, Associate Professor, NTUA[2]

Stratis Ioannis, Professor, NKUA

Tyros Konstantinos, Associate Professor, NKUA

[1] National and Kapodistrian University of Athens
[2] National Technical University of Athens

This page is left intentionally blank.

Contents

Introduction

The aim of this dissertation is threefold. At first, we develop a technique that provides regularity results for L_p and L_p regular random variables (parts I and II respectively). Next, we define a class of weighted hypergraphs that satisfy relative counting and removal lemmas (part III). Finally, we present number theoretical (part IV) and algorithmic (part V) applications of the aforementioned results.

In part I, which is based on [DKK16], we introduce the concepts of semirings and uniformity norms, and prove a regularity result for L_p random variables with $p > 1$. This result extends the previous work that dealt with the case $p = 2$, (see e.g. [Tao06b, Tao06c, Tao11]) and its proof is implemented by developing a technique which is based on an inequality about martingale difference sequences and may be seen as an L_p analogue of the energy increment strategy. Moreover, we give applications of this result in the context of martingale convergence and graphon regularity.

In part II, which is based mainly on [DKK18], we define the class of L_p regular random variables; a class of random variables that was introduced in [BCCZ14] and originates from the work of Kohayakawa and Rödl [Koh97, KR03]. For this class, we show that a Hölder-type inequality is satisfied and we use the techniques introduced in the previous part to obtain a regularity result.

Part III is based on the work we did in [DKK15, DKK18]. After we introduce some variants of the well known box norms, we proceed to define a class of weighted hypergraphs. The most important property of this class is that it is the largest class of weighted hypergraphs that we know of, which satisfies relative counting and removal lemmas. This answers a question that was posed in [BCCZ14] and extends similar results already known for smaller classes of weighted hypergraphs (see e.g. [Tao06c, CFZ15, DK16]).

In part IV we give a number theoretical application of part III results. More precisely, we prove a special case of the multidimensional Green–Tao theorem (see [CM12]) using an arithmetic version of the relative removal lemma.

Finally, part V, which is based on [BK17], contains an algorithmic consequence of the technique we developed in parts I and II. More precisely, we construct an algorithm that approximates L_p regular matrices ($p > 1$) by a finite sum of matrices of rank 1. This approximation is done in the cut norm and extends the already existing results about L_∞ regular matrices (see e.g. [COCF10]).

Aknowlegments. I would like to thank my advisors P. Dodos and V. Kanellopoulos for their constant guidance and useful suggestions.

Basic Concepts & General notation

1. By $\mathbb{N}, \mathbb{Z}, \mathbb{R}$ and $\mathbb{C}$ we denote the sets of natural numbers (including 0), integers, real numbers and complex numbers respectively. Moreover, for every positive integer n we set $[n] \coloneqq \{1, 2, \dots, n\}$. For every set X by $|X|$ we shall denote its cardinality and by $P(X)$ we shall denote its powerset. If $k \in \mathbb{N}$ and $k \leqslant |X|$ then by $\binom{X}{k}$ we shall denote the set of all subsets of X of cardinality k, i.e.

$$\binom{X}{k} = \{Y \subseteq X \colon |Y| = k\}.$$

2. By $\mathbf{P}$ we shall denote the set of prime numbers. Also for every positive integer n, by $\mathbf{P}_n$ we shall denote the set of prime numbers which are lower or equal to n. Also, by $\pi(n)$ we shall denote the number of elements in $\mathbf{P}_n$, i.e.

$$\pi(n) = |\mathbf{P}_n|.$$

3. If X is a nonempty set and $\mathcal{F} \subseteq P(X)$ we write

$$\bigcup \mathcal{F} = \bigcup_{F \in \mathcal{F}} F.$$

Also, if k is a positive integer and $\mathcal{A}_1, \dots, \mathcal{A}_k$ are families of subsets of X we write

$$\bigcap_{i=1}^{k} \mathcal{A}_i = \big\{A_1 \cap \dots A_k \colon A_i \in \mathcal{A}_i \text{ for every } i \in [k].\big\}.$$

Finally, if d is a positive integer, $X_1, \dots, X_d$ are nonempty sets and $\mathcal{A}_i$ is a family of subsets of X_i, for every $i \in [d]$ then we write

$$\bigtimes_{i=1}^{d} \mathcal{A}_i = \big\{A_1 \times \dots \times A_d \colon A_i \in \mathcal{A}_i \text{ for every } i \in [d]\big\}.$$

4. If (X, Σ, μ) is a probability space and $f \colon X \to \mathbb{R}$ is a random variable we will write

$$\int_X f(x)\, d\mu(x) \equiv \mathbb{E}_X(f) \equiv \mathbb{E}[f(x) \,|\, x \in X]$$

to denote the mean value of f in X.

5. If (X,Σ,μ) is a probability space and $\mathcal{P}\subseteq\Sigma$ is a partition of X by $\mathcal{A}_{\mathcal{P}}$ we will denote the σ-algebra produced by the cells of $\mathcal{P}$ and by $\iota(\mathcal{P})$ we will denote the measure of the "smallest" cell of $\mathcal{P}$, i.e. $\iota(\mathcal{P})=\min\{\mu(P)\colon P\in\mathcal{P}\}$. Also, if $f\colon X\to\mathbb{R}$ is a random variable we will write $\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})$ to denote the conditional probability of f on the σ-algebra $\mathcal{P}$. Moreover, if $\mathcal{P}$ is finite then recall that

$$\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})=\sum_{P\in\mathcal{P}}\frac{\int_P f\,d\mu}{\mu(P)}\mathbf{1}_P,$$

where for every $A\subseteq X$, $\mathbf{1}_A$ stands for the characteristic function of A, that is,

$$\mathbf{1}_A=\begin{cases}1, & \text{if } x\in A\\ 0, & \text{otherwise.}\end{cases}$$

6. For every function $f\colon\mathbb{N}\to\mathbb{N}$ and every $\ell\in[n]$ by $f^{(\ell)}:=\mathbb{N}\to\mathbb{N}$ we will denote the *ℓ-th iteration* of f defined recursively by the rule

$$\begin{cases}f^{(0)}(n)=n\\ f^{(\ell+1)}(n)=f(f^{(\ell)}(n).\end{cases}$$

7. Recall that a *hypergraph* is a pair $\mathcal{H}=(V,E)$ where V is a non-empty set and $E\subseteq P(V)$. The elements of V are called *vertices* and the elements of E are called edges. If E is a nonempty subset of $\binom{V}{r}$ for some $r\in\mathbb{N}$, then the hypergraph $\mathcal{H}$ is called r-uniform. Therefore, a 2-uniform hypergraph is a graph with at least one edge.

8. Let (X,Σ,μ) be a probability space and recall that a *graphon* is an integrable random variable $W\colon X\times X\to\mathbb{R}$ which is symmetric, that is,

$$W(x,y)=W(y,x)$$

for every $x,y\in X$. If $p>1$ and W is graphon which belongs to L_p, then W is said to be an *L_p graphon*.

9. Let (X,Σ,μ) be a probability space. Recall that a set $A\in\Sigma$ is called an *atom* if $\mu(A)>0$ and for every $B\subseteq A$ with $B\in\Sigma$, $\mu(B)=0$. The set of atoms of the probability space X will be denoted by $\mathrm{Atoms}(X)$.

10. Let (X,Σ,μ) be a probability space and $\eta>0$. The probability space X will be called η-nonatomic if $\mu(A)\leqslant\eta$ for every $A\in\mathrm{Atoms}(X)$.

11. Let n,m be two positive integers. Then, by $\gcd(n,m)$ we denote the greatest common divisor of n and m and by $\mathrm{lcm}(n,m)$ we denote their least common multiple.

12. For every complex number s by $\mathrm{Re}(s)$ we shall denote the real part of s and by $\mathrm{Im}(s)$ we shall denote its imaginary part.

13. By $O(X)$ we shall denote a quantity Y for which there exists some constant $C > 0$ such that $|Y| \leqslant C|X|$. If this constant depends on some parameters, say $a_1, \ldots, a_t$ we will write $Y = O_{a_1,\ldots,a_t}(X)$.

14. By $o(1)$ we shall denote a quantity that can be made arbitrarily close to 0. If this quantity depends on some parameters, say $a_1, \ldots, a_t$ we will write $o_{a_1,\ldots,a_t}(1)$.

15. For every positive integer d and for every $K \subseteq \mathbb{Z}^d$ we denote the volume of K by $\mathrm{vol}_d(K)$. If d is implied then we just write $\mathrm{vol}(K)$.

16. For every positive integer d and for every $x = (x_1, \ldots, x_d) \in \mathbb{Z}^d$ by $\|x\|_\infty$ we denote its infinity norm as usual, i.e.

$$\|x\|_\infty = \max_{1 \leqslant i \leqslant d} |x_i|.$$

Part I

Decomposition of random variables

CHAPTER 1

Semirings and uniformity norms

We first introduce the following slight strengthening of the classical concept of a semiring of sets (see also [BN08]).

Definition 1.1. *Let X be a nonempty set and k a positive integer. Also let $\mathcal{S}$ be a collection of subsets of X. We say that $\mathcal{S}$ is a* k-semiring on X *if the following properties are satisfied.*

(P1) *We have that $\emptyset, X \in \mathcal{S}$.*
(P2) *For every $S, T \in \mathcal{S}$ we have that $S \cap T \in \mathcal{S}$.*
(P3) *For every $S, T \in \mathcal{S}$ there exist $\ell \in [k]$ and $R_1, \dots, R_\ell \in \mathcal{S}$ which are pairwise disjoint and such that $S \setminus T = R_1 \cup \dots \cup R_\ell$.*

From now on we view every element of a k-semiring $\mathcal{S}$ as a "structured" set and a linear combination of few characteristic functions of elements of $\mathcal{S}$ as a "simple" function. We will use the following norm in order to quantify how far from being "simple" a given function is.

Definition 1.2. *Let (X, Σ, μ) be a probability space, k a positive integer and $\mathcal{S}$ a k-semiring on X with $\mathcal{S} \subseteq \Sigma$. For every $f \in L_1(X, \Sigma, \mu)$ we set*

$$\|f\|_{\mathcal{S}} = \sup\Big\{ \big| \int_S f \, d\mu \big| : S \in \mathcal{S} \Big\}. \tag{1.1}$$

The quantity $\|f\|_{\mathcal{S}}$ will be called the $\mathcal{S}$-uniformity norm *of f.*

The $\mathcal{S}$-uniformity norm is, in general, a seminorm. Note, however, that if the k-semiring $\mathcal{S}$ is sufficiently rich, then the function $\|\cdot\|_{\mathcal{S}}$ is indeed a norm. More precisely, the function $\|\cdot\|_{\mathcal{S}}$ is a norm if and only if the family $\{\mathbf{1}_S : S \in \mathcal{S}\}$ separates points in $L_1(X, \Sigma, \mu)$, that is, for every $f, g \in L_1(X, \Sigma, \mu)$ with $f \neq g$ there exists $S \in \mathcal{S}$ with $\int_S f \, d\mu \neq \int_S g \, d\mu$.

The simplest example of a k-semiring on a nonempty set X, is an algebra of subsets of X. Indeed, observe that a family of subsets of X is a 1-semiring if and only if it is an algebra. Another basic example is the collection of all intervals of a linearly ordered set, a family which is easily seen to be a 2-semiring. More interesting (and useful) k-semirings can be constructed with the following lemma.

LEMMA 1.3. *Let X be a nonempty set. Also let $m, k_1, \dots, k_m$ be positive integers and set $k = \sum_{i=1}^m k_i$. If $\mathcal{S}_i$ is a k_i-semiring on X for every $i \in [m]$, then the family*

$$\mathcal{S} = \Big\{ \bigcap_{i=1}^m S_i : S_i \in \mathcal{S}_i \text{ for every } i \in [m] \Big\} \tag{1.2}$$

is a k-semiring on X.

PROOF. Clearly we may assume that $m \geqslant 2$. Notice, first, that the family $\mathcal{S}$ satisfies properties (P1) and (P2) in Definition 1.1. To see that property (P3) is also satisfied, fix $S, T \in \mathcal{S}$ and write $S = \bigcap_{i=1}^m S_i$ and $T = \bigcap_{i=1}^m T_i$ where $S_i, T_i \in \mathcal{S}_i$ for every $i \in [m]$. We set $P_1 = X \setminus T_1$ and $P_j = T_1 \cap \dots \cap T_{j-1} \cap (X \setminus T_j)$ if $j \in \{2, \dots, m\}$. Observe that the sets $P_1, \dots, P_m$ are pairwise disjoint. Moreover,

$$X \setminus \Big(\bigcap_{i=1}^m T_i \Big) = \bigcup_{j=1}^m P_j \tag{1.3}$$

and so

$$S \setminus T = \Big(\bigcap_{i=1}^m S_i \Big) \setminus \Big(\bigcap_{i=1}^m T_i \Big) = \bigcup_{j=1}^m \Big(\bigcap_{i=1}^m S_i \cap P_j \Big). \tag{1.4}$$

Let $j \in [m]$ be arbitrary. Since $\mathcal{S}_j$ is a k_j-semiring, there exist $\ell_j \in [k_j]$ and pairwise disjoint sets $R_1^j, \dots, R_{\ell_j}^j \in \mathcal{S}_j$ such that $S_j \setminus T_j = R_1^j \cup \dots \cup R_{\ell_j}^j$. Thus, setting

(a) $B_1 = X$ and $B_j = \bigcap_{1 \leqslant i < j} (S_i \cap T_i)$ if $j \in \{2, \dots, m\}$,
(b) $C_j = \bigcap_{j < i \leqslant m} S_i$ if $j \in \{1, \dots, m-1\}$ and $C_m = X$,

and invoking the definition of the sets $P_1, \dots, P_m$ we obtain that

$$S \setminus T = \bigcup_{j=1}^m \Big(\bigcup_{n=1}^{\ell_j} \big(B_j \cap R_n^j \cap C_j \big) \Big). \tag{1.5}$$

Now set $I = \bigcup_{j=1}^m \big(\{j\} \times [\ell_j] \big)$ and observe that $|I| \leqslant k$. For every $(j, n) \in I$ let $U_n^j = B_j \cap R_n^j \cap C_j$ and notice that $U_n^j \in \mathcal{S}$, $U_n^j \subseteq R_n^j$ and $U_n^j \subseteq P_j$. It follows that the family $\{U_n^j : (j, n) \in I\}$ is contained in $\mathcal{S}$ and consists of pairwise disjoint sets. Moreover, by (1.5), we have

$$S \setminus T = \bigcup_{(j,n) \in I} U_n^j. \tag{1.6}$$

Hence, the family $\mathcal{S}$ satisfies property (P3) in Definition 1.1, as desired. □

By Lemma 1.3, we have the following corollary.

COROLLARY 1.4. *The following hold.*

(a) *Let X be a nonempty set. Also let k be a positive integer and for every $i\in[k]$ let $\mathcal{A}_i$ be an algebra on X. Then the family*

$$\{A_1\cap\dots\cap A_k : A_i\in\mathcal{A}_i \text{ for every } i\in[k]\} \tag{1.7}$$

is a k-semiring on X.

(b) *Let $d,k_1,\dots,k_d$ be a positive integers and set $k=\sum_{i=1}^d k_i$. Also let $X_1,\dots,X_d$ be nonempty sets and for every $i\in[d]$ let $\mathcal{S}_i$ be a k_i-semiring on X_i. Then the family*

$$\{S_1\times\dots\times S_d : S_i\in\mathcal{S}_i \text{ for every } i\in[d]\} \tag{1.8}$$

is k-semiring on $X_1\times\dots\times X_d$.

Next we isolate some basic properties of the $\mathcal{S}$-uniformity norm.

Lemma 1.5. *Let (X,Σ,μ) be a probability space, k a positive integer and $\mathcal{S}$ a k-semiring on X with $\mathcal{S}\subseteq\sigma$. Also let $f\in L_1(X,\Sigma,\mu)$. Then the following hold.*

(a) *We have $\|f\|_{\mathcal{S}}\leqslant\|f\|_{L_1}$.*
(b) *If $\mathcal{B}$ is a σ-algebra on X with $\mathcal{B}\subseteq\mathcal{S}$, then $\|\mathbb{E}(f\,|\,\mathcal{B})\|_{\mathcal{S}}\leqslant\|f\|_{\mathcal{S}}$.*
(c) *If $\mathcal{S}$ is a σ-algebra, then $\|f\|_{\mathcal{S}}\leqslant\|\mathbb{E}(f\,|\,\mathcal{S})\|_{L_1}\leqslant 2\|f\|_{\mathcal{S}}$.*

Proof. Part (a) is straightforward. For part (b), fix a σ-algebra $\mathcal{B}$ on X with $\mathcal{B}\subseteq\mathcal{S}$ and set $P=\{x\in X:\mathbb{E}(f\,|\,\mathcal{B})(x)\geqslant 0\}$ and $N=X\setminus P$. Notice that $P,N\in\mathcal{B}\subseteq\mathcal{S}$. Hence, for every $S\in\mathcal{S}$ we have

$$\begin{aligned}\Big|\int_S\mathbb{E}(f\,|\,\mathcal{B})\,d\mathbb{P}\Big| &\leqslant \max\Big\{\int_{P\cap S}\mathbb{E}(f\,|\,\mathcal{B})\,d\mathbb{P},-\int_{N\cap S}\mathbb{E}(f\,|\,\mathcal{B})\,d\mathbb{P}\Big\} \\ &\leqslant \max\Big\{\int_{P}\mathbb{E}(f\,|\,\mathcal{B})\,d\mathbb{P},-\int_{N}\mathbb{E}(f\,|\,\mathcal{B})\,d\mathbb{P}\Big\} \\ &= \max\Big\{\int_P f\,d\mathbb{P},-\int_N f\,d\mathbb{P}\Big\}\leqslant\|f\|_{\mathcal{S}}\end{aligned} \tag{1.9}$$

which yields that $\|\mathbb{E}(f\,|\,\mathcal{B})\|_{\mathcal{S}}\leqslant\|f\|_{\mathcal{S}}$.

Finally, assume that $\mathcal{S}$ is a σ-algebra and notice that $\int_S f\,d\mathbb{P}=\int_S\mathbb{E}(f\,|\,\mathcal{S})\,d\mathbb{P}$ for every $S\in\mathcal{S}$. In particular, we have $\|f\|_{\mathcal{S}}\leqslant\|\mathbb{E}(f\,|\,\mathcal{S})\|_{L_1}$. Also let, as above, $P=\{x\in X:\mathbb{E}(f\,|\,\mathcal{S})(x)\geqslant 0\}$ and $N=X\setminus P$. Since $P,N\in\mathcal{S}$ we obtain that

$$\|\mathbb{E}(f\,|\,\mathcal{S})\|_{L_1}\leqslant 2\cdot\max\Big\{\int_P\mathbb{E}(f\,|\,\mathcal{S})\,d\mathbb{P},-\int_N\mathbb{E}(f\,|\,\mathcal{S})\,d\mathbb{P}\Big\}\leqslant 2\|f\|_{\mathcal{S}} \tag{1.10}$$

and the proof is completed. □

We close this chapter by presenting some examples of k-semirings which are relevant from a combinatorial perspective. In the first example the underlying space is the Cartesian product of a finite sequence of nonempty finite sets. The corresponding semirings are related to the development of Szemerédi's regularity method for hypergraphs as we shall see in Part II.

EXAMPLE 1. Let $d \in \mathbb{N}$ with $d \geqslant 2$ and $V_1, \dots, V_d$ nonempty finite sets. We view the Cartesian product $V_1 \times \dots \times V_d$ as a discrete probability space equipped with the uniform probability measure. For every nonempty subset F of $[d]$ let $\pi_F \colon \prod_{i\in[d]} V_i \to \prod_{i\in F} V_i$ be the natural projection and set

$$\mathcal{A}_F = \Big\{ \pi_F^{-1}(A) : A \subseteq \prod_{i\in F} V_i \Big\}. \tag{1.11}$$

The family $\mathcal{A}_F$ is an algebra of subsets of $V_1 \times \dots \times V_d$ and consists of those sets which depend only on the coordinates determined by F.

More generally, let $\mathcal{F}$ be a family of nonempty subsets of $[d]$. Set $k = |\mathcal{F}|$ and observe that, by Corollary 1.4, we may associate with the family $\mathcal{F}$ a k-semiring $\mathcal{S}_{\mathcal{F}}$ on $V_1 \times \dots \times V_d$ defined by the rule

$$S \in \mathcal{S}_{\mathcal{F}} \Leftrightarrow S = \bigcap_{F\in\mathcal{F}} A_F \text{ where } A_F \in \mathcal{A}_F \text{ for every } F \in \mathcal{F}. \tag{1.12}$$

Notice that if the family $\mathcal{F}$ satisfies $[d] \notin \mathcal{F}$ and $\cup\mathcal{F} = [d]$, then it gives rise to a non-trivial semiring whose corresponding uniformity norm is a genuine norm.

It turns out that there is a minimal non-trivial semiring $\mathcal{S}_{\min}$ one can obtain in this way. It corresponds to the family $\mathcal{F}_{\min} = \binom{[d]}{1}$ and is particularly easy to grasp since it consists of all rectangles of $V_1 \times \dots \times V_d$. The $\mathcal{S}_{\min}$-uniformity norm is known as the *cut norm* and was introduced by Frieze and Kannan [FK99].

At the other extreme, this construction also yields a maximal non-trivial semiring $\mathcal{S}_{\max}$ on $V_1 \times \dots \times V_d$. It corresponds to the family $\mathcal{F}_{\max} = \binom{[d]}{d-1}$ and consists of those subsets of the product which can be written as $A_1 \cap \dots \cap A_d$ where for every $i \in [d]$ the set A_i does not depend on the i-th coordinate. The $\mathcal{S}_{\max}$-uniformity norm is known as the *Gowers box norm* and was introduced by Gowers [Gow06, Gow07]. This norm should not be confused with the box norms that are discussed in Chapter 7.

In the second example the underlying space is of the form $X \times X$ where X is the sample space of a probability space (X, Σ, μ). The corresponding semirings are related to the theory of convergence of graphs (see, e.g., [BCL$^+$08, Lov12]).

EXAMPLE 2. Let (X, Σ, μ) be a probability space and define

$$\mathcal{S}_{\Box} = \big\{ S \times T : S, T \in \Sigma \big\}. \tag{1.13}$$

That is, $\mathcal{S}_\Box$ is the family of all measurable rectangles of $X \times X$. By Corollary 1.4, we see that $\mathcal{S}_\Box$ is a 2-semiring on $X \times X$. The $\mathcal{S}_\Box$-uniformity norm is also referred to as the *cut norm* and is usually denoted by $\|\cdot\|_\Box$. In particular, for every integrable random variable $f\colon X \times X \to \mathbb{R}$ we have

$$\|f\|_\Box = \sup\Big\{\Big|\int_{S\times T} f\, d\mu\Big| : S, T \in \mathcal{F}\Big\}. \tag{1.14}$$

There is another natural semiring in this context which was introduced by Bollobás and Nikiforov [BN08] and can be considered as the "symmetric" version of $\mathcal{S}_\Box$. Specifically, let

$$\Sigma_\Box = \big\{S \times T : S, T \in \Sigma \text{ and either } S = T \text{ or } S \cap T = \emptyset\big\} \tag{1.15}$$

and observe that $\Sigma_\Box$ is a 4-semiring which is contained, of course, in $\mathcal{S}_\Box$. On the other hand, note that the family $\mathcal{S}_\Box$ is not much larger than $\Sigma_\Box$ since every element of $\mathcal{S}_\Box$ can be written as the disjoint union of at most 4 elements of $\Sigma_\Box$. Therefore, for every integrable random variable $f\colon X \times X \to \mathbb{R}$ we have

$$\|f\|_{\Sigma_\Box} \leqslant \|f\|_\Box \leqslant 4\|f\|_{\Sigma_\Box}. \tag{1.16}$$

CHAPTER 2

Regularity lemma via martingales

2.1. Backround material

A main ingredient towards the proof of the Regularity Lemma is the following martingale differences inequality.

2.1.1. A martingale difference sequence inequality. 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,

$$\Big(\sum_{i=0}^n \|d_i\|_{L_2}^2\Big)^{1/2} = \big\|\sum_{i=0}^n d_i\big\|_{L_2}. \tag{2.1}$$

We will need the following extension of this basic fact.

Proposition 2.1. *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*

$$\Big(\sum_{i=0}^n \|d_i\|_{L_p}^2\Big)^{1/2} \leqslant \Big(\frac{1}{p-1}\Big)^{1/2}\big\|\sum_{i=0}^n d_i\big\|_{L_p}. \tag{2.2}$$

It is a remarkable fact that the constant $(p-1)^{-1/2}$ appearing in the right-hand side of (A.5) is best possible. This sharp estimate was recently proved by Ricard and Xu [RX16]. The proof is presented in Appendix A.

2.1.2. Some pieces of notation. We now introduce some pieces of notation that we need in the statement and proof of the Regularity lemma that follows. For every pair k,ℓ of positive integers, every $0<\sigma\leqslant 1$, every $1<p\leqslant 2$ and every growth function $F\colon\mathbb{N}\to\mathbb{R}$ we define $h\colon\mathbb{N}\to\mathbb{N}$ recursively by the rule

$$\begin{cases} h(0)=0, \\ h(i+1)=h(i)+\lceil \sigma^2\,\ell\,F^{(h(i)+2)}(0)^2(p-1)^{-1}\rceil \end{cases} \tag{2.3}$$

and we set

$$R = h\big(\lceil \ell\,\sigma^{-2}(p-1)^{-1}\rceil - 1\big). \tag{2.4}$$

Finally, we define

$$\mathrm{Reg}(k,\ell,\sigma,p,F) = F^{(R)}(0). \tag{2.5}$$

Note that if $F\colon \mathbb{N}\to\mathbb{N}$ is a primitive recursive growth function which belongs to the class $\mathcal{E}^n$ of Grzegorczyk's hierarchy for some $n\in\mathbb{N}$ (see, e.g., [Ros84]), then the numbers $\mathrm{Reg}(k,\ell,\sigma,p,F)$ are controlled by a primitive recursive function belonging to the class $\mathcal{E}^m$ where $m=\max\{4,n+2\}$[1] .

2.2. Regularity Lemma

We are now ready to state the main result of this chapter.

THEOREM 2.2. *Let k,ℓ be positive integers, $0<\sigma\leqslant 1$, $1<p\leqslant 2$ and $F\colon\mathbb{N}\to\mathbb{R}$ a growth function. Also let (X,Σ,μ) be a probability space and $(\mathcal{S}_i)$ an increasing sequence of k-semirings on X with $\mathcal{S}_i\subseteq\Sigma$ for every $i\in\mathbb{N}$. Finally, let $\mathcal{C}$ be a family in $L_p(X,\Sigma,\mu)$ such that $\|f\|_{L_p}\leqslant 1$ for every $f\in\mathcal{C}$ and with $|\mathcal{C}|=\ell$. Then there exist*

(a) *a natural number N with $N\leqslant \mathrm{Reg}(k,\ell,\sigma,p,F)$,*
(b) *a partition $\mathcal{P}$ of X with $\mathcal{P}\subseteq\mathcal{S}_N$ and $|\mathcal{P}|\leqslant (k+1)^N$, and*
(c) *a finite refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q}\subseteq\mathcal{S}_i$ for some $i\geqslant N$*

such that for every $f\in\mathcal{C}$, writing $f=f_{\mathrm{str}}+f_{\mathrm{err}}+f_{\mathrm{unf}}$ where

$$f_{\mathrm{str}}=\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}}),\quad f_{\mathrm{err}}=\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\ \ \textit{and}\ \ f_{\mathrm{unf}}=f-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}}), \tag{2.6}$$

we have the estimates

$$\|f_{\mathrm{err}}\|_{L_p}\leqslant\sigma \quad \textit{and}\quad \|f_{\mathrm{unf}}\|_{\mathcal{S}_i}\leqslant\frac{1}{F(i)} \tag{2.7}$$

for every $i\in\{0,\dots,F(N)\}$.

The case "$p=2$" in Theorem 2.2 is essentially due to Tao [Tao06b, Tao06c, Tao11]. His approach, however, is somewhat different since he works with σ-algebras instead of k-semirings.

The increasing sequence $(\mathcal{S}_i)$ of k-semirings can be thought of as the higher-complexity analogue of the classical concept of a filtration in the theory of martingales. In fact, this is more than an analogy since, by applying Theorem 2.2 to appropriately selected filtrations, one is able to recover the fact that, for any $1<p\leqslant 2$, every L_p bounded martingale is L_p convergent. We discuss these issues in section 4.1.

[1]For more information about primitive recursive functions see [DK16, Appendix A]

We also note that the idea to obtain "uniformity" estimates with respect to an arbitrary growth function has been considered by several authors. This particular feature is essential when one wishes to iterate this structural decomposition (this is the case, for instance, in the context of hypergraphs – see, e.g., [Tao06c]). On the other hand, the need to "regularize", simultaneously, a finite family of random variables appears frequently in extremal combinatorics and related parts of Ramsey theory (see, e.g., [DKT14, DKK18]). Nevertheless, in most applications one deals with a single random variable and with a single semiring. Hence, we will isolate this special case in order to facilitate future references.

To this end, for every positive integer k, every $0<\sigma\leqslant 1$, every $1<p\leqslant 2$ and every growth function $F\colon \mathbb{N}\to\mathbb{R}$ we set

$$\mathrm{Reg}'(k,\sigma,p,F)=(k+1)^{\mathrm{Reg}(k,1,\sigma,p,F')} \tag{2.8}$$

where $F'\colon\mathbb{N}\to\mathbb{R}$ is the growth function defined by the rule $F'(n)=F\big((k+1)^n\big)$ for every $n\in\mathbb{N}$. We have the following corollary.

Corollary 2.3. *Let k be a positive integer, $0<\sigma\leqslant 1$, $1<p\leqslant 2$ and $F\colon\mathbb{N}\to\mathbb{R}$ a growth function. Also let (X,Σ,μ) be a probability space and let $\mathcal{S}$ be a k-semiring on X with $\mathcal{S}\subseteq\Sigma$. Finally, let $f\in L_p(X,\Sigma,\mu)$ with $\|f\|_{L_p}\leqslant 1$. Then there exist*

(a) *a positive integer M with $M\leqslant \mathrm{Reg}'(k,\sigma,p,F)$,*
(b) *a partition $\mathcal{P}$ of X with $\mathcal{P}\subseteq\mathcal{S}$ and $|\mathcal{P}|=M$, and*
(c) *a finite refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q}\subseteq\mathcal{S}$*

such that, writing $f=f_{\mathrm{str}}+f_{\mathrm{err}}+f_{\mathrm{unf}}$ where

$$f_{\mathrm{str}}=\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}}),\quad f_{\mathrm{err}}=\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\quad\textit{and}\quad f_{\mathrm{unf}}=f-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}}), \tag{2.9}$$

we have the estimates

$$\|f_{\mathrm{err}}\|_{L_p}\leqslant\sigma\quad\textit{and}\quad \|f_{\mathrm{unf}}\|_{\mathcal{S}}\leqslant\frac{1}{F(M)}. \tag{2.10}$$

Finally, we notice that the assumption that $1<p\leqslant 2$ in the above results is not restrictive, since the case of random variables in L_p for $p>2$ is reduced to the case $p=2$. On the other hand, we remark that Theorem 2.2 does not hold true for $p=1$ (see Section 4.1). Thus, the range of p in Theorem 2.2 is optimal.

2.2.1. Proof of Theorem 2.2. We start with the following lemma.

Lemma 2.4. *Let k be a positive integer, $p\geqslant 1$ and $0<\delta\leqslant 1$. Also let (X,Σ,μ) be a probability space, $\mathcal{S}$ a k-semiring on X with $\mathcal{S}\subseteq\Sigma$, $\mathcal{Q}$ a finite partition of X with $\mathcal{Q}\subseteq\mathcal{S}$ and $f\in L_p(X,\Sigma,\mu)$ with $\|f-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})\|_{\mathcal{S}}>\delta$. Then there exists a* *refinement $\mathcal{R}$ of $\mathcal{Q}$ with $\mathcal{R}\subseteq\mathcal{S}$ and $|\mathcal{R}|\leqslant|\mathcal{Q}|(k+1)$, and such that $\|\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{R}})-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})\|_{L_p}>\delta$.*

Proof. By our assumptions, there exists $S\in\mathcal{S}$ such that

$$\Big|\int_S \big(f-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})\big)\,d\mu\Big|>\delta. \tag{2.11}$$

Since $\mathcal{S}$ is a k-semiring on X, there exists a refinement $\mathcal{R}$ of $\mathcal{Q}$ such that: (i) $\mathcal{R}\subseteq\mathcal{S}$, (ii) $|\mathcal{R}|\leqslant|\mathcal{Q}|(k+1)$, and (iii) $S\in\mathcal{A}_{\mathcal{R}}$. It follows, in particular, that

$$\int_S \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{R}})\,d\mu=\int_S f\,d\mu. \tag{2.12}$$

Hence, by (2.11) and the monotonicity of the L_p norms, we obtain that

$$\begin{aligned}\delta &< \Big|\int_S \big(\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{R}})-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})\big)\,d\mu\Big| \\ &\leqslant \|\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{R}})-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})\|_{L_1}\leqslant\|\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{R}})-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})\|_{L_p}\end{aligned} \tag{2.13}$$

and the proof is completed. □

We proceed with the following lemma.

Lemma 2.5. *Let k,ℓ be positive integers, $0<\delta,\sigma\leqslant 1$ and $1<p\leqslant 2$, and set*

$$n=\Big\lceil\frac{\sigma^2\ell}{\delta^2(p-1)}\Big\rceil. \tag{2.14}$$

Also let (X,Σ,μ) be a probability space and let $(\mathcal{S}_i)$ be an increasing sequence of k-semirings on X with $\mathcal{S}_i\subseteq\Sigma$ for every $i\in\mathbb{N}$. Finally, let $m\in\mathbb{N}$ and $\mathcal{P}$ a partition of X with $\mathcal{P}\subseteq\mathcal{S}_m$ and $|\mathcal{P}|\leqslant(k+1)^m$. Then for every family $\mathcal{C}$ in $L_p(X,\Sigma,\mu)$ with $|\mathcal{C}|=\ell$ there exist $j\in\{m,\dots,m+n\}$ and a refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q}\subseteq\mathcal{S}_j$ and $|\mathcal{Q}|\leqslant(k+1)^j$, and such that either

(a) *$\|\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\|_{L_p}>\sigma$ for some $f\in\mathcal{C}$, or*
(b) *$\|\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\|_{L_p}\leqslant\sigma$ and $\|f-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})\|_{\mathcal{S}_{j+1}}\leqslant\delta$ for every $f\in\mathcal{C}$.*

The case "$p=2$" in Lemma 2.5 can be proved with an "energy increment strategy" which ultimately depends upon the fact that martingale difference sequences are orthogonal in L_2 (see, e.g., [Tao06b, Theorem 2.11]). In the non-Hilbertian case (that is, when $1<p<2$) the geometry is more subtle and we will rely, instead, on Proposition 2.1. The argument can therefore be seen as the L_p-version of the "energy increment strategy". More applications of this method are given in the next chapter 6.

Proof of Lemma 2.5. Assume that the first part of the lemma is not satisfied. Note that this is equivalent to saying that

(H1) for every $j \in \{m, \dots, m+n\}$, every refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q} \subseteq \mathcal{S}_j$ and $|\mathcal{Q}| \leqslant (k+1)^j$ and every $f \in \mathcal{C}$ we have $\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} \leqslant \sigma$.

We will use hypothesis (H1) to show that part (b) is satisfied.

To this end we will argue by contradiction. Let $j \in \{m, \dots, m+n\}$ and let $\mathcal{Q}$ be a refinement of $\mathcal{P}$ with $\mathcal{Q} \subseteq \mathcal{S}_j$ and $|\mathcal{Q}| \leqslant (k+1)^j$. Observe that hypothesis (H1) and our assumption that part (b) does not hold true, imply that there exists $f \in \mathcal{C}$ (possibly depending on the partition $\mathcal{Q}$) such that $\|f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}})\|_{\mathcal{S}_{j+1}} > \delta$. Since the sequence $(\mathcal{S}_i)$ is increasing, Lemma 2.4 can be applied to the k-semiring $\mathcal{S}_{j+1}$, the partition $\mathcal{Q}$ and the random variable f. Hence, we obtain that

(H2) for every $j \in \{m, \dots, m+n\}$ and every refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q} \subseteq \mathcal{S}_j$ and $|\mathcal{Q}| \leqslant (k+1)^j$ there exist $f \in \mathcal{C}$ and a refinement $\mathcal{R}$ of $\mathcal{Q}$ with $\mathcal{R} \subseteq \mathcal{S}_{j+1}$ and $|\mathcal{R}| \leqslant (k+1)^{j+1}$, and such that $\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{R}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}})\|_{L_p} > \delta$.

Recursively and using hypothesis (H2), we select a finite sequence $\mathcal{P}_0, \dots, \mathcal{P}_n$ of partitions of X with $\mathcal{P}_0 = \mathcal{P}$ and a finite sequence $f_1, \dots, f_n$ in $\mathcal{C}$ such that for every $i \in [n]$ we have: (P1) $\mathcal{P}_i$ is a refinement of $\mathcal{P}_{i-1}$, (P2) $\mathcal{P}_i \subseteq \mathcal{S}_{m+i}$ and $|\mathcal{P}_i| \leqslant (k+1)^{m+i}$, and (P3) $\|\mathbb{E}(f_i \,|\, \mathcal{A}_{\mathcal{P}_i}) - \mathbb{E}(f_i \,|\, \mathcal{A}_{\mathcal{P}_{i-1}})\|_{L_p} > \delta$. It follows, in particular, that $(\mathcal{A}_{\mathcal{P}_i})_{i=0}^n$ is an increasing sequence of finite sub-σ-algebras of Σ. Also note that, by the classical pigeonhole principle and the fact that $|\mathcal{C}| = \ell$, there exist $g \in \mathcal{C}$ and $I \subseteq [n]$ with $|I| \geqslant n/\ell$ and such that $g = f_i$ for every $i \in I$.

Next, set $f = g - \mathbb{E}(g \,|\, \mathcal{A}_{\mathcal{P}})$ and let $(d_i)_{i=0}^n$ be the difference sequence associated with the finite martingale $\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_0}), \dots, \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_n})$. Observe that for every $i \in I$ we have $d_i = \mathbb{E}(g \,|\, \mathcal{A}_{\mathcal{P}_i}) - \mathbb{E}(g \,|\, \mathcal{A}_{\mathcal{P}_{i-1}})$ and so, by the choice of I and property (P3), we obtain that $\|d_i\|_{L_p} > \delta$ for every $i \in I$. Therefore, by Proposition 2.1, we have

$$\begin{aligned} \sigma \ &\overset{(2.14)}{\leqslant} \ \sqrt{p-1}\,\delta\Big(\frac{n}{\ell}\Big)^{1/2} \leqslant \sqrt{p-1}\,\delta|I|^{1/2} \\ &< \ \sqrt{p-1}\cdot\Big(\sum_{i=0}^n \|d_i\|_{L_p}^2\Big)^{1/2} \\ &\overset{(A.5)}{\leqslant} \ \Big\|\sum_{i=0}^n d_i\Big\|_{L_p} = \|\mathbb{E}(g \,|\, \mathcal{A}_{\mathcal{P}_n}) - \mathbb{E}(g \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p}. \end{aligned} \tag{2.15}$$

On the other hand, by properties (P1) and (P2), we see that $\mathcal{P}_n$ is a refinement of $\mathcal{P}$ with $\mathcal{P}_n \subseteq \mathcal{S}_{m+n}$ and $|\mathcal{P}_n| \leqslant (k+1)^{m+n}$. Therefore, by hypothesis (H1), we must have $\|\mathbb{E}(g \,|\, \mathcal{A}_{\mathcal{P}_n}) - \mathbb{E}(g \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} \leqslant \sigma$ which contradicts, of course, the estimate in (2.15). The proof of Lemma 2.5 is thus completed. □

The following lemma is the last step of the proof of Theorem 2.2.

LEMMA 2.6. *Let k,ℓ be positive integers, $0<\sigma\leqslant 1$, $1<p\leqslant 2$ and $H\colon \mathbb{N}\to\mathbb{R}$ a growth function. Set $L=\lceil \ell\,\sigma^{-2}(p-1)^{-1}\rceil$ and define (n_i) recursively by the rule*

$$\begin{cases} n_0=0, \\ n_{i+1}=n_i+\lceil \sigma^2\,\ell\, H(n_i)^2(p-1)^{-1}\rceil. \end{cases} \tag{2.16}$$

Also let (X,Σ,μ) be a probability space and let $(\mathcal{S}_i)$ be an increasing sequence of k-semirings on X with $\mathcal{S}_i\subseteq\Sigma$ for every $i\in\mathbb{N}$. Finally, let $\mathcal{C}$ be a family in $L_p(X,\Sigma,\mu)$ such that $\|f\|_{L_p}\leqslant 1$ for every $f\in\mathcal{C}$ and with $|\mathcal{C}|=\ell$. Then there exist $j\in\{0,\dots,L-1\}$, $J\in\{n_j,\dots,n_{j+1}\}$ and two partitions $\mathcal{P},\mathcal{Q}$ of X with the following properties: (i) *$\mathcal{P}\subseteq\mathcal{S}_{n_j}$ and $\mathcal{Q}\subseteq\mathcal{S}_J$,* (ii) *$|\mathcal{P}|\leqslant (k+1)^{n_j}$ and $|\mathcal{Q}|\leqslant (k+1)^J$,* (iii) *$\mathcal{Q}$ is a refinement of $\mathcal{P}$, and* (iv) *$\|\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\|_{L_p}\leqslant\sigma$ and $\|f-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})\|_{\mathcal{S}_{J+1}}\leqslant 1/H(n_j)$ for every $f\in\mathcal{C}$.*

PROOF. It is similar to the proof of Lemma 2.5. Indeed, assume, towards a contradiction, that the lemma is false. Recursively and using Lemma 2.5, we select a finite sequence $J_0,\dots,J_L$ in $\mathbb{N}$ with $J_0=0$, a finite sequence $\mathcal{P}_0,\dots,\mathcal{P}_L$ of partitions of X with $\mathcal{P}_0=\{X\}$ and a finite sequence $f_1,\dots,f_L$ in $\mathcal{C}$ such that for every $i\in[L]$ we have that: (P1) $J_i\in\{n_{i-1},\dots,n_i\}$, (P2) the partition $\mathcal{P}_i$ is a refinement of $\mathcal{P}_{i-1}$, (P3) $\mathcal{P}_i\subseteq\mathcal{S}_{J_i}$ with $|\mathcal{P}_i|\leqslant(k+1)^{J_i}$, and (P4) $\|\mathbb{E}(f_i\,|\,\mathcal{A}_{\mathcal{P}_i})-\mathbb{E}(f_i\,|\,\mathcal{A}_{\mathcal{P}_{i-1}})\|_{L_p}>\sigma$. As in the proof of Lemma 2.5, we observe that $(\mathcal{A}_{\mathcal{P}_i})_{i=0}^L$ is an increasing sequence of finite sub-σ-algebras of Σ, and we select $g\in\mathcal{C}$ and $I\subseteq[L]$ with $|I|\geqslant L/\ell$ and such that $g=f_i$ for every $i\in I$. Let $(d_i)_{i=0}^L$ be the difference sequence associated with the finite martingale $\mathbb{E}(g\,|\,\mathcal{A}_{\mathcal{P}_0}),\dots,\mathbb{E}(g\,|\,\mathcal{A}_{\mathcal{P}_L})$. Notice that, by property (P4), we have $\|d_i\|_{L_p}>\sigma$ for every $i\in I$. Hence, by the choice of L, Proposition 2.1 and the fact that $\|g\|_{L_p}\leqslant 1$, we conclude that

$$\begin{aligned} 1 \quad &\leqslant \quad \sqrt{p-1}\,\sigma|I|^{1/2} < \sqrt{p-1}\cdot\Big(\sum_{i=0}^{L}\|d_i\|_{L_p}^2\Big)^{1/2} \\ &\overset{(A.5)}{\leqslant} \quad \Big\|\sum_{i=0}^{L} d_i\Big\|_{L_p} = \|\mathbb{E}(g\,|\,\mathcal{A}_{\mathcal{P}_L})\|_{L_p}\leqslant\|g\|_{L_p}\leqslant 1 \end{aligned} \tag{2.17}$$

which is clearly a contradiction. The proof of Lemma 2.6 is completed. □

We are ready to complete the proof of Theorem 2.2.

PROOF OF THEOREM 2.2. Fix the data k,ℓ,σ,p, the growth function F, the sequence $(\mathcal{S}_i)$ and the family $\mathcal{C}$. We define $H\colon\mathbb{N}\to\mathbb{R}$ by the rule $H(n)=F^{(n+2)}(0)$ and we observe that H is a growth function. Moreover, for every $i\in\mathbb{N}$ let $m_i=F^{(i)}(0)$ and set $\mathcal{S}'_i=\mathcal{S}_{m_i}$. Notice that $(\mathcal{S}'_i)$ is an increasing sequence of k-semirings of X with $\mathcal{S}'_i\subseteq\Sigma$ for every $i\in\mathbb{N}$.

Let $j, J, \mathcal{P}$ and $\mathcal{Q}$ be as in Lemma 2.6 when applied to k, ℓ, σ, p, H, the sequence (Σ_i) and the family $\mathcal{C}$. We set

$$N = m_{n_j} = F^{(n_j)}(0) \tag{2.18}$$

and we claim that the natural number N and the partitions $\mathcal{P}$ and $\mathcal{Q}$ are as desired.

Indeed, notice first that $n_j \leqslant n_{L-1}$. Since F is a growth function, by the choice of h and R in (2.3) and (2.4) respectively, we have

$$N \leqslant F^{(n_{L-1})}(0) = F^{(R)}(0) \overset{(2.5)}{=} \mathrm{Reg}(k, \ell, \sigma, p, F). \tag{2.19}$$

On the other hand, note that $n_j \leqslant F^{(n_j)}(0) = N$ and so $|\mathcal{P}| \leqslant (k+1)^{n_j} \leqslant (k+1)^N$ and $\mathcal{P} \subseteq \mathcal{S}'_{n_j} = \mathcal{S}_N$. Moreover, by Lemma 2.6, we see that $\mathcal{Q}$ is a finite refinement of $\mathcal{P}$ with $\mathcal{Q} \subseteq \mathcal{S}_i$ for some $i \geqslant N$. It follows that $N, \mathcal{P}$ and $\mathcal{Q}$ satisfy the requirements of the theorem. Finally, let $f \in \mathcal{C}$ be arbitrary and write $f = f_{\mathrm{str}} + f_{\mathrm{err}} + f_{\mathrm{unf}}$ where $f_{\mathrm{str}} = \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})$, $f_{\mathrm{err}} = \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})$ and $f_{\mathrm{unf}} = f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}})$. Invoking Lemma 2.6, we obtain that

$$\|f_{\mathrm{err}}\|_{L_p} = \|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} \leqslant \sigma. \tag{2.20}$$

Also observe that $n_j + 1 \leqslant J + 1$ which is easily seen to imply that $\mathcal{S}_{F(N)} \subseteq \mathcal{S}'_{J+1}$. Therefore, using Lemma 2.6 once again, for every $i \in \{0, \dots, F(N)\}$ we have

$$\begin{aligned} \|f_{\mathrm{unf}}\|_{\mathcal{S}_i} &= \|f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}})\|_{\mathcal{S}_i} \leqslant \|f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}})\|_{\mathcal{S}'_{J+1}} \\ &\leqslant \frac{1}{H(n_j)} = \frac{1}{F\big(F(N)\big)} \leqslant \frac{1}{F(i)}. \end{aligned} \tag{2.21}$$

The proof of Theorem 2.2 is completed. $\square$

CHAPTER 3

Applications of the regularity lemma

In this chapter we present two applications of Theorem 2.2. More applications of Theorem 2.2, such as the well-known Szemerédi's regularity lemma ([Sze78, Tao06b]) may be found in [DK16].

3.1. Martingale convergence theorem

Our goal in this section is to use Theorem 2.2 to show the well-known fact that, for any $1 < p \leqslant 2$, every L_p bounded martingale is L_p convergent (see, e.g., [Dur10]). Besides its intrinsic interest, this result also implies that Theorem 2.2 does not hold true for the end-point case $p = 1$. In fact, based on the argument below, one can easily construct a counterexample to Theorem 2.2 using any L_1 bounded martingale which is not L_1 convergent.

We will need the following known approximation result (see, e.g., [Pis16]). We recall the proof for the convenience of the reader.

Lemma 3.1. *Let (X,Σ,μ) be a probability space and $p \geqslant 1$. Also let (g_i) be a martingale in $L_p(X,\Sigma,\mu)$ and $\delta > 0$. Then there exist an increasing sequence (Σ_i) of finite sub-σ-algebras of Σ and a martingale (f_i) adapted to the filtration (Σ_i) such that $\|g_i - f_i\|_{L_p} \leqslant \delta$ for every $i \in \mathbb{N}$.*

Proof. Fix a filtration $(\mathcal{B}_i)$ such that (g_i) is adapted to $(\mathcal{B}_i)$ and let (Δ_i) be the martingale difference sequence associated with (g_i). Recursively and using the fact that the set of simple functions is dense in L_p, we select an increasing sequence (Σ_i) of finite sub-σ-algebras of Σ and a sequence (s_i) of simple functions such that for every $i \in \mathbb{N}$ we have that: (i) Σ_i is contained in $\mathcal{B}_i$, (ii) $\|\Delta_i - s_i\|_{L_p} \leqslant \delta/2^{i+2}$, and (iii) $s_i \in L_p(X,\Sigma,\mu)$. For every $i \in \mathbb{N}$ let $d_i = \mathbb{E}(\Delta_i \,|\, \Sigma_i)$ and notice that the sequence (d_i) is a martingale difference sequence since, by (i),

$$\begin{aligned} \mathbb{E}(d_{i+1} \,|\, \Sigma_i) &= \mathbb{E}\big(\mathbb{E}(\Delta_{i+1} \,|\, \mathcal{F}_{i+1}) \,|\, \Sigma_i\big) \\ &= \mathbb{E}(\Delta_{i+1} \,|\, \Sigma_i) = \mathbb{E}\big(\mathbb{E}(\Delta_{i+1} \,|\, \mathcal{B}_i) \,|\, \Sigma_i\big) = 0. \end{aligned} \tag{3.1}$$

Thus, setting $f_i = d_0 + \cdots + d_i$, we see that (f_i) is a martingale adapted to the filtration (Σ_i). Moreover, by (ii) and (iii), for every $i \in \mathbb{N}$ we have

$$\begin{aligned} \|g_i - f_i\|_{L_p} &\leqslant \sum_{k=0}^{i} \|\Delta_k - d_k\|_{L_p} \leqslant \frac{\delta}{2} + \sum_{k=0}^{i} \|s_k - d_k\|_{L_p} \\ &= \frac{\delta}{2} + \sum_{k=0}^{i} \|\mathbb{E}(s_k - \Delta_k \,|\, \Sigma_k)\|_{L_p} \leqslant \frac{\delta}{2} + \sum_{k=0}^{i} \|s_k - \Delta_k\|_{L_p} \leqslant \delta \end{aligned} \tag{3.2}$$

and the proof is completed. □

We will also need the following well known fact that martingale difference sequences are monotone basic sequence in L_p, for $p \geqslant 1$, i.e. if $(d_i)_{i=0}^n$ is a martingale difference sequence in L_p for some $p \geqslant 1$, then for every $0 \leqslant k \leqslant n$ and every $a_0, \ldots, a_n \in \mathbb{R}$ we have

$$\|\sum_{i=0}^{k} a_i d_i\|_{L_p} \leqslant \|\sum_{i=0}^{n} a_i d_i\|_{L_p}. \tag{3.3}$$

In particular,

$$\|\sum_{i=k}^{\ell} d_i\|_{L_p} \leqslant 2\|\sum_{i=0}^{n} d_i\|_{L_p}, \tag{3.4}$$

for every $0 \leqslant k \leqslant \ell \leqslant n$.[1] We pass now to the main theorem of this section.

THEOREM 3.2. *Let $1 < p \leqslant 2$ and (X, Σ, μ) be a probability space. Then any $L_p(X, \Sigma, \mu)$ bounded martingale is L_p convergent.*

Assume, towards a contradiction, that there exists a bounded martingale (g_i) in $L_p(X, \Sigma, \mu)$ which is not norm convergent. By (3.4), we see that (g_i) has no convergent subsequence whatsoever. Therefore, by passing to a subsequence of (g_i) and rescaling, we may assume that there exists $0 < \varepsilon \leqslant 1/3$ such that: (i) $\|g_i\|_{L_p} \leqslant 1/2$ for every $i \in \mathbb{N}$, and (ii) $\|g_i - g_j\|_{L_p} \geqslant 3\varepsilon$ for every $i, j \in \mathbb{N}$ with $i \neq j$. By Lemma 3.1 applied to the martingale (g_i) and the constant "$\delta = \varepsilon$", there exist

(P1) an increasing sequence (Σ_i) of finite sub-σ-algebras of Σ, and
(P2) a martingale (f_i) adapted to the filtration (Σ_i)

such that $\|g_i - f_i\|_{L_p} \leqslant \varepsilon$ for every $i \in \mathbb{N}$. Hence,

(P3) $\|f_i\|_{L_p} \leqslant 1$ for every $i \in \mathbb{N}$, and
(P4) $\|f_i - f_j\|_{L_p} \geqslant \varepsilon$ for every $i, j \in \mathbb{N}$ with $i \neq j$.

---

[1] For further properties of martingale difference sequences see Appendix A.

Notice that, by (P1), for every $i \in \mathbb{N}$ the space $L_p(X, \Sigma, \mu)$ is finite-dimensional. Since $\|\cdot\|_{\Sigma_i}$ is a norm on $L_p(X, \Sigma, \mu)$, there exists a constant $C_i \geqslant 1$ such that

$$\|f\|_{\mathcal{F}_i} \leqslant \|f\|_{L_p} \leqslant C_i \|f\|_{\Sigma_i} \tag{3.5}$$

for every $f \in L_p(X, \Sigma, \mu)$.

Define $F \colon \mathbb{N} \to \mathbb{R}$ by the rule

$$F(i) = (i+1) + (8/\varepsilon) \sum_{j=0}^{i} C_i \tag{3.6}$$

and observe that F is a growth function. Next, set

$$n = F\big(\mathrm{Reg}(1, 1, \varepsilon/8, p, F)\big) + 1 \tag{3.7}$$

and let $(\mathcal{S}_i)$ be defined by $\mathcal{S}_i = \Sigma_i$ if $i \leqslant n$ and $\mathcal{S}_i = \Sigma_n$ if $i > n$. Clearly, $(\mathcal{S}_i)$ is an increasing sequence of 1-semirings on X. We apply Theorem 2.2 to the probability space (X, Σ_n, μ), the sequence $(\mathcal{S}_i)$ and the random variable f_n, and we obtain a natural number $N \leqslant \mathrm{Reg}(1, 1, \varepsilon/8, p, F)$, a finite partition $\mathcal{P}$ of X with $\mathcal{P} \subseteq \mathcal{S}_N$ and a finite refinement $\mathcal{Q}$ of $\mathcal{P}$ such that, writing $f_n = f_{\mathrm{str}} + f_{\mathrm{err}} + f_{\mathrm{unf}}$ where

$$f_{\mathrm{str}} = \mathbb{E}(f_n \,|\, \mathcal{A}_{\mathcal{P}}), \quad f_{\mathrm{err}} = \mathbb{E}(f_n \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f_n \,|\, \mathcal{A}_{\mathcal{P}}) \;\text{ and }\; f_{\mathrm{unf}} = f_n - \mathbb{E}(f_n \,|\, \mathcal{A}_{\mathcal{Q}}),$$

we have that $\|f_{\mathrm{err}}\|_{L_p} \leqslant \varepsilon/8$ and $\|f_{\mathrm{unf}}\|_{\mathcal{S}_i} \leqslant 1/F(i)$ for every $i \in \{0, \dots, F(N)\}$. In particular, by the choice of n and $(\mathcal{S}_i)$, we see that

$$\|f_{\mathrm{err}}\|_{L_p} \leqslant \frac{\varepsilon}{8} \;\text{ and }\; \|f_{\mathrm{unf}}\|_{\Sigma_{N+1}} \leqslant \frac{1}{F(N+1)}. \tag{3.8}$$

Now observe that, by property (P2),

$$f_N = \mathbb{E}(f_n \,|\, \Sigma_N) = \mathbb{E}(f_{\mathrm{str}} \,|\, \Sigma_N) + \mathbb{E}(f_{\mathrm{err}} \,|\, \Sigma_N) + \mathbb{E}(f_{\mathrm{unf}} \,|\, \Sigma_N) \tag{3.9}$$

and, similarly,

$$f_{N+1} = \mathbb{E}(f_n \,|\, \Sigma_{N+1}) = \mathbb{E}(f_{\mathrm{str}} \,|\, \Sigma_{N+1}) + \mathbb{E}(f_{\mathrm{err}} \,|\, \Sigma_{N+1}) + \mathbb{E}(f_{\mathrm{unf}} \,|\, \Sigma_{N+1}). \tag{3.10}$$

The fact that $\mathcal{P} \subseteq \mathcal{S}_N$ yields that $\mathcal{A}_{\mathcal{P}} \subseteq \Sigma_N \subseteq \Sigma_{N+1}$ and so

$$f_{\mathrm{str}} = \mathbb{E}(f_{\mathrm{str}} \,|\, \Sigma_N) = \mathbb{E}(f_{\mathrm{str}} \,|\, \Sigma_{N+1}). \tag{3.11}$$

On the other hand, by (3.8), we have

$$\|\mathbb{E}(f_{\mathrm{err}} \,|\, \Sigma_N)\|_{L_p} \leqslant \frac{\varepsilon}{8} \;\text{ and }\; \|\mathbb{E}(f_{\mathrm{err}} \,|\, \Sigma_{N+1})\|_{L_p} \leqslant \frac{\varepsilon}{8}. \tag{3.12}$$

Finally, notice that $\mathbb{E}(f_{\mathrm{unf}} \,|\, \Sigma_N) \in L_p(X, \Sigma, \mu)$. Thus, by (3.5) and Lemma 1.5, we obtain that

$$\begin{aligned} \|\mathbb{E}(f_{\mathrm{unf}} \,|\, \Sigma_N)\|_{L_p} &\leqslant C_N \|\mathbb{E}(f_{\mathrm{unf}} \,|\, \Sigma_N)\|_{\Sigma_N} \leqslant C_N \|f_{\mathrm{unf}}\|_{\Sigma_N} \\ &\leqslant C_N \|f_{\mathrm{unf}}\|_{\Sigma_{N+1}} \overset{(3.8)}{\leqslant} \frac{C_N}{F(N+1)} \overset{(3.6)}{\leqslant} \frac{\varepsilon}{8}. \end{aligned} \tag{3.13}$$

With identical arguments we see that

$$\|\mathbb{E}(f_{\mathrm{unf}} \,|\, \Sigma_{N+1})\|_{L_p} \leqslant \frac{\varepsilon}{8}. \tag{3.14}$$

Combining (3.9)–(3.14), we conclude that $\|f_N - f_{N+1}\|_{L_p} \leqslant \varepsilon/2$ which contradicts, of course, property (P4). Hence, every bounded martingale in $L_p(X, \Sigma, \mu)$ is norm convergent, as desired.

3.2. Weak and strong regularity lemmas for graphons

We now extend the, so-called, *strong regularity lemma for L_2 graphons* (see, e.g., [Lov12, LS07]).
Let (X, Σ, μ) and W be an L_p *graphon.*[2] Also, let $\mathcal{R}$ be a finite partition of X with $\mathcal{R} \subseteq \Sigma$ and notice that the family

$$\mathcal{R}^2 = \{S \times T : S, T \in \mathcal{R}\} \tag{3.15}$$

is a finite partition of $X \times X$. As in Chapter 1, let $\mathcal{A}_{\mathcal{R}^2}$ be the σ-algebra on $X \times X$ generated by $\mathcal{R}^2$ and observe that $\mathcal{A}_{\mathcal{R}^2}$ consists of measurable sets. If $W \colon X \times X \to \mathbb{R}$ is a graphon, then the conditional expectation of W with respect to $\mathcal{A}_{\mathcal{R}^2}$ is usually denoted by $W_{\mathcal{R}}$. Note that $W_{\mathcal{R}}$ is also a graphon and satisfies (see, e.g., [Lov12])

$$\|W_{\mathcal{R}}\|_{\square} \leqslant \|W\|_{\square} \tag{3.16}$$

where $\|\cdot\|_{\square}$ is the cut norm defined in (1.14). On the other hand, by standard properties of the conditional expectation (see, e.g., [Dur10]), we have $\|W_{\mathcal{R}}\|_{L_p} \leqslant \|W\|_{L_p}$ for any $p \geqslant 1$. It follows, in particular, that $W_{\mathcal{R}}$ is an L_p graphon provided, of course, that $W \in L_p$.

We have the following Proposition.

Proposition 3.3 (Strong regularity lemma for L_p graphons). *For every $0 < \varepsilon \leqslant 1$, every $1 < p \leqslant 2$ and every positive function $h \colon \mathbb{N} \to \mathbb{R}$ there exists a positive integer $\mathrm{s}(\varepsilon, p, h)$ with the following property. If (X, Σ, μ) is a probability space and $W \colon X \times X \to \mathbb{R}$ is an L_p graphon with $\|W\|_{L_p} \leqslant 1$, then there exist a partition $\mathcal{R}$*

[2] For the definition of a graphon see Basic Concepts & General Notation in the beging of the thesis. Also for further results about L_p graphons see [BR09, BCCZ14].

of X with $\mathcal{R}\subseteq\Sigma$ and $|\mathcal{R}|\leqslant \mathrm{s}(\varepsilon,p,h)$, and an L_p graphon $U\colon X\times X\to\mathbb{R}$ such that $\|W-U\|_{L_p}\leqslant\varepsilon$ and $\|U-U_{\mathcal{R}}\|_\square\leqslant h\big(|\mathcal{R}|\big)$.

Proof. Fix the constants ε,p and the function h, and define $F\colon\mathbb{N}\to\mathbb{R}$ by the rule

$$F(n)=(n+1)+\sum_{i=0}^{n}\frac{8}{h(i)}. \tag{3.17}$$

Notice that F is a growth function. We set

$$\mathrm{s}(\varepsilon,p,h)=\mathrm{Reg}'(4,\varepsilon,p,F) \tag{3.18}$$

and we claim that with this choice the result follows.

Indeed, let (X,Σ,μ) be a probability space and fix an L_p graphon $W\colon X\times X\to\mathbb{R}$ with $\|W\|_{L_p}\leqslant 1$. Also let $\Sigma_\square$ be the 4-semiring on $X\times X$ which is defined via formula (1.15) for the given probability space (X,Σ,μ). We apply Corollary 2.3 to $\Sigma_\square$ and the random variable W and we obtain

(a) a partition $\mathcal{P}$ of $X\times X$ with $\mathcal{P}\subseteq\Sigma_\square$ and $|\mathcal{P}|\leqslant\mathrm{Reg}'(4,\varepsilon,p,F)$, and
(b) a finite refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q}\subseteq\Sigma_\square$

such that, writing the graphon W as $W_{\mathrm{str}}+W_{\mathrm{err}}+W_{\mathrm{str}}$ where $W_{\mathrm{str}}=\mathbb{E}(W\,|\,\mathcal{A}_{\mathcal{P}})$, $W_{\mathrm{err}}=\mathbb{E}(W\,|\,\mathcal{A}_{\mathcal{Q}})-\mathbb{E}(W\,|\,\mathcal{A}_{\mathcal{P}})$ and $W_{\mathrm{unf}}=W-\mathbb{E}(W\,|\,\mathcal{A}_{\mathcal{Q}})$, we have the estimates $\|W_{\mathrm{err}}\|_{L_p}\leqslant\varepsilon$ and $\|W_{\mathrm{unf}}\|_{\Sigma_\square}\leqslant 1/F\big(|\mathcal{P}|\big)$. Note that, by (a) and (b) and the definition of the 4-semiring $\Sigma_\square$ in (1.15), there exist two finite partitions $\mathcal{R},\mathcal{Z}$ of X with $\mathcal{R},\mathcal{Z}\subseteq\Sigma$ and such that $\mathcal{P}=\mathcal{R}^2$ and $\mathcal{Q}=\mathcal{Z}^2$. It follows, in particular, that the random variables $W_{\mathrm{str}},W_{\mathrm{err}}$ and W_{unf} are all L_p graphons.

We will show that the partition $\mathcal{R}$ and the L_p graphon $U:=W_{\mathrm{str}}+W_{\mathrm{unf}}$ are as desired. To this end notice first that

$$|\mathcal{R}|\leqslant|\mathcal{R}^2|=|\mathcal{P}|\leqslant\mathrm{Reg}'(4,\varepsilon,p,F)\overset{(3.18)}{=}\mathrm{s}(\varepsilon,p,h). \tag{3.19}$$

Next observe that

$$\|W-U\|_{L_p}=\|W_{\mathrm{err}}\|_{L_p}\leqslant\varepsilon. \tag{3.20}$$

Finally note that, by (3.16), we have $\|(W_{\mathrm{unf}})_{\mathcal{R}}\|_\square\leqslant\|W_{\mathrm{unf}}\|_\square$. Moreover, the fact that $\mathcal{P}=\mathcal{R}^2$ and the choice of W_{str} yield that $(W_{\mathrm{str}})_{\mathcal{R}}=W_{\mathrm{str}}$. Therefore,

$$\begin{aligned}\|U-U_{\mathcal{R}}\|_\square &\leqslant 2\|W_{\mathrm{unf}}\|_\square\overset{(1.16)}{\leqslant}8\|W_{\mathrm{unf}}\|_{\Sigma_\square}\leqslant\frac{8}{F\big(|\mathcal{P}|\big)}\\ &\overset{(3.19)}{\leqslant}\frac{8}{F\big(|\mathcal{R}|\big)}\overset{(3.17)}{\leqslant}h\big(|\mathcal{R}|\big)\end{aligned} \tag{3.21}$$

and the proof of Corollary 3.3 is completed. $\square$

We pass now to the so called weak regularity lemma. In [BCCZ14] Borgs, Chayes, Cohn and Zhao extended the weak regularity lemma that already existed for L_2 graphons (see, e.g., [Lov12]) to L_p graphons for any $p > 1$. Their extension follows, of course, from Proposition 3.3, but this reduction is rather ineffective since the bound obtained by Proposition 3.3 is quite poor. However, this estimate can be significantly improved if instead of invoking Corollary 2.3, one argues directly as in the proof of Lemma 2.5. More precisely, we have the following result.

PROPOSITION 3.4 (Weak regularity lemma for L_p graphons.). *For every $0 < \varepsilon \leqslant 1$, every $1 < p \leqslant 2$, every probability space (X, Σ, μ) and every L_p graphon $W : X \times X \to \mathbb{R}$ with $\|W\|_{L_p} \leqslant 1$ there exists a partition $\mathcal{R}$ of X with $\mathcal{R} \subseteq \Sigma$ and*

$$|\mathcal{R}| \leqslant 4^{(p-1)^{-1}\varepsilon^{-2}} \tag{3.22}$$

and such that $\|W - W_{\mathcal{R}}\|_{\square} \leqslant \varepsilon$.

The estimate in (3.22) matches the bound for the weak regularity lemma for the case of L_2 graphons (see, e.g., [Lov12]) and is essentially optimal.

Part II

L_p regular random variables

CHAPTER 4

Hypergraph systems

We introduce the concept of a hypergraph system (see [Tao06c, DK16, DKK15, DKK18])

Definition 4.1. *A hypergraph system is a triple*

$$\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n]\rangle, \mathcal{H}) \tag{4.1}$$

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 for every $i \in [n]$, (X_i, Σ_i, μ_i) is η-nonatomic, then $\mathcal{H}$ will be called η-nonatomic.*

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

$$\mathcal{B}_e = \left\{\pi_e^{-1}(\mathbf{A}) : \mathbf{A} \in \boldsymbol{\Sigma}_e\right\} \tag{4.2}$$

where $\pi_e \colon \boldsymbol{X} \to \boldsymbol{X}_e$ is the natural projection. Observe that if $f \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$, then there exists a unique random variable $\boldsymbol{f} \in L_1(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ such that

$$f = \boldsymbol{f} \circ \pi_e \tag{4.3}$$

and note that the map $L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu}) \ni f \to \boldsymbol{f} \in L_1(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ is a linear isometry.

Now, when $|e| \geqslant 2$, let $\partial e = \{e' \subseteq e : |e'| = |e| - 1\}$ and define

$$\mathcal{S}_{\partial e} = \bigcap_{e' \in \partial e} \mathcal{B}_{e'} \subseteq \mathcal{B}_e. \tag{4.4}$$

Observe that for every $|e| \geqslant 2$, $\mathcal{S}_{\partial e}$ is a $|e|-1$ semiring. Hence, if $f \in L_1(\boldsymbol{X}_e, \Sigma_e, \mu_e)$ is a random variable its uniformity norm on the previous semiring is

$$\|f\|_{\mathcal{S}_{\partial_e}} = \sup\{\big|\int_A f \, d\boldsymbol{\mu}\big| \colon A \in \mathcal{S}_{\partial_e}\}. \tag{4.5}$$

From now on, we will refer to this norm as the cut norm of f. Also, observe that every $A \in \mathcal{S}_{\partial e}$ 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.

We present now a Sierpiński type result in the context of η-nonatomic hypergraph systems which will be very useful.

Proposition 4.2. *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 $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) \colon i \in [n] \rangle, \mathcal{H})$ 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 $a \leqslant \boldsymbol{\mu}(B) < a + 2\eta$.*

Before we proceed to the proof of the previous Proposition we need some preliminary work. To this end, recall that the 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 $\mu(C) = c$. This result may be extended on η-nonatomic probability spaces in the following way.

Lemma 4.3. *Let $\eta > 0$ and (X, Σ, μ) be an η-nonatomic probability space. Also let $B \in \Sigma$ with $\mu(B) > \eta$ and $\eta \leqslant c < \mu(B)$. Then, there exist $C \in \Sigma$ with $C \subseteq B$ and $c \leqslant \mu(C) < c + \eta$.*

Lemma 4.3 is straightforward for discrete probability spaces. The general case follows from the aforementioned result of Sierpiński and a transfinite exhaustion argument. More precicely,

Proof of Lemma 4.3. 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 a family $\langle Z_\alpha : \alpha < \omega_1 \rangle$ of measurable events of (X, Σ, μ) such that $\mu(Z_\alpha) < c$ 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 see that $\mu(Z_\alpha) < c$. Finally, let $\alpha = \beta + 1$ be a successor ordinal. By Sierpiński's result and hypothesis (H), the set $B \setminus Z_\alpha$ must contain a set $A \in \mathrm{Atoms}(X)$. We set $Z_{\alpha+1} = Z_\alpha \cup A$ and we observe that $\mu(Z_{\alpha+1} \setminus Z_\alpha) = \mu(A) > 0$. Also notice that $\mu(Z_{\alpha+1}) < c + \eta$. Thus, invoking hypothesis (H) once again, we conclude that $\mu(Z_{\alpha+1}) < c$ and the proof of Lemma 4.3 is completed. ☐

We are ready now to prove Proposition 4.2

Proof of Proposition 4.2. We argue as in the proof of Lemma 4.3. 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

$$\sum_{e' \in \partial e} \boldsymbol{\mu}(\boldsymbol{X} \setminus A_{e'}) \geqslant \boldsymbol{\mu}\Big(\boldsymbol{X} \setminus \bigcap_{e' \in \partial e} A_{e'}\Big) > 1 - a \geqslant r\eta. \tag{4.6}$$

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 $\boldsymbol{\mu}(A) \leqslant \eta^{r-1} \leqslant \eta$ for every atom A of $(\boldsymbol{X}, \mathcal{B}_{e'_1}, \boldsymbol{\mu})$. Hence, by Lemma 4.3 applied for "$A = \boldsymbol{X} \setminus A_{e'_1}$" 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 $A^1_{e'_1} = A_{e'_1} \cup B_{e'_1}$ and $A^1_{e'} = A_{e'}$ if $e' \in \partial e \setminus \{e'_1\}$. Notice that: (i) $\boldsymbol{\mu}(A^1_{e'_1}) \geqslant \boldsymbol{\mu}(A_{e'_1}) + \eta$, (ii) $\bigcap_{e' \in \partial e} A^1_{e'} \in \mathcal{S}_{\partial e}$, and (iii) $A \subseteq \bigcap_{e' \in \partial e} A^1_{e'}$. Moreover, $\boldsymbol{\mu}\big(\bigcap_{e' \in \partial e} A^1_{e'}\big) \leqslant \boldsymbol{\mu}(A) + 2\eta < a + 2\eta$ and so, by hypothesis (H), we obtain that $\boldsymbol{\mu}\big(\bigcap_{e' \in \partial e} A^1_{e'}\big) < a$. It follows, in particular, that the estimate in (4.6) is satisfied for the family $\langle A^1_{e'} : e' \in \partial e \rangle$.

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 $(A^m_{e'})_{m=0}^M$ in $\mathcal{B}_{e'}$ with $A^0_{e'} = A_{e'}$, such that for every $m \in [M]$ the following hold.

(C1) For every $e' \in \partial e$ we have $A^{m-1}_{e'} \subseteq A^m_{e'}$. Moreover, $\boldsymbol{\mu}(A^m_{e'_m}) \geqslant \boldsymbol{\mu}(A^{m-1}_{e'_m}) + \eta$.

(C2) We have $\boldsymbol{\mu}\big(\bigcap_{e' \in \partial e} A^m_{e'}\big) < a$.

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 $\boldsymbol{\mu}(A^\ell_g) \geqslant 2$ which is clearly a contradiction. □

CHAPTER 5

L_p regular random variables

5.1. The class of L_p regular random variables

We describe now a generalisation of L_p random variables in the context of hypergraph systems, the class of L_p **regular** random variables (see [BCCZ14, DKK18]). These random variables satisfy a Hölder-type inequality, a property which will play a crucial role in what follows.

Before we introduce the aforementioned family of random variables it is useful to recall one of the most well known pseudorandomness conditions for graphs, introduced in [Koh97, KR03]. 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 $|X|, |Y| \geqslant \gamma|V|$, 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., [BR09]).

Without further redue we proceed to the definition of L_p regular random variables.

Definition 5.1. *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*

$$\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} \leqslant C. \tag{5.1}$$

The main point in Definition 5.1 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 graphs, however, L_∞ regularity is a more subtle property. It is implied by the pseudorandomness conditions appearing in the work of Green and Tao [GT08, GT10], though closer to the spirit of this work is the work of Tao in [Tao06a].

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.

5.2. A Hölder-type inequality for L_p regular random variables

A useful inequality when studying L_p random variables is the Hölder inequality. The following proposition asserts that a similar inequality holds for L_p regular random variables.

Proposition 5.2 (Hölder-type inequality). *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 be the conjugate exponent of p. Finally, let $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) \colon i \in [n] \rangle, \mathcal{H})$ 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*

$$\Big(\int_A f \, d\boldsymbol{\mu} \Big)^q \leqslant C^q \big(\boldsymbol{\mu}(A) + (r+3)\eta \big). \tag{5.2}$$

(b) *On the other hand, if (5.2) 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 5.2 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 (this form is related to the semiring defined on (1.7) and is needed in the next chapter). Recall that for every probability space (X, Σ, μ) and every finite partition $\mathcal{P}$ of X with $\mathcal{P} \subseteq \Sigma$, $\iota(\mathcal{P}) = \min\{\mu(P) : P \in \mathcal{P}\}$. Then, we have the following lemma.

LEMMA 5.3. *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 $B_1,\dots,B_r\in\mathcal{S}$ 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 $C_{[r],1},\dots,C_{[r],r}$ satisfy the requirements of the lemma. So, assume that G is nonempty and let

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

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 $B_1,\dots,B_r$ 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 $B_1,\dots,B_r\in\mathcal{S}$ 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 5.3 is completed. □

We are ready now to prove Proposition 5.2.

PROOF OF PROPOSITION 5.2. (a) Fix $A \in \mathcal{S}_{\partial e}$. If $\eta \leqslant \boldsymbol{\mu}(A)$, then we claim that

$$\Big(\int_A f\, d\boldsymbol{\mu}\Big)^q \leqslant C^q\big(\boldsymbol{\mu}(A) + r\eta\big). \tag{5.3}$$

Indeed, by Lemma 5.3, 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 Proposition 4.2, there exists $B \in \mathcal{S}_{\partial e}$ with $A \subseteq B$ and $\eta \leqslant \boldsymbol{\mu}(B) < 3\eta$. Therefore,

$$\Big(\int_A f\, d\boldsymbol{\mu}\Big)^q \leqslant \Big(\int_B f\, d\boldsymbol{\mu}\Big)^q \overset{(5.3)}{\leqslant} C^q\big(\boldsymbol{\mu}(B) + r\eta\big) \leqslant C^q\big(\boldsymbol{\mu}(A) + (r+3)\eta\big) \tag{5.4}$$

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 (5.2) 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 $1 < p < \infty$,

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

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

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

as desired. $\square$

CHAPTER 6

Regularity lemma for L_p regular random variables

In this chapter we present a decomposition of L_p regular random variables which first appeared in [DKK18]. The proof proceeds via an "energy"-type increment argument and is close in the spirit of the proof of Theorem 2.2. More precicely, our interest is to prove the following result.

THEOREM 6.1 (Regularity Lemma). *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 $f_e = f^e_{\mathrm{str}} + f^e_{\mathrm{err}} + f^e_{\mathrm{unf}}$ with

$$f^e_{\mathrm{str}} = \mathbb{E}(f_e \,|\, \mathcal{A}_{\mathcal{P}_e}),\ \ f^e_{\mathrm{err}} = \mathbb{E}(f_e \,|\, \mathcal{A}_{\mathcal{Q}_e}) - \mathbb{E}(f_e \,|\, \mathcal{A}_{\mathcal{P}_e}),\ \ f^e_{\mathrm{unf}} = f_e - \mathbb{E}(f_e \,|\, \mathcal{A}_{\mathcal{Q}_e}), \tag{6.1}$$

we have the estimates

$$\|f^e_{\mathrm{str}}\|_{L_p} \leqslant C, \quad \|f^e_{\mathrm{err}}\|_{L_{p^\dagger}} \leqslant \sigma \quad \text{and} \quad \|f^e_{\mathrm{unf}}\|_{\mathcal{S}_{\partial e}} \leqslant \frac{1}{F(M)} \tag{6.2}$$

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

Note that, unless $p = \infty$, the structured part of the above decomposition (namely, the function f^e_{str}) is not uniformly bounded. This is an intrinsic feature of L_p regular hypergraphs, and is an important difference between Theorem 6.1 and several related results (see, e.g., [BR09] ,[COCF10], [CFZ15], [Gow10], [GT08], [Koh97], [RTTV08], [TZ08]). Observe, however, that, by part (b) and (6.2), one has a very good control on the correlation between f^e_{str} and $f^{e'}_{\mathrm{unf}}$ for every $e, e' \in \mathcal{H}$. Hence, by appropriately selecting the growth function F, we can force the function

f^e_{str} to behave like a bounded function for many practical purposes. The main part of the proof of Theorem 6.1 will be given in section 6. Before we proceed to it we will need some preparatory work.

A partition lemma. Let $n, r \in \mathbb{N}$ with $n \geqslant r \geqslant 2$, let $C > 0$ and $1 < p \leqslant \infty$. Let q denote the conjuggate exponent of p, i.e. $1/p + 1/q = 1$ 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.*

The following result is a refinement of Lemma 5.3. Recall that for every probability space (X, Σ, μ) and every partition $\mathcal{P}$ of X with $\mathcal{P} \subseteq \Sigma$ we write $\iota(\mathcal{P}) = \min\{\mu(P) \colon P \in \mathcal{P}\}$.

Lemma 6.2. *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*

$$\eta \leqslant \big(\vartheta \cdot \iota(\mathcal{P})\big)^q \tag{6.3}$$

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*

$$\int_{A \triangle B} \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\, d\boldsymbol{\mu} \leqslant Cr\vartheta \quad \textit{and} \quad \int_{A \triangle B} f\, d\boldsymbol{\mu} \leqslant 5Cr^2\vartheta. \tag{6.4}$$

Proof. We fix $A \in \mathcal{S}_{\partial e}$ and we set

$$\theta = \vartheta^q \cdot \iota(\mathcal{P})^{q-1}. \tag{6.5}$$

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 $\partial e = \{e'_1, \dots, e'_r\}$ 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 5.3 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 $B_1^P, \dots, B_r^P \in \mathcal{S}$ 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

$$\mathcal{Q} = \bigcup_{P \in \mathcal{P}} \mathcal{Q}_P \quad \text{and} \quad B = \bigcup_{P \in \mathcal{P}} B_P. \tag{6.6}$$

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

$$A \triangle B = \Big(\bigcup_{P \in \mathcal{P} \setminus \mathcal{P}^*} (A \cap P) \Big) \cup \Big(\bigcup_{P \in \mathcal{P}^*} \big(\bigcup_{i=1}^{r} B_i^P \big) \Big) \tag{6.7}$$

where for every $P \in \mathcal{P}^*$ the sets $B_1^P, \dots, B_r^P$ 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

$$\boldsymbol{\mu}(A \triangle B) \leqslant r\theta. \tag{6.8}$$

On the other hand, by (6.3), we have $\iota(\mathcal{P}) \geqslant \eta$, and so $\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} \leqslant C$ since f is (C, η, p)-regular. Hence, by Hölder's inequality, we obtain that

$$\int_{A \triangle B} \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}}) \, d\boldsymbol{\mu} \leqslant \|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} \cdot \boldsymbol{\mu}(A \triangle B)^{1/q} \overset{(6.8)}{\leqslant} C r^{1/q} \theta^{1/q} \overset{(6.5)}{\leqslant} C r \vartheta.$$

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

$$\int_{A \triangle B} f \, d\boldsymbol{\mu} = \sum_{P \in \mathcal{P} \setminus \mathcal{P}^*} \int_{A \cap P} f \, d\boldsymbol{\mu} + \sum_{P \in \mathcal{P}^*} \sum_{i=1}^{r} \int_{B_i^P} f \, d\boldsymbol{\mu}. \tag{6.9}$$

By (6.3) and (6.5), 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 5.2,

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

and so

$$\sum_{P \in \mathcal{P} \setminus \mathcal{P}^*} \int_{A \cap P} f \, d\boldsymbol{\mu} \leqslant 5 C r \theta^{1/q} \sum_{P \in \mathcal{P} \setminus \mathcal{P}^*} \boldsymbol{\mu}(P)^{1/q}. \tag{6.10}$$

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

$$\sum_{P \in \mathcal{P}^*} \sum_{i=1}^{r} \int_{B_i^P} f \, d\boldsymbol{\mu} \leqslant 5 C r^2 \theta^{1/q} \sum_{P \in \mathcal{P}^*} \boldsymbol{\mu}(P)^{1/q}. \tag{6.11}$$

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

$$\sum_{P \in \mathcal{P}} \boldsymbol{\mu}(P)^{1/q} \leqslant |\mathcal{P}|^{1/p} \leqslant \iota(\mathcal{P})^{-1/p}. \tag{6.12}$$

Combining (6.9)–(6.12) we conclude that

$$\int_{A\triangle B} f\,d\boldsymbol{\mu} \leqslant 5Cr^2\theta^{1/q}\sum_{P\in\mathcal{P}}\boldsymbol{\mu}(P)^{1/q} \leqslant 5Cr^2\theta^{1/q}\cdot\iota(\mathcal{P})^{-1/p} \overset{(6.5)}{=} 5Cr^2\vartheta$$

and the proof of Lemma 6.2 is completed. □

Proof of Theorem 6.1. We begin the proof of the Regularity Lemma with the following lemma. It asserts (roughly speaking) that if a given approximation of an L_p regular random variable is not sufficiently close to f in the cut norm, then we can find a much nicer approximation.

Lemma 6.3. *Let $0<\delta,\eta<1$ and set $\vartheta=\delta(12Cr^2)^{-1}$. 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*

$$\eta\leqslant\big(\vartheta\cdot\iota(\mathcal{P})\big)^q \tag{6.13}$$

and that $\mathscr{H}$ 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

$$\Big|\int_A\big(f-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\,d\boldsymbol{\mu}\Big|>\delta. \tag{6.14}$$

Next, we apply Lemma 6.2 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

$$\begin{aligned}\Big|\int_A\big(f-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\,d\boldsymbol{\mu}-\int_B\big(f-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\,d\boldsymbol{\mu}\Big|&\leqslant\\ \leqslant\int_{A\triangle B}f\,d\boldsymbol{\mu}+\int_{A\triangle B}\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\,d\boldsymbol{\mu}&\leqslant 5Cr^2\vartheta+Cr\vartheta\leqslant 6Cr^2\vartheta=\delta/2\end{aligned}$$

and so, by (6.14),

$$\Big|\int_B\big(f-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\,d\boldsymbol{\mu}\Big|\geqslant\delta/2. \tag{6.15}$$

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

$$\int_B\big(f-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\,d\boldsymbol{\mu}=\int_B\big(\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\,d\boldsymbol{\mu}. \tag{6.16}$$

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

$$\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_{p^\dagger}} \geqslant \|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_1} \geqslant \\ \Big| \int_B \big(\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\big)\, d\boldsymbol{\mu}\Big| \overset{(6.16)}{=} \Big| \int_B \big(f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\big)\, d\boldsymbol{\mu}\Big| \overset{(6.15)}{\geqslant} \delta/2$$

as desired. □

The previous lemma has the following dichotomy as a consequence.

Lemma 6.4. *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*

$$\eta \leqslant \big(\vartheta^N \cdot \iota(\mathcal{P})\big)^{q^N} \tag{6.17}$$

and that $\mathscr{H}$ 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 proof follows similar steps with the proof of Lemma 2.5.

Proof. 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 6.3, we see that

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

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 $\iota(\mathcal{P}_i) \geqslant \big(\vartheta \cdot \iota(\mathcal{P}_{i-1})\big)^q$, and (P2) $\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_i}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_{i-1}})\|_{L_{p^\dagger}} > \delta/2$.

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

Therefore, by the choice of N and Proposition 2.1,

$$\sigma \leqslant (p^\dagger - 1)^{1/2} \frac{\delta}{2} N^{1/2} \;<\; (p^\dagger - 1)^{1/2} \Big(\sum_{i=0}^{N} \|d_i\|^2_{L_{p^\dagger}} \Big)^{1/2}$$
$$\leqslant \; \Big\| \sum_{i=0}^{N} d_i \Big\|_{L_{p^\dagger}} = \|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_N}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_{p^\dagger}}.$$

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 $\iota(\mathcal{Q}) \geqslant \big(\vartheta^N \cdot \iota(\mathcal{P})\big)^{q^N}$ and so, by (H1), $\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_N}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_{p^\dagger}} \leqslant \sigma$ 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 $N_0 = 0$, $\eta_0 = 1$, $\theta_0 = \big(12Cr^2F(1)\big)^{-1}$ and

$$\begin{cases} N_{m+1} = \lceil 4(p^\dagger - 1)^{-1} \sigma^2 F(\lceil \eta_m^{-1} \rceil)^2 \rceil, \\ \eta_{m+1} = (\vartheta_m^{N_{m+1}} \cdot \eta_m)^{q^{N_{m+1}}}, \\ \vartheta_{m+1} = \big(12Cr^2 F(\lceil \eta_{m+1}^{-1} \rceil)\big)^{-1}. \end{cases} \tag{6.18}$$

The following lemma is the last step of the proof of Theorem 6.1 and is similar to Lemma 2.6.

Lemma 6.5. *Let $0 < \sigma \leqslant 1$ and $F\colon \mathbb{N} \to \mathbb{R}$ a growth function. Set*

$$L = \lceil C^2 (p^\dagger - 1)^{-1} \sigma^{-2} n^r \rceil \tag{6.19}$$

and let (η_m) be as in (6.18). 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 $\|\mathbb{E}(f_e \,|\, \mathcal{A}_{\mathcal{Q}_e}) - \mathbb{E}(f_e \,|\, \mathcal{A}_{\mathcal{P}_e})\|_{L_{p^\dagger}} \leqslant \sigma$ and $\|f_e - \mathbb{E}(f_e \,|\, \mathcal{A}_{\mathcal{Q}_e})\|_{\mathcal{S}_{\partial e}} \leqslant 1/F(\lceil \eta_m^{-1} \rceil)$.*

Proof. It is similar to the proof of Lemma 6.4 and so we will briefly sketch the argument. If the lemma is false, then using Lemma 6.4 we select for every $e \in \mathcal{H}$ partitions $\mathcal{P}^e_0, \dots, \mathcal{P}^e_L$ of $\boldsymbol{X}$ with $\mathcal{P}^e_0 = \{\boldsymbol{X}\}$ as well as $e_1, \dots, e_L \in \mathcal{H}$ such that for every $m \in [L]$ we have: (P1) $\mathcal{P}^e_m$ is a refinement of $\mathcal{P}^e_{m-1}$ with $\mathcal{P}^e_m \subseteq \mathcal{S}_{\partial e}$ and $\iota(\mathcal{P}^e_m) \geqslant \eta_m$ for every $e \in \mathcal{H}$, and (P2) $\|\mathbb{E}(f_{e_m} \,|\, \mathcal{A}_{\mathcal{P}^{e_m}_m}) - \mathbb{E}(f_{e_m} \,|\, \mathcal{A}_{\mathcal{P}^{e_m}_{m-1}})\|_{L_{p^\dagger}} > \sigma$. 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 $\mathbb{E}(f_{\boldsymbol{e}} \,|\, \mathcal{A}_{\mathcal{P}^{\boldsymbol{e}}_0}), \dots, \mathbb{E}(f_{\boldsymbol{e}} \,|\, \mathcal{A}_{\mathcal{P}^{\boldsymbol{e}}_L})$ 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 $\iota(\mathcal{P}^{\boldsymbol{e}}_m) \geqslant \eta_m \geqslant \eta_L \geqslant \eta$

we see that $\|\mathbb{E}(f_{\boldsymbol{e}}\,|\,\mathcal{A}_{\mathcal{P}^{\boldsymbol{e}}_m})\|_{L_{p^\dagger}} \leqslant \|\mathbb{E}(f_{\boldsymbol{e}}\,|\,\mathcal{A}_{\mathcal{P}^{\boldsymbol{e}}_m})\|_{L_p} \leqslant C$ for every $m \in [L]$. Hence, by the choice of L in (6.19) and Proposition 2.1, we conclude that

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

which is clearly a contradiction. □

We are ready to complete the proof of Theorem 6.1.

Proof of Theorem 6.1. 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 (6.19) and (6.18) respectively. We set $\mathrm{Reg} = \lceil \eta_L^{-1} \rceil$ and we claim that with this choice the result follows. Indeed, set $\eta \coloneqq 1/\mathrm{Reg} \leqslant \eta_L$ 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 $m \in \{0,\dots,L-1\}$, $\langle \mathcal{P}_e : e \in \mathcal{H}\rangle$ and $\langle \mathcal{Q}_e : e \in \mathcal{H}\rangle$ be as in Lemma 6.5 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. □

Part III

Pseudorandomness

CHAPTER 7

Box norms

We begin by introducing some pieces of notation. Let $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i)\colon i \in [n]\rangle, \mathcal{H})$ be a hypergraph system and $e \subseteq [n]$ be nonempty. Then, recall that by $\pi_e\colon \boldsymbol{X} \to \boldsymbol{X}_e$ we denote the natural projection. If the set $[n] \setminus e$ is nonempty, then for every $\mathbf{x}_e \in \boldsymbol{X}_e$ and every $\mathbf{x}_{[n]\setminus e} \in \boldsymbol{X}_{[n]\setminus e}$ we denote the unique element $\mathbf{x}$ of $\boldsymbol{X}$ such that $\mathbf{x}_e = \pi_e(\mathbf{x})$ and $\mathbf{x}_{[n]\setminus e} = \pi_{[n]\setminus e}(\mathbf{x})$. Moreover, 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, $f_{\mathbf{x}_e}(\mathbf{x}_{[n]\setminus e}) = f(\mathbf{x}_e, \mathbf{x}_{[n]\setminus e})$. Finally, let $\ell \in \mathbb{N}$ with $\ell \geqslant 2$. For every $\mathbf{x}_e^{(0)} = (x_i^{(0)})_{i\in e}, \dots, \mathbf{x}_e^{(\ell-1)} = (x_i^{(\ell-1)})_{i\in e}$ in $\boldsymbol{X}_e$ and every $\omega = (\omega_i)_{i\in e} \in \{0, \dots, \ell-1\}^e$ we set

$$\mathbf{x}_e^{(\omega)} = (x_i^{(\omega_i)})_{i\in e} \in \boldsymbol{X}_e. \tag{7.1}$$

Notice that if $\omega = m^e$ for some $m \in \{0, \dots, \ell-1\}$ (that is, $\omega = (\omega_i)_{i\in e}$ with $\omega_i = m$ for every $i \in e$), then $\mathbf{x}_e^{(\omega)} = \mathbf{x}_e^{(m)}$.

Recall now, that the *box norm* of a random variable $f\colon \boldsymbol{X}_e \to \mathbb{R}$ is the quantity

$$\|f\|_{\square^e} \coloneqq \mathbb{E}\Big[\prod_{\omega\in\{0,1\}^e} f(\mathbf{x}_e^{(\omega)}) \,\Big|\, \mathbf{x}_e^{(0)}, \mathbf{x}_e^{(1)} \in \boldsymbol{X}_e\Big]^{1/2^{|e|}}. \tag{7.2}$$

These norms were introduced by Gowers [Gow01], [Gow07] and are a fundamental tool in additive and extremal combinatorics.

7.1. ℓ-Box norms

Throughout this section let $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n]\rangle, \mathcal{H})$ denote a hypergraph system. The variant of the box norm that interests us in this work is the following one, which first appeared[1] in [Hat09]. Let $\ell \geqslant 2$ be an even integer and $e \in \mathcal{H}$. Then the *ℓ-box norm* of a random variable $f\colon \boldsymbol{X}_e \to \mathbb{R}$ is defined by

$$\|f\|_{\square_\ell^e} \coloneqq \mathbb{E}\Big[\prod_{\omega\in\{0,\dots,\ell-1\}^e} f(\mathbf{x}_e^{(\omega)}) \,\Big|\, \mathbf{x}_e^{(0)}, \dots, \mathbf{x}_e^{(\ell-1)} \in \boldsymbol{X}_e\Big]^{1/\ell^{|e|}}. \tag{7.3}$$

Observe that when $\ell = 2$ then the previous norm coincides with the classic box norm of (7.2).

[1] Actually, the framework in [Hat09] is more general and includes several other variants of (7.2).

7.1.1. Basic properties. Let $e \subseteq [n]$ be nonempty and let $\ell \geqslant 2$ be an even integer. Also let $f \in L_1(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$. We first observe that the ℓ-box norm of f can be recursively defined as follows. If $|e| = 1$, then by (7.3) we have

$$\|f\|_{\square_\ell^e} = \mathbb{E}\Big[\prod_{\omega=0}^{\ell-1} f(x_j^{(\omega)}) \,\Big|\, x_j^{(0)}, \dots, x_j^{(\ell-1)} \in X_j\Big]^{1/\ell^{|e|}} = \big(\mathbb{E}[f]^\ell\big)^{1/\ell} = |\mathbb{E}[f]|. \tag{7.4}$$

On the other hand, if $|e| \geqslant 2$, then for every $j \in e$ we have

$$\|f\|_{\square_\ell^e} = \mathbb{E}\Big[\|\prod_{\omega=0}^{\ell-1} f(\,\cdot\,, x_j^{(\omega)})\|_{\square_\ell^{e\setminus\{j\}}}^{\ell^{|e|-1}} \,\Big|\, x_j^{(0)}, \dots, x_j^{(\ell-1)} \in X_j\Big]^{1/\ell^{|e|}}. \tag{7.5}$$

In the following proposition we gather further properties of the ℓ-box norms.

Proposition 7.1. *Let $e \subseteq [n]$ be nonempty and let $\ell \geqslant 2$ be an even integer.*

(a) (Gowers–Cauchy–Schwarz inequality) *For every $\omega \in \{0, \dots, \ell-1\}^e$ let $f_\omega \in L_1(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$. Then we have*

$$\Big|\mathbb{E}\Big[\prod_{\omega\in\{0,\dots,\ell-1\}^e} f_\omega(\mathbf{x}_e^{(\omega)}) \,\Big|\, \mathbf{x}_e^{(0)}, \dots, \mathbf{x}_e^{(\ell-1)} \in \boldsymbol{X}_e\Big]\Big| \leqslant \prod_{\omega\in\{0,\dots,\ell-1\}^e} \|f_\omega\|_{\square_\ell^e}. \tag{7.6}$$

(b) *Let $f \in L_1(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$. Then we have $|\mathbb{E}[f]| \leqslant \|f\|_{\square_\ell^e}$. Moreover, if $\ell_1 \leqslant \ell_2$ are even positive integers, then $\|f\|_{\square_{\ell_1}^e} \leqslant \|f\|_{\square_{\ell_2}^e}$.*

(c) *If $|e| \geqslant 2$, then $\|\cdot\|_{\square_\ell^e}$ is a norm on the vector subspace of $L_1(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ consisting of all $f \in L_1(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ with $\|f\|_{\square_\ell^e} < \infty$.*

(d) *Let $1 < p \leqslant \infty$ and let q denote the conjugate exponent of p. Assume that $\ell \geqslant q$ and that $e = \{i, j\}$ is a doubleton. Then for every $f \in L_1(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$, every $u \in L_p(X_i, \Sigma_i, \mu_i)$ and every $v \in L_p(X_j, \Sigma_j, \mu_j)$ we have*

$$|\mathbb{E}[f(x_i, x_j)u(x_i)v(x_j) \,|\, x_i \in X_i, x_j \in X_j]| \leqslant \|f\|_{\square_\ell^e} \|u\|_{L_p} \|v\|_{L_p}. \tag{7.7}$$

Proof. (a) We follow the proof from [GT10, Lemma B.2] which proceeds by induction on the cardinality of e. The case "$|e| = 1$" is straightforward, and so let $r \geqslant 2$ and assume that the result has been proved for every $e' \subseteq [n]$ with $1 \leqslant |e'| \leqslant r-1$. Let $e \subseteq [n]$ with $|e| = r$ be arbitrary. Fix $j \in e$, set $e' = e \setminus \{j\}$ and for every $\omega \in \{0, \dots, \ell-1\}^e$ let $f_\omega \in L_1(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$. Moreover, for every $\omega_j \in \{0, \dots, \ell-1\}$ we define $G_{\omega_j} \colon \boldsymbol{X}_{e'}^\ell \to \mathbb{R}$ by

$$G_{\omega_j}(\mathbf{x}_{e'}^{(0)}, \dots, \mathbf{x}_{e'}^{(\ell-1)}) = \mathbb{E}\Big[\prod_{\omega_{e'}\in\{0,\dots,\ell-1\}^{e'}} f_{(\omega_{e'},\omega_j)}(\mathbf{x}_{e'}^{(\omega_{e'})}, x_j) \,\Big|\, x_j \in X_j\Big] \tag{7.8}$$

where $(\omega_{e'},\omega_j)$ is the unique element ω of $\{0,\dots,\ell-1\}^e$ such that $\omega(j)=\omega_j$ and $\omega(i)=\omega_{e'}(i)$ for every $i\in e'$. Observe that

$$\Big|\mathbb{E}\Big[\prod_{\omega\in\{0,\dots,\ell-1\}^e} f_\omega(\mathbf{x}_e^{(\omega)})\,\Big|\,\mathbf{x}_e^{(0)},\dots,\mathbf{x}_e^{(\ell-1)}\in \boldsymbol{X}_e\Big]\Big| = \Big|\mathbb{E}\Big[\prod_{\omega_j=0}^{\ell-1} G_{\omega_j}\Big]\Big|$$

and, by Hölder's inequality, $|\mathbb{E}\big[\prod_{\omega_j=0}^{\ell-1} G_{\omega_j}\big]|\leqslant \prod_{\omega_j=0}^{\ell-1}\mathbb{E}[G_{\omega_j}^\ell]^{1/\ell}$. Therefore, it is enough to show that for every $\omega_j\in\{0,\dots,\ell-1\}$ we have

$$\mathbb{E}[G_{\omega_j}^\ell]\leqslant \prod_{\omega_{e'}\in\{0,\dots,\ell-1\}^{e'}} \|f_{(\omega_{e'},\omega_j)}\|_{\square_\ell^e}^\ell. \tag{7.9}$$

Indeed, fix $\omega_j\in\{0,\dots,\ell-1\}$ and notice that, by (7.8),

$$G_{\omega_j}^\ell(\mathbf{x}_{e'}^{(0)},\dots,\mathbf{x}_{e'}^{(\ell-1)}) = \mathbb{E}\Big[\prod_{\omega_{e'}\in\{0,\dots,\ell-1\}^{e'}}\prod_{\omega=0}^{\ell-1} f_{(\omega_{e'},\omega_j)}(\mathbf{x}_{e'}^{(\omega_{e'})},x_j^{(\omega)})\Big] \tag{7.10}$$

where the expectation is over all $x_j^{(0)},\dots,x_j^{(\ell-1)}\in X_j$. By (7.10) and Fubini's theorem, we see that

$$\mathbb{E}[G_{\omega_j}^\ell] = \mathbb{E}\Big[\mathbb{E}\big[\prod_{\omega_{e'}\in\{0,\dots,\ell-1\}^{e'}}\prod_{\omega=0}^{\ell-1} f_{(\omega_{e'},\omega_j)}(\mathbf{x}_{e'}^{(\omega_{e'})},x_j^{(\omega)})\,\big|\,\mathbf{x}_{e'}^{(0)},\dots,\mathbf{x}_{e'}^{(\ell-1)}\in\boldsymbol{X}_{e'}\big]\Big]$$

where the outer expectation is over all $x_j^{(0)},\dots,x_j^{(\ell-1)}\in X_j$. Thus, applying the induction hypothesis and Hölder's inequality, we obtain that

$$\begin{aligned}\mathbb{E}\big[G_{\omega_j}^\ell\big] &\leqslant \mathbb{E}\Big[\prod_{\omega_{e'}\in\{0,\dots,\ell-1\}^{e'}} \big\|\prod_{\omega=0}^{\ell-1} f_{(\omega_{e'},\omega_j)}(\,\cdot\,,x_j^{(\omega)})\big\|_{\square_\ell^{e'}}\Big] \\ &\leqslant \prod_{\omega_{e'}\in\{0,\dots,\ell-1\}^{e'}} \mathbb{E}\Big[\big\|\prod_{\omega=0}^{\ell-1} f_{(\omega_{e'},\omega_j)}(\,\cdot\,,x_j^{(\omega)})\big\|_{\square_\ell^{e'}}^{\ell^{|e'|}}\Big]^{1/\ell^{|e'|}}.\end{aligned} \tag{7.11}$$

By (7.5) and (7.11), we conclude that (7.9) is satisfied.

(b) It is a consequence of the Gowers–Cauchy–Schwarz inequality. Specifically, for every $\omega\in\{0,\dots,\ell-1\}^e$ let $f_\omega=f$ if $\omega=\{0\}^e$ and $f_\omega=1$ otherwise. By (7.6), we see that $|\mathbb{E}[f]|\leqslant\|f\|_{\square_\ell^e}$. Next, let $\ell_1\leqslant\ell_2$ be even positive integers. As before, for every $\omega\in\{0,\dots,\ell_2-1\}^e$ let $f_\omega=f$ if $\omega\in\{0,\dots,\ell_1-1\}^e$; otherwise, let $f_\omega=1$.

Then we have

$$\begin{aligned}\|f\|_{\square^e_{\ell_1}}^{\ell_1^{|e|}} &= \mathbb{E}\Big[\prod_{\omega\in\{0,\dots,\ell_1-1\}^e} f(\mathbf{x}_e^{(\omega)})\,\Big|\,\mathbf{x}_e^{(0)},\dots,\mathbf{x}_e^{(\ell_1-1)}\in \boldsymbol{X}_e\Big] \\ &= \mathbb{E}\Big[\prod_{\omega\in\{0,\dots,\ell_2-1\}^e} f_\omega(\mathbf{x}_e^{(\omega)})\,\Big|\,\mathbf{x}_e^{(0)},\dots,\mathbf{x}_e^{(\ell_2-1)}\in \boldsymbol{X}_e\Big] \overset{(??)}{\leqslant} \|f\|_{\square^e_{\ell_2}}^{\ell_1^{|e|}}\end{aligned}$$

which implies that $\|f\|_{\square^e_{\ell_1}} \leqslant \|f\|_{\square^e_{\ell_2}}$.

(c) Absolute homogeneity is straightforward. The triangle inequality

$$\|f+g\|_{\square^e_\ell} \leqslant \|f\|_{\square^e_\ell} + \|g\|_{\square^e_\ell}$$

follows by raising both sides to the power $\ell^{|e|}$ and then applying (7.6). Finally, let $f\in L_1(\boldsymbol{X}_e,\boldsymbol{\Sigma}_e,\boldsymbol{\mu}_e)$ with $\|f\|_{\square^e_\ell}=0$ and observe that it suffices to show that $f=0$ $\boldsymbol{\mu}_e$-almost everywhere. First we note that using (7.6) and arguing precisely as in [GT10, Corollary B.3] we have that $\mathbb{E}[f\cdot \mathbf{1}_R]=0$ for every measurable rectangle R of $\boldsymbol{X}_e$ (that is, every set R of the form $\prod_{i\in e} A_i$ where $A_i\in\Sigma_i$ for every $i\in e$). We claim that this implies that $\mathbb{E}[f\cdot\mathbf{1}_A]=0$ for every $A\in\boldsymbol{\Sigma}_e$; this is enough to complete the proof. Indeed, fix $A\in\boldsymbol{\Sigma}_e$ and let $\varepsilon>0$ be arbitrary. Since f is integrable, there exists $\delta>0$ such that $\mathbb{E}[\,|f|\cdot\mathbf{1}_C]<\varepsilon$ for every $C\in\boldsymbol{\Sigma}_e$ with $\boldsymbol{\mu}_e(C)<\delta$. Moreover, by Caratheodory's extension theorem, there exists a finite family $R_1,\dots,R_m$ of pairwise disjoint measurable rectangles of $\boldsymbol{X}_e$ such that, setting $B=\bigcup_{k=1}^m R_k$, we have $\boldsymbol{\mu}_e(A\triangle B)<\delta$ (see, e.g., [Bil08, Theorem 11.4]). Hence, $\mathbb{E}[f\cdot\mathbf{1}_B]=0$ and so

$$|\mathbb{E}[f\cdot\mathbf{1}_A]| = |\mathbb{E}[f\cdot\mathbf{1}_A]-\mathbb{E}[f\cdot\mathbf{1}_B]| \leqslant \mathbb{E}[\,|f|\cdot\mathbf{1}_{A\triangle B}]<\varepsilon.$$

Since ε was arbitrary, we conclude that $\mathbb{E}[f\cdot\mathbf{1}_A]=0$.

(d) Set $I=\mathbb{E}[f(x_i,x_j)u(x_i)v(x_j)\,|\,x_i\in X_i, x_j\in X_j]$ and let ℓ' denote the conjugate exponent of ℓ. Notice that $1<\ell'\leqslant p$. By Hölder's inequality, we have

$$\begin{aligned}|I| = {} & \big|\mathbb{E}\big[\mathbb{E}[f(x_i,x_j)v(x_j)\,|\,x_j\in X_j]u(x_i)\,\big|\,x_i\in X_i\big]\big| \\ \leqslant {} & \mathbb{E}\big[\mathbb{E}[f(x_i,x_j)v(x_j)\,|\,x_j\in X_j]^\ell\,\big|\,x_i\in X_i\big]^{1/\ell}\cdot\|u\|_{L_{\ell'}} \leqslant I_1^{1/\ell}\cdot\|u\|_{L_p}\end{aligned} \tag{7.12}$$

where $I_1 = \mathbb{E}\big[\prod_{\omega=0}^{\ell-1} f(x_i, x_j^{(\omega)}) v(x_j^{(\omega)}) \,\big|\, x_i \in X_i, x_j^{(0)}, \dots, x_j^{(\ell-1)} \in X_j\big]$. Moreover,

$$\begin{aligned} I_1 = &\quad \mathbb{E}\Big[\mathbb{E}\big[\prod_{\omega=0}^{\ell-1} f(x_i, x_j^{(\omega)}) \,\big|\, x_i \in X_i\big] \cdot \prod_{\omega=0}^{\ell-1} v(x_j^{(\omega)}) \,\Big|\, x_j^{(0)}, \dots, x_j^{(\ell-1)} \in X_j\Big] \\ \leqslant &\quad \mathbb{E}\Big[\mathbb{E}\big[\prod_{\omega=0}^{\ell-1} f(x_i, x_j^{(\omega)}) \,\big|\, x_i \in X_i\big]^{\ell} \,\Big|\, x_j^{(0)}, \dots, x_j^{(\ell-1)} \in X_j\Big]^{1/\ell} \cdot \|v\|_{L_{\ell'}}^{\ell} \\ \overset{(7.5)}{=} &\quad \|f\|_{\square_\ell^e}^{\ell} \cdot \|v\|_{L_{\ell'}}^{\ell} \leqslant \|f\|_{\square_\ell^e}^{\ell} \cdot \|v\|_{L_p}^{\ell}. \end{aligned}$$

By (7.12) and the previous expression the result follows. □

7.1.2. The (ℓ, p)-box norms. We will need the following L_p versions of the ℓ-box norms. We remark that closely related norms appear in [Can]. Recall that by $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n] \rangle, \mathcal{H})$ we denote a hypergraph system.

DEFINITION 7.2. *Let $e \subseteq [n]$ be nonempty and let $\ell \geqslant 2$ be an even integer. Also let $1 \leqslant p < \infty$ and $f \in L_p(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$. The (ℓ, p)-*box norm *of f is defined by*

$$\|f\|_{\square_{\ell,p}^e} \coloneqq \big\| |f|^p \big\|_{\square_\ell^e}^{1/p}. \tag{7.13}$$

*Moreover, for every $f \in L_\infty(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ we define the (ℓ, ∞)-*box norm *of f by*

$$\|f\|_{\square_{\ell,\infty}^e} \coloneqq \|f\|_{L_\infty}. \tag{7.14}$$

We have the following analogue of Proposition 7.1.

PROPOSITION 7.3. *Let $e \subseteq [n]$ be nonempty and let $\ell \geqslant 2$ be an even integer.*

(a) *Let $1 \leqslant p < \infty$. If $f_\omega \in L_p(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ for every $\omega \in \{0, \dots, \ell-1\}^e$, then*

$$\mathbb{E}\Big[\prod_{\omega \in \{0,\dots,\ell-1\}^e} |f_\omega|^p(\mathbf{x}_e^{(\omega)}) \,\Big|\, \mathbf{x}_e^{(0)}, \dots, \mathbf{x}_e^{(\ell-1)} \in \boldsymbol{X}_e\Big] \leqslant \prod_{\omega \in \{0,\dots,\ell-1\}^e} \|f_\omega\|_{\square_{\ell,p}^e}^p. \tag{7.15}$$

(b) *Let $1 < p, q < \infty$ be conjugate exponents, that is, $1/p + 1/q = 1$. Then for every $f \in L_p(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ and every $g \in L_q(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ we have*

$$\|fg\|_{\square_\ell^e} \leqslant \|f\|_{\square_{\ell,p}^e} \cdot \|g\|_{\square_{\ell,q}^e}. \tag{7.16}$$

(c) *Assume that $|e| \geqslant 2$ and let $1 \leqslant p < \infty$. Then $\|\cdot\|_{\square_{\ell,p}^e}$ is a norm on the vector subspace of $L_p(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ consisting of all $f \in L_p(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ with $\|f\|_{\square_{\ell,p}^e} < \infty$. Moreover, the following hold.*
   - (i) *For every $f \in L_p(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ we have $\|f\|_{L_p} \leqslant \|f\|_{\square_{\ell,p}^e}$.*
   - (ii) *For every $1 \leqslant p_1 \leqslant p_2 < \infty$ and every $f \in L_{p_2}(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ we have $\|f\|_{\square_{\ell,p_1}^e} \leqslant \|f\|_{\square_{\ell,p_2}^e}$.*
   - (iii) *For every $f \in L_\infty(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ we have $\lim_{p\to\infty} \|f\|_{\square_{\ell,p}^e} = \|f\|_{\square_{\ell,\infty}^e}$.*

PROOF. Part (a) follows immediately by (7.6). For part (b) fix a pair $1 < p, q < \infty$ of conjugate exponents, and let $f \in L_p(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ and $g \in L_q(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$ be arbitrary. We define $F, G\colon \boldsymbol{X}_e^\ell \to \mathbb{R}$ by $F(\mathbf{x}_e^{(0)}, \dots, \mathbf{x}_e^{(\ell-1)}) = \prod_{\omega \in \{0,\dots,\ell-1\}^e} f(\mathbf{x}_e^{(\omega)})$ and $G(\mathbf{x}_e^{(0)}, \dots, \mathbf{x}_e^{(\ell-1)}) = \prod_{\omega \in \{0,\dots,\ell-1\}^e} g(\mathbf{x}_e^{(\omega)})$. By Hölder's inequality, we have

$$\|fg\|_{\square_\ell^e}^{\ell^{|e|}} \leqslant \mathbb{E}[\,|F \cdot G|\,] \leqslant \mathbb{E}[\,|F|^p]^{1/p} \cdot \mathbb{E}[\,|G|^q]^{1/q}.$$

Noticing that $\mathbb{E}[\,|F|^p]^{1/p} = \|f\|_{\square_{\ell,p}^e}$ and $\mathbb{E}[\,|G|^q]^{1/q} = \|g\|_{\square_{\ell,q}^e}$ we conclude that (7.16) is satisfied.

We proceed to show part (c). Arguing as in the proof of the classical Minkowski's inequality we see that the (ℓ, p)-box norm satisfies the triangle inequality. Absolute homogeneity is clear and so, by Proposition 7.1, we conclude that $\|\cdot\|_{\square_{\ell,p}^e}$ is indeed a norm. Next, observe that part (c.i) follows by (7.15) applied for $f_\omega = f$ if $\omega = \{0\}^e$ and $f_\omega = 1$ otherwise. For part (c.ii) set $p = p_2/p_1$ and notice that

$$\|f\|_{\square_{\ell,p_1}^e}^{p_1} = \big\||f|^{p_1}\big\|_{\square_\ell^e} \overset{(7.16)}{\leqslant} \big\||f|^{p_1}\big\|_{\square_{\ell,p}^e} = \|f\|_{\square_{\ell,p_2}^e}^{p_2/p_1}.$$

Finally, let $f \in L_\infty(\boldsymbol{X}_e, \boldsymbol{\Sigma}_e, \boldsymbol{\mu}_e)$. By part (c.i), we have $\|f\|_{L_p} \leqslant \|f\|_{\square_{\ell,p}^e} \leqslant \|f\|_{L_\infty}$. Since $\lim_{p\to\infty} \|f\|_{L_p} = \|f\|_{L_\infty}$, we obtain that $\lim_{p\to\infty} \|f\|_{\square_{\ell,p}^e} = \|f\|_{L_\infty} = \|f\|_{\square_{\ell,\infty}^e}$ and the proof is completed. $\square$

7.2. A counting lemma for L_p graphons

Let n be a positive integer and let $\mathcal{G}$ be a nonempty graph on $[n]$. Recall that the *maximum degree* of $\mathcal{G}$ is the number $\Delta(\mathcal{G}) \coloneqq \max\big\{|\{e \in \mathcal{G} : i \in e\}| : i \in [n]\big\}$. Given two graphons W and U, a natural problem (which is of particular importance in the context of graph limits – see [Lov12]) is to estimate the quantity

$$\Big|\mathbb{E}\Big[\prod_{\{i,j\}\in\mathcal{G}} W(x_i, x_j)\,\Big|\, x_1, \dots, x_n \in X\Big] - \mathbb{E}\Big[\prod_{\{i,j\}\in\mathcal{G}} U(x_i, x_j)\,\Big|\, x_1, \dots, x_n \in X\Big]\Big|.$$

If W and U are uniformly bounded, then this problem has a very satisfactory answer (see, e.g., [Lov12]). The unbounded case, however, is quite involved. Recently, there was progress in this direction in [BCCZ14, Theorem 2.20] where effective estimates were obtained provided that W and U are L_p graphons for some $p > \Delta(\mathcal{G})$. It is important to note that this integrability restriction is necessary at this level of generality. Indeed, if $p < \Delta(\mathcal{G})$, then the above difference may not even be defined.

Nevertheless, we have the following theorem which has the advantage of being applicable to L_p graphons for any $p > 1$ but requires a rather different type of integrability assumption.

Theorem 7.4. *Let Δ be a positive integer, $C \geqslant 1$ and $1 < p \leqslant \infty$. We set $\ell = 2$ if either $p = \infty$ or $\Delta = 1$; otherwise, let*

$$\ell = \min\big\{2n : n \in \mathbb{N} \text{ and } 2n \geqslant p^{(\Delta-1)^{-1}}(p^{(\Delta-1)^{-1}} - 1)^{-1}\big\}. \tag{7.17}$$

Also let $\mathscr{G} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n]\rangle, \mathcal{G})$ be a 2-uniform hypegraph system with $\Delta(\mathcal{G}) = \Delta$. For every $e \in \mathcal{G}$, let $f_e, g_e \in L_p(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ such that

$$\|\mathbf{f}_e\|_{\square^e_{\ell,p}} \leqslant 1 \quad \text{and} \quad \|\mathbf{g}_e\|_{\square^e_{\ell,p}} \leqslant 1 \tag{7.18}$$

where $\mathbf{f}_e$ and $\mathbf{g}_e$ are as in (4.3) for f_e and g_e respectively. Assume that for every $\mathcal{G}_1, \mathcal{G}_2 \subseteq \mathcal{G}$ with $\mathcal{G}_1 \cap \mathcal{G}_2 = \emptyset$ we have

$$\Big\| \prod_{e\in\mathcal{G}_1} f_e \prod_{e\in\mathcal{G}_2} g_e \Big\|_{L_p} \leqslant C. \tag{7.19}$$

(*Here, we follow the convention that the product of an empty family of functions is equal to the constant function 1.*) *Then we have*

$$\Big|\mathbb{E}\Big[\prod_{e\in\mathcal{G}} f_e\Big] - \mathbb{E}\Big[\prod_{e\in\mathcal{G}} g_e\Big]\Big| \leqslant C \cdot \sum_{e\in\mathcal{G}} \|\mathbf{f}_e - \mathbf{g}_e\|_{\square^e_\ell}. \tag{7.20}$$

Proof. Set $M = |\mathcal{G}|$ and write $\mathcal{G} = \{e_1, \dots, e_M\}$. Since

$$\mathbb{E}\Big[\prod_{e\in\mathcal{G}} f_e\Big] - \mathbb{E}\Big[\prod_{e\in\mathcal{G}} g_e\Big] = \sum_{k=1}^{m} \mathbb{E}\Big[\prod_{s<k} g_{e_s}(f_{e_k} - g_{e_k}) \prod_{s>k} f_{e_s}\Big]$$

it suffices to show that for every $k \in [M]$ we have

$$\Big|\mathbb{E}\Big[\prod_{s<k} g_{e_s}(f_{e_k} - g_{e_k}) \prod_{s>k} f_{e_s}\Big]\Big| \leqslant C \cdot \|\mathbf{f}_{e_k} - \mathbf{g}_{e_k}\|_{\square^{e_k}_\ell}. \tag{7.21}$$

So, fix $k \in [M]$, and set $e = e_k$ and $H_e = f_{e_k} - g_{e_k} \in L_p(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$. Moreover, for every $e' \in \mathcal{G} \setminus \{e\}$ let $s \in [M] \setminus \{k\}$ be such that $e' = e_s$ and set $h_{e'} = g_{e_s}$ if $s < k$ and $h_{e'} = f_{e_s}$ if $k < s$; notice that $h_{e'} \in L_p(\boldsymbol{X}, \mathcal{B}_{e'}, \boldsymbol{\mu})$. Thus, setting

$$I = \mathbb{E}\Big[H_e \prod_{e'\in\mathcal{G}\setminus\{e\}} h_{e'}\Big],$$

we need to show that $|I| \leqslant C \cdot \|\mathbf{H}_e\|_{\square^e_\ell}$ where $\mathbf{H}_e$ is as in (4.3) for H_e.

To this end, we first observe that if $\Delta = 1$, then the result is straightforward. Indeed, in this case we have $\ell = 2$, and the edges of $\mathcal{G}$ are pairwise disjoint. Hence, by part (b) of Proposition 7.1 and part (c.ii) of Proposition 7.3, we see that

$$|I| = |\mathbb{E}[\mathbf{H}_e]| \cdot \prod_{e'\in\mathcal{G}\setminus\{e\}} |\mathbb{E}[\mathbf{h}_{e'}]| \leqslant \|\mathbf{H}_e\|_{\square^e_2} \cdot \prod_{e'\in\mathcal{G}\setminus\{e\}} \|\mathbf{h}_e\|_{\square^e_{2,p}} \overset{(7.18)}{\leqslant} C \cdot \|\mathbf{H}_e\|_{\square^e_2}.$$

Therefore, in what follows we will assume that $\Delta \geqslant 2$. To simplify the exposition we will also assume that $p \neq \infty$. (The proof for the case $p = \infty$ is similar.) Write

$e=\{i,j\}$, and set $\mathcal{G}(i)=\{e'\in\mathcal{G}\setminus\{e\}: i\in e'\}$ and $\mathcal{G}^*(i)=\{e'\in\mathcal{G}\setminus\{e\}: i\notin e'\}$; notice that $\mathcal{G}\setminus\{e\}=\mathcal{G}(i)\cup\mathcal{G}^*(i)$. Let ℓ' be the conjugate exponent of ℓ and observe that, by (7.17), we have $\ell\geqslant q'$ where q' is the conjugate exponent of $p^{(\Delta-1)^{-1}}$. Hence,

$$1<\ell'\leqslant p^{(\Delta-1)^{-1}}\leqslant p. \tag{7.22}$$

We set

$$I_{e,\mathcal{G}(i)}=\mathbb{E}\Big[\prod_{\omega=0}^{\ell-1}\mathbf{H}_e(x_i^{(\omega)},x_j)\prod_{e'\in\mathcal{G}(i)}\mathbf{h}_{e'}(x_i^{(\omega)},x_{e'\setminus\{i\}})\Big] \tag{7.23}$$

and

$$I_{\mathcal{G}(i)}=\mathbb{E}\Big[\prod_{e'\in\mathcal{G}(i)}\prod_{\omega=0}^{\ell-1}|\mathbf{h}_{e'}|^{\ell'}(x_i^{(\omega)},x_{e'\setminus\{i\}})\Big] \tag{7.24}$$

where both expectations are over all $x_i^{(0)},\dots,x_i^{(\ell-1)}\in X_i$ and $\mathbf{x}_{[n]\setminus\{i\}}\in\boldsymbol{X}_{[n]\setminus\{i\}}$.

Claim 7.5. *We have $|I|\leqslant C\cdot I_{e,\mathcal{G}(i)}^{1/\ell}$.*

Proof of Claim 7.5. Since $i\notin e'$ for every $e'\in\mathcal{G}^*(i)$, we have

$$I=\mathbb{E}\Big[\mathbb{E}\big[\mathbf{H}_e(x_i,x_j)\prod_{e'\in\mathcal{G}(i)}\mathbf{h}_{e'}(x_i,x_{e'\setminus\{i\}})\,\big|\,x_i\in X_i\big]\cdot\prod_{e'\in\mathcal{G}^*(i)}\mathbf{h}_{e'}(\mathbf{x}_{e'})\Big].$$

By Hölder's inequality, (7.19), (7.22) and (7.23), we obtain that

$$\begin{aligned}|I| &\leqslant \mathbb{E}\Big[\mathbb{E}\big[\mathbf{H}_e(x_i,x_j)\prod_{e'\in\mathcal{G}(i)}\mathbf{h}_{e'}(x_i,x_{e'\setminus\{i\}})\,\big|\,x_i\in X_i\big]^{\ell}\Big]^{1/\ell}\cdot\Big\|\prod_{e'\in\mathcal{G}^*(i)}h_{e'}\Big\|_{L_{\ell'}}\\ &\leqslant I_{e,\mathcal{G}(i)}^{1/\ell}\cdot\Big\|\prod_{e'\in\mathcal{G}^*(i)}h_{e'}\Big\|_{L_p}\leqslant C\cdot I_{e,\mathcal{G}(i)}^{1/\ell}\end{aligned}$$

as desired. □

We proceed with the following claim.

Claim 7.6. *We have $I_{e,\mathcal{G}(i)}\leqslant\|\mathbf{H}_e\|_{\square_\ell^e}^{\ell}\cdot I_{\mathcal{G}(i)}^{1/\ell'}$.*

Proof of Claim 7.6. Note that $j\notin e'$ for every $e'\in\mathcal{G}(i)$, and so

$$I_{e,\mathcal{G}(i)}=\mathbb{E}\Big[\mathbb{E}\big[\prod_{\omega=0}^{\ell-1}\mathbf{H}_e(x_i^{(\omega)},x_j)\,\big|\,x_j\in X_j\big]\cdot\prod_{e'\in\mathcal{G}(i)}\prod_{\omega=0}^{\ell-1}\mathbf{h}_{e'}(x_i^{(\omega)},x_{e'\setminus\{i\}})\Big].$$

Using this observation the claim follows by Hölder's inequality and arguing precisely as in the proof of Claim 7.5. □

The following claim is the last step of the proof.

Claim 7.7. *We have $I_{\mathcal{G}(i)}\leqslant 1$.*

Proof of Claim 7.7. We may assume, of course, that $\mathcal{G}(i)$ is nonempty. We set $m = |\mathcal{G}(i)|$ and we observe that $1 \leqslant m \leqslant \Delta - 1$. Therefore, by (7.22), we see that

$$1 < (\ell')^r \leqslant (\ell')^{\Delta-1} \leqslant p \tag{7.25}$$

for every $r \in [m]$. Write $\mathcal{G}(i) = \{e'_1, \dots, e'_m\}$ and for every $r \in [m]$ let $j_r \in [n]$ such that $e'_r = \{i, j_r\}$. For every $d \in [m]$ set

$$Q_d = \mathbb{E}\Big[\prod_{r=d}^{m}\prod_{\omega=0}^{\ell-1} |\mathbf{h}_{e'_r}|^{(\ell')^d}(x_i^{(\omega)}, x_{j_r})\Big] \tag{7.26}$$

and note that

$$Q_1 = I_{\mathcal{G}(i)} \ \text{ and } \ Q_m = \mathbb{E}\Big[\prod_{\omega=0}^{\ell-1} |\mathbf{h}_{e'_m}|^{(\ell')^m}(x_i^{(\omega)}, x_{j_m})\Big]. \tag{7.27}$$

(Here, the expectation is over all $x_i^{(0)}, \dots, x_i^{(\ell-1)} \in X_i$ and $\mathbf{x}_{[n]\setminus\{i\}} \in \boldsymbol{X}_{[n]\setminus\{i\}}$.) Now observe that it is enough to show that for every $d \in [m-1]$ we have

$$Q_d \leqslant Q_{d+1}^{1/\ell'}. \tag{7.28}$$

Indeed, by (7.28), we see that $Q_1 \leqslant Q_m^{1/(\ell')^{m-1}}$. Hence, by (7.27), the monotonicity of the L_p norms and part (a) of Proposition 7.3, we obtain that

$$\begin{aligned} I_{\mathcal{G}(i)} &\leqslant \mathbb{E}\Big[\prod_{\omega=0}^{\ell-1} |\mathbf{h}_{e'_m}|^{(\ell')^m}(x_i^{(\omega)}, x_{j_m})\Big]^{\ell'/(\ell')^m} \\ &\overset{(7.25)}{\leqslant} \mathbb{E}\Big[\prod_{\omega=0}^{\ell-1} |\mathbf{h}_{e'_m}|^{p}(x_i^{(\omega)}, x_{j_m})\Big]^{\ell'/p} \leqslant \big\|\mathbf{h}_{e'_m}\big\|_{\square^{e'_m}_{\ell,p}}^{\ell\ell'} \overset{(7.18)}{\leqslant} 1. \end{aligned} \tag{7.29}$$

It remains to show (7.28). Fix $d \in [m-1]$ and notice that $j_d \notin e'_r$ for every $r \in \{d+1, \dots, m\}$. Thus,

$$Q_d = \mathbb{E}\Big[\mathbb{E}\big[\prod_{\omega=0}^{\ell-1} |\mathbf{h}_{e'_d}|^{(\ell')^d}(x_i^{(\omega)}, x_{j_d}) \,\big|\, x_{j_d} \in X_{j_d}\big] \cdot \prod_{r=d+1}^{m}\prod_{\omega=0}^{\ell-1} |\mathbf{h}_{e'_r}|^{(\ell')^d}(x_i^{(\omega)}, x_{j_r})\Big].$$

By Hölder's inequality and arguing as in the proof of (7.29), we see that

$$Q_d \leqslant \mathbb{E}\Big[\prod_{\omega \in \{0,\dots,\ell-1\}^{e'_d}} |\mathbf{h}_{e'_d}|^{(\ell')^d}(\mathbf{x}_{e'_d}^{(\omega)})\Big]^{1/\ell} \cdot Q_{d+1}^{1/\ell'} \leqslant \big\|\mathbf{h}_{e'_d}\big\|_{\square^{e'_d}_{\ell,p}}^{\ell(\ell')^d} \cdot Q_{d+1}^{1/\ell'}$$

as desired. □

By Claims 7.5, 7.6 and 7.7, we conclude that (7.21) is satisfied, and so the entire proof of Theorem 7.4 is completed. □

CHAPTER 8

Pseudorandom families

8.1. Definition and basic properties

We introduce a class of weighted hypergraphs which first appeared in [DKK18, Definition 6.1]. Closely related definitions appear in [CFZ15, Tao06a]. As we have already noted in the introduction, the most important property of this class is that it satisfies relative versions of the counting and removal lemmas, as we will see in the following two chapters. We follow the notation[1] described in the beginning of Chapter 7.

Definition 8.1. *Let $n, r \in \mathbb{N}$ with $n \geqslant r \geqslant 2$, and let $C \geqslant 1$ and $0 < \eta < 1$. Also let $1 < p \leqslant \infty$ 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 r-uniform hypergraph system. 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 *if the following hold.*

(C1) (Copies of sub-hypergraphs of $\mathcal{H}$) *For every nonempty $\mathcal{G} \subseteq \mathcal{H}$ we have $\mathbb{E}\big[\prod_{e \in \mathcal{G}} \nu_e\big] \geqslant 1 - \eta$.*

(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 condition) *For every $e' \in \mathcal{H} \setminus \{e\}$ and every $\omega \in \{0,1\}$ let $g^{(\omega)}_{e'} \in L_1(\boldsymbol{X}, \mathcal{B}_{e'}, \boldsymbol{\mu})$ such that either $0 \leqslant g^{(\omega)}_{e'} \leqslant \nu_{e'}$ or $0 \leqslant g^{(\omega)}_{e'} \leqslant 1$. Let $\boldsymbol{\nu}_e$ and $\boldsymbol{\psi}_e$ be as in (4.3) for ν_e and ψ_e respectively. Then we have*

$$\Big|\mathbb{E}\Big[(\boldsymbol{\nu}_e - \boldsymbol{\psi}_e)(\mathbf{x}_e) \prod_{\omega \in \{0,1\}} \mathbb{E}\big[\prod_{e' \in \mathcal{H} \setminus \{e\}} g^{(\omega)}_{e'}(\mathbf{x}_e, \mathbf{x}_{[n]\setminus e}) \,\big|\, \mathbf{x}_{[n]\setminus e} \in \boldsymbol{X}_{[n]\setminus e}\big] \Big|\, \mathbf{x}_e \in \boldsymbol{X}_e\Big]\Big| \leqslant \eta. \tag{8.1}$$

(C3) (Integrability of the 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 $\boldsymbol{\nu}_{e,\mathcal{G}}(\mathbf{x}_e) = \mathbb{E}\big[\prod_{e' \in \mathcal{G}} \nu_{e'}(\mathbf{x}_e, \mathbf{x}_{[n]\setminus e}) \,\big|\, \mathbf{x}_{[n]\setminus e} \in$*

[1] Recall, that if (X, Σ, μ) is a probability space and $f \colon X \to \mathbb{R}$ is a random variable then the mean value of f in X is denoted by $\int_X f(x)\, d\mu(x) = \mathbb{E}\big[f(x) \big| x \in X\big]$.

$\boldsymbol{X}_{[n]\setminus e}\big]$. *Then, setting*

$$\ell := \min\Big\{2n : n \in \mathbb{N} \text{ and } 2n \geqslant 2q + \Big(1 - \frac{1}{C}\Big) + \frac{1}{p}\Big\}, \tag{8.2}$$

we have

$$\mathbb{E}[\boldsymbol{\nu}^{\ell}_{e,\mathcal{G}}] \leqslant C + \eta. \tag{8.3}$$

Definition 8.1 looks rather technical at first sight, but it is possible to justify combinatorially conditions (C1)–(C3). 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 $\mathcal{H}$. 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 8.2. *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. □

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[2] by Conlon, Fox and Zhao in [CFZ15, Lemma 6.3], though closely related variants appear in the work of Green and Tao [GT08]. 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 $g^{(0)}_{e'} = \mathbf{1}_{A_f}$ if $e' \cap e = f$; otherwise, let $g^{(j)}_{e'} = 1$. By (8.1), 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 $\|\nu_e - \psi_e\|_{\mathcal{S}_{\partial e}} \leqslant \eta$. Condition (C3) can be seen as an instance of the general fact that by taking averages we improve integrability. It will be used in the following form.

[2]Note that in [CFZ15] condition (C2.b) is referred to as the “strong linear forms” condition.

LEMMA 8.3. *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 $\mathbf{A} \in \mathbf{\Sigma}_e$ we have*

$$\int_{\mathbf{A}} \boldsymbol{\nu}_{e,\mathcal{G}}^{2q}\, d\boldsymbol{\mu}_e \leqslant (C+1)\, \boldsymbol{\mu}_e(\mathbf{A})^{e(C,p)} \tag{8.4}$$

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

(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 $\|\boldsymbol{\nu}_{e,\mathcal{G}} - 1\|_{L_2} \leqslant 4\eta^{1/2}$. In particular, for every $\mathbf{A} \in \mathbf{\Sigma}_e$ we have*

$$\int_{\mathbf{A}} \boldsymbol{\nu}_{e,\mathcal{G}}^{2}\, d\boldsymbol{\mu}_e \leqslant 2\boldsymbol{\mu}_e(\mathbf{A}) + 8\eta^{1/2}. \tag{8.5}$$

PROOF. (a) The fact that $\ell > 2q$ follows immediately by (8.2). Next, fix $\mathbf{A} \in \mathbf{\Sigma}_e$. By Hölder's inequality, we have

$$\int_{\mathbf{A}} \boldsymbol{\nu}_{e,\mathcal{G}}^{2q}\, d\boldsymbol{\mu}_e \leqslant \|\boldsymbol{\nu}_{e,\mathcal{G}}\|_{L_\ell}^{2q} \cdot \boldsymbol{\mu}_e(\mathbf{A})^{1-\frac{2q}{\ell}} \overset{(8.3)}{\leqslant} (C+1)\, \boldsymbol{\mu}_e(\mathbf{A})^{1-\frac{2q}{\ell}}. \tag{8.6}$$

On the other hand, by (8.2) and the choice of $e(C,p)$, we see that $1 - \frac{2q}{\ell} \geqslant e(C,p)$. By (8.6), 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,

$$\begin{aligned} \|\boldsymbol{\nu}_{e,\mathcal{G}} - 1\|_{L_2}^2 &= \int (\boldsymbol{\nu}_{e,\mathcal{G}}^2 - 2\,\boldsymbol{\nu}_{e,\mathcal{G}} + 1)\, d\boldsymbol{\mu}_e \\ &\leqslant \Big| \int (\boldsymbol{\nu}_{e,\mathcal{G}}^2 - 1)\, d\boldsymbol{\mu}_e \Big| + 2 \Big| \int (\boldsymbol{\nu}_{e,\mathcal{G}} - 1)\, d\boldsymbol{\mu}_e \Big| \leqslant 4\eta^{1/2}. \end{aligned} \tag{8.7}$$

Now let $\mathbf{A} \in \mathbf{\Sigma}_e$ and note that $\|\boldsymbol{\nu}_{e,\mathcal{G}} \cdot \mathbf{1}_{\mathbf{A}} - \mathbf{1}_{\mathbf{A}}\|_{L_2} \leqslant \|\boldsymbol{\nu}_{e,\mathcal{G}} - 1\|_{L_2}$. Hence, by (8.7) and the triangle inequality, we have $\|\boldsymbol{\nu}_{e,\mathcal{G}} \cdot \mathbf{1}_{\mathbf{A}}\|_{L_2} \leqslant \|\mathbf{1}_{\mathbf{A}}\|_{L_2} + (4\eta^{1/2})^{1/2}$ 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. $\square$

8.2. 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 (8.1) 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 (8.1) to these conditions by repeated applications of the Cauchy–Schwarz inequality. This method was developed extensively by Green and Tao [GT08, GT10] and has become standard in the field. As such, we will not present the proof of the following proposition here, see e.g [DKK15, CFZ15].

PROPOSITION 8.4. *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 (4.3) for ν_e. Assume that the following properties are satisfied.*

(P1) *If ℓ is as in (8.2), then*
$$1-\eta \leqslant \mathbb{E}\big[\prod_{e\in\mathcal{H}}\prod_{\omega\in\{0,\dots,\ell-1\}^e} \boldsymbol{\nu}_e^{n_{e,\omega}}(\mathbf{x}_e^{(\omega)}) \,\big|\, \mathbf{x}_e^{(0)},\dots,\mathbf{x}_e^{(\ell-1)} \in \boldsymbol{X}_e\big] \leqslant C+\eta$$
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*
$$\big|\mathbb{E}\big[\prod_{\omega\in\{0,1\}^e}(\boldsymbol{\nu}_e-\boldsymbol{\psi}_e)(\mathbf{x}_e^{(\omega)}) \prod_{e'\in\mathcal{H}\setminus\{e\}}\prod_{\omega\in\{0,1\}^{e'}} \boldsymbol{\nu}_{e'}^{n_{e',\omega}}(\mathbf{x}_{e'}^{(\omega)}) \,\big|\, {\substack{\mathbf{x}_e^{(0)},\mathbf{x}_e^{(1)}\in\boldsymbol{X}_e \\ \mathbf{x}_{e'}^{(0)},\mathbf{x}_{e'}^{(1)}\in\boldsymbol{X}_{e'}}}\big]\big| \leqslant \eta$$
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 $\eta' = (C+1)\,\eta^{1/2^r}$.

8.3. The linear forms condition

We isolate now a special subclass of pseudorandom families that will play an important role in the arithmetic applications of the relative removal lemma in Part 4.

DEFINITION 8.5 (Linear forms condition for hypergraphs). *Let $n, r \in \mathbb{N}$ with $n \geqslant r \geqslant 2$ and $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n]\rangle, \mathcal{H})$ be an r-uniform hypergraph system. Also 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 $\langle \nu_e : e \in \mathcal{H}\rangle$ satisfies the linear forms condition if*
$$\mathbb{E}\big[\prod_{e\in\mathcal{H}}\prod_{\omega\in\{0,1\}^e} \boldsymbol{\nu}_e^{n_{e,\omega}}(\mathbf{x}_e^{(\omega)}) \,\big|\, \mathbf{x}_e^{(0)},\mathbf{x}_e^{(1)} \in \boldsymbol{X}_e\big] = 1+o(1) \tag{8.8}$$

for any choice of $n_{e,\omega} \in \{0,1\}$. *In the previous expression* $\boldsymbol{\nu}_e$ *is as in* (4.3) .

Taking $C = 1$, $p = \infty$, $\ell = 1$ and $\psi_e = 1$ for every $e \in \mathcal{H}$ then a family of measures that satisfies (8.8) we see that it also satisfies properties (P1) and (P2) in Proposition 8.4, see [CFZ15, Lemma 6.3].

8.4. Examples of Pseudorandom families

We present now two examples of pseudorandom families. The proofs that the following examples are indeed pseudorandom families are omitted and may be found in [DKK15].
Our first example is the following theorem.

THEOREM 8.6. *Let* $n \in \mathbb{N}$ *with* $n \geqslant 3$, $C \geqslant 1$ *and* $1 < p \leqslant \infty$, *and let* ℓ *be as in* (8.2). *Also let* $0 < \eta \leqslant (4C)^{-n\ell^n}$ *and let* $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n]\rangle, \mathcal{H})$ *be a hypergraph system with* $\mathcal{H} = \binom{n}{n-1}$. *(In particular,* $\mathscr{H}$ *is* $(n-1)$*-uniform.) For every* $e \in \mathcal{H}$ *let* $\lambda_e \in L_1(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ *and* $\varphi_e \in L_p(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ *be nonnegative random variables, and let* $\boldsymbol{\lambda}_e$ *and* $\boldsymbol{\varphi}_e$ *be as in* (4.3) *for* λ_e *and* φ_e *respectively. Assume that the following conditions are satisfied.*

(I) *We have*

$$1 - \eta \leqslant \mathbb{E}\Big[\prod_{e\in\mathcal{H}} \prod_{\omega\in\{0,\dots,\ell-1\}^e} \boldsymbol{\lambda}_e^{n_{e,\omega}}(\mathbf{x}_e^{(\omega)}) \,\Big|\, \mathbf{x}^{(0)}, \dots, \mathbf{x}^{(\ell-1)} \in \boldsymbol{X}\Big] \leqslant 1 + \eta \tag{8.9}$$

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

(II) *For every* $e \in \mathcal{H}$ *we have* $\|\boldsymbol{\varphi}_e\|_{\square^e_{\ell,p}} \leqslant C$.

Then the family $\langle \lambda_e + \varphi_e : e \in \mathcal{H}\rangle$ *is* (C', η', p)*-pseudorandom where* $C' = (4C)^{n\ell}$ *and* $\eta' = (4C)^{n\ell}\, \eta^{1/\ell^{n-1}}$.

We will briefly comment on the assumptions of Theorem 8.6. We first observe that condition (I) is a modification of the "linear forms condition". It expresses the fact that the weighted hypergraph $\langle \lambda_e : e \in \mathcal{H}\rangle$ contains roughly the expected number of copies of the ℓ-blow-up of $\mathcal{H}$ and its sub-hypergraphs; as such, it is a rather strong independence-type assumption. On the other hand, note that condition (II) is just an integrability assumption for the function φ_e. Thus, we see that the family $\langle \lambda_e + \varphi_e : e \in \mathcal{H}\rangle$ is a perturbation of $\langle \lambda_e : e \in \mathcal{H}\rangle$ where only integrability conditions are imposed on each "noise" φ_e.

The second example is the following theorem. This theorem was motivated by [CFZ13, Lemmas 5 and 6] which dealt with the case $C = 1$, $p = \infty$ and $\psi_e = 1$ for every $e \in \mathcal{H}$.

THEOREM 8.7. *Let $n \in \mathbb{N}$ with $n \geqslant 3$, $C \geqslant 1$ and $1 < p \leqslant \infty$, and let ℓ be as in (8.2). Also let $0 < \eta \leqslant 1/(n\ell)$ and let $\mathscr{H} = (n, \langle (X_i, \Sigma_i, \mu_i) : i \in [n]\rangle, \mathcal{H})$ be a hypergraph system with $\mathcal{H} = \binom{n}{n-1}$. (Again observe that $\mathscr{H}$ is $(n-1)$-uniform.) For every $e \in \mathcal{H}$ let $\nu_e, \psi_e \in L_p(\boldsymbol{X}, \mathcal{B}_e, \boldsymbol{\mu})$ be nonnegative random variables, and let $\boldsymbol{\nu}_e$ and $\boldsymbol{\psi}_e$ be as in (4.3) for ν_e and ψ_e respectively. Assume that the following conditions are satisfied.*

(I) *We have*

$$1 - \eta \leqslant \mathbb{E}\Big[\prod_{e\in\mathcal{H}} \prod_{\omega\in\{0,\dots,\ell-1\}^e} \boldsymbol{\psi}_e^{n_{e,\omega}}(\mathbf{x}_e^{(\omega)}) \,\Big|\, \mathbf{x}^{(0)}, \dots, \mathbf{x}^{(\ell-1)} \in \boldsymbol{X}\Big] \leqslant C + \eta \tag{8.10}$$

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

(II) *We have $1 \leqslant \|\boldsymbol{\nu}_e\|_{\square^e_{\ell,p}} < \infty$, $\|\boldsymbol{\psi}_e\|_{\square^e_{\ell,p}} \leqslant C$ and*

$$\|\boldsymbol{\nu}_e - \boldsymbol{\psi}_e\|_{\square^e_\ell} \leqslant \eta\,(C \cdot M)^{-(n-1)\ell} \tag{8.11}$$

where $M = \max\{\|\boldsymbol{\nu}_e\|_{\square^e_{\ell,p}} : e \in \mathcal{H}\}$.

Then the family $\langle \nu_e : e \in \mathcal{H}\rangle$ is (C, η', p)-pseudorandom where $\eta' = n\ell\eta$.

CHAPTER 9

Relative counting lemma for pseudorandom families

We present now a relative counting lemma for pseudorandom families. Similar results may be found in several places, see e.g. [Tao06c, NRS06, GT08, CFZ15].

THEOREM 9.1 (Relative Counting lemma). *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*

$$\Big| \int \prod_{e \in \mathcal{H}} g_e \, d\boldsymbol{\mu} - \int \prod_{e \in \mathcal{H}} h_e \, d\boldsymbol{\mu} \Big| \leqslant \gamma. \tag{9.1}$$

The hypotheses of Theorem 9.1 might appear rather strong: on the one hand the function g_e is dominated by ν_e (and so, by Lemma 8.2 , 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, as we will see in Proposition 10.3. The rest of this chapter is devoted to the proof of Theorem 9.1.

Proof of Theorem 9.1. 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 9.1 only for the case "$\zeta = 1$". Indeed, if the numbers $\eta(n, r, C, p, 1)$ and $\alpha(n, r, C, p, 1)$ have been determined, then it is easy to see that for every $\zeta \geqslant 1$ Theorem 9.1 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

$$\beta(\gamma) = \big(10(C+1)^2\gamma^{-1}\big)^{2q/x(C,p)} \ \text{ and } \ \theta(\gamma) = \big(20(C+1)\beta(\gamma)\big)^{-2q}\gamma^{2q}, \tag{9.2}$$

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 $m \in \{0,\dots,n^r\}$ and every $0 < \gamma \leqslant 1$ we define $\alpha_m(\gamma)$ and $\eta_m(\gamma)$ in $(0,1]$ recursively by the rule

$$\alpha_0(\gamma) = \gamma/5 \ \text{ and } \ \alpha_{m+1}(\gamma) = \alpha_m\big(\theta(\gamma)\big) \tag{9.3}$$

and

$$\eta_0(\gamma) = \big(30(C+1)\big)^{-4q}\gamma^{4q} \text{ and } \eta_{m+1}(\gamma) = \eta_m(\theta(\gamma)). \tag{9.4}$$

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 $m \in \{0,\dots,|\mathcal{H}|\}$ we will show that for every $0 < \gamma \leqslant 1$ if the family $\langle \nu_e : e \in \mathcal{H}\rangle$ is $(C,\eta(\gamma),p)$-pseudorandom where $\eta_m(\gamma)$ is as in (9.4), then the estimate (9.1) 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 (9.3), and (P3) $|\{e \in \mathcal{H} : g_e \neq h_e\}| \leqslant m$.

The initial case "$m = 0$" is straightforward, and so let $m \in \{1,\dots,|\mathcal{H}|\}$ 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

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

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 $G, H \colon \boldsymbol{X}_{e_0} \to \mathbb{R}$ 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 8.1. On the other hand, notice that $G = H$ if $\mathcal{G} = \emptyset$.

We are ready present two claims which are the main steps towards the proof of Theorem 9.1. Their proof will be given after we see how they are used in the proof of this Theorem. In the following claim we obtain a first estimate for $|\Delta|$. As we said earlier it is the first step of the proof of Theorem 9.1 and is important to note that its proof does not use the inductive assumptions and relies, instead, on the local linear forms condition (condition (C2.b) in Definition 8.1) and Hölder's inequality. Closely related estimates appear in [CFZ15, Tao06a].

CLAIM 9.2. *We have*

$$|\Delta| \leqslant 2(C+1)\big(\|G-H\|_{L_{2q}} + \eta_m(\gamma)^{1/2}\big) + \|g_{e_0} - h_{e_0}\|_{\mathcal{S}_{\partial e_0}}. \tag{9.5}$$

The next claim is the second step of the proof.

CLAIM 9.3. *If $\beta(\gamma)$ and $\theta(\gamma)$ are as in (9.2), then we have*

$$\int (G-H)^{2q}\, d\boldsymbol{\mu}_{e_0} \leqslant 2\beta(\gamma)^{2q}\theta(\gamma) + (C+1)^2\beta(\gamma)^{-x(C,p)} + 8\eta_m(\gamma)^{1/2}. \tag{9.6}$$

Granting Claims 9.2 and 9.3, the proof of the inductive step (and, consequently, of Theorem 9.1) is completed as follows. First observe that, by (9.4) 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 9.3 and the choice of $\beta(\gamma)$ and $\theta(\gamma)$ in (9.2), it is easy to see that $\|G-H\|_{L_{2q}} \leqslant 3\big(10(C+1)\big)^{-1}\gamma$. Therefore, by Claim 9.2 and property (P2)

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

It remains to prove Claims 9.2 and 9.3.

Proof of Claim 9.2. Let $\mathbf{g}_{e_0}$ be as in (4.3) 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

$$|I_2| \leqslant \|g_{e_0} - h_{e_0}\|_{\mathcal{S}_{\partial e_0}}. \tag{9.7}$$

This follows by Fubini's theorem and the following well-known fact (see, e.g., [Gow07]). We recall the proof for the convenience of the reader.

FACT 9.4. *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 $\{f_1, \dots, f_k\}$ be an enumeration of ∂e. We define $Z\colon [0,1]^k \to \mathbb{R}$ by the rule $Z(t_1, \dots, t_k) = \int g_e \prod_{i=1}^k \mathbf{1}_{[u_{f_i} > t_i]}\, d\boldsymbol{\mu}$. Notice that $\bigcap_{i=1}^k [u_{f_i} > t_i] \in \mathcal{S}_{\partial e}$ for every $(t_1, \dots, t_k) \in [0,1]^k$ and so $\|G\|_{L_\infty} \leqslant \|g_e\|_{\mathcal{S}_{\partial e}}$. On the other hand, denoting by $\boldsymbol{\lambda}$ the Lebesgue measure on $[0,1]^k$, by Fubini's theorem we have $\int g_e \prod_{f\in\partial e} u_f\, d\boldsymbol{\mu} = \int G\, d\boldsymbol{\lambda}$ and the result follows. □

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 8.1 and notice that by condition (C2.a) we have $|\int(\nu_{e_0} - \psi_{e_0})\, d\boldsymbol{\mu}| \leqslant \eta(\gamma)$. 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

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

where $\boldsymbol{\psi}_{e_0}$ is as in (4.3) for ψ_{e_0}. Next, writing $(G-H)^2 = G^2 - 2GH + H^2$ and applying (8.1), we see that $|\int(\boldsymbol{\nu}_{e_0} - \boldsymbol{\psi}_{e_0})(G-H)^2\, d\boldsymbol{\mu}_{e_0}| \leqslant 4\eta(\gamma)$. On the other hand, by Hölder's inequality, $|\int \boldsymbol{\psi}_{e_0}(G-H)^2\, d\boldsymbol{\mu}_{e_0}| \leqslant C\|G-H\|^2_{L_{2q}}$. Therefore,

$$|I_1| \leqslant 2(C+1)\big(\|G-H\|_{L_{2q}} + \eta_m(\gamma)^{1/2}\big). \tag{9.8}$$

Combining (9.7) and (9.8) we conclude that the estimate in (9.5) is satisfied, as desired.

Before we pass to the proof of Claim 9.3 we make the following comments. Estimates of this form are usually obtained for stronger norms than the cut norm, and as such, they depend on stronger pseudorandomness conditions. In fact, so far the only general method available in this context was developed by Conlon, Fox and Zhao [CFZ15]. 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 [Sha16, TZ15b]).

We introduce a new method to deal with these types of problems which 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 8.3 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 (9.6) thus completing the proof of Claim 9.3. More precicely

Proof of Claim 9.3. 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 $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbf{\Sigma}_{e_0}$. 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 $h'_{e_0} = (H^{2q-1}\mathbf{1}_{\mathbf{A}}) \circ \pi_{e_0} \in L_1(\boldsymbol{X}, \mathcal{B}_{e_0}, \boldsymbol{\mu})$ and notice that $0 \leqslant h'_{e_0} \leqslant 1$. Moreover, by the definition of G and H, we see that

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

On the other hand, by (9.3) and property (P2), we have $\|g_e - h_e\|_{\mathcal{S}_{\partial e}} \leqslant \alpha_{m-1}\big(\theta(\gamma)\big)$ for every $e \in \mathcal{H} \setminus \{e_0\}$. Hence, by our inductive assumptions, we obtain that

$$I_1 \leqslant \theta(\gamma). \tag{9.9}$$

The estimation of I_2 is similar. Indeed, observe that

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

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

$$I_2 \leqslant \beta(\gamma)^{2q-1} \theta(\gamma). \tag{9.10}$$

We proceed to estimate I_3. Let $\boldsymbol{\nu}_{e_0,\mathcal{G}}$ and ℓ be as in Definition 8.1, 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,

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

Thus, by Lemma 8.3 and the choice of $x(C,p)$, we have

$$\begin{aligned} I_3 &\leqslant \int_{\mathbf{C}} \boldsymbol{\nu}_{e_0,\mathcal{G}}^{2q} \, d\boldsymbol{\mu}_{e_0} \leqslant (C+1)\, \boldsymbol{\mu}_{e_0}(\mathbf{C})^{x(C,p)} + 8\eta_m(\gamma)^{1/2} \\ &\leqslant (C+1)^2 \beta(\gamma)^{-x(C,p)} + 8\eta_m(\gamma)^{1/2}. \end{aligned} \tag{9.11}$$

Combining (9.9)–(9.11) we conclude that the estimate in (9.6) is satisfied. The proof of Claim 9.3 is completed.

CHAPTER 10

Relative removal lemma for pseudorandom families

THEOREM 10.1 (Relative Removal lemma). *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*

$$\int \prod_{e \in \mathcal{H}} f_e \, d\boldsymbol{\mu} \leqslant \delta. \tag{10.1}$$

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

$$\int_{\boldsymbol{X} \setminus F_e} f_e \, d\boldsymbol{\mu} \leqslant \varepsilon \quad \text{and} \quad \bigcap_{e \in \mathcal{H}} F_e = \emptyset. \tag{10.2}$$

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 $\bigcup_{e' \subsetneq e} \mathcal{P}_{e'}$.*

Before we proceed to the proof of the previous theorem we need some preparatory work.

10.1. Preliminary tools

The first key ingredient towards the proof of Theorem 10.1 is the following version of the removal lemma for hypergraph systems which is due to Tao [Tao06c] (see also [DK16] for an exposition). Closely related discrete analogues were obtained earlier by Gowers [Gow07] and , independently, by Nagle, Rödl, Schacht and Shokan [NRS06, RS04].

THEOREM 10.2 (Removal lemma). *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

$$\boldsymbol{\mu}\Big(\bigcap_{e\in\mathcal{H}} E_e\Big) \leqslant \Delta(n,r,\varepsilon). \tag{10.3}$$

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

$$\boldsymbol{\mu}(E_e \setminus F_e) \leqslant \varepsilon \quad \text{and} \quad \bigcap_{e\in\mathcal{H}} F_e = \emptyset. \tag{10.4}$$

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 $\bigcup_{e' \subsetneq e} \mathcal{P}_{e'}$.*

Another key ingredient for the proof of Theorem 10.1 is the following proposition.

PROPOSITION 10.3. *Let n, r, C, p and $\mathscr{H}$ be as in Theorem 9.1, 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 $f_e = f^e_{\mathrm{str}} + f^e_{\mathrm{err}} + f^e_{\mathrm{unf}}$ where f^e_{str}, f^e_{err} and f^e_{unf} are as in (6.1). Assume that the estimates in (6.2) 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 $\|f_e \cdot \mathbf{1}_A - f^e_{\mathrm{str}} \cdot \mathbf{1}_A\|_{\mathcal{S}_{\partial e}} \leqslant \alpha$.*

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

$$\int_{\boldsymbol{X}\setminus A} f_e \, d\boldsymbol{\mu} \leqslant C^p \zeta^{1-p} + \alpha \quad \text{and} \quad \int_{\boldsymbol{X}\setminus A} f^e_{\mathrm{str}} \, d\boldsymbol{\mu} \leqslant C^p \zeta^{1-p}. \tag{10.5}$$

PROOF. 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

$$f_e \cdot \mathbf{1}_A - f^e_{\mathrm{str}} \cdot \mathbf{1}_A = f^e_{\mathrm{err}} \cdot \mathbf{1}_A + \sum_{P \in \mathcal{P}'} f^e_{\mathrm{unf}} \cdot \mathbf{1}_P.$$

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

$$\begin{aligned} \Big|\int_B (f_e \cdot \mathbf{1}_A - f^e_{\mathrm{str}} \cdot \mathbf{1}_A)\, d\boldsymbol{\mu}\Big| &\leqslant \Big|\int_{B\cap A} f^e_{\mathrm{err}}\, d\boldsymbol{\mu}\Big| + \sum_{P\in\mathcal{P}'} \Big|\int_{B\cap P} f^e_{\mathrm{unf}}\, d\boldsymbol{\mu}\Big| \\ &\leqslant \|f^e_{\mathrm{err}}\|_{L_{p^\dagger}} + M \cdot \|f^e_{\mathrm{unf}}\|_{\mathcal{S}_{\partial e}} \leqslant \sigma + \frac{M}{F(M)} \leqslant \alpha \end{aligned}$$

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

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

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

and so, by Hölder's inequality,

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

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

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

and the proof of Proposition 10.3 is completed. □

10.2. Proof of the Relative Removal lemma

We begin by introducing some numerical invariants. First, we set

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

where q is the conjugate exponent of p. Also let $\Delta(n, r, \frac{\varepsilon}{6\zeta})$ and $K(n, r, \frac{\varepsilon}{6\zeta})$ be as in Theorem 10.2 and note that we may assume that $\Delta(n, r, \frac{\varepsilon}{6\zeta}) \leqslant \frac{\varepsilon}{6\zeta}$. We define

$$\delta = \delta(n, r, C, p, \varepsilon) = \frac{\Delta(n, r, \frac{\varepsilon}{6\zeta})^{n^r}}{2} \ \text{ and } \ k = k(n, r, C, p, \varepsilon) = K\Big(n, r, \frac{\varepsilon}{6\zeta}\Big). \tag{10.6}$$

Next, let $\alpha(n, r, C, p, \zeta, \delta)$ and $\eta(n, r, C, p, \zeta, \delta)$ be as in Theorem 9.1 and set

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

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 6.1. Finally, we define

$$\eta = \eta(n, r, C, p, \varepsilon) = \min\{1/\mathrm{Reg}, \eta(n, r, C, p, \zeta, \delta)\}. \tag{10.7}$$

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

$$\int \prod_{e\in\mathcal{H}} f_e \, d\boldsymbol{\mu} \leqslant \delta. \tag{10.8}$$

By Lemma 8.2, 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 (10.7), we may apply Theorem 6.1 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 $f_e = f^e_{\mathrm{str}} + f^e_{\mathrm{err}} + f^e_{\mathrm{unf}}$ where f^e_{str}, f^e_{err} and f^e_{unf} are as in (6.1), we have the estimates

$$\|f^e_{\mathrm{str}}\|_{L_p} \leqslant C+1, \quad \|f^e_{\mathrm{err}}\|_{L_{p^\dagger}} \leqslant \alpha/2 \quad \text{and} \quad \|f^e_{\mathrm{unf}}\|_{\mathcal{S}_{\partial e}} \leqslant \frac{1}{F(M)} \tag{10.9}$$

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

$$A_e = [f^e_{\mathrm{str}} \leqslant \zeta], \quad g_e = f_e \cdot \mathbf{1}_{A_e} \quad \text{and} \quad h_e = f^e_{\mathrm{str}} \cdot \mathbf{1}_{A_e} \tag{10.10}$$

and notice that $0 \leqslant g_e \leqslant \nu_e$ and $0 \leqslant h_e \leqslant \zeta$. Moreover, by Proposition 10.3, we see that $\|g_e - h_e\|_{\mathcal{S}_{\partial e}} \leqslant \alpha$.

Claim 10.4. *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 9.1,

$$\Big|\int \prod_{e\in\mathcal{H}} g_e \, d\boldsymbol{\mu} - \int \prod_{e\in\mathcal{H}} h_e \, d\boldsymbol{\mu}\Big| \leqslant \delta. \tag{10.11}$$

On the other hand, we have $0 \leqslant g_e \leqslant f_e$ for every $e \in \mathcal{H}$. Hence, by (10.8) and (10.11),

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

Finally, by (10.6), we have $2\delta \leqslant \Delta(n,r,\frac{\varepsilon}{6\zeta})^{n^r}$ and the proof is completed. □

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 10.4 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 10.2, for every $e \in \mathcal{H}$ there exists $F_e \in \mathcal{B}_e$ with

$$\boldsymbol{\mu}(E_e \setminus F_e) \leqslant \frac{\varepsilon}{6\zeta} \quad \text{and} \quad \bigcap_{e\in\mathcal{H}} F_e = \emptyset. \tag{10.12}$$

Moreover, by (10.6), 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

$$\int_{\boldsymbol{X}\setminus F_e} f_e \, d\boldsymbol{\mu} \leqslant \varepsilon \tag{10.13}$$

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

$$\int_{\boldsymbol{X}\setminus F_e} f_e\, d\boldsymbol{\mu} \leqslant \int_{\boldsymbol{X}\setminus F_e} h_e\, d\boldsymbol{\mu} + \Big|\int_{\boldsymbol{X}\setminus F_e} (g_e - h_e)\, d\boldsymbol{\mu}\Big| + \Big|\int_{\boldsymbol{X}\setminus F_e} (f_e - g_e)\, d\boldsymbol{\mu}\Big|. \quad (10.14)$$

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

$$\begin{aligned} \int_{\boldsymbol{X}\setminus F_e} h_e\, d\boldsymbol{\mu} &\leqslant \int_{\boldsymbol{X}\setminus E_e} h_e\, d\boldsymbol{\mu} + \int_{E_e\setminus F_e} h_e\, d\boldsymbol{\mu} \qquad (10.15)\\ &\leqslant \Delta\Big(n, r, \frac{\varepsilon}{6\zeta}\Big) + \zeta\, \boldsymbol{\mu}(E_e \setminus F_e) \overset{(10.12)}{\leqslant} \varepsilon/3. \end{aligned}$$

To estimate the second term in the right-hand side of (10.14), 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

$$\Big|\int_{\boldsymbol{X}\setminus F_e} (g_e - h_e)\, d\boldsymbol{\mu}\Big| \leqslant \sum_{A\in\mathcal{F}} \Big|\int_A (g_e - h_e)\, d\boldsymbol{\mu}\Big| \leqslant |\mathcal{F}|\, \alpha \leqslant k^{2^r} \alpha \leqslant \varepsilon/3. \quad (10.16)$$

Finally, to estimate the last term in the right-hand side of (10.14), 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 10.3 and the choice of ζ and α, we obtain that

$$\begin{aligned} \Big|\int_{\boldsymbol{X}\setminus F_e} (f_e - g_e)\, d\boldsymbol{\mu}\Big| &= \int_{\boldsymbol{X}\setminus F_e} f_e \cdot \mathbf{1}_{\boldsymbol{X}\setminus A_e}\, d\boldsymbol{\mu} \leqslant \int_{\boldsymbol{X}\setminus A_e} f_e\, d\boldsymbol{\mu} \qquad (10.17)\\ &\leqslant (C+1)^p \zeta^{1-p} + \alpha \leqslant \varepsilon/3. \end{aligned}$$

Combining (10.14)–(10.17) we conclude that (10.13) is satisfied, and so the entire proof of Theorem 10.1 is completed.

Part IV

Arithmetic consequences of the Relative Removal lemma

CHAPTER 11

An arithmetic version of the Relative Removal lemma

In this chapter we present a Szemerédi-type result for sparse preudorandom subsets of finite additive groups. (Recall that an *additive group* is an abelian group written additively.) The argument for deducing this result is well-known – see , e.g., [Gow07, RTST06, Sol04, Tao06a] – and originates from the work of Ruzsa and Szemerédi [RS78]. It follows from Theorem 10.1 arguing precisely as in the proof of [Tao06a, Theorem 2.18].

THEOREM 11.1. *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 $\langle \varphi_i : i \in [k] \rangle$ be a collection of group homomorphisms from Z into Z' such that the set $\{\varphi_i(d) - \varphi_j(d) : i, j \in [k] \text{ and } d \in Z\}$ 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) *$(X_i, \mu_i)$ 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*

$$\nu_{[k]\setminus\{j\}}\big((x_i)_{i\in[k]}\big) = \nu\Big(\sum_{i\in[k]} \big(\varphi_i(x_i) - \varphi_j(x_i)\big)\Big). \tag{11.1}$$

(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

$$\mathbb{E}\Big[\prod_{j\in[k]} f\big(a + \varphi_j(d)\big) \,\Big|\, a \in Z', d \in Z\Big] \geqslant c. \tag{11.2}$$

PROOF. Let k, C, p and δ be as in the statement of the theorem and set $r = k-1$. Also let $\eta(k, r, C, p, \frac{\delta}{2k^2})$ and $\delta(k, r, C, p, \frac{\delta}{2k^2})$ be as in Theorem 10.1 and define

$$N = N(k, C, p, \delta) = \Big\lceil \frac{1}{\eta(k, r, C, p, \frac{\delta}{2k^2})} \Big\rceil \text{ and } c = c(k, C, p, \delta) = \delta\Big(k, r, C, p, \frac{\delta}{2k^2}\Big).$$

We will show that N and c are as desired.

To this end, fix the data $Z, Z', \langle \varphi_i : i \in [k]\rangle, \mathscr{H}, \nu$ and $\langle \nu_{[n]\setminus\{j\}} : j \in [k]\rangle$. Moreover, let $f\colon Z' \to \mathbb{R}$ with $0 \leqslant f \leqslant \nu$ and $\mathbb{E}[f] \geqslant \delta$ and assume, towards a contradiction, that (11.2) is not satisfied. First, we introduce some families of group homomorphisms between the additive groups $\boldsymbol{X}, Z'\times Z$ and Z' as follows. We begin by defining $Q\colon \boldsymbol{X} \to Z' \times Z$ by the rule

$$Q\big((x_i)_{i\in[k]}\big) = \Big(\sum_{i\in[k]} \varphi_i(x_i), -\sum_{i\in[k]} x_i\Big). \tag{11.3}$$

Using the fact that the set $\{\varphi_i(d)-\varphi_j(d) : i,j \in [k] \text{ and } d \in Z\}$ generates Z', we see that Q is an onto homomorphism. Next, for every $j \in [k]$ we define $s_j\colon Z'\times Z \to Z'$ and $Q_j\colon \boldsymbol{X} \to Z'$ by setting

$$s_j(a,d) = a + \varphi_j(d) \ \text{ and } \ Q_j(\mathbf{x}) = s_j\big(Q(\mathbf{x})\big). \tag{11.4}$$

Observe that for every $j \in [k]$ the maps s_j and Q_j are onto homomorphisms. Also notice that, by (11.3) and (11.4), we have

$$Q_j\big((x_i)_{i\in[k]}\big) = \sum_{i\in[k]} \big(\varphi_i(x_i) - \varphi_j(x_i)\big) = \sum_{i\in[k]\setminus\{j\}} \big(\varphi_i(x_i) - \varphi_j(x_i)\big) \tag{11.5}$$

and so $Q_j \in L_1(\boldsymbol{X}, \mathcal{B}_{[k]\setminus\{j\}}, \boldsymbol{\mu})$. Finally, for every $j \in [k]$ we set $e_j = [k]\setminus\{j\}$ and we define $f_{e_j}\colon \boldsymbol{X} \to \mathbb{R}$ by

$$f_{e_j} = f \circ Q_j. \tag{11.6}$$

Note that, by (11.1) and (11.5), we also have $\nu_{e_j} = \nu \circ Q_j$ for every $j \in [k]$.

We claim that the hypergraph system $\mathscr{H}$ and the families $\langle \nu_{e_j} : j \in [k]\rangle$ and $\langle f_{e_j} : j \in [k]\rangle$ satisfy the assumptions of Theorem 10.1. Indeed, by the choice of N and the fact that $|X_i| = |Z| \geqslant N$ for every $i \in [k]$, the hypergraph system $\mathscr{H}$ is $\eta(k,r,C,p,\frac{\delta}{2k^2})$-nonatomic and r-uniform. It is also clear that for every $j \in [k]$ we have $f_{e_j}, \nu_{e_j} \in L_1(\boldsymbol{X}, \mathcal{B}_{e_j}, \boldsymbol{\mu})$ and $0 \leqslant f_{e_j} \leqslant \nu_{e_j}$. Hence, it is enough to show that

$$\mathbb{E}\Big[\prod_{j\in[k]} f_{e_j}(\mathbf{x}) \,\Big|\, \mathbf{x} \in \boldsymbol{X}\Big] \leqslant \delta\Big(k,r,C,p,\frac{\delta}{2k^2}\Big). \tag{11.7}$$

To see that (11.7) is satisfied notice first that $|Q^{-1}(a_1,d_1)| = |Q^{-1}(a_2,d_2)|$ for every $(a_1,d_1),(a_2,d_2) \in Z'\times Z$ since $Q\colon \boldsymbol{X} \to Z'\times Z$ is an onto homomorphism. Therefore, the map Q is a measure preserving transformation. (Here, we view $\boldsymbol{X}$ and $Z'\times Z$ as discrete probability spaces equipped with the corresponding uniform probability measures.) By (11.4), (11.6), the choice of c and our assumption that (11.2) is not satisfied, we conclude that

$$\begin{aligned}\mathbb{E}\Big[\prod_{j\in[k]} f_{e_j}(\mathbf{x})\,\Big|\,\mathbf{x}\in\boldsymbol{X}\Big] &= \mathbb{E}\Big[\prod_{j\in[k]} f\big(s_j(a,d)\big)\,\Big|\,a\in Z', d\in Z\Big]\\ &= \mathbb{E}\Big[\prod_{j\in[k]} f\big(a+\varphi_j(d)\big)\,\Big|\,a\in Z', d\in Z\Big] < \delta\Big(k,r,C,p,\frac{\delta}{2k^2}\Big).\end{aligned}$$

It follows from the previous discussion that we may apply Theorem 10.1 and we obtain a family $\langle \mathbf{F}_{e_j} : j\in[k]\rangle$ with $\mathbf{F}_{e_j}\subseteq \prod_{i\in e_j} X_i$ for every $j\in[k]$ such that, setting $F_{e_j}=\mathbf{F}_{e_j}\times X_j$, we have

$$\bigcap_{j\in[n]} F_{e_j}=\emptyset \ \text{ and } \ \mathbb{E}[f_{e_j}\cdot \mathbf{1}_{\boldsymbol{X}\setminus F_{e_j}}]\leqslant \frac{\delta}{2k^2}. \tag{11.8}$$

Now for every $j\in[k]$ we set

$$A_j=\Big\{a\in Z' : |Q_j^{-1}(a)\cap(\boldsymbol{X}\setminus F_{e_j})|<\frac{1}{k}\cdot|Q_j^{-1}(a)|\Big\}. \tag{11.9}$$

Claim 11.2. *The following hold.*

(a) *For every $a\in Z'$ and every $d\in Z$ we have $\prod_{j\in[k]}\mathbf{1}_{A_j}\big(a+\varphi_j(d)\big)=0$.*
(b) *For every $j\in[k]$ we have $\mathbb{E}[f\cdot\mathbf{1}_{Z'\setminus A_j}]<\delta/k$.*

Granting the above claim, the proof of the theorem is completed as follows. By part (a) of Claim 11.2 applied for "$d=0$", we see that $\bigcap_{j\in[k]}A_j=\emptyset$ and as such $Z'=\bigcup_{j\in[k]}(Z'\setminus A_j)$. Therefore, invoking part (b) of Claim 11.2, we get that $\mathbb{E}[f]\leqslant\sum_{j\in[k]}\mathbb{E}[f\cdot\mathbf{1}_{Z'\setminus A_j}]<\delta$ which is clearly a contradiction.

We proceed to the proof of Claim 11.2. First we argue for part (a). Assume that there exists a pair $(a_0,d_0)\in Z'\times Z$ such that $a_0+\varphi_j(d_0)\in A_j$ for every $j\in[k]$. Set $E_0=Q^{-1}\big(\{(a_0,d_0)\}\big)$. Note that $E_0=Q_j^{-1}\big(\{a_0+\varphi_j(d_0)\}\big)$ for every $j\in[k]$ and so, by (11.9), we have $|E_0\cap(\boldsymbol{X}\setminus F_{e_j})|<|E_0|/k$. But this is impossible by (11.8) and the classical pigeonhole principle. Thus, we conclude that $\prod_{j\in[k]}\mathbf{1}_{A_j}\big(a+\varphi_j(d)\big)=0$ for every $a\in Z'$ and every $d\in Z$. For part (b), fix $j\in[k]$. Since $Q_j\colon\boldsymbol{X}\to\mathbf{Z}'$ is an onto homomorphism, we have $|Q_j^{-1}(a)|=|\boldsymbol{X}|/|Z'|$ for every $a\in Z'$. Therefore,

$$\begin{aligned}\mathbb{E}[f\cdot\mathbf{1}_{Z'\setminus A_j}] &= \frac{1}{|\boldsymbol{X}|}\sum_{a\in Z'\setminus A_j}|Q_j^{-1}(a)|\cdot f(a)\\ &\overset{(11.9)}{\leqslant} \frac{1}{|\boldsymbol{X}|}\sum_{a\in Z'}k\,|Q_j^{-1}(a)\cap(\boldsymbol{X}\setminus F_{e_j})|\cdot f(a)\\ &\overset{(11.6)}{=} k\,\mathbb{E}[f_{e_j}\cdot\mathbf{1}_{\boldsymbol{X}\setminus F_{e_j}}]\overset{(11.8)}{\leqslant}\frac{\delta}{2k}.\end{aligned}$$

This completes the proof of Claim 11.2, and as we have already indicated, the proof of Theorem 11.1 is also completed. □

CHAPTER 12

"Pseudorandom" functions in the primes

In this chapter we introduce the appropriate arithmetic setting in order to apply Theorem 11.1. So, we first define a function in $\mathbf{P}^d$ that is majorized by a function that obeys certain pseudorandomness conditions. In order to do so we use the W-trick-see e.g. [Tao06a, GT10, CFZ14, FZ15, TZ15a]- which originates from the work of B. Green [Gre05]. The W-trick is very useful since it states that if we want to find certain "structures" in $\mathbf{P}^d$ by the Dirichlet theorem we may restrict our selves to primes that belong to an arithmetic progression.

In the second section we discuss the form the majorizing function should have and in the last two sections we define this majorant and prove that it obeys certain pseudorandomness conditions.

Before we begin we need to fix some notation and prove some preliminary results. At first, we define the functions $w, W\colon \mathbb{N} \to [0,\infty)$ by the rule

$$w(n) = \log^{(4)}(n) \text{ and } W(n) = \prod_{p\in\mathbf{P}\colon\, p\leqslant w(n)} p, \tag{12.1}$$

for every $n \in \mathbb{N}$. For these functions we will need the following lemma.

LEMMA 12.1. *Let N be a large positive integer*[1]*, $w = w(N)$ and $W = W(N)$. Then,*

$$W \leqslant \sqrt{\log N}. \tag{12.2}$$

PROOF. The prime number theorem (see Appendix B, Theorem B.1) suggests that

$$\log W = \sum_{p\leqslant w} \log p = \big(1 + o_{N\to\infty}(1)\big)w = O(w)$$

and thus $W = e^{O(w)}$. This implies of course (12.2). $\square$

Now, for every $\gamma > 0$ we define the function $R_\gamma\colon \mathbb{N} \to [1,\infty)$ by the rule

$$R_\gamma(n) = n^{\gamma/2}. \tag{12.3}$$

[1]From now on when we say that N is a large positive integer we practically see this N as tending to ∞, i.e. is sufficiently large for the purpose at hand.

Finally, from now on ϕ will denote the Euler totient function, μ will denote the Möbius function and $\widetilde{\Lambda}$ will denote the restriction of the Von Mangoldt function in the primes, i.e. $\widetilde{\Lambda}(n) = \mathbf{1}_{\mathbf{P}}(n)\log n$, for every $n \in \mathbb{Z}$. For more details about these functions see Appendix B.

12.1. The W-trick

We begin with the one dimensional case. Let N be large positive integer and $w = w(N), W = W(N)$ be as in (12.1). Then, for every $b \in \{0, \dots, W-1\}$ such that $\gcd(b, W) = 1$ the modified Von Mangoldt function $\widetilde{\Lambda}_{b,W} \colon \mathbb{Z} \to [0, \infty)$ is defined by the rule

$$\widetilde{\Lambda}_{b,W}(n) = \begin{cases} \frac{\phi(W)}{W}\log(Wn+b), & \text{when } Wn+b \in \mathbf{P} \\ \quad 0 & , \text{ otherwise,} \end{cases} \tag{12.4}$$

for every $n \in \mathbb{Z}$. A very important fact about this function is the following[2].

Proposition 12.2. *For every large positive integer N and for every $b \in \{0, \dots, W(N)-1\}$ such that $\gcd\big(b, W(N)\big) = 1$ we have that*

$$\sum_{n \in [N]} \widetilde{\Lambda}_{b,W(N)}(n) = \big(1 + o_{N\to\infty}(1)\big)N. \tag{12.5}$$

Proof. Let N be a large positive integer, $w = w(N), W = W(N)$ and $b \in \{0, \dots, W-1\}$ with $\gcd(b, W) = 1$. The main ingredient for the proof is the Siegel-Walfisz theorem (Theorem B.8 in Appendix B). By (12.2) we see that this theorem may be applied and thus

$$\frac{\phi(W)}{W} \sum_{\substack{n \in [WN+b] \\ n \equiv b \bmod W}} \widetilde{\Lambda}(n) = \big(1 + o_{N\to\infty}(1)\big)N.$$

But then

$$\begin{aligned} \sum_{n \in [N]} \widetilde{\Lambda}_{b,W}(n) &= \sum_{n \in [N]} \frac{\phi(W)}{W}\widetilde{\Lambda}(Wn+b) = \frac{\phi(W)}{W} \sum_{\substack{n \in [WN+b] \\ n \equiv b \bmod W}} \widetilde{\Lambda}(n) \\ &= \big(1 + o_{N\to\infty}(1)\big)N. \end{aligned}$$

and the proof is completed. □

We extend now the previous function to higher dimensions. To this end, let N be a large positive integer, $w = w(N), W = W(N)$ be as in (12.1) and d be a positive integer. Then, for every $\mathbf{b} = (b_1, \dots, b_d) \in \mathbb{Z}^d$ such that $b_i \in \{0, \dots, W-1\}$

[2] Using the terminology of [GT08, GT10] the following proposition states that the function defined on (12.4) constitutes a "measure" on the set $\mathbb{Z}_N$.

and $\gcd(b_i, W) = 1$ for every $i \in [d]$ we define the multidimensional modified Von Mangoldt function $\widetilde{\Lambda}_{\mathbf{b},W,d} \colon \mathbb{Z}^d \to [0,\infty)$ by the rule

$$\widetilde{\Lambda}_{\mathbf{b},W,d}(\mathbf{n}) = \widetilde{\Lambda}_{b_1,W}(n_1)\dots\widetilde{\Lambda}_{b_d,W}(n_d), \tag{12.6}$$

for every $\mathbf{n} = (n_1,\dots,n_d) \in \mathbb{Z}^d$. As a straightforward consequence of Proposition 12.2 we have the following similar result for the multidimensional modified Von Mangoldt function.

PROPOSITION 12.3. *For every large positive integer N and $\mathbf{b} = (b_1,\dots,b_d) \in \mathbb{Z}^d$ such that $b_i \in \{0,\dots,W(N)-1\}$ and $\gcd\big(b_i, W(N)\big) = 1$ for every $i \in [d]$ we have*

$$\sum_{\mathbf{n}\in[N]^d} \widetilde{\Lambda}_{\mathbf{b},W(N),d}(\mathbf{n}) = (1 + o_{N\to\infty}(1))N^d. \tag{12.7}$$

12.2. Truncated divisor sums

We define now a function that as we will see later on gives rise to a pointwise “majorant” of the modified Von Magoldt function with the additional property that this “majorant” has “good” pseudorandom properties. In [GT08], motivated by [GY03, GY], this function was defined as

$$\Lambda_R(n) = \sum_{\substack{d|n\\ d\leqslant R}} \mu(d)\log(R/d),$$

for $R > 0$.[3] In the works that followed (see e.g [GT10]), the previous function was modified to take eventually the following form. Let $\chi\colon \mathbb{R} \to [0,1]$ be a smooth and compactly supported function, a be a positive integer and $R > 0$. Then, we define the function $\Lambda_{\chi,R,a} \colon \mathbb{Z} \to [0,\infty)$ by the rule

$$\Lambda_{\chi,R,a}(n) = \log R\Big(\sum_{d|n} \mu(d)\chi\big(\frac{\log d}{\log R}\big)\Big)^a, \tag{12.8}$$

for every $n \in \mathbb{Z}$.[4]

REMARK 1. Note that $\Lambda_R = \Lambda_{\chi,R,1}$, where $\chi(x) = \max(1-|x|, 0)$, although we have abused notation since χ is not smooth in this case.

Observe now that if χ is supported on $[-1,1]$, $n = p^k$ for some prime p and some k and $\gcd(n, \prod_{p\leqslant R} p) = 1$ then $\Lambda_{\chi,R,a}(n) = \chi(0)^a \log R$. Thus, $\Lambda_{\chi,R,a}$ may be seen as weights on the “almost” primes, although it also give weights to other numbers as well. When $a = 1$ we have the disadvantage that $\Lambda_{\chi,R,1}$ can be negative. Therefore in what follows we will take $a = 2$.

[3] For intuition about what reason could lead to this function see Proposition B.4.

[4] This function for $a = 2$ is closely related to the Λ^2 Selberg sieve (see [IK04, Chapter 6]).

We will need the following result about the functions that where defined (12.8). This result is known as the linear forms condition estimate, see [GT08, GT10, CFZ14] and is an immediate consequence of Theorem C.20 in Appendix C.

PROPOSITION 12.4. *Let D be a positive integer, $\chi\colon \mathbb{R} \to [0,1]$ be a smooth and supported on $[-1,1]$ function such that $\chi(0) = 1$ and $\int |\chi'(x)|^2 dx = 1$ and N be a large positive integer. Also, let $w = w(N), W = W(N)$ as in (12.1) and $\widetilde{N} = \lfloor N/W \rfloor$. Then, there exists a constant $\gamma = \gamma(D,\chi) > 0$ such that if $R = R_\gamma(\widetilde{N})$ is as in (12.3) the following holds. Let $1 \leqslant d, t \leqslant D$ and $\psi_1, \ldots, \psi_t \colon \mathbb{Z}^d \to \mathbb{Z}$ be non zero affine linear forms with no two of them be rational multiples of each other and with coefficients bounded by D. Also, let $B = \prod_{i\in[d]} I_i$ where for every $i \in [d]$, I_i is a set of $\widetilde{N}$ consecutive integers and $b_1, \ldots, b_t \in \{0, \ldots, W-1\}$ with $\gcd(b_i, W) = 1$ for every $i \in [t]$. Then,*

$$\mathbb{E}\big[\big(\frac{\phi(W)}{W}\big)^t \prod_{i\in[t]} \Lambda_{\chi,R,2}(W\psi_i(\mathbf{n}) + b_i) \,|\, \mathbf{n} \in B\big] = 1 + o_{D,N\to\infty}(1). \tag{12.9}$$

12.3. Construction of the majorants

From now on we fix a positive integer D, a large integer $N, w = w(N), W = W(N)$ and $\widetilde{N} = \lfloor N/W \rfloor$. Moreover, we fix a smooth and supported on $[-1,1]$ function $\chi\colon \mathbb{R} \to [0,1]$ with the additional properties that $\chi(0) = 1$ and $\int_{\mathbb{R}} |\chi'(x)|^2 dx = 1$. Finally, we fix the constant $\gamma = \gamma(D,\chi)$ that arises from Proposition 12.4 for the previous choice of D and χ and also fix $R = R_\gamma(\widetilde{N})$.

We will first construct a majorant for the one dimensional modified Von Mangoldt function and then we will do the same for higher dimensions.

For the one dimensional case we have the following. For every $0 < \varepsilon_1, \varepsilon_2 < 1$ with $\varepsilon_1 < \varepsilon_2$ and every $b \in \{0, \ldots, W-1\}$ with $\gcd(b, W) = 1$ we define the function

$$\nu_{\varepsilon_1,\varepsilon_2,b}\colon \mathbb{Z}_{\widetilde{N}} \to [0,\infty)$$

by the rule

$$\nu_{\varepsilon_1,\varepsilon_2,b}(n) = \begin{cases} \frac{\phi(W)}{W}\Lambda_{\chi,R,2}(Wn+b), & \text{when } n \in [\varepsilon_1\widetilde{N}, \varepsilon_2\widetilde{N}] \\ 1 & , \text{ otherwise,} \end{cases} \tag{12.10}$$

for every $n \in \mathbb{Z}_{\widetilde{N}}$. Let's show first that the previous function bounds pointwise $\widetilde{\Lambda}_{b,W}$.

PROPOSITION 12.5. *Let $0 < \varepsilon_1, \varepsilon_2 < 1$ with $\varepsilon_1 < \varepsilon_2$, $b \in \{0, \ldots, W-1\}$ such that $\gcd(b, W) = 1$ and $\widetilde{\Lambda}_{b,W}$ be as in (12.4). Then, there exists $\delta_\gamma > 0$ such that*

$$\delta_\gamma \cdot \widetilde{\Lambda}_{b,W}(n) \leqslant \nu_{\varepsilon_1,\varepsilon_2,b}(n),$$

for every positive integer $n \in [\varepsilon_1\widetilde{N}, \varepsilon_2\widetilde{N}]$.

Proof. Let $\delta_\gamma = \gamma/6$ and $n \in [\varepsilon_1 \widetilde{N}, \varepsilon_2 \widetilde{N}]$. It suffices to consider the case $Wn+b$ is prime since otherwise $\widetilde{\Lambda}_{b,W}(n) = 0$. Then, by the definition of $\widetilde{N}$, Lemma 12.1 and the fact that $\sqrt{\log N} \leqslant \widetilde{N}$ for large N we have

$$Wn + b \leqslant \sqrt{\log N}\widetilde{N} + \sqrt{\log N} \leqslant \widetilde{N}^3.$$

and hence by the definition of R

$$\delta_\gamma \log(Wn + b) \leqslant \frac{\gamma}{6} \log \widetilde{N}^3 \leqslant \log \widetilde{N}^{\gamma/2} = \log R.$$

On the other hand, by the discussion after (12.8) and since $\chi(0) = 1$ we have that $\Lambda_{\chi,R,2}(Wn + b) = \log R$, when $Wn + b$ is prime. Thus,

$$\delta_\gamma \widetilde{\Lambda}_{b,W}(n) = \delta_\gamma \frac{\phi(W)}{W} \log(Wn + b) \leqslant \frac{\phi(W)}{W} \log R = \nu_{\varepsilon_1,\varepsilon_2,b}(n)$$

and the proof is completed. □

We proceed to the higher dimensions. For every $d \leqslant D$, every $0 < \varepsilon_1, \varepsilon_2 < 1$ with $\varepsilon_1 < \varepsilon_2$ and every $\mathbf{b} = (b_1, \ldots, b_d) \in \mathbb{Z}^d$ with $b_i \in \{0, \ldots, W-1\}$ and $\gcd(b_i, W) = 1$ we define the function $\nu_{\varepsilon_1,\varepsilon_2,\mathbf{b},d} \colon \mathbb{Z}^d_{\widetilde{N}} \to [0, \infty)$ by the rule

$$\nu_{\varepsilon_1,\varepsilon_2,\mathbf{b},d}(\mathbf{n}) = \nu_{\varepsilon_1,\varepsilon_2,b_1}(n_1) \ldots \nu_{\varepsilon_1,\varepsilon_2,b_d}(n_d) \tag{12.11}$$

for every $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{Z}^d_{\widetilde{N}}$. Using Proposition 12.5 we see that the following proposition holds.

Proposition 12.6. *For every $d, \varepsilon_1, \varepsilon_2$ and $\mathbf{b}$ as before the following holds. If δ_γ is as in Proposition 12.5 and $\delta_{\gamma,d} = \delta_\gamma^d > 0$ we have that*

$$\delta_{\gamma,d} \cdot \widetilde{\Lambda}_{\mathbf{b},W,d}(\mathbf{n}) \leqslant \nu_{\varepsilon_1,\varepsilon_2,\mathbf{b},d}(\mathbf{n}),$$

for every $\mathbf{n} \in [\varepsilon_1 \widetilde{N}, \varepsilon_2 \widetilde{N}]^d$.

Remark 2. The quantities $\varepsilon_1, \varepsilon_2$ will be chosen in the proof of the multidimensional Green–Tao theorem in order to extend constellations of $\mathbb{Z}^d_{\widetilde{N}}$ that arise from the use of Theorem 11.1 to genuine constellations of $\mathbb{Z}^d$.

12.4. Pseudorandomness conditions for the majorants.

Our task now is to show that $\nu_{\varepsilon_1,\varepsilon_2,\mathbf{b},d}$ obeys a certain pseudorandomness condition. We will prove this result for the one dimensional case and then as a consequence we will have a similar result for higher dimensions also. More precisely, we have

Proposition 12.7. *Let $0 < \varepsilon_1, \varepsilon_2 < 1$ with $\varepsilon_1 < \varepsilon_2$, $b \in \{0, \ldots, W-1\}$ with* $\gcd(b, W) = 1$ *and $1 \leqslant d, t \leqslant D$ be positive integers. Also let $\psi_1, \ldots, \psi_t \colon \mathbb{Z}^d \to \mathbb{Z}$ be* *non constant affine-linear forms where no two of them are rational multiples of one another and where their coefficients are bounded by D. Then,*

$$\mathbb{E}_{\mathbf{n}\in\mathbb{Z}^d_{\widetilde{N}}}\Big[\prod_{i\in[t]}\nu_{\varepsilon_1,\varepsilon_2,b}\big(\psi_i(\mathbf{n})\big)\Big] = 1 + o_{D,N\to\infty}(1).$$

In the last expression[5] *we induce the affine linear forms $\psi_j\colon \mathbb{Z}^d_{\widetilde{N}} \to \mathbb{Z}_{\widetilde{N}}$ from their global counterparts $\psi_j\colon \mathbb{Z}^d \to \mathbb{Z}$ in the obvious manner.*

Proof. The idea is to split $\mathbb{Z}^d_{\widetilde{N}}$ into smaller boxes and then apply Proposition 12.4. For notational simplicity we set $\nu_{\varepsilon_1,\varepsilon_2,b} = \nu$. So, let $Q = Q(N)$ be the largest prime that is lower or equal to $\widetilde{N}^{1/2}$ and observe that by the Bertrand-Chebysev theorem (Theorem B.3) we have that $Q \geqslant \widetilde{N}^{1/2}/2$. Then,

$$\widetilde{N}^{1/2} \leqslant \widetilde{N}/Q \leqslant 2\widetilde{N}^{1/2}. \tag{12.12}$$

Consider now the boxes

$$B_{u_1,\dots,u_d} \coloneqq \Big\{\mathbf{n}\in\mathbb{Z}^d_{\widetilde{N}}\colon n_j \in \big[\lfloor u_j\frac{\widetilde{N}}{Q}\rfloor, \lfloor (u_j+1)\frac{\widetilde{N}}{Q}\rfloor\big)\Big\},$$

for every $(u_1,\dots,u_d)\in\mathbb{Z}^d_Q$. Then, we have the following result.

Claim 12.8. *The following holds true.*

$$\mathbb{E}_{\mathbf{n}\in\mathbb{Z}^d_{\widetilde{N}}}\Big[\prod_{i\in[t]}\nu\big(\psi_i(\mathbf{n})\big)\Big] = (1+o_d(1))\mathbb{E}_{u_1,\dots,u_d\in\mathbb{Z}_Q}\Big[\mathbb{E}_{\mathbf{n}\in B_{u_1,\dots,u_t}}\big[\prod_{i\in[t]}\nu\big(\psi_i(\mathbf{n})\big)\big]\Big].$$

Proof of the claim. First observe that for every $u_1,\dots,u_d\in\mathbb{Z}_Q$ we have

$$\big(\frac{\widetilde{N}}{Q}-1\big)^d \leqslant |B_{u_1,\dots,u_d}| \leqslant \big(\frac{\widetilde{N}}{Q}+1\big)^d$$

and since $Q\leqslant \widetilde{N}^{1/2}$ we have,

$$|B_{u_1,\dots,u_d}| = \big(1+o_d(1)\big)\big(\frac{\widetilde{N}}{Q}\big)^d.$$

Hence,

$$\begin{aligned}\mathbb{E}_{\mathbf{n}\in B_{u_1,\dots,u_d}}\big[\prod_{i\in[t]}\nu\big(\psi_i(\mathbf{n})\big)\big] &= \frac{1}{Q^d}\sum_{u_1,\dots,u_d\in\mathbb{Z}_Q}\;\sum_{\mathbf{n}\in B_{u_1,\dots,u_d}}\frac{\prod_{i\in[t]}\nu(\psi_i(\mathbf{n}))}{|B_{u_1,\dots,u_d}|}\\ &= \big(1+o_d(1)\big)\mathbb{E}_{\mathbf{n}\in\mathbb{Z}^d_{\widetilde{N}}}\Big[\prod_{i\in[t]}\nu\big(\psi_i(\mathbf{n})\big)\Big]\end{aligned}$$

and the proof is completed. □

[5]This expression is usually referred to as the linear forms condition for functions in $\mathbb{Z}_{\widetilde{N}}$, see [GT10]

In order to apply Proposition 12.4 we do the following dichotomy. We call a box *good* if for every $i \in [t]$ the set $\{\psi_i(\mathbf{n})\colon \mathbf{n} \in B_{u_1,\dots,u_d}\}$ either lies in the subset $[\varepsilon_1 \widetilde{N}, \varepsilon_2 \widetilde{N}]$ of $\mathbb{Z}_{\widetilde{N}}$ or it is completely outside of this subset. If a box is not good we call it *bad*. For the good boxes using (12.12) and since N is sufficiently large we may apply Proposition 12.4 and obtain that

$$\mathbb{E}_{\mathbf{n}\in B_{u_1,\dots,u_d}}\Big[\prod_{i\in[t]} \nu\big(\psi_i(\mathbf{n})\big)\Big] = 1 + o_{D,N\to\infty}(1). \tag{12.13}$$

For the bad boxes we take the trivial bound

$$\nu(n) \leqslant 1 + \frac{\phi(W)}{W}\Lambda_{\chi,R,2}(Wn+1)$$

which by expansion and the use of Proposition 12.4 yields that

$$\mathbb{E}_{\mathbf{n}\in B_{u_1,\dots,u_d}}\Big[\prod_{i\in[t]} \nu\big(\psi_i(\mathbf{n})\big)\Big] \leqslant \big(2^t + o_{D,N\to\infty}(1)\big)$$

and thus

$$\mathbb{E}_{\mathbf{n}\in B_{u_1,\dots,u_d}}\Big[\prod_{i\in[t]} \nu\big(\psi_i(\mathbf{n})\big)\Big] = O_D(1). \tag{12.14}$$

Therefore it suffices to show that the number of bad boxes is at most $O_D(Q^{d-1})$. Indeed, assuming the previous bound we have

$$\mathbb{E}_{\mathbf{n}\in\mathbb{Z}^d_{\widetilde{N}}}\Big[\prod_{i\in[t]} \nu\big(\psi_i(\mathbf{n})\big)\Big] = \big(1+o_d(1)\big)\frac{1}{Q^d}\sum_{u_1,\dots,u_d\in\mathbb{Z}^d_Q} \mathbb{E}_{B_{u_1,\dots,u_t}}\Big[\prod_{i\in[t]} \nu\big(\psi_i(\mathbf{n})\big)\Big]$$

$$= \big(1+o_D(1)\big)\frac{1}{Q^d}\Big(\big(Q^d - O_D(Q^{d-1})\big)\big(1+o_D(1)\big) + O_D(Q^{d-1})\Big) = 1 + o_{D,N\to\infty}(1),$$

since Q increases with N. It remains to show the bound about the number of bad boxes. Before we do so we need the following result.

CLAIM 12.9. *Assume that for some $u_1,\dots,u_d \in \mathbb{Z}_Q$ and for some $i \in [t]$ there exists $\mathbf{n} \in B_{u_1,\dots,u_d}$ and $\ell \in \mathbb{Z}$ such that*

$$\varepsilon_1 \widetilde{N} \leqslant \psi_i(\mathbf{n}) + \ell\widetilde{N} \leqslant \varepsilon_2 \widetilde{N}.$$

. Then, for every $\mathbf{n}' \in B_{u_1,\dots,u_d}$ we have that

$$1 \leqslant \psi_i(\mathbf{n}') + \ell\widetilde{N} \leqslant \widetilde{N}.$$

PROOF OF THE CLAIM. Since ψ_i is an affine linear form there exist $L_{i,1},\dots,L_{i,d}, c_i \in \mathbb{Z}$ such that

$$\psi_i(\mathbf{x}) = \sum_{j\in[d]} L_{i,j}x_j + c_i,$$

for every $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{Z}^d$. Then if $\mathbf{n} = (n_1, \ldots, n_d)$ we have that

$$(\varepsilon_1 - \ell)\widetilde{N} - c_i \leqslant \sum_{j\in[d]} L_{i,j} n_j \leqslant (\varepsilon_2 - \ell)\widetilde{N} - c_i. \tag{12.15}$$

Let now $\mathbf{n}' = (n'_1, \ldots, n'_d) \in B_{u_1,\ldots,u_d}$ and observe that $|n_j - n'_j| \leqslant \widetilde{N}/Q$ for every $j \in [d]$. Then we have that

$$\begin{aligned}\sum_{j\in[d]} L_{i,j} n'_j = \sum_{j\in[d]} L_{i,j}(n'_j - n_j) + \sum_{j\in[d]} L_{i,j} n_j &\overset{(12.15)}{\leqslant} (\varepsilon_2 - \ell)\widetilde{N} - c_i + 2Dt\widetilde{N}^{1/2} \\ &\leqslant (\varepsilon_2 - \ell)\widetilde{N} - c_i + 2D^2\widetilde{N}^{1/2} \leqslant (1-\ell)\widetilde{N} - c_i,\end{aligned}$$

since D, ε_2 are fixed and N is large enough. Working similarly we also obtain that

$$\sum_{j\in[d]} L_{i,j} n'_j \geqslant -\ell\widetilde{N} - c_i$$

and thus we have proved the desired result. □

We are ready now to bound the number of bad boxes.

Claim 12.10. *The number of bad boxes is bounded by $O_D(Q^{d-1})$.*

Proof of the claim. Assume that for every $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{Z}^d$

$$\psi_i(\mathbf{n}) = \sum_{j\in[d]} L_{i,j} x_j + c_i,$$

for some $L_{i,1}, \ldots, L_{i,d}, c_i \in \mathbb{Z}$. Also assume that $B_{u_1,\ldots,u_d}$ is bad. Then, by the definition of bad boxes there exist $i \in [t]$ and $\mathbf{n}, \mathbf{n}' \in B_{u_1,\ldots,u_d}$ such that $\psi_i(\mathbf{n})$ lies in $[\varepsilon_1\widetilde{N}, \varepsilon_2\widetilde{N}]$ while $\psi_i(\mathbf{n}')$ does not. Then, by the Claim 12.9 we may find integer ℓ such that either

$$1 \leqslant \psi_i(\mathbf{n}') + \ell\widetilde{N} < \varepsilon_1\widetilde{N} \leqslant \psi_i(\mathbf{n}) + \ell\widetilde{N} \leqslant \varepsilon_2\widetilde{N} \tag{12.16}$$

either

$$\varepsilon_1\widetilde{N} \leqslant \psi_i(\mathbf{n}) + \ell\widetilde{N} \leqslant \varepsilon_2\widetilde{N} < \psi_i(\mathbf{n}') + \ell\widetilde{N} \leqslant \widetilde{N}. \tag{12.17}$$

But from the definition of $B_{u_1,\ldots,u_d}$ and since $L_{i,j}$s, c_i are bounded by D we also have

$$\psi_i(\mathbf{n}), \psi_i(\mathbf{n}') = \sum_{j\in[d]} L_{i,j} \lfloor u_j \frac{\widetilde{N}}{Q} \rfloor + c_i + O_D(\frac{\widetilde{N}}{Q})$$

which together with (12.16) and (12.17) yields that either

$$\varepsilon_1\widetilde{N} = \sum_{j\in[d]} L_{i,j} \lfloor u_j \frac{\widetilde{N}}{Q} \rfloor + c_i + \ell\widetilde{N} + O_D(\frac{\widetilde{N}}{Q})$$

either

$$\varepsilon_2\widetilde{N} = \sum_{j\in[d]} L_{i,j}\lfloor u_j\frac{\widetilde{N}}{Q}\rfloor + c_i + \ell\widetilde{N} + O_D(\frac{\widetilde{N}}{Q}).$$

Since now $\lfloor u_j\frac{\widetilde{N}}{Q}\rfloor = u_j\frac{\widetilde{N}}{Q} + O(1)$ we have

$$\sum_{j\in[d]} L_{i,j}u_j = \Big(\varepsilon_1 Q - c_i\frac{Q}{\widetilde{N}} + O_D(1)\Big)\mathrm{mod}Q$$

or

$$\sum_{j\in[d]} L_{i,j}u_j = \Big(\varepsilon_2 Q - c_i\frac{Q}{\widetilde{N}} + O_D(1)\Big)\mathrm{mod}Q.$$

Since $(L_{i,j})_{j\in[d]}$ is non-zero, the number of d-tuples $u_1,\dots,u_d$ which satisfy these equations is $O_D(Q^{d-1})$, which happens because we have $d-1$ degrees of freedom in the choice of u_j's. Therefore, letting i vary and taking into account that the previous should hold for ε_1 or ε_2 we have that the number of bad boxes is bounded by

$$2DO_D(Q^{d-1}) = O_D(Q^{d-1})$$

which completes the proof of the claim. □

With the completion of the proof the previous claim we also have completed the proof of Proposition 12.7. □

As an immediate consequence we have the following proposition for the function $\nu_{\varepsilon_1,\varepsilon_2,\mathbf{b},d}$.

Proposition 12.11. *Let $0<\varepsilon_1,\varepsilon_2<1$ with $\varepsilon_1<\varepsilon_2$ and d,t be positive integers with $1\leqslant dt\leqslant D$. Also, for every $i\in[t]$ and every $j\in[d]$ let $\psi_{ij}\colon\mathbb{Z}^d\to\mathbb{Z}$ be non constant affine-linear forms where no two of them are rational multiples of one another and where their coefficients are bounded by D. Then,*

$$\mathbb{E}_{\mathbf{n}\in\mathbb{Z}^d_{\widetilde{N}}}\Big[\prod_{i\in[t]}\nu_{\varepsilon_1,\varepsilon_2,\mathbf{b},d}\big(\psi_{i1}(\mathbf{n}),\dots,\psi_{id}(\mathbf{n})\big)\Big] = 1 + o_{D,N\to\infty}(1),$$

where as in Proposition 12.7 we induce the affine linear forms $\psi_j\colon\mathbb{Z}^d_{\widetilde{N}}\to\mathbb{Z}_{\widetilde{N}}$ from their global counterparts $\psi_j\colon\mathbb{Z}^d\to\mathbb{Z}$ in the obvious manner.

CHAPTER 13

A multidimensional Green–Tao theorem

In this section we prove a special case of the multidimensional Green–Tao theorem. More specifically we will show that every "large" subset of $\mathbf{P}_N^d$, where N is large, contains at least one constellation of every finite set of $\mathbb{Z}^d$ that is in *general position*. This result was proved by B. Cook and Á. Magyar in [CM12]. For the general case the arguments that we use here don't work and in order to give a complete proof one needs to take a completely different approach passing through some deep results, [GT10, GTZ12, GT12, FZ15].

13.1. Shapes in $\mathbb{Z}^d$

This section contains definitions about special types of shapes in $\mathbb{Z}^d$ and a technical lemma concerning one of these types. First recall that a *shape* in $\mathbb{Z}^d$ is just a finite set of vectors $u_1, \dots, u_k \in \mathbb{Z}^d$ and a constellation of this shape is called every homeothetic copy of it, i.e. a constellation of $u_1, \dots, u_k \in \mathbb{Z}^d$ is a set of the form $x + tu_1, \dots, x + tu_k \in \mathbb{Z}^d$, for some $x \in \mathbb{Z}^d$ and $t \in \mathbb{Z} \setminus \{0\}$.

Definition 13.1 (General Position). *A shape* $\{u_1, \dots, u_k\} \subseteq \mathbb{Z}^d \setminus \{\mathbf{0}\}$ *is in* general position *if for every* $i, j \in [k]$ *with* $i \neq j$ *and every* $l \in [d]$ *we have that* $u_{i,l} \neq u_{j,l}$*, where* $u_{i,l}$ *and* $u_{j,l}$ *are the* l*th coordinates of* u_i *and* u_j *respectively.*

For example, the shape of $\mathbb{Z}^2$, $\{(1, 2), (2, 1)\}$ is in general position while the shape of $\mathbb{Z}^3$, $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ is not.
Observe now that a shape $\{u_1, \dots, u_k\} \subseteq \mathbb{Z}^d$ may be seen as a vector $\mathbf{u} = (u_1, \dots, u_k) \in \mathbb{Z}^{dk}$. Having this in mind we have the following definition.

Definition 13.2 (Primitive shapes). *A shape* $\mathbf{u} = (u_1, \dots, u_k) \in \mathbb{Z}^{dk}$ *is called primitive if*

$$\{\mathbf{x} \in \mathbb{Z}^{dk} \colon \mathbf{x} = \lambda\mathbf{u}, \text{ for some } 0 < \lambda < 1\} = \emptyset,$$

i.e. the line segment $[\mathbf{0}, \mathbf{u}]$ *does not contain other point of* $\mathbb{Z}^{dk}$ *other than* $\mathbf{0}$ *and* $\mathbf{u}$.

Remark 3. Observe that if $\mathbf{u} \in \mathbb{Z}^{dk}$ is primitive then

$$\inf_{\substack{\mathbf{m} \in \mathbb{Z}^{d^2} \setminus \{\mathbf{0}, \mathbf{u}\}, \\ \mathbf{y} \in [\mathbf{0}, \mathbf{u}]}} \|\mathbf{m} - \mathbf{y}\|_\infty = 1.$$

From now on, for every $\mathbf{x} \in \mathbb{Z}^{dk}$ *for some* $d, k \geqslant 1$, *by* $[\mathbf{0}, \mathbf{x}]$ *we will denote the line segment with boundary points* $\mathbf{0}$ *and* $\mathbf{x}$.

We proceed now to a lemma concerning affine linear forms defined on shapes that are in general position.[1].

LEMMA 13.3. *Let* d *be a positive integer and* $\{u_1, \dots, u_d\} \subseteq \mathbb{Z}^d$ *be a shape in general position such that* $u_i s$ *are linearly independent for every* $i \in [d]$. *Also, for every* $j, l \in [d]$ *let* $\Psi_{j,l} \colon \mathbb{Z}^d \to \mathbb{Z}$ *be the function defined by the rule*

$$\Psi_{j,l}(\mathbf{x}) = \sum_{i \neq j} x_i u_{i,l} - \Big(\sum_{i \neq j} x_i\Big) u_{j,l},$$

where $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{Z}^d$ *and* $u_{i,l}$ *is the* l*th coordinate of* u_i *for every* $i, l \in [d]$. *Then, no two of the functions* $\Psi_{j,l}$ *are rational multiples of each other.*

PROOF. Let $j, j', l, l' \in [d]$. We distinguish the following two cases. The first case is when $j = j'$. Assume on the contrary that there exists some rational λ such that

$$\Psi_{j,l}(\mathbf{x}) = \lambda \Psi_{j,l'}(\mathbf{x})$$

for all $\mathbf{x} \in \mathbb{Z}^d$. Then, by the definition of $\Psi_{i,j}$s this λ would satisfy the following equation

$$\lambda = \frac{\sum_{i \neq j} x_i (u_{i,l'} - u_{j,l'})}{\sum_{i \neq j} x_i (u_{i,l} - u_{j,l})},$$

for all $\mathbf{x}$. But by comparing the coefficients of the x_is this would imply one of the following two cases in turn

- either for every $i, i' \neq j$, with $i \neq i'$ we have that $u_{i,l} - u_{j,l} = u_{i',l} - u_{j,l}$ and $u_{i,l'} - u_{j,l'} = u_{i',l'} - u_{j,l'}$
- either there exists some b such that for every $i \neq j$, $u_{i,l'} - u_{j,l'} = b(u_{i,l} - u_{j,l})$

But the first case contradicts the fact that $u_1, \dots, u_d$ are in general position while the second case contradicts the fact that $u_1, \dots, u_d$ are linearly independent. Therefore the case $j = j'$ is proved. For the case $j \neq j'$ we work similarly [2]. Thus, the proof of the lemma is completed. □

13.2. A special case of the multidimensional Green–Tao Theorem

Our interest in this section is to prove the multidimensional Green–Tao Theorem for shapes in general position. Before we proceed to the precise statement and proof of this theorem we have the following preparatory lemma.

[1]This lemma along with Proposition 12.11 and Theorem 11.1 plays an important role in the proof of the special case of the multidimensional Green–Tao Theorem.

[2]In fact, in this case only the linear independency of the u_is is needed.

LEMMA 13.4. *Let $\delta > 0$ and d, k be positive integers with $k \geqslant 3$. Also let N be a large positive integer, $w = w(N), W = W(N)$ be as in (12.1) and $\widetilde{N} = \lfloor N/W \rfloor$. Then, for every $A \subseteq \mathbf{P}_N^d$ with $|A| \geqslant \delta |\mathbf{P}_N|^d$ there exists $\mathbf{b} = (b_1, \ldots, b_d) \in \mathbb{Z}^d$ with $b_i \in \{0, \ldots, W-1\}$ and $\gcd(b_i, W) = 1$, for every $i \in [d]$ such that*

$$\sum_{\mathbf{n} \in [\widetilde{N}]^d} \mathbf{1}_{A,\mathbf{b},W}(\mathbf{n}) \widetilde{\Lambda}_{\mathbf{b},W,d}(\mathbf{n}) \geqslant \frac{\delta}{2^{d+1}} \widetilde{N}^d, \tag{13.1}$$

where for every $\mathbf{b}$

$$\mathbf{1}_{A,\mathbf{b},W}(\mathbf{n}) = \mathbf{1}_A(Wn_1 + b_1, \ldots, Wn_d + b_d),$$

for every $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{Z}^d$.

PROOF. We begin with the following result.

CLAIM 13.5. *The following expression holds true*

$$|A \cap [\sqrt{N}, N]^d| \geqslant \frac{\delta}{2} |\mathbf{P}_N^d|. \tag{13.2}$$

PROOF OF THE CLAIM. Recall that for every positive integer n, $\pi(n) = |\mathbf{P}_n|$. Assume now that (13.2) does not hold true. Then we should have that

$$\frac{\delta}{2}\pi(N)^d = \frac{\delta}{2}|\mathbf{P}_N|^d \leqslant |A \cap \big([1,N] \setminus [\sqrt{N}, N]\big)^d| \leqslant \pi(N)^d - \big(\pi(N) + \pi(\sqrt{N})\big)^d. \tag{13.3}$$

By the binomial theorem and the prime number theorem we have that

$$\begin{aligned} \pi(N)^d - \big(\pi(N) + \pi(\sqrt{N})\big)^d &= \sum_{k=0}^{d-1} \binom{d}{k} \pi(N)^k \pi(\sqrt{N})^{d-k} \\ &= \big(1 + o_{N\to\infty}(1)\big) \frac{1}{\log^d N} \sum_{k=0}^{d-1} \binom{d}{k} 2^{d-k} N^{\frac{d+k}{2}} \\ &\leqslant \frac{1}{\log^d N} 2^{d+1}(d-1)dN^{d-\frac{1}{2}}. \end{aligned}$$

But then (13.3) and the prime number theorem would imply that

$$N^d \leqslant \frac{2}{\delta} 2^{d+1}(d-1)dN^{d-\frac{1}{2}},$$

which is clearly a contradiction since N is sufficiently large. Therefore we have completed the proof of Claim 13.5. □

Using now the previous claim and the prime number theorem once again we have that

$$\sum_{\mathbf{n}=(n_1,\ldots,n_d) \in A \cap [\sqrt{N}, N]^d} \widetilde{\Lambda}(n_1) \ldots \widetilde{\Lambda}(n_d) \geqslant \frac{\delta}{2} \pi(N)^d \log(\sqrt{N})^d \geqslant \frac{\delta}{2^{d+1}} N^d. \tag{13.4}$$

Set

$$\mathrm{Co}(W)=\{(b_1,\dots,b_d)\in\{0,\dots,W-1\}^d\colon\ \gcd(b_i,W)=1,\ \text{for every } i\in[d]\}\quad (13.5)$$

and observe that there exists $\mathbf{b}=(b_1,\dots,b_d)\in\mathrm{Co}(W)$ such that

$$\begin{aligned}
&\Big(\frac{\phi(W)}{W}\Big)^d \sum_{\mathbf{n}=(n_1,\dots,n_d)\in[\widetilde{N}]^d}\mathbf{1}_{A,\mathbf{b},W}(\mathbf{n})\prod_{i\in[d]}\log(Wn_i+b_i)\\
&=\Big(\frac{\phi(W)}{W}\Big)^d \max_{\mathbf{b}'=(b_1',\dots,b_d')\in\mathrm{Co}(W)}\ \sum_{\mathbf{n}=(n_1,\dots,n_d)\in[\widetilde{N}]^d}\mathbf{1}_{A,\mathbf{b}',W}(\mathbf{n})\prod_{i\in[d]}\log(Wn_i+b_i') \qquad (13.6)\\
&\geqslant \frac{1}{W^d}\sum_{\mathbf{b}'=(b_1',\dots,b_d')\in\mathrm{Co}(W)}\ \sum_{\mathbf{n}=(n_1,\dots,n_d)\in[\widetilde{N}]^d}\mathbf{1}_{A,\mathbf{b}',W}(\mathbf{n})\prod_{i\in[d]}\log(Wn_i+b_i')
\end{aligned}$$

We will show that the previous $\mathbf{b}$ is the desired one. To this end, for this choice of $\mathbf{b}$ we have

$$\begin{aligned}
&\sum_{\mathbf{n}\in[\widetilde{N}]^d}\mathbf{1}_{A,\mathbf{b},W}(\mathbf{n})\widetilde{\Lambda}_{\mathbf{b},W,d}(\mathbf{n})\\
&\overset{(13.6)}{\geqslant}\frac{1}{W^d}\sum_{\mathbf{b}=(b_1,\dots,b_d)\in\{0,\dots,W-1\}^d}\ \sum_{\mathbf{n}=(n_1,\dots,n_d)\in[\widetilde{N}]^d}\mathbf{1}_{A,\mathbf{b},W}(\mathbf{n})\prod_{i\in[d]}\log(Wn_i+b_i)\\
&\geqslant\frac{1}{W^d}\sum_{\mathbf{n}=(n_1,\dots,n_d)\in A\cap[\sqrt{N},N]^d}\widetilde{\Lambda}(n_1)\dots\widetilde{\Lambda}(n_d)\overset{(13.4)}{\geqslant}\frac{\delta}{2^{d+1}}\widetilde{N}^d
\end{aligned}$$

which completes the proof of the Lemma. □

We are ready now to prove the main result of this chapter and of this part in general.

THEOREM 13.6. *Let d,k be positive integers with $k\geqslant 3$, $\mathbf{u}=(u_1,\dots,u_k)\in\mathbb{Z}^{kd}\setminus\{\mathbf{0}\}$ be a shape in general position and $\delta>0$. Also, let N be sufficiently large. Then for every $A\subseteq\mathbf{P}_N^d$ with $|A|\geqslant\delta|\mathbf{P}_N^d|$ there exist $x\in\mathbb{Z}^d$ and $t\in\mathbb{Z}\setminus\{0\}$ such that*

$$x+tu_1,\dots,x+tu_k\in A.$$

PROOF. Our aim is to form the proper setting in order to apply Theorem 11.1. So, let

$$D=\max\{kd2^{d(k-1)},\max_{i\in[k]}\|u_i\|_\infty\},$$

$w=w(N), W=W(N)$ be as in (12.1), $\widetilde{N}=\lfloor N/W\rfloor$, $Z=\mathbb{Z}_{\widetilde{N}}$, $Z'=\mathbb{Z}_{\widetilde{N}}^d$ and $\mathbf{b}=(b_1,\dots,b_d)\in\mathbb{Z}^d$ be as in Lemma 13.4. Moreover, let $\gamma=\gamma(D,\chi)$ be as in

Proposition 12.4 and $R = R_\gamma(\widetilde{N})$ be as in (12.3). Finally, let

$$\varepsilon_2 = \big(1 - \frac{\delta}{10^{d+1}d^2}\big),\ \varepsilon_1 = \frac{\delta}{10^{d+1}d^2}\varepsilon_2, \tag{13.7}$$

$\delta_{\gamma,d}$ be as in Proposition 12.6 and

$$N_0 = N(k, 1, \infty, \delta\delta_{\gamma,D}/2^{d+2}),\quad c_0 = c(k, 1, \infty, \delta\delta_{\gamma,D}/2^{d+2}) \tag{13.8}$$

be as in Theorem 11.1. We begin with the following claim.

Claim 13.7. *We may assume that $k = d$ and also that the vectors $u_1, \dots, u_k$ are linearly independent. We may also assume that the shape $\mathbf{u} = (u_1, \dots, u_k)$ is primitive.*

Proof of Claim 13.7. For the first part of the claim let $w_1, \dots, w_k \in \mathbb{Z}^k$ be independent vectors. For every $i \in [k]$ set $u'_i = (u_i, w_i) \in \mathbb{Z}^{d+k}$. Expand the linearly independent u'_i's to form a basis $u'_1, \dots, u'_k, u'_{k+1}, \dots, u'_{k+d}$ of $\mathbb{Z}^{d+k}$ and observe that this expansion can be done in order for the basis $\langle u'_i \colon i \in [k+d]\rangle$ of $\mathbb{Z}^{d+k}$ to be in general position. Set $A' := A \times \mathbf{P}^k$ and observe that if there exist $x_1 \in \mathbb{Z}^d$, $x_2 \in \mathbb{Z}^k$ and $t \in \mathbb{Z} \setminus \{0\}$ such that $(x_1, x_2) + tu'_i \in A'$ for every $i \in [k+d]$ then $x_1 + tu_i \in A$ for every $i \in [k]$.

For the second part of the claim observe that it suffices to show that there exists a primitive shape $\mathbf{u}' = (u'_1, \dots, u'_k) \in \mathbb{Z}^{dk}$ and a positive integer s such that $\mathbf{u} = s\mathbf{u}'$. To this end we assume that $\mathbf{u}$ is not primitive since otherwise we take $s = 1$ and $\mathbf{u}' = \mathbf{u}$. Then, there exist finite $\lambda \in (0, 1)$ such that $\lambda^{-1} | u_{i,j}$ for every $i, j \in [d]$, where as usual $u_{i,j}$ is the jth coordinate of u_i. Setting λ_0 to be the minimum λ that has the previous property we have that $\mathbf{u}' = \lambda_0\mathbf{u}$ is primitive. Thus, if we take $s = \lambda_0^{-1}$ we have the desired result. The proof of the claim is completed. □

Hence in what follows we assume that $k = d$, that $u_1, \dots, u_d$ form a basis of $\mathbb{Z}^d$ and that $\{u_1, \dots, u_d\}$ is primitive.
We define now the functions $\varphi_1, \dots, \varphi_d, \colon \mathbb{Z}_{\widetilde{N}} \to \mathbb{Z}^d_{\widetilde{N}}$ by the rule

$$\varphi_i(m) = m \cdot u_i,$$

for every $m \in \mathbb{Z}_{\widetilde{N}}$ and every $i \in [d]$ and observe that the set

$$\{\varphi_i(m) - \varphi_j(m) \colon i, j \in [d] \text{ and } m \in \mathbb{Z}_{\widetilde{N}}\}$$

generates $\mathbb{Z}^d_{\widetilde{N}}$, since $\langle u_i \colon i \in [d]\rangle$ is a basis of $\mathbb{Z}^d$.
We consider further the $(d-1)$-uniform hypergraph system $\mathscr{H} = (d, \langle (X_i, \mu_i) \colon i \in [d]\rangle, \mathcal{H})$ where: (a) $\mathcal{H} = \binom{d}{d-1}$, and (b) (X_i, μ_i) is the discrete probability space with $X_i = \mathbb{Z}_{\widetilde{N}}$ and μ_i the uniform probability measure on $\mathbb{Z}_{\widetilde{N}}$ for every $i \in [d]$.

Moreover, we set $\nu = \nu_{\varepsilon_1,\varepsilon_2,\mathbf{b},d} \colon \mathbb{Z}_{\widetilde{N}}^d \to \mathbb{R}$ to be the function defined in (12.11) and define

$$\nu_{[d]\setminus\{j\}} \colon \boldsymbol{X} \to \mathbb{R}$$

by the rule

$$\nu_{[d]\setminus\{j\}}\big((x_i)_{i\in[k]}\big) = \nu\Big(\sum_{i\in[d]} \big(\varphi_i(x_i) - \varphi_j(x_i)\big)\Big).$$

By Lemma 13.3, the choice of D and Corollary 12.11 we see that

$$\mathbb{E}\big[\prod_{j\in[d]}\prod_{\omega\in\{0,1\}^{[d]\setminus\{j\}}} \nu\big(\sum_{i\in[d]}\varphi_i(x_i^{(\omega_i)}) - \varphi_j(x_i^{(\omega_i)})\big)^{n_{j,\omega}} \,\big|\, {\textstyle x_1^{(0)},x_1^{(1)} \atop {\cdots \atop x_d^{(0)},x_d^{(1)}}} \in \mathbb{Z}_{\widetilde{N}}\big] = 1 + o_{N\to\infty}(1),$$

for any choice of $n_{j,\omega} \in \{0,1\}$. Therefore, the family $\langle \nu_{[d]\setminus\{j\}} \colon j \in [d]\rangle$ satisfies the linear forms condition defined in (8.8) and thus since N is sufficiently large we see that the previous family is $(1, N_0^{-1}, \infty)$ pseudorandom and $|Z| = \widetilde{N} \geqslant N_0$, where N_0 is as in (13.8).

We set now $f \colon \mathbb{Z}_{\widetilde{N}}^d \to [0,\infty)$ to be the function defined by the rule

$$f(\mathbf{n}) = \delta_{\gamma,d} \cdot \widetilde{\Lambda}_{\mathbf{b},W,d}(\mathbf{n}) \cdot \mathbf{1}_{A\cap[\varepsilon_1\widetilde{N},\varepsilon_2\widetilde{N}]^d}(Wn_1 + b_1, \dots, Wn_d + b_d),$$

where $\mathbf{n} = (n_1,\dots,n_d) \in \mathbb{Z}^d$ and $\widetilde{\Lambda}_{\mathbf{b},W,d}$ is as in (12.6) and by Theorem 12.6 we see that $f \leqslant \nu$. For the function f we also have the following Claim.

Claim 13.8. *The following inequality holds true.*

$$\mathbb{E}\big[f(\mathbf{n}) \,|\, \mathbf{n} \in \mathbb{Z}_{\widetilde{N}}^d\big] \geqslant \frac{\delta\delta_{\gamma,d}}{2^{d+2}}. \tag{13.9}$$

Proof of claim 13.8. For the choice of ε_1 and ε_2 in (13.7) we use Proposition 12.6 and obtain that

$$\sum_{\mathbf{n}\in[\widetilde{N}]^d\setminus[\varepsilon_1\widetilde{N},\varepsilon_2\widetilde{N}]^d} \widetilde{\Lambda}_{\mathbf{b},W,d}(\mathbf{n}) \leqslant \frac{\delta}{10^{d+1}}\widetilde{N}^d$$

and thus

$$\sum_{\mathbf{n}\in[\widetilde{N}]^d\setminus[\varepsilon_1\widetilde{N},\varepsilon_2\widetilde{N}]^d} f(\mathbf{n}) \leqslant \frac{\delta\delta_{\gamma,d}}{10^{d+1}}\widetilde{N}^d.$$

Combining now the previous expression with Lemma 13.4 we see that

$$\sum_{\mathbf{n}\in[\widetilde{N}]^d} f(\mathbf{n}) \geqslant \delta\delta_{\gamma,d}\big(\frac{1}{2^{d+2}} - \frac{1}{10^{d+2}}\big)\widetilde{N} \geqslant \frac{\delta\delta_{\gamma,d}}{2^{d+2}}\widetilde{N}.$$

Using now the identification $\mathbb{Z}_{\widetilde{N}} = [\widetilde{N}]$ the proof of the Claim is completed. $\square$

Therefore, by Theorem 11.1 we have that

$$\mathbb{E}\big[\prod_{j\in[d]} f\big(\mathbf{x}+\varphi_j(t)\big)\,\big|\,\mathbf{x}\in\mathbb{Z}^d_{\widetilde{N}}, t\in\mathbb{Z}_{\widetilde{N}}\big] \geqslant c_0. \tag{13.10}$$

We observe now that by the prime number theorem the contribution of the trivial term $t=0$ is $O(\log^{-d}\widetilde{N})$ and thus since N is large we see that there exist $\mathbf{x}\in\mathbb{Z}^d_{\widetilde{N}}$ and $t\in\mathbb{Z}_{\widetilde{N}}\setminus\{0\}$ such that

$$\mathbf{x}+tu_1,\dots,\mathbf{x}+tu_d\in A\cap[\varepsilon_1\widetilde{N},\varepsilon_2\widetilde{N}]^d. \tag{13.11}$$

It remains to show that the previous expression gives rise to a genuine constellation. More precisely we have the following claim.

Claim 13.9. *There exist $\mathbf{x}'\in\mathbb{Z}^d$ and $t'\in\mathbb{Z}\setminus\{0\}$ such that*

$$\mathbf{x}'+t'u_1,\dots,\mathbf{x}'+t'u_d\in A \tag{13.12}$$

Proof of Claim 13.9. By (13.11) there exist $\mathbf{x}_1,\dots,\mathbf{x}_d\in\mathbb{Z}^d$ and $t_1,\dots,t_d\in\mathbb{Z}\setminus\{0\}$ such that

$$\mathbf{x}_1+t_1u_1,\dots,\mathbf{x}_d+t_du_d\in A\cap[\varepsilon_1\widetilde{N},\varepsilon_2\widetilde{N}]^d,$$

with $\mathbf{x}_i=\mathbf{x}$ in $\mathbb{Z}^d_{\widetilde{N}}$ and $t_i=t$ in $\mathbb{Z}_{\widetilde{N}}$ for every $i\in[d]$. Thus, our task is to show that there exist $\mathbf{x}'\in\mathbb{Z}^d$ and $t'\in\mathbb{Z}\setminus\{0\}$ such that $\mathbf{x}_i=\mathbf{x}'$ and $t_i=t'$ for every $i\in[d]$. Assume first that we have found a t' such that $t_i=t'$ in $\mathbb{Z}^d$ and $t'u_i\in[\varepsilon_1\widetilde{N},\varepsilon_2\widetilde{N}]^d$ for every $i\in[d]$. Then, for every $i\in[d]$, $\mathbf{x}_i+t'u_i\in[\varepsilon_1\widetilde{N},\varepsilon_2\widetilde{N}]^d$ and thus by the choice of $\varepsilon_1,\varepsilon_2$ we have $\mathbf{x}_i=\mathbf{x}$ in $\mathbb{Z}^d$ for every $i\in[d]$. Thus, we may set $\mathbf{x}'=\mathbf{x}$.

It remains to show that there exists a t' such that $t_i=t'$ in $\mathbb{Z}^d$ and $t'u_i\in[\varepsilon_1\widetilde{N},\varepsilon_2\widetilde{N}]^d$, for every $i\in[d]$. To this end we will use the fact that $\mathbf{u}=(u_1,\dots,u_d)$ is a primitive shape and more precisely that by Remark 3 we have

$$\inf_{\substack{\mathbf{m}\in\mathbb{Z}^{d^2}\setminus\{\mathbf{0},\mathbf{u}\},\\ \mathbf{y}\in[\mathbf{0},\mathbf{u}]}}\|\mathbf{m}-\mathbf{y}\|_\infty=1. \tag{13.13}$$

We do the identification $\mathbb{Z}_{\widetilde{N}}=[\widetilde{N}]$ and observe that for every $i\in[d]$ there exists $k_i\in\mathbb{Z}$ and $m_i\in\mathbb{Z}^d$ such that

$$t_i=t+k_iN \text{ and } t_iu_i-m_i\widetilde{N}\in[\varepsilon_1\widetilde{N},\varepsilon_2\widetilde{N}]^d.$$

Thus there exist $m'_1,\dots,m'_d\in\mathbb{Z}^d$ such that

$$tu_i-m'_i\widetilde{N}\in[\varepsilon_1\widetilde{N},\varepsilon_2\widetilde{N}]^d, \tag{13.14}$$

for every $i\in[d]$. More especially we have that for every $i\in[d]$

$$\|tu_i-m'_i\widetilde{N}\|_\infty<\widetilde{N}$$

and thus

$$\|\frac{t}{\widetilde{N}}\mathbf{u}-\mathbf{m}\|_\infty < 1,$$

where $\mathbf{m} = (m'_1, \ldots, m'_d) \in \mathbb{Z}^{d^2}$. Since $t \in [\widetilde{N}-1]$ by (13.13) we see that $\mathbf{m} = \mathbf{0}$ or $\mathbf{m} = \mathbf{u}$. Therefore by (13.14), taking $t' = t$ in the first case and $t' = t - N$ in the second one we have completed the proof of the claim. □

With the proof of the previous claim the proof of Theorem 13.6 is completed also. □

Theorem 13.6 provides us with the following corollaries.

COROLLARY 13.10. *Let d, k be positive integers with $k \geqslant 3$, $\mathbf{u} = (u_1, \ldots, u_k) \in \mathbb{Z}^{dk} \setminus \{\mathbf{0}\}$ be a shape in general position and $\delta > 0$. Then, every $A \subseteq \mathbf{P}^d$ with*

$$\limsup_{N\to\infty} \frac{|A \cap [1,N]^d|}{|\mathbf{P}^d_N|} > \delta^{3} \tag{13.15}$$

contains infinitely many constellations of $\mathbf{u}$.

PROOF. By (13.15) there exists a sequence $(N_j)_{j=1}^\infty$ of large positive integers such that

$$|A \cap [1, N_j]^d| \geqslant \delta |\mathbf{P}^d_{N_j}|,$$

for every j. Therefore applying Theorem 13.6 successively to those N_j's gives us the desired result. □

Furthermore, by the previous corollary we obtain the following result.

COROLLARY 13.11. *For every positive integers d, k with $k \geqslant 3$, the set $\mathbf{P}^d$ contains infinitely many constellations of every shape $\mathbf{u} \in \mathbb{Z}^{dk} \setminus \{\mathbf{0}\}$ that is in general position.*

Finally, as a corollary we obtain the Green–Tao theorem, [GT08]. More precisely,

COROLLARY 13.12 (Green-Tao theorem). *Let k be a positive integer with $k \geqslant 3$, N be sufficiently large and $\delta > 0$. Also, let $A \subseteq \mathbf{P}_N$ with $|A| \geqslant \delta|\mathbf{P}_N|$. Then there exist $x, t \in \mathbb{Z}$ with $t \neq 0$ such that*

$$x + t, \ldots, x + kt \in A.$$

PROOF. Just observe that the set $\{1, \ldots, k\}$ is in general position in $\mathbb{Z}$ and apply Theorem 13.6. □

[3] The (LHS) of this expression is usually referred to as the upper density of the set A.

Part V

Algorithmic consequences of the regularity method

CHAPTER 14

An algorithmic regularity lemma for L_p regular sparse matrices

In this chapter we discuss an algorithmic regularity lemma for L_p regular sparse matrices. This result is based on the techniques described in Parts I and II.

To proceed with our discussion it is useful at this point to introduce some pieces of notation and some terminology. Unless otherwise stated, in the rest of this chapter by n_1 and n_2 we denote two positive integers. Now, if X is a nonempty finite set, then by μ_X we denote the uniform probability measure on X, that is, $\mu_X(A) = |A|/|X|$, for every $A \subseteq X$. For notational simplicity, the probability measures $\mu_{[n_1]}, \mu_{[n_2]}$ and $\mu_{[n_1]\times[n_2]}$ will be denoted by μ_1, μ_2 and $\boldsymbol{\mu}$. If $\mathcal{P}$ is a partition of $[n_1] \times [n_2]$, then by $\mathcal{A}_{\mathcal{P}}$ we denote the (finite) σ-algebra on $[n_1] \times [n_2]$ generated by $\mathcal{P}$.

Next, let X_1, X_2 be nonempty finite sets and set

$$\mathcal{S}_{X_1\times X_2} = \{A_1 \times A_2 \colon A_1 \subseteq X_1 \text{ and } A_2 \subseteq X_2\}.$$

If X_1 and X_2 are understood from the context (in particular, if $X_1 = [n_1]$ and $X_2 = [n_2]$), then we shall denote $\mathcal{S}_{X_1\times X_2}$ simply by $\mathcal{S}$. Moreover, fro every partition $\mathcal{P}$ of $X_1 \times X_2$ with $\mathcal{P} \subseteq S_{X_1\times X_2}$ we set

$$\iota(\mathcal{P}) = \min\big\{ \min\{\mu_{X_1}(P_1), \mu_{X_2}(P_2)\} : P = P_1 \times P_2 \in \mathcal{P}\big\}.$$

That is, the quantity $\iota(\mathcal{P})$ is the minimal density of each side of each rectangle $P_1 \times P_2$ belonging to the partition $\mathcal{P}$.

Now recall that a *cut matrix* $g \colon [n_1]\times[n_2] \to \mathbb{R}$ is a matrix for which there exist two sets $S \subseteq [n_1]$ and $T \subseteq [n_2]$, and a real number c such that $g = c \cdot \mathbf{1}_{S\times T}$; the set $S \times T$ is called the *support* of the matrix g. Also recall that for every matrix $f \colon [n_1]\times[n_2] \to \mathbb{R}$ the *cut norm* of f (see also Chapter 1, Example 1) is the quantity

$$\|f\|_\square = \max_{\substack{S\subseteq[n_1]\\ T\subseteq[n_2]}} \Big| \sum_{(x_1,x_2)\in S\times T} f(x_1,x_2)\Big| = (n_1\, n_2)\cdot \max_{\substack{S\subseteq[n_1]\\ T\subseteq[n_2]}} \Big| \int_{S\times T} f\, d\boldsymbol{\mu}\Big|.$$

We are now ready to introduce the class of L_p regular matrices (see also Definition 5.1).

Definition 14.1 (L_p regular matrices). *Let $0 < \eta \leqslant 1$, $C \geqslant 1$ and $1 \leqslant p \leqslant \infty$. A matrix $f\colon [n_1]\times[n_2] \to \{0,1\}$ is called (C,η,p)-regular (or simply L_p* regular *if C and η are understood) if for every partition $\mathcal{P}$ of $[n_1]\times[n_2]$ with $\mathcal{P}\subseteq\mathcal{S}$ and $\iota(\mathcal{P}) \geqslant \eta$ we have*

$$\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} \leqslant C\,\|f\|_{L_1}. \tag{14.1}$$

The following theorem is the main result of this chapter.

Theorem 14.2 (Algorithmic Regularity Lemma). *There exist absolute constants $a_1, a_2 > 0$, an algorithm and a polynomial Π_0 such that the following holds. Let $0 < \varepsilon < 1/2$ and $C \geqslant 1$. Also let $1 < p \leqslant \infty$, set $p^\dagger = \min\{2,p\}$ and let q denote the conjugate exponent of $p^\dagger$ (that is, $1/p^\dagger + 1/q = 1$). We set*

$$\tau = \Big\lceil \frac{a_1 \cdot C^2}{(p^\dagger - 1)\,\varepsilon^2} \Big\rceil \quad \text{and} \quad \eta = \Big(\frac{a_2\cdot\varepsilon}{C}\Big)^{\sum_{i=1}^{\tau+1}(\frac{2}{p^\dagger}+1)^{i-1}q^i}. \tag{14.2}$$

If we input

`INP`: *a (C,η,p)-regular matrix $f\colon [n_1]\times[n_2]\to\{0,1\}$,*

then the algorithm outputs

`OUT`: *a partition $\mathcal{P}$ of $[n_1]\times[n_2]$ with $\mathcal{P}\subseteq\mathcal{S}$, $|\mathcal{P}| \leqslant 4^\tau$ and $\iota(\mathcal{P}) \geqslant \eta$, such that*

$$\|f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\|_\square \leqslant \varepsilon \|f\|_\square. \tag{14.3}$$

Moreover, this algorithm has running time $(\tau\, 4^\tau)\cdot \Pi_0(n_1\, n_2)$.

Theorem 14.2 extends [COCF10, Theorem 1] which corresponds to the case $p=\infty$[1]. Note that, by (14.2) and (14.3), the matrix f is well approximated by a sum of at most 4^τ cut matrices with disjoint supports and, moreover, the positive integer τ is independent of the size of f and its density. Also observe that, as expected, the running time of the algorithm in Theorem 14.2 increases as p decreases to 1.

14.1. Backround material

The proof of Theorem 14.2 will be based on Proposition 2.1 and the following algorithmic version of Grothendieck's inequality. This result is due to Alon and Naor [AN06].

Proposition 14.3. *There exist a constant $a_0 > 0$, an algorithm and a polynomial Π_{AN} such that the following holds. If we input*

`INP`: *a matrix $f\colon [n_1]\times[n_2]\to\mathbb{R}$,*

[1] Actually, the argument in [COCF10] works for the more general case $p \geqslant 2$. We also remark that the cut matrices obtained by [COCF10, Theorem 1] do not necessarily have disjoint supports, but this can be easily arranged—see [COCF10, Corollary 1] for more details.

then the algorithm outputs

`OUT`: *a set* $A \in \mathcal{S}$ *such that* $(n_1 n_2)\big|\int_A f\, d\boldsymbol{\mu}\big| \geqslant a_0 \|f\|_\square$.

Moreover, this algorithm has running time $\Pi_{\mathrm{AN}}(n_1 n_2)$.

The constant a_0 in Proposition 14.3 is closely related to Grothendieck's constant K_G (see, e.g., [Pis12]); in particular, we have $a_0 \geqslant K_G^{-1}$.

14.2. Preparatory Lemmas

In this section we prove some preparatory results concerning L_p regular matrices. We begin with the following lemma.

Lemma 14.4. *There exist an algorithm and a polynomial* Π_1 *such that the following holds. Let* X_1, X_2 *be two nonempty finite sets, let* ν_1, ν_2 *denote the uniform measures on* X_1 *and* X_2 *respectively, and let* $\boldsymbol{\nu}$ *denote the uniform probability measure on* $X_1 \times X_2$. *Also let* $0 < \vartheta < 1/2$. *If we input*

`INP`: *two sets* $A_1 \subseteq X_1$ *and* $A_2 \subseteq X_2$ *with* $\nu_1(A_1) \geqslant \vartheta$ *and* $\nu_2(A_2) \geqslant \vartheta$,

then the algorithm outputs

`OUT1`: *a partition* $\mathcal{Q} \subseteq \mathcal{S}$ *with* $|\mathcal{Q}| \leqslant 4$ *and* $\iota(\mathcal{Q}) \geqslant \vartheta$, *and*

`OUT2`: *a set* $B \in \mathcal{Q}$ *such that* $A_1 \times A_2 \subseteq B$ *and* $\boldsymbol{\nu}\big(B \setminus (A_1 \times A_2)\big) \leqslant 2\vartheta$.

Moreover, this algorithm has running time $\Pi_1(|X_1| \cdot |X_2|)$.

Proof. We distinguish the following four (mutually exclusive) cases.

Case 1: $\nu_1(A_1) < 1 - \vartheta$ *and* $\nu_2(A_2) < 1 - \vartheta$. In this case the algorithm outputs $\mathcal{Q} = \{A_1 \times A_2, (X_1 \setminus A_1) \times A_2, A_1 \times (X_2 \setminus A_2), (X_1 \setminus A_1) \times (X_2 \setminus A_2)\}$ and $B = A_1 \times A_2$. Notice that $\mathcal{Q}$ and B satisfy the requirements of the lemma.

Case 2: $\nu_1(A_1) < 1 - \vartheta$ *and* $\nu_2(A_2) \geqslant 1 - \vartheta$. In this case the algorithm outputs $\mathcal{Q} = \{A_1 \times X_2, (X_1 \setminus A_1) \times X_2\}$ and $B = A_1 \times X_2$. Again, it is easy to see that $\mathcal{Q}$ and B satisfy the requirements of the lemma.

Case 3: $\nu_1(A_1) \geqslant 1 - \vartheta$ *and* $\nu_2(A_2) < 1 - \vartheta$. This case is similar to Case 2. In particular, we set $\mathcal{Q} = \{X_1 \times A_2, X_1 \times (X_2 \setminus A_2)\}$ and $B = X_1 \times A_2$.

Case 4: $\nu_1(A_1) \geqslant 1 - \vartheta$ *and* $\nu_2(A_2) \geqslant 1 - \vartheta$. In this case the algorithm outputs $\mathcal{Q} = \{X_1 \times X_2\}$ and $B = X_1 \times X_2$. As before, it is easy to see that $\mathcal{Q}$ and B are as desired.

Finally, notice that the most costly part of this algorithm is to estimate the quantities $\nu_1(A_1)$ and $\nu_2(A_2)$, but of course this can be done in polynomial time of $|X_1| \cdot |X_2|$. Thus, this algorithm will stop in polynomial time of $|X_1| \cdot |X_2|$. $\square$

The following lemma is a Hölder-type inequality for L_p regular matrices (see also Proposition 5.2).

Lemma 14.5. *Let $0<\eta<1/2$ and $C\geqslant 1$. Also let $1<p\leqslant 2$ and let q denote its conjugate exponent. Finally, let $f\colon [n_1]\times[n_2]\to\{0,1\}$ be (C,η,p)-regular. Then for every $A\subseteq [n_1]\times[n_2]$ with $A\in\mathcal{S}$ we have*

$$\int_A f\, d\boldsymbol{\mu} \leqslant C\,\|f\|_{L_1}(\boldsymbol{\mu}(A)+6\eta)^{1/q}. \tag{14.4}$$

Proof. Fix a nonempty subset A of $[n_1]\times[n_2]$ with $A\in\mathcal{S}$, and let $A_1\subseteq[n_1]$ and $A_2\subseteq[n_2]$ such that $A=A_1\times A_2$. If $\mu_1(A_1)\geqslant\eta$ and $\mu_2(A_2)\geqslant\eta$, then we claim that

$$\int_A f\, d\boldsymbol{\mu} \leqslant C\,\|f\|_{L_1}(\boldsymbol{\mu}(A)+2\eta)^{1/q}. \tag{14.5}$$

Indeed, by Lemma 14.4 applied for $X_1=[n_1]$ and $X_2=[n_2]$, we obtain a partition $\mathcal{Q}$ of $[n_1]\times[n_2]$ with $\mathcal{Q}\in\mathcal{S}$ and $\iota(\mathcal{Q})\geqslant\eta$, and a set $B\in\mathcal{Q}$ such that $A\subseteq B$ and $\boldsymbol{\mu}(B\setminus A)\leqslant 2\eta$. By the L_p regularity of f, we have

$$\frac{\int_B f\,d\boldsymbol{\mu}}{\boldsymbol{\mu}(B)}\,\boldsymbol{\mu}(B)^{1/p}\leqslant \|\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})\|_{L_p}\leqslant C\,\|f\|_{L_1}$$

and so

$$\int_A f\,d\boldsymbol{\mu}\leqslant\int_B f\,d\boldsymbol{\mu}\leqslant C\,\|f\|_{L_1}\boldsymbol{\mu}(B)^{1/q}\leqslant C\,\|f\|_{L_1}(\boldsymbol{\mu}(A)+2\eta)^{1/q}.$$

Next, we assume that $\mu_1(A_1)\geqslant\eta$ and $\mu_2(A_2)<\eta$ and observe that we may select a set $B\subseteq[n_2]$ with $\eta<\mu_2(B)\leqslant 2\eta$. Then, we have

$$\begin{aligned}\int_A f\,d\boldsymbol{\mu} &\leqslant \int_{A_1\times(A_2\cup B)} f\,d\boldsymbol{\mu} \overset{(14.5)}{\leqslant} C\,\|f\|_{L_1}\big(\boldsymbol{\mu}\big(A_1\times(A_2\cup B)\big)+2\eta\big)^{1/q}\\ &\leqslant C\,\|f\|_{L_1}(\boldsymbol{\mu}(A)+2\eta\,\mu_1(A_1)+2\eta)^{1/q}\leqslant C\,\|f\|_{L_1}(\boldsymbol{\mu}(A)+4\eta)^{1/q}.\end{aligned}$$

The case $\mu_1(A_1)<\eta$ and $\mu_2(A_2)\geqslant\eta$ is identical.

Finally, assume that $\mu_1(A_1)<\eta$ and $\mu_2(A_2)<\eta$, and observe that there exist $B_1\subseteq[n_1]$ and $B_2\subseteq[n_2]$ such that $\eta<\mu_1(B_1)\leqslant 2\eta$ and $\eta<\mu_2(B_2)\leqslant 2\eta$. Then,

$$\begin{aligned}\int_A f\,d\boldsymbol{\mu} &\leqslant \int_{(A_1\cup B_1)\times(A_2\cup B_2)} f\,d\boldsymbol{\mu}\\ &\overset{(14.5)}{\leqslant} C\,\|f\|_{L_1}\big(\boldsymbol{\mu}\big((A_1\cup B_1)\times(A_2\cup B_2)\big)+2\eta\big)^{1/q}\\ &\leqslant C\,\|f\|_{L_1}(\boldsymbol{\mu}(A)+8\eta^2+2\eta)^{1/q}\leqslant C\,\|f\|_{L_1}(\boldsymbol{\mu}(A)+6\eta)^{1/q}\end{aligned}$$

and the proof of the lemma is completed. □

Lemmas 14.4 and 14.5 will be used in the proof of the following result.

LEMMA 14.6. *There exist an algorithm and a polynomial Π_2 such that the following holds. Let $0<\varepsilon<1/2$ and $C\geqslant 1$. Let $1<p\leqslant\infty$, set $p^\dagger=\min\{2,p\}$ and let q denote the conjugate exponent of $p^\dagger$. Also let a_0 be as in Proposition 14.3, and set*

$$\vartheta=\frac{a_0\,\varepsilon}{16C}\quad\textit{and}\quad \eta\leqslant\Big(\vartheta\cdot\iota(\mathcal{P})^{\frac{2}{p^\dagger}+1}\Big)^q.$$

If we input

`INP1`: *a partition $\mathcal{P}$ of $[n_1]\times[n_2]$ with $\mathcal{P}\subseteq\mathcal{S}$,*
`INP2`: *a subset A of $[n_1]\times[n_2]$ with $A\in\mathcal{S}$, and*
`INP3`: *a (C,η,p)-regular matrix $f\colon[n_1]\times[n_2]\to\{0,1\}$,*

then the algorithm outputs

`OUT1`: *a refinement $\mathcal{Q}$ of $\mathcal{P}$ with $\mathcal{Q}\subseteq\mathcal{S}$, $|\mathcal{Q}|\leqslant 4|\mathcal{P}|$ and $\iota(\mathcal{Q})\geqslant(\vartheta\cdot\iota(\mathcal{P})^{\frac{2}{p^\dagger}+1})^q$, and*
`OUT2`: *a set $B\in\mathcal{A}_{\mathcal{Q}}$ such that*

$$\int_{A\triangle B}\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\,d\boldsymbol{\mu}\leqslant 2C\,\|f\|_{L_1}\vartheta\quad\textit{and}\quad\int_{A\triangle B}f\,d\boldsymbol{\mu}\leqslant 6C\,\|f\|_{L_1}\vartheta. \tag{14.6}$$

If we additionally assume that the matrix f in `INP3` *satisfies*

$$\Big|\int_A\big(f-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\,d\boldsymbol{\mu}\Big|\geqslant a_0\,\varepsilon\,\|f\|_{L_1}, \tag{14.7}$$

then the partition $\mathcal{Q}$ in `OUT2` *satisfies*

$$\|\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}})-\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\|_{L_{p^\dagger}}\geqslant\frac{a_0\,\varepsilon\,\|f\|_{L_1}}{2}. \tag{14.8}$$

Finally, this algorithm has running time $|\mathcal{P}|\cdot\Pi_2(n_1n_2)$.

Lemma 14.6 is an algorithmic version of Lemmas 6.2 and 6.3. We also notice that if the matrix f satisfies the estimate in (14.7), then inequality (14.8) implies that the partition $\mathcal{Q}$ is a genuine refinement of $\mathcal{P}$. We proceed to the proof.

PROOF OF LEMMA 14.6. We may (and we will) assume that A is nonempty. We select $A_1\subseteq[n_1]$ and $A_2\subseteq[n_2]$ such that $A=A_1\times A_2$, and we set

$$\theta=\vartheta^q\cdot\iota(\mathcal{P})^{\frac{2q}{p^\dagger}}.$$

Also let

$$\mathcal{P}^1=\{P=P_1\times P_2\in\mathcal{P}\colon \mu_1(A_1\cap P_1)<\theta\mu_1(P_1)\text{ and }\mu_2(A_2\cap P_2)<\theta\mu_2(P_2)\},$$
$$\mathcal{P}^2=\{P=P_1\times P_2\in\mathcal{P}\colon \mu_1(A_1\cap P_1)<\theta\mu_1(P_1)\text{ and }\mu_2(A_2\cap P_2)\geqslant\theta\mu_2(P_2)\},$$
$$\mathcal{P}^3=\{P=P_1\times P_2\in\mathcal{P}\colon \mu_1(A_1\cap P_1)\geqslant\theta\mu_1(P_1)\text{ and }\mu_2(A_2\cap P_2)<\theta\mu_2(P_2)\},$$
$$\mathcal{P}^4=\{P=P_1\times P_2\in\mathcal{P}\colon \mu_1(A_1\cap P_1)\geqslant\theta\mu_1(P_1)\text{ and }\mu_2(A_2\cap P_2)\geqslant\theta\mu_2(P_2)\}.$$

Clearly, the family $\{\mathcal{P}^1,\mathcal{P}^2,\mathcal{P}^3,\mathcal{P}^4\}$ is a partition of $\mathcal{P}$.

Now for every $P \in \mathcal{P}$ we perform the following subroutine. First, assume that $P \in \mathcal{P}^1 \cup \mathcal{P}^2 \cup \mathcal{P}^3$ and notice that in this case we have $\boldsymbol{\mu}(A \cap P) \leqslant \theta \boldsymbol{\mu}(P)$. Then we set $B_P = \emptyset$ and $\mathcal{Q}_P = \{P\}$. On the other hand, if $P = P_1 \times P_2 \in \mathcal{P}^4$, then we apply Lemma 14.4 for $X_1 = P_1$ and $X_2 = P_2$, and we obtain[2] a partition $\mathcal{Q}_P$ of P with $\mathcal{Q} \in \mathcal{S}$, $|\mathcal{Q}_P| \leqslant 4$ and $\iota(\mathcal{Q}_P) \geqslant \theta \cdot \iota(\mathcal{P})$, and a set $B_P \in \mathcal{Q}_P$ such that $A \cap P \subseteq B_P$ and $\boldsymbol{\mu}(B_P \setminus (A \cap P)) \leqslant 2\theta \boldsymbol{\mu}(P)$.

Once this is done, the algorithm outputs

$$\mathcal{Q} = \bigcup_{P \in \mathcal{P}} \mathcal{Q}_P \quad \text{and} \quad B = \bigcup_{P \in \mathcal{P}} B_P.$$

Notice that there exists a polynomial Π_2 such that this algorithm has running time $|\mathcal{P}| \cdot \Pi_2(n_1 n_2)$. Indeed, recall that the algorithm in Lemma 14.4 runs in polynomial time and observe that we have applied Lemma 14.4 at most $|\mathcal{P}|$ times.

We proceed to show that the partition $\mathcal{Q}$ and the set B satisfy the requirements of the lemma. To this end, we first observe that $\mathcal{Q}$ satisfies the requirements in `OUT1`. Moreover, we have $B \in \mathcal{A}_{\mathcal{Q}}$ and

$$A \triangle B = \Big(\bigcup_{i=1}^{3} \bigcup_{P \in \mathcal{P}^i} (A \cap P) \Big) \cup \Big(\bigcup_{P \in \mathcal{P}^4} \big(B_P \setminus (A \cap P) \big) \Big). \tag{14.9}$$

Therefore,

$$\boldsymbol{\mu}(A \triangle B) \leqslant 2\theta \tag{14.10}$$

and so, by the L_p regularity of f, Hölder's inequality, the monotonicity of the L_p norms and the fact that $p^\dagger \leqslant p$, we obtain that

$$\begin{aligned}\int_{A \triangle B} \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\, d\boldsymbol{\mu} &\leqslant \|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_{p^\dagger}} \cdot \boldsymbol{\mu}(A \triangle B)^{1/q} \leqslant \|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} \cdot \boldsymbol{\mu}(A \triangle B)^{1/q} \\ &\leqslant C\, \|f\|_{L_1} (2\theta)^{1/q} \leqslant 2C\, \|f\|_{L_1} \vartheta\end{aligned}$$

which proves the first inequality in (14.6). For the second inequality, by (14.9), we have

$$\int_{A \triangle B} f\, d\boldsymbol{\mu} = \sum_{P \in \mathcal{P}^1 \cup \mathcal{P}^2 \cup \mathcal{P}^3} \int_{A \cap P} f\, d\boldsymbol{\mu} + \sum_{P \in \mathcal{P}^4} \int_{B_P \setminus (A \cap P)} f\, d\boldsymbol{\mu} \tag{14.11}$$

and, by the definition of θ and the fact that $\eta \leqslant (\vartheta \cdot \iota(\mathcal{P})^{\frac{2}{p^\dagger}+1})^q$, we have $\eta \leqslant \theta \boldsymbol{\mu}(P)$ for every $P \in \mathcal{P}$. Thus, if $P \in \mathcal{P}^1 \cup \mathcal{P}^2 \cup \mathcal{P}^3$, then, by Lemma 14.5 and our assumption that f is (C, η, p)-regular (and, consequently, $(C, \eta, p^\dagger)$-regular), we have

$$\int_{A \cap P} f\, d\boldsymbol{\mu} \leqslant C\, \|f\|_{L_1} (\boldsymbol{\mu}(A \cap P) + 6\eta)^{1/q} \leqslant 3C\, \|f\|_{L_1} \big(\theta \boldsymbol{\mu}(P)\big)^{1/q}$$

---

[2]Notice that if ν_1 is the uniform probability measure on X_1, then for every $A \subseteq X_1$ we have $\nu_1(A) = \mu_1(A)/\mu_1(X_1)$, and similarly for X_2.

which yields that

$$\sum_{P\in\mathcal{P}^1\cup\mathcal{P}^2\cup\mathcal{P}^3} \int_{A\cap P} f\, d\boldsymbol{\mu} \leqslant 3C\, \|f\|_{L_1}\, \theta^{1/q} \sum_{P\in\mathcal{P}^1\cup\mathcal{P}^2\cup\mathcal{P}^3} \boldsymbol{\mu}(P)^{1/q}. \tag{14.12}$$

On the other hand, by the choice of the family $\{B_P : P \in \mathcal{P}^4\}$ and Lemma 14.5,

$$\sum_{P\in\mathcal{P}^4} \int_{B_P\setminus(A\cap P)} f\, d\boldsymbol{\mu} \leqslant 6C\, \|f\|_{L_1}\, \theta^{1/q} \sum_{P\in\mathcal{P}^4} \boldsymbol{\mu}(P)^{1/q}. \tag{14.13}$$

Moreover, since $q \geqslant 2$ we have that $x^{1/q}$ is concave on $\mathbb{R}_+$, and so

$$\sum_{P\in\mathcal{P}} \boldsymbol{\mu}(P)^{1/q} \leqslant |\mathcal{P}|^{\frac{1}{p^\dagger}} \leqslant \iota(\mathcal{P})^{-\frac{2}{p^\dagger}}. \tag{14.14}$$

Combining (14.12)–(14.14), we see that the second inequality in (14.6) is satisfied.

Finally, assume that the matrix f satisfies (14.7). By (14.6) and the choice of ϑ,

$$\begin{aligned}\Big| \int_A \big(f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\, d\boldsymbol{\mu} &- \int_B \big(f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\, d\boldsymbol{\mu}\Big| \\ &\leqslant \int_{A\triangle B} \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\, d\boldsymbol{\mu} + \int_{A\triangle B} f\, d\boldsymbol{\mu} \leqslant \frac{a_0\,\varepsilon\, \|f\|_{L_1}}{2}\end{aligned}$$

and so, by (14.7), we have

$$\Big| \int_B \big(f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\, d\boldsymbol{\mu}\Big| \geqslant \frac{a_0\,\varepsilon\, \|f\|_{L_1}}{2}. \tag{14.15}$$

Moreover, the fact that $B \in \mathcal{A}_{\mathcal{Q}}$ yields that

$$\int_B \big(f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\, d\boldsymbol{\mu} = \int_B \big(\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\, d\boldsymbol{\mu}. \tag{14.16}$$

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

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

$$\geqslant \Big| \int_B \big(\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{Q}}) - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\, d\boldsymbol{\mu}\Big| \overset{(14.16)}{=} \Big| \int_B \big(f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}})\big)\, d\boldsymbol{\mu}\Big| \overset{(14.15)}{\geqslant} \frac{a_0\,\varepsilon\, \|f\|_{L_1}}{2}$$

and the proof of Lemma 14.6 is completed. □

14.3. Proof of the algorithmic regularity lemma

We will describe a recursive algorithm that performs the following steps. Starting from the trivial partition of $[n_1] \times [n_2]$ and using Lemma 14.6 as a subroutine, the algorithm will produce an increasing family of partitions of $[n_1] \times [n_2]$. Simultaneously, using Proposition 14.3 as a subroutine, the algorithm will be checking if the partition that is produced at each step satisfies the requirements in `OUT` of Theorem 14.2. The fact that this algorithm will eventually terminate is based on Proposition 2.1.

Proof of Theorem 14.2. Let a_0 be as in Proposition 14.3, and set

$$\vartheta = \frac{a_0\,\varepsilon}{16C}, \quad \tau = \Big\lceil \frac{4C^2}{(p^\dagger - 1)\,\varepsilon^2\, a_0^2} \Big\rceil \ \text{ and } \ \eta = \vartheta^{\sum_{i=1}^{\tau+1} (\frac{2}{p^\dagger}+1)^{i-1} q^i}. \tag{14.17}$$

Also fix a (C,η,p)-regular matrix $f\colon [n_1]\times[n_2]\to\{0,1\}$. The algorithm performs the following steps.

`InitialStep`: We set $\mathcal{P}_0 = \{[n_1]\times[n_2]\}$ and we apply the algorithm in Proposition 14.3 for the matrix $f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}_0})$. Thus, we obtain a set $A_0 \subseteq [n_1]\times[n_2]$ with $A_0 \in \mathcal{S}$ and such that $(n_1\, n_2)\big|\int_{A_0} \big(f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}_0})\big)\,d\boldsymbol{\mu}\big| \geqslant a_0 \|f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}_0})\|_\Box$. If $\big|\int_{A_0} \big(f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}_0})\big)\,d\boldsymbol{\mu}\big| \leqslant a_0\,\varepsilon\,\|f\|_{L_1}$, then the algorithm outputs the partition $\mathcal{P}_0$ and `Halts`. Otherwise, the algorithm sets $m=1$ and enters into the following loop.

`GeneralStep`: The algorithm will have as an input a positive integer $m \in [\tau-1]$, a partition[3] $\mathcal{P}_{m-1} \subseteq \mathcal{S}$ and a set $A_{m-1} \subseteq [n_1]\times[n_2]$ with $A_{m-1}\in\mathcal{S}$, such that

(a) $|\mathcal{P}_{m-1}| \leqslant 4^m$,
(b) $(\vartheta\cdot\iota(\mathcal{P}_{m-1})^{\frac{2}{p^\dagger}+1})^q \geqslant \vartheta^{\sum_{i=1}^{m}(\frac{2}{p^\dagger}+1)^{i-1}q^i}$, and
(c) $\big|\int_{A_{m-1}} \big(f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}_{m-1}})\big)\,d\boldsymbol{\mu}\big| > a_0\,\varepsilon\,\|f\|_{L_1}$.

By (b) and the choice of η in (14.17), we have $\eta \leqslant (\vartheta\cdot\iota(\mathcal{P}_{m-1})^{\frac{2}{p^\dagger}+1})^q$. This fact together with the choice of ϑ in (14.17) allows us to perform the algorithm in Lemma 14.6 for the matrix f, the partition $\mathcal{P}_{m-1}$ and the set A_{m-1}. Thus, we obtain a refinement $\mathcal{P}_m$ of $\mathcal{P}_{m-1}$ with $\mathcal{P}_m\subseteq\mathcal{S}$, $|\mathcal{P}_m| \leqslant 4|\mathcal{P}_{m-1}|$, $\iota(\mathcal{P}_m) \geqslant (\vartheta\cdot\iota(\mathcal{P}_{m-1})^{\frac{2}{p^\dagger}+1})^q$, such that

$$\|\mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}_m}) - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}_{m-1}})\|_{L_{p^\dagger}} \geqslant \frac{a_0\,\varepsilon\,\|f\|_{L_1}}{2}.$$

Next, we apply the algorithm in Proposition 14.3 for the matrix $f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}_m})$, and we obtain a set $A_m \subseteq [n_1]\times[n_2]$ with $A_m\in\mathcal{S}$ and such that

$$(n_1\, n_2)\Big|\int_{A_m} \big(f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}_m})\big)\,d\boldsymbol{\mu}\Big| \geqslant a_0\,\|f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}_m})\|_\Box.$$

If $\big|\int_{A_m} \big(f - \mathbb{E}(f\,|\,\mathcal{A}_{\mathcal{P}_m})\big)\,d\boldsymbol{\mu}\big| \leqslant a_0\,\varepsilon\,\|f\|_{L_1}$, then the algorithm outputs the partition $\mathcal{P}_m$ and `Halts`. Otherwise, if $m < \tau-1$, then the algorithm reruns the loop we described above for the positive integer $m+1$, the partition $\mathcal{P}_m$ and the set A_m, while if $m = \tau-1$, then the algorithm proceeds to the following step.

`FinalStep`: The algorithm will have as an input a partition $\mathcal{P}_{\tau-1}\subseteq\mathcal{S}$ and a set $A_{\tau-1}\subseteq[n_1]\times[n_2]$ with $A_{\tau-1}\in\mathcal{S}$, such that

(d) $|\mathcal{P}_{\tau-1}| \leqslant 4^{\tau-1}$,

[3] Notice that $\mathcal{P}_0\subseteq\mathcal{S}$ and $\iota(\mathcal{P}_0) = 1$.

(e) $(\vartheta \cdot \iota(\mathcal{P}_{\tau-1})^{\frac{2}{p^\dagger}+1})^q \geqslant \vartheta^{\sum_{i=1}^{\tau}(\frac{2}{p^\dagger}+1)^{i-1}q^i}$, and

(f) $|\int_{A_{\tau-1}}\big(f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_{\tau-1}})\big)\, d\boldsymbol{\mu}| > a_0\,\varepsilon\,\|f\|_{L_1}$.

Again observe that, by (e) and the choice of η in (14.17), we have $\eta \leqslant (\vartheta \cdot \iota(\mathcal{P}_{\tau-1})^{\frac{2}{p^\dagger}+1})^q$. Using this fact and the choice of ϑ in (14.17), we may apply the algorithm in Lemma 14.6 for the matrix f, the partition $\mathcal{P}_{\tau-1}$ and the set $A_{\tau-1}$. Therefore, we obtain a refinement $\mathcal{P}_\tau$ of $\mathcal{P}_{\tau-1}$ with $\mathcal{P}_\tau \subseteq \mathcal{S}$, $|\mathcal{P}_\tau| \leqslant 4|\mathcal{P}_{\tau-1}|$, $\iota(\mathcal{P}_\tau) \geqslant (\vartheta \cdot \iota(\mathcal{P}_{\tau-1})^{\frac{2}{p^\dagger}+1})^q$, and such that

$$\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_\tau}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_{\tau-1}})\|_{L_{p^\dagger}} \geqslant \frac{a_0\,\varepsilon\,\|f\|_{L_1}}{2}.$$

The algorithm outputs the partition $\mathcal{P}_\tau$ and `Halts`.

Notice that there exists a polynomial Π_0 such that the previous algorithm has running time $(\tau\, 4^\tau)\cdot \Pi_0(n_1 n_2)$. Indeed, by Proposition 14.3, there exists a polynomial Π_0' such that the `InitialStep` runs in time $\Pi_0'(n_1\, n_2)$. Moreover, by the running times of the algorithms in Lemma 14.6 and Proposition 14.3, there exists a polynomial Π_0'' such that each of the `GeneralStep` runs in time $4^\tau \cdot \Pi_0''(n_1\, n_2)$. Finally, invoking again Lemma 14.6, we see that there exists a polynomial Π_0''' such that the `FinalStep` runs in time $\Pi_0'''(n_1\, n_2)$. Therefore, the algorithm we described above runs in time

$$\Pi_0'(n_1\, n_2) + (\tau-1)\, 4^\tau\, \Pi_0''(n_1\, n_2) + \Pi_0'''(n_1\, n_2)$$

which in turn yields that there exists a polynomial Π_0 such that the algorithm has running time $(\tau\, 4^\tau)\cdot \Pi_0(n_1\, n_2)$.

It remains to verify that the previous algorithm will produce a partition that satisfies the requirements in `OUT` of Theorem 14.2. As we have noted, the argument is based on Proposition 2.1.

We proceed to the details. First assume that the algorithm has stopped before the `FinalStep`. Then the output of the algorithm is one of the partitions we described in `InitialStep` and in `GeneralStep`, say $\mathcal{P}_m$ for some $m \in \{0,\dots,\tau-1\}$. Observe that $\mathcal{P}_m$ satisfies $\mathcal{P}_m \subseteq \mathcal{S}$, $|\mathcal{P}_m| \leqslant 4^m$, and $\iota(\mathcal{P}_m) \geqslant \eta$; in other words, $\mathcal{P}_m$ satisfies the first three requirements in `OUT` of Theorem 14.2. Moreover, recall that there exists a set $A_m \subseteq [n_1]\times[n_2]$ with $A_m \in \mathcal{S}$, and such that

$$(n_1\, n_2)\Big|\int_{A_m}\big(f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_m})\big)\, d\boldsymbol{\mu}\Big| \geqslant a_0\, \|f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_m})\|_\square.$$

On the other hand, since the output of the algorithm is the partition $\mathcal{P}_m$, we have $|\int_{A_m}\big(f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_m})\big)\, d\boldsymbol{\mu}| \leqslant a_0\,\varepsilon\,\|f\|_{L_1}$. Combining these estimates, we conclude that $\|f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_m})\|_\square \leqslant \varepsilon\|f\|_\square$.

Next, assume that the algorithm reaches the `FinalStep`. Recall that $\mathcal{P}_\tau \subseteq \mathcal{S}$ and observe that, by (d) above and the fact that $|\mathcal{P}_\tau| \leqslant 4|\mathcal{P}_{\tau-1}|$, we have $|\mathcal{P}_\tau| \leqslant 4^\tau$. Moreover, by (e) and the choice of η in (14.17),

$$\iota(\mathcal{P}_\tau) \geqslant (\vartheta \cdot \iota(\mathcal{P}_{\tau-1})^{\frac{2}{p^\dagger}+1})^q \geqslant \vartheta^{\sum_{i=1}^{\tau}(\frac{2}{p^\dagger}+1)^{i-1}q^i} \geqslant \eta. \tag{14.18}$$

Thus, we only need to show that $\|f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_\tau})\|_\square \leqslant \varepsilon \|f\|_\square$. To this end assume, towards a contradiction, that $\|f - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_\tau})\|_\square > \varepsilon \|f\|_\square$. Notice that, by the choice of η in (14.17) and (14.18), we have $(\vartheta \cdot \iota(\mathcal{P}_\tau)^{\frac{2}{p^\dagger}+1})^q \geqslant \eta$. Using the previous two estimates, Proposition 14.3, Lemma 14.6 and arguing precisely as in the `GeneralStep`, we may select a refinement $\mathcal{P}_{\tau+1}$ of $\mathcal{P}_\tau$ with $\mathcal{P}_{\tau+1} \subseteq \mathcal{S}$ and $\iota(\mathcal{P}_{\tau+1}) \geqslant \eta$, and such that $\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_{\tau+1}}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_\tau})\|_{L_{p^\dagger}} \geqslant (a_0\, \varepsilon\, \|f\|_{L_1})/2$. It follows that there exists an increasing finite sequence $(\mathcal{P}_i)_{i=0}^{\tau+1}$ of partitions with $\mathcal{P}_0 = \{[n_1] \times [n_2]\}$ and such that for every $i \in [\tau+1]$ we have $\mathcal{P}_i \subseteq \mathcal{S}$, $\iota(\mathcal{P}_i) \geqslant \eta$, and

$$\|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_i}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_{i-1}})\|_{L_{p^\dagger}} \geqslant \frac{a_0\, \varepsilon\, \|f\|_{L_1}}{2}. \tag{14.19}$$

Now set $d_0 = \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_0})$ and $d_i = \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_i}) - \mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_{i-1}})$ for every $i \in [\tau+1]$, and observe that the sequence $(d_i)_{i=0}^{\tau+1}$ is a martingale difference sequence. Therefore, by Proposition 2.1 and the fact that the matrix f is (C, η, p)-regular, we have

$$\begin{aligned}
\frac{a_0\, \varepsilon\, \|f\|_{L_1}}{2} \cdot \sqrt{\tau+1} \;&\overset{(14.19)}{\leqslant}\; \Big(\sum_{i=1}^{\tau+1} \|d_i\|^2_{L_{p^\dagger}}\Big)^{1/2} \leqslant \Big(\sum_{i=0}^{\tau+1} \|d_i\|^2_{L_{p^\dagger}}\Big)^{1/2} \\
&\overset{(A.5)}{\leqslant}\; \frac{1}{\sqrt{p^\dagger - 1}} \Big\| \sum_{i=0}^{\tau+1} d_i \Big\|_{L_{p^\dagger}} = \frac{1}{\sqrt{p^\dagger - 1}} \|\mathbb{E}(f \,|\, \mathcal{A}_{\mathcal{P}_{\tau+1}})\|_{L_{p^\dagger}} \\
&\leqslant\; \frac{C}{\sqrt{p^\dagger - 1}} \|f\|_{L_1}
\end{aligned}$$

which clearly contradicts the choice of τ in (14.17). The proof of Theorem 14.2 is thus completed. □

CHAPTER 15

Applications

15.1. Tensor approximation algorithms

Throughout this chapter let $k \geqslant 2$ be an integer. Also let $n_1, \dots, n_k$ be positive integers, and let $\boldsymbol{\mu}_k$ denote the uniform probability measure on $[n_1] \times \dots \times [n_k]$.

Recall that a *k-dimensional tensor* is a function $F\colon [n_1]\times\dots\times[n_k] \to \mathbb{R}$. (Notice, in particular, that a 2-dimensional tensor is just a matrix.) Also recall, that a tensor $G\colon [n_1] \times \dots \times [n_k] \to \mathbb{R}$ is called a *cut tensor* if there exist a real number c and for every $i \in [k]$ a subset S_i of $[n_i]$ such that $G = c \cdot \mathbf{1}_{S_1\times\dots\times S_k}$. Finally, recall that for every tensor $F\colon [n_1] \times \dots [n_k] \to \mathbb{R}$ its *cut norm* is defined as

$$\|F\|_\square = \Big(\prod_{i=1}^{k} n_i\Big) \cdot \max\Big\{\Big|\int_{S_1\times\dots\times S_k} F\, d\boldsymbol{\mu}_k\Big| : S_i \subseteq [n_i] \text{ for every } i \in [k]\Big\}.$$

Next, let

$$k_1 = \lfloor k/2 \rfloor, \quad A_k = [n_1] \times \dots \times [n_{k_1}] \ \text{ and } \ B_k = [n_{k_1+1}] \times \dots \times [n_k], \tag{15.1}$$

and for every tensor $F\colon [n_1] \times \dots \times [n_k] \to \{0,1\}$ let the *respective matrix* f_F of F be the matrix $f_F\colon A_k \times B_k \to \{0,1\}$ defined by the rule

$$f_F\big((i_1,\dots,i_{k_1}),(i_{k_1+1},\dots,i_k)\big) = F(i_1,\dots,i_k) \tag{15.2}$$

for every $\big((i_1,\dots,i_{k_1}),(i_{k_1+1},\dots,i_k)\big) \in A_k \times B_k = [n_1] \times \dots \times [n_k]$.

As in [COCF10], we extend the notion of L_p regularity from matrices to tensors as follows.

Definition 15.1 (L_p-weakly regular tensors). *Let $0 < \eta \leqslant 1, C \geqslant 1$ and $1 \leqslant p \leqslant \infty$. A tensor $F\colon [n_1]\times\dots\times[n_k]$ is called* (C,η,p)-weakly regular *if its respective matrix f_F is (C,η,p)-regular, that is, if for every partition $\mathcal{P}$ of $A_k \times B_k$ with $\mathcal{P} \subseteq \mathcal{S}_{A_k\times B_k}$ and $\iota(\mathcal{P}) \geqslant \eta$ we have $\|\mathbb{E}(f_F \,|\, \mathcal{A}_{\mathcal{P}})\|_{L_p} \leqslant C$.*

To state our main result about L_p regular tensors we need to introduce some numerical invariants. Specifically, let $\varepsilon > 0$ and $C \geqslant 1$. Also let $1 < p \leqslant \infty$, set $p^\dagger = \min\{2,p\}$ and let q denote the conjugate exponent of $p^\dagger$. Finally, let a_1, a_2 be as in Theorem 14.2, and define

$$\tau(\varepsilon, C, p) = \left\lceil \frac{a_1\, C^2}{(p^\dagger - 1)\,\varepsilon^2} \right\rceil \quad\text{and}\quad \eta(\varepsilon, C, p) = \Big(\frac{a_2\,\varepsilon}{C}\Big)^{\sum_{i=1}^{\tau(\varepsilon,C,p)+1} (\frac{2}{p^\dagger}+1)^{i-1} q^i}. \tag{15.3}$$

We have the following theorem.

THEOREM 15.2. *There exist a constant b, an algorithm and a polynomial Π_3 such that the following holds. Let $0 < \varepsilon < 1/2$ and $C \geqslant 1$. Also let $1 < p \leqslant \infty$, and let $\tau = \tau(\varepsilon/2, C, p)$ and $\eta = \eta(\varepsilon/2, C, p)$ be as in (15.3). If we input*

`INP`: *a (C, η, p)-weakly regular tensor $F\colon [n_1] \times \cdots \times [n_k] \to \{0, 1\}$,*

then the algorithm outputs

`OUT`: *cut tensors $G_1, \ldots, G_s$ with $s \leqslant \Big(\dfrac{2b\,C}{\varepsilon\,\eta^2}\Big)^{2(k-1)}$ and such that*

$$\Big\|F - \sum_{i=1}^{s} G_i\Big\|_\square \leqslant \varepsilon \|F\|_\square \quad \textit{and} \quad \sum_{i=1}^{s} \|G_i\|_{L_\infty}^2 \leqslant \Big(\frac{C\,\|F\|_{L_1}}{\eta^2}\Big)^2 b^{2k}. \tag{15.4}$$

Moreover, this algorithm has running time $\big(\tau\, 4^\tau + \big(\frac{2C}{\varepsilon\eta^2}\big)^{3k}\big) \cdot \Pi_3\big(\prod_{i=1}^k n_i\big)$.

Theorem 15.2 can be proved arguing precisely as in the proof of [COCF10, Theorem 2] and using Theorem 14.2 instead of [COCF10, Corollary 1]. We leave the details to the interested reader.

15.2. MAX-CSP instances approximation

It is well known that it is NP-hard not only to compute the optimal solution for the MAX-CSP problem, but also to find "good" approximations of this optimal solution (see, e.g., [Hås01, KKMO07, TSSW00]). We will show that such approximations may be computed in polynomial time if we assume some additional properties for the given MAX-CSP problem (see also [FK99, COCF10]). In what follows let n, k denote two positive integers with $k \leqslant n$.

Let $V = \{x_1, \ldots, x_n\}$ be a set of Boolean variables, and recall that an *assignment* σ on V is a map $\sigma\colon V \to \{0, 1\}$. Notice that if σ is an assignment on V and $W \subseteq V$, then $\sigma|_W\colon W \to \{0, 1\}$ is an assignment on W. Also recall that a *k-constraint* is a pair (ϕ, V_ϕ) where $V_\phi \subseteq V$ with $|V_\phi| = k$ and $\phi\colon \{0, 1\}^{V_\phi} \to \{0, 1\}$ is a not identically zero map. Finally, recall that a *k-CSP instance* over V is a family $\mathcal{F}$ of k-constraints over V.

For every k-CSP instance $\mathcal{F}$ we define

$$\mathrm{OPT}(\mathcal{F}) = \max_{\sigma \in \{0,1\}^V} \sum_{(\phi, V_\phi) \in \mathcal{F}} \phi(\sigma|_{V_\phi}). \tag{15.5}$$

Moreover, let Ψ_k be the set of all non-zero maps $\{0,1\}^k \to \{0,1\}$. We have the following definition.

DEFINITION 15.3. *Let $\psi \in \Psi_k$. Also let (ϕ, V_ϕ) be a k-constraint over V where $V_\phi = \{x_{i_1}, \dots, x_{i_k}\}$ for some $1 \leqslant i_1 < \cdots < i_k \leqslant n$. We say that (ϕ, V_ϕ) is of* type ψ *if for every assignment $\sigma \colon V \to \{0,1\}$ we have*

$$\psi\big(\sigma(x_{i_1}), \dots, \sigma(x_{i_k})\big) = \phi(\sigma|_{V_\phi}).$$

Observe that every k-CSP instance $\mathcal{F}$ can be represented by a family $(F^\psi_{\mathcal{F}})_{\psi \in \Psi_k}$ of $2^{2^k} - 1$ tensors where for every $\psi \in \Psi_k$ the tensor $F^\psi_{\mathcal{F}} \colon [n]^k \to \{0,1\}$ is defined by the rule

$$F^\psi_{\mathcal{F}}(i_1, \dots, i_k) = \begin{cases} 1 & \text{if there is } (\phi, V_\phi) \in \mathcal{F} \text{ of type } \psi \\ & \text{with } V_\phi = \{x_{i_1}, \dots, x_{i_k}\}, \\ 0 & \text{otherwise.} \end{cases} \tag{15.6}$$

Having this representation in mind, we say that a k-constraint $\mathcal{F}$ is (C, η, p)-weakly regular for some $0 < \eta \leqslant 1$, $C \geqslant 1$ and $1 \leqslant p \leqslant \infty$, provided that for every $\psi \in \Psi_k$ the tensor $F^\psi_{\mathcal{F}}$ defined above is (C, η, p)-weakly regular.

We have the following theorem which extends [COCF10, Theorem 3]. It follows from Theorem 15.2 using the arguments in the proof of [COCF10, Theorem 3]; as such, its proof is left to the reader.

THEOREM 15.4. *There exist an algorithm, a constant $\gamma > 0$ and a polynomial Π_4 such that the following holds. Let k be a positive integer, and let $0 < \varepsilon < 1/2$, $C \geqslant 1$ and $1 < p \leqslant \infty$. Set $a = \varepsilon\, 2^{-(2^k + 2k + 2)}$, and let $\tau = \tau(a, C, p)$ and $\eta = \eta(a, C, p)$ be as in (15.3). If we input*

`INP`: *a (C, η, p)-weakly regular k-CSP instance $\mathcal{F}$ over a set $V = \{x_1, \dots, x_n\}$ of Boolean variables,*

then the algorithm outputs

`OUT`: *an assignment $\sigma \colon V \to \{0,1\}$ such that*

$$\sum_{(\phi, V_\phi) \in \mathcal{F}} \phi(\sigma|_{V_\phi}) \geqslant (1 - \varepsilon) \cdot \mathrm{OPT}(\mathcal{F}).$$

Moreover, this algorithm has running time

$$\Pi_4\bigg(n^k \cdot \exp\Big(k\, 2^k\, 2^{2^k} \big(\frac{2C}{\varepsilon\, \eta^2}\big)^{2k} \ln\big(\frac{2C}{\varepsilon\, \eta^2}\big)\Big)\bigg).$$

Appendices

APPENDIX A

Analytic inequalities

Through the rest of this chapter (X, Σ, μ) will denote a probability space and L_p will denote the space $L_p(X, \Sigma, \mu)$, for every $1 < p \leqslant \infty$.

A.1. A uniform convexity inequality

Our aim in this section is to show the following proposition (see, e.g. [Nao04]).

Proposition A.1. *Let $1 < p \leqslant 2$ and $f, g \in L_p$. Then*

$$\|f\|_{L_p}^2 + (p-1)\|g\|_{L_p}^2 \leqslant \frac{\|f+g\|_{L_p}^2 + \|f-g\|_{L_p}^2}{2}. \tag{A.1}$$

The proof of the previous inequality is a straightforward consequence of two well known analytic inequalities, the Bonami-Beckner "two point" inequality ([Gar07, Proposition 13.1.1]) and Hanner's inequality ([Nao04]). We present them here for the convenience of the reader.

Theorem A.2 (Bonami-Beckner inequality). *Let $1 < p_1 \leqslant p_2 < \infty$ and $x, y \in \mathbb{R}$. Then,*

$$\Big(\frac{1}{2}(|x + r_{p_2} y|^{p_2} + |x - r_{p_2} y|^{p_2})\Big)^{1/p_2} \leqslant \Big(\frac{1}{2}(|x + r_{p_1} y|^{p_1} + |x - r_{p_1} y|^{p_1})\Big)^{1/p_1}, \tag{A.2}$$

where for every $1 < p < \infty$, $r_p = 1/\sqrt{p-1}$. More specifically, for every $1 < p \leqslant 2$, we have that

$$\big(x^2 + (p-1)y^2\big)^{1/2} \leqslant \Big(\frac{|x+y|^p + |x-y|^p}{2}\Big)^{1/p}. \tag{A.3}$$

Theorem A.3 (Hanner's inequality). *Let $1 < p \leqslant 2$ and $f, g \in L_p$. Then*

$$\big|\|f\|_{L_p} - \|g\|_{L_p}\big|^p + \big(\|f\|_{L_p} + \|g\|_{L_p}\big)^p \leqslant \|f+g\|_{L_p}^p + \|f-g\|_{L_p}^p. \tag{A.4}$$

We are now ready to prove Theorem A.1.

Proof of Theorem A.1. We have that

$$\begin{aligned}\left(\frac{\|f+g\|_{L_p}^2+\|f-g\|_{L_p}^2}{2}\right)^{1/2} &\geqslant \left(\frac{\|f+g\|_{L_p}^p+\|f-g\|_{L_p}^p}{2}\right)^{1/p}\\ &\overset{(A.4)}{\geqslant} \left(\frac{\left(\|f\|_{L_p}+\|g\|_{L_p}\right)^p+\left|\,\|f\|_{L_p}-\|g\|_{L_p}\right|^p}{2}\right)^{1/p}\\ &\overset{(A.3)}{\geqslant} \left(\|f\|_{L_p}^2+(p-1)\,\|g\|_{L_p}^2\right)^{1/2}\end{aligned}$$

and the proof of Theorem A.1 is completed. □

A.2. A martingale difference sequence inequality

We will now prove Proposition 2.1[1]. We restate it here for the convenience of the reader.

Proposition 2.1. 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

$$\Big(\sum_{i=0}^n \|d_i\|_{L_p}^2\Big)^{1/2} \leqslant \Big(\frac{1}{p-1}\Big)^{1/2}\Big\|\sum_{i=0}^n d_i\Big\|_{L_p}. \tag{A.5}$$

The proof of Proposition 2.1 follows directly from the following lemma whose proof is based on an elegant pseudo-differentiation argument and is due to Ricard and Xu (see [RX16]).

Lemma A.4. *Let $f\in L_p$ and $\mathcal{G}$ be a sub-σ-algebra of Σ. Then,*

$$\|\mathbb{E}(f\,|\,\mathcal{G})\|_{L_p}^2+(p-1)\|f-\mathbb{E}(f\,|\,\mathcal{G})\|_{L_p}^2\leqslant\|f\|_{L_p}^2. \tag{A.6}$$

Let's see first how this Lemma implies Proposition 2.1.

Proof of Proposition 2.1. By iteration of (A.6) we obtain that

$$\|d_0\|_{L_p}^2+(p-1)\sum_{i=1}^n\|d_i\|_{L_p}^2\leqslant\Big\|\sum_{i=0}^n d_i\Big\|_{L_p}^2.$$

But $p\leqslant 2$ and thus we have

$$(p-1)\sum_{i=0}^n\|d_i\|_{L_p}^2\leqslant\|d_0\|_{L_p}^2+(p-1)\sum_{i=1}^n\|d_i\|_{L_p}^2,$$

which completes the proof of Proposition 2.1. □

[1]For a more instructive yet far more lengthy and with a worst contant proof of the previous inequality the reader may refer (and is encouraged to do so) to [Pis16].

It remains to prove Lemma A.4.

Proof of Lemma A.4. The proof is based on a pseudo-differentiation argument. Set $a = \mathbb{E}(f \,|\, \mathcal{G})$ and $b = f - \mathbb{E}(f \,|\, \mathcal{G})$. Define the function $F\colon [0,1] \to \mathbb{R}$ by the rule $F(t) = \|a + tb\|_{L_p}^2 + (p-1)t^2\|b\|_{L_p}^2$, for every $t \in [0,1]$. Also, for every real continuous function ϕ defined on an interval I of $\mathbb{R}$ recall that its pseudo-derivative of second order at $t \in I$ is

$$D^2\phi(t) = \liminf_{h \to 0^+} \frac{\phi(t+h) + \phi(t-h) - 2\phi(t)}{h^2}.$$

Also recall that if $D^2\phi \geqslant 0$, then ϕ is convex.

Fact A.5. *The function F is convex.*

Proof of Fact A.5. Let $h > 0$ and $t \in \mathbb{R}$. Applying Proposition A.2, for $f = a/h + tb/h$ and $g = b$ we obtain that

$$\frac{F(t+h) + F(t-h) - 2F(t)}{h^2} \geqslant 0.$$

Hence, $D^2F \geqslant 0$ and thus F is convex. □

Define the function $G(t) = \|a + tb\|_{L_p}^2$, for every $t \in [0,1]$. Since $\mathbb{E}(\cdot \,|\, \mathcal{G})$ is a contraction on L_p we have that

$$\|a + tb\|_{L_p} \geqslant \|\mathbb{E}(a + tb \,|\, \mathcal{G})\|_{L_p} = \|a\|_{L_p}.$$

Also, since G is convex its right-derivative G'_+ exists and by the previous inequality we have that $G'_+(0) \geqslant 0$ and thus $F'_+(0) = G'_+(0) \geqslant 0$ too. Thus, F is increasing and hence $F(0) \leqslant F(1)$. This completes the proof of Lemma A.4. □

APPENDIX B

Analytic number theory backround

B.1. Prime number theorems

Recall that $\pi(n) = |\{p \in \mathbf{P} \colon p \leqslant n\}|$ for every positive integer n. Then,

THEOREM B.1 (Prime number theorem). *Let N be a large positive integer, then*

$$\pi(N) = (1 + o_{N\to\infty}(1))\frac{N}{\log N}.$$

The previous theorem is a celebrated result first proved in 1896 independently by J.Hadamard and C.J. de la Valle-Poussin. For a proof of this result see [Apo76, Chapter 13].

The following result was first proved by P.G.L Dirichlet and is sometimes referred to as the Dirichlet's prime number theorem.

THEOREM B.2 (Dirichlet's theorem). *Let a, q be coprime. Then there exist infinitely many primes of the form $a + nq$.*

For a proof see [Apo76, Chapter 7]. Observe that if for some a, q we have that $\gcd(a, q) > 1$ then there is no prime of the form $a + nq$, for $n \geqslant 1$.

Closing this section we present the Bertrand-Chebysev theorem, see [AZHE10, Chapter 2]. It states the following

THEOREM B.3 (Bertrand-Chebysev theorem). *For every $n \geqslant 1$, there exists at least one $p \in \mathbf{P}$ such that $n \leqslant p \leqslant 2n$.*

B.2. Arithmetic functions

An arithmetic (or arithmetical) function is a real (or complex) function defined on the set of natural numbers. An arithmetic function f is called *multiplicative* if $f(nm) = f(n) \cdot f(m)$, for all coprime natural numbers n, m. If, $f(1) = 1$ and $f(nm) = f(n) \cdot f(m)$ for all natural numbers n, m, regardless if they are coprime or not, then the function f will be called *completely mulptiplicative*. In the rest of this section we will present some well known arithmetic functions.

B.2.1. The Möbius function μ. The Möbius function μ is the arithmetic function defined by the rule

$$\mu(n) = \begin{cases} 1, & \text{if } n \text{ is square-free with an even number of prime factors,} \\ -1, & \text{if } n \text{ is square-free with an odd number of prime factors,} \\ 0, & \text{otherwise.} \end{cases}$$

Observe that μ is a multiplicative function. For the Möbius function we have the following proposition (for a full proof see [Apo76]) which is known as the Möbius inversion formula.

PROPOSITION B.4 (Möbius inversion formula). *Let f, g be two arithmetic functions such that*

$$g(n) = \sum_{d|n} f(d),$$

for every positive integer n. Then, for every positive integer n

$$f(n) = \sum_{d|n} \mu(d) g(n/d).$$

B.2.2. The Euler totient function ϕ. The Euler totient function ϕ is the arithmetic multiplicative function defined by the rule

$$\phi(n) = |\{k \colon 1 \leqslant k \leqslant n \text{ and } \gcd(k,n) = 1\}|,$$

for every positive integer n. The following identity (for a proof see [Apo76]) is known as the Euler's product formula.

PROPOSITION B.5 (Euler's product formula). *For every positive integer n we have*

$$\phi(n) = n \prod_{p|n} \big(1 - \frac{1}{p}\big).$$

B.2.3. The Von Mangoldt function Λ. The Von Mangoldt function Λ is the *non muptiplicative* arithmetic function defined by the rule

$$\Lambda(n) = \begin{cases} \log p, & \text{if } n = p^k \text{ for some } p \in \mathbf{P} \text{ and some positive integer } k, \\ 0 \;, & \text{otherwise.} \end{cases}$$

By a straightforward calculation one may see that for every positive integer n

$$\log(n) = \sum_{d|n} \Lambda(n)$$

and thus by Proposition B.4,

$$\Lambda(n) = \sum_{d|n} \mu(d) \log(n/d). \tag{B.1}$$

Another important property of the Von Mangoldt function is the following proposition (for a proof see [Apo76]), which is in fact equivalent to the prime number theorem.

Proposition B.6. *Let N be a large positive integer. Then*

$$\sum_{n\in[N]} \Lambda(n) = (1+o(1))N.$$

B.2.4. The restriction of the Von Mangoldt function in the primes $\widetilde{\Lambda}$. The restriction of the Von Mangoldt function Λ in the primes is the arithmetical function $\widetilde{\Lambda}$ defined by the rule

$$\widetilde{\Lambda}(n) = \begin{cases} \log n, & \text{if } n \in \mathbf{P}, \\ \quad 0 \quad, & \text{otherwise}, \end{cases}$$

i.e. $\widetilde{\Lambda}(n) = \mathbf{1}_{\mathbf{P}}(n)\Lambda(n)$. This function has similar properties with Λ. For example we have the following proposition

Proposition B.7. *Let N be a large positive integer. Then*

$$\sum_{n\in[N]} \widetilde{\Lambda}(n) = (1+o_{N\to\infty}(1))N.$$

A quantitative version of the previous proposition is the following theorem of C.L.Siegel and A.Walfisz.

Theorem B.8 (Siegel-Walfisz Theorem). *Let $\varepsilon > 0$ and m be a positive integer. Also, let q, a be positive integers with $\gcd(q,a) = 1$ and $q \leqslant (\log m)^{1-\varepsilon}$. Then, there exists a constant c such that*

$$\sum_{\substack{n\in[m] \\ n\equiv a \mathrm{mod} q}} \widetilde{\Lambda}(n) = \frac{m}{\phi(q)} + mO_\varepsilon\big(\exp(-c\sqrt{\log m})\big).$$

For a proof of the previous result see [Dav00, Chapter 20].

B.3. Euler products

We present now a useful result about arithmetic functions which proof can be found in many textbooks, see e.g. [Apo76, Theorem 11.6]

THEOREM B.9. *Let f be a multiplicative function, $s \in \mathbb{C}$ and assume that the*

$$\sum_{n=1}^{\infty} \left|\frac{f(n)}{n^s}\right| < \infty.$$

Then,

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \sum_{k \in \mathbb{N}} \frac{f(p^k)}{p^{ks}}. \tag{B.2}$$

Expressions of the form of the (RHS) of (B.2) are known as Euler products. A straightforward consequence of the previous theorem is the following proposition.

PROPOSITION B.10. *Let $f\colon \mathbb{N}^d \to \mathbb{C}$ be a multiplicative function in each coordinate, i.e.*

$$f(n_1, \ldots, n_i m_i, \ldots, n_d) = f(n_1, \ldots, n_i, \ldots, n_d) \cdot f(n_1, \ldots, m_i, \ldots, n_d)$$

for every $i \in [d]$ and for every n_i, m_i such that $\gcd(n_i, m_i) = 1$. Also, let $s \in \mathbb{C}$. Then, assuming that

$$\sum_{n_1, \ldots, n_d = 1}^{\infty} \left|\frac{f(n_1, \ldots, n_d)}{(n_1 \ldots n_d)^s}\right| < \infty$$

we have

$$\sum_{n_1, \ldots, n_d = 1}^{\infty} \frac{f(n_1, \ldots, n_d)}{(n_1 \ldots n_d)^s} = \prod_p \sum_{m_1, \ldots, m_d \in \mathbb{N}} \frac{f(p^{m_1}, \ldots, p^{m_d})}{p^{\sum_{i=1}^d m_i s}}.$$

B.4. The Chinese remainder theorem

The classical Chinese remainder theorem states that for every positive integers $m_1, \ldots, m_t$ and every $a_1, \ldots, a_t \in \mathbb{Z}$ the system of equations

$$\begin{cases} x \equiv a_1 \mathrm{mod} m_1, \\ \vdots \\ x \equiv a_t \mathrm{mod} m_t \end{cases}$$

is solvable if and only if $a_i \equiv a_j \mathrm{mod}\, \gcd(m_i, m_j)$ for every $i, j \in [t]$ with $i \neq j$. Furthermore, this solution is unique modulo $\mathrm{lcm}(m_1, \ldots, m_t)$. This theorem implies the following result which is known as the Chinese remainder theorem of group theory.

THEOREM B.11 (Chinese remainder theorem-Group theory). *Let $p_1, \ldots, p_s$ be distinct primes and $m = \prod_{l=1}^s p_l$. Then, there exists an group isomorphism between $\mathbb{Z}_m$ and $\bigoplus_{l=1}^s \mathbb{Z}_{p_l}$, where $\bigoplus$ denotes the direct sum of groups.*

From the previous theorem we obtain the following proposition.

Proposition B.12. *Let d be a positive integer, $L_0, \dots, L_d \in \mathbb{Z}$ and $p_1, \dots, p_s$ be distinct primes. Also, let $a_1, \dots, a_d \in \mathbb{Z}$. Finally, let $\psi \colon \mathbb{Z}^d \to \mathbb{Z}$ be an affine linear form defined by the rule*

$$\psi(\mathbf{x}) = \sum_{i=1}^{d} L_i x_i + L_0, \tag{B.3}$$

for every $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{Z}^d$ and for every $i \in [s]$ let $\mathbf{x}_i \in \mathbb{Z}_{p_i}^d$ such that

$$\psi(\mathbf{x}_i) \equiv a_i \mathrm{mod} p_i. \tag{B.4}$$

Then, if $D = \prod_{i\in[s]} p_i$ there exists a unique $\mathbf{y} \in \mathbb{Z}_D^d$ such that

$$\psi(\mathbf{y}) \equiv a_i \mathrm{mod} p_i,$$

for every $i \in [s]$. If in addition, $a_1 = \cdots = a_s = 0$ then there exists a unique $\mathbf{y} \in \mathbb{Z}_D^d$ such that

$$\psi(\mathbf{y}) \equiv 0 \mathrm{mod} D.$$

Proof. For every $i \in [s]$, there exist $\mathbf{x}_{i,1}, \dots, \mathbf{x}_{i,d} \in \mathbb{Z}_{p_i}^d$ such that $\mathbf{x}_i = (\mathbf{x}_{i,1}, \dots, \mathbf{x}_{i,d})$. and thus by (B.3) and (B.4) we have that

$$\begin{cases} \sum_{i=1}^d L_i \cdot x_{1,i} + L_0 \equiv a_1 \mathrm{mod} p_1 \\ \sum_{i=1}^d L_i \cdot x_{2,i} + L_0 \equiv a_2 \mathrm{mod} p_2 \\ \vdots \\ \sum_{i=1}^d L_i \cdot x_{s,i} + L_0 \equiv a_s \mathrm{mod} p_s. \end{cases}$$

Moreover, by Theorem B.11 there exist unique $y_1, \dots, y_d \in \mathbb{Z}_D$ such that for every $i \in [d], j \in [s]$

$$y_i = x_{j,i} \mathrm{mod} p_j.$$

Thus, setting $\mathbf{y} = (y_1, \dots, y_d)$ and using the linearity of the modulo operation we see that

$$\psi(\mathbf{y}) \equiv a_i \mathrm{mod} p_i,$$

for every $i \in [s]$. If now $a_1 = \cdots = a_s = 0$ for the previous $\mathbf{y}$ we have that $\psi(\mathbf{y}) \equiv 0 \mathrm{mod} p_i$, for every $i \in [s]$ and hence $\psi(\mathbf{y}) \equiv 0 \mathrm{mod} D$. This completes the proof of the lemma. $\square$

B.5. The Riemann ζ function

Recall that the Riemman ζ function is defined for every $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1/2$ by the rule

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Although this function has been studied a lot we will only need two basic properties. The first one is the following lemma which shows that ζ has a simple pole at 1 with residue 1.

PROPOSITION B.13. *If* $\mathrm{Re}(s) > 1$ *and* $s = O(1)$, *then* $\zeta(s) = \frac{1}{s-1} + O(1)$.

PROOF. Since,

$$\frac{1}{s-1} = \int_1^\infty \frac{dx}{x^s} = \sum_{n=1}^\infty \int_n^{n+1} \frac{dx}{x^s}$$

we have that

$$\zeta(s) - \frac{1}{s-1} = \sum_{n=1}^\infty \int_n^{n+1} (\frac{1}{n^s} - \frac{1}{x^s})dx.$$

From the mean value theorem and the hypotheses that $s = O(1), \mathrm{Re}(s) > 1$ we have $\frac{1}{n^s} - \frac{1}{x^s} = O(\frac{1}{n^2})$. Indeed, for every $n \in \mathbb{N}$ and every $x \in [n, n+1]$ we have

$$|n^{-s} - x^{-s}| = \big|s \int_n^x y^{-1-s} dy\big| \leqslant |s| n^{-1-\Re(s)} \leqslant |s| n^{-2}$$

and thus $\frac{1}{n^s} - \frac{1}{x^s} = O(\frac{1}{n^2})$. Therefore,

$$\zeta(s) - \frac{1}{s-1} = \sum_{n=1}^\infty O(\frac{1}{n^2}) = O(\sum_{n=1}^\infty \frac{1}{n^2}) = O(1)$$

and the proof of the lemma is complete. □

The second basic property of the Riemann ζ function is the following amalgamation of Proposition B.13 and Theorem B.9.

PROPOSITION B.14. *Let* $s \in \mathbb{C}$ *with* $\mathrm{Re}(s) > 1$ *and* $s = O(1)$. *Then,*

$$\prod_p (1 - \frac{1}{p^s})^{-1} = \frac{1}{s-1} + O(1).$$

PROOF. By Lemma B.13 we have

$$\zeta(s) = \frac{1}{s-1} + O(1)$$

and by Theorem B.9 we have

$$\zeta(s) = \prod_p (1 - \frac{1}{p^s})^{-1}.$$

Putting the two previous together we have the desired result. □

APPENDIX C

The Goldston–Yildirim estimate

C.1. Backround material

C.1.1. Sieve factors. Throughout this subsection let $\chi\colon \mathbb{R} \to \mathbb{R}$ be a smooth and compactly supported function.

Recall that the modified Fourier transform φ of χ is defined by the formula

$$e^{x}\chi(x) = \int_{-\infty}^{+\infty} \varphi(\xi) e^{-\mathrm{i}x\xi} d\xi. \tag{C.1}$$

One important property of the modified Fourier transform is the fact that it decreases rapidly. More precisely we have the following proposition (see [SS03, Chapter 5, Theorem 1.3])

Proposition C.1. *Let φ be the modified Fourier transform of χ. Then, for every $\xi \in \mathbb{R}$ and every $A > 0$ we have*

$$|\varphi(\xi)| = O_A\big((1+\xi)^{-A}\big).$$

We are about now to define the notion of *sieve factors.*

Definition C.2 (Sieve factors,[GT10]). *Let $a \in \mathbb{N}$ with $a \geqslant 1$. Then, the sieve factor of χ with parameter a is the quantity*

$$c_{\chi,a} = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \prod_{B \subseteq [a]} \big(\sum_{j \in B} (1 + \mathrm{i}\xi_j) \big)^{(-1)^{|B|-1}} \prod_{j=1}^{a} \varphi(\xi_j) d\xi_j,$$

where φ is the modified Fourier transform of χ.

Despite the fact that sieve factors look very complicated to estimate, for some particular choices of a they take a rather simple form. To be more specific, we have (see [GT10, Lemma D.2])

$$c_{\chi,1} = -\chi'(0) \tag{C.2}$$

and

$$c_{\chi,2} = \int_{-\infty}^{+\infty} |\chi'(x)|^2 dx. \tag{C.3}$$

Moreover, for every choice of a we have that $c_{\chi,a} \in \mathbb{R}$.

C.1.2. Systems of affine linear forms. Recall now that an affine linear form (or affine linear map) on $\mathbb{Z}^d$, for some positive integer d is a function $\psi\colon \mathbb{Z}^d \to \mathbb{Z}$ of the form $\psi = \overline{\psi} + \psi(\mathbf{0})$, where $\overline{\psi}\colon \mathbb{Z}^d \to \mathbb{Z}$ is a linear form and $\psi(\mathbf{0}) \in \mathbb{Z}$. Also recall that two affine linear forms $\psi_1, \psi_2\colon \mathbb{Z}^d \to \mathbb{Z}$ are called affinely related if the linear maps $\psi_1 - \psi_1(\mathbf{0})$ and $\psi_2 - \psi_2(\mathbf{0})$ are parallel.

Next recall that a system of affine linear forms $\Psi = (\psi_1, \dots, \psi_t)$ is a t-tuple of affine linear forms for some positive integer t. The previous system of affine linear forms may be seen as an affine linear map from $\mathbb{Z}^d$ to $\mathbb{Z}^t$, i.e. $\Psi = \overline{\Psi} + \Psi(\mathbf{0})$, where $\overline{\Psi}\colon \mathbb{Z}^d \to \mathbb{Z}^t$ is a linear map and $\Psi(\mathbf{0}) \in \mathbb{Z}^t$. From now on, in order to avoid degeneracies, we will assume that if we have a system Ψ as before none of the ψ_is is constant. For any system of affine linear forms we define its *size* as follows.

Definition C.3. *Let d, t, N be positive integers and $\Psi\colon \mathbb{Z}^d \to \mathbb{Z}^t$ be a system of affine linear forms. Then, we define the* size $\|\Psi\|_N$ *of Ψ with the respect to the scale parameter N by the rule*

$$\|\Psi\|_N = \sum_{i=1}^{t} \sum_{j=1}^{d} |\overline{\psi}_i(e_j)| + \sum_{i=1}^{t} \Big|\frac{\psi_i(0)}{N}\Big|, \tag{C.4}$$

where $e_1, \dots, e_d$ is the standard basis of $\mathbb{Z}^d$.

Observe that the size of a system of affine linear forms Ψ is a decreasing function of the scale parameter. More precisely, if N_1, N_2 are positive integers with $N_1 \leqslant N_2$ then $\|\Psi\|_{N_1} \geqslant \|\Psi\|_{N_2}$. We also have the following definition.

Definition C.4. *Let d, t, q be positive integers and $\Psi = (\psi_1, \dots, \psi_t)\colon \mathbb{Z}^d \to \mathbb{Z}^t$ be a system of affine linear forms. Then, we define the set*

$$\mathrm{C}(\Psi, q) = \Big\{n \in \mathbb{Z}_q^d\colon \prod_{i \in [t]} \gcd\big(\psi_i(n), q\big) = 1\Big\}.$$

In the previous expression we induce the affine forms $\psi_i\colon \mathbb{Z}_q^d \to \mathbb{Z}$ from their global counterparts $\psi_i\colon \mathbb{Z}^d \to \mathbb{Z}$ in the obvious way.

C.1.3. Local factors. We define now the so-called *local factors* and isolate some of their basic properties.

Definition C.5 (Local factors, [GT10]). *Let d, t be two positive integers. Also let $\Psi = (\psi_1, \dots, \psi_t)\colon \mathbb{Z}^d \to \mathbb{Z}^t$ be a system of affine linear forms. For every positive integer q the q-local factor of Ψ is defined by the rule*

$$\beta_{\Psi,q} = \Big(\frac{q}{\phi(q)}\Big)^t\ \mathbb{E}_{n \in \mathbb{Z}_q^d} \mathbf{1}_{C(\Psi,q)},$$

where $C(\Psi, q)$ is as in Definition C.4 and ϕ is the Euler totient function (see Appendix B, Subsection B.2.2.). More specifically, if $q \in \mathbf{P}$ then

$$\beta_{\Psi,q} = \Big(\frac{q}{q-1}\Big)^t \, \mathbb{E}_{n \in \mathbb{Z}_q^d} \mathbf{1}_{C(\Psi,q)}. \tag{C.5}$$

Notice at this point that if q is a positive integer and Ψ is a system of affine linear forms by the Chinese remainder theorem (see Appendix B, Section B.4) we have

$$\beta_{\Psi,q} = \prod_{\substack{p \in \mathbf{P}, \\ p \mid q}} \beta_{\Psi,p}. \tag{C.6}$$

We finally have the following lemma.

Lemma C.6. *Let t, d, L be positive integers and $\Psi = (\psi_1, \dots, \psi_t)$ be a system of affine linear forms from $\mathbb{Z}^d$ to $\mathbb{Z}$ with $\|\Psi\|_1 \leqslant L$. Also let $p \in \mathbf{P}$. Then $\beta_{\Psi,p} = 1 + O(1/p)$. If in addition no two of the forms $\psi_1, \dots, \psi_t$ are affinely related then we have that $\beta_{\Psi,p} = 1 + O(1/p^2)$. The implied constants depend on d, t, L.*

Proof. Let n be selected uniformly at random from $\mathbb{Z}_p^d$. Then, $\mathbf{1}_{C(\Psi,q)}(n) = 1$ with probability $1 - O_t(1/p)$. Moreover it is easy to observe that

$$\Big(\frac{p}{p-1}\Big)^t = 1 + O_t(1/p).$$

Combining the previous two estimations and (C.5) we have $\beta_{\Psi,p} = 1 + O(1/p)$. For the second part of the lemma assume that no two of the forms $\psi_1, \dots, \psi_t$ are affinely related. Then, it is easy to see that for every $1 \leqslant i < j \leqslant t$, ψ_i, ψ_j are not multiple of each other modulo p. Therefore , if n is selected uniformly at random from $\mathbb{Z}_p^d$ then p divides both $\psi_i(n), \psi_j(n)$ with probability $O(1/p^2)$. Then, using the inclusion-exclusion principle and working as in the proof of the first part of the Lemma we obtain that $\beta_{\Psi,p} = 1 + O(1/p^2)$. □

C.2. The Goldston–Yildirim correlation estimates

The following theorem is due to Green and Tao[GT10] who where based on the work of Goldston and Yildirim (see ,e.g,[GY, GY03, GY07]). Similar results may be found in [GT08, Tao06a, CFZ14].

Theorem C.7 (Goldston–Yildirim correlation estimate). *Let t, d, L be positive integers, N be a large positive integer and $\Psi = (\psi_1, \dots, \psi_t) \colon \mathbb{Z}^d \to \mathbb{Z}^t$ be a system of non-constant affine linear forms with $\|\Psi\|_1 \leqslant L$.*[1] *Let $a = (a_1, \dots, a_t) \in \mathbb{N}^t$ be a t-tuple of positive integers, $K \subseteq [-N, N]^d$ be a convex body and $\chi_1, \dots, \chi_t \colon \mathbb{R} \to [0, 1]$*

[1]In the original statement of Green–Tao they assume that $\|\Psi\|_N \leqslant L$. Since, this affects only the constants that arise, for argument clarity we take the size of Ψ with scale parameter 1.

be smooth and compactly supported functions. Also let $R = N^{\gamma}$, for some sufficiently small $\gamma = \gamma(t, d, L, \chi_1, \dots, \chi_t, \alpha) > 0$. Additionally for every $i \in [t]$ let Λ_{χ_i,R,a_i} be as in (12.8), let c_{χ_i,a_i} be the sieve factor of χ_i with parameter a_i, and for every $p \in \mathbf{P}$ let $\beta_{\Psi,p}$ be the p-local factor of Ψ. Finally, set

$$P_{\Psi} = \{p \in \mathbf{P} \colon \psi_i, \psi_j \textit{ are linearly dependent } \mathrm{mod} p \textit{ for some } i, j \in [t]\} \tag{C.7}$$

and $X = \sum_{p \in P_{\Psi}} p^{-1/2}$. Then,

$$\sum_{n \in K \cap \mathbb{Z}^d} \prod_{i \in [t]} \Lambda_{\chi_i,R,a_i}(\psi_i(n)) = \prod_{i \in [t]} c_{\chi_i,a_i} \cdot \mathrm{vol}(K) \cdot \prod_{p} \beta_{\Psi,p} + O\big(\frac{N^d}{\log^{1/20} R} e^{O(X)}\big), \tag{C.8}$$

where the implied constants depend on $t, d, L, \chi_1, \dots, \chi_t$ and a.

In subsection C.2.1 we present a sceleton of the proof of the previous theorem and in subsection C.2.2 we prove all the intermediate results which we used in this sceleton.

From now on all the implied constants will depend on the parameters t, d, L, $\chi_1, \dots, \chi_t$ and a or a subset of these parameters. Moreover μ will denote the Möbius function and ϕ will denote the Euler totient function (see Appendix B).

C.2.1. Sceleton of the proof of Theorem C.7. Before we enter the main part of the proof we need to write the (LHS) of (C.8) in a more manageable form. To this end for every $i \in [t]$ we set the *fibre* of i to be the set $\mathcal{F}_i = \{i\} \times [a_i]$ and define

$$\Omega = \{(i, j) \colon i \in [t], j \in [a_i]\} = \bigcup_{i \in [t]} \mathcal{F}_i \subseteq \mathbb{N}^2.$$

Then, we see that the (LHS) of (C.8) equals

$$\log^t R \sum_{\substack{(m_{i,j})_{(i,j) \in \Omega} \in \mathbb{N}^{\Omega} \\ m_{i,j} \text{ square-free}}} \Big(\prod_{(i,j) \in \Omega} \mu(m_{i,j}) \chi_i\big(\frac{\log m_{i,j}}{\log R}\big)\Big) \sum_{n \in K \cap \mathbb{Z}^d} \prod_{(i,j) \in \Omega} \mathbf{1}_{m_{i,j} | \psi_i(n)}.$$

Moreover, for every $i \in [t]$ we set $m_i = \mathrm{lcm}(m_{i,1}, \dots, m_{i,a_i})$ and observe that the previous expression may be rewritten as

$$\log^t R \sum_{\substack{(m_{i,j})_{(i,j) \in \Omega} \in \mathbb{N}^{\Omega} \\ m_{i,j} \text{ square-free}}} \Big(\prod_{(i,j) \in \Omega} \mu(m_{i,j}) \chi_i\big(\frac{\log m_{i,j}}{\log R}\big)\Big) \sum_{n \in K \cap \mathbb{Z}^d} \prod_{i \in [t]} \mathbf{1}_{m_i | \psi_i(n)}. \tag{C.9}$$

We enter now the main part of the proof.

Step 1: Elimination of the role of K and N. For every $m_1, \ldots, m_t$ as before we set $m = \operatorname{lcm}(m_1, \ldots, m_t)$ (also square-free) and also we set

$$\alpha_{m_1,\ldots,m_t} = \mathbb{E}_{n\in\mathbb{Z}_m^d} \prod_{i\in[t]} \mathbf{1}_{m_i|\psi_i(n)}. \tag{C.10}$$

for which we have the following claim.

Claim C.8.

$$\alpha_{m_1,\ldots,m_t} = \prod_{p\in\mathbf{P}} \alpha_{p^{r_{p,1}},\ldots,p^{r_{p,t}}}$$

where for every p and for every $i \in [t]$ we have $r_{p,i} = 1$ if $p|m_i$ and $r_{p,i} = 0$ otherwise.

We now have the following lemma.

Lemma C.9. *For every square-free integers $(m_{i,j})_{(i,j)\in\Omega}$ we have*

$$\sum_{n\in K\cap\mathbb{Z}^d} \prod_{i\in[t]} \mathbf{1}_{m_i|\psi_i(n)} = \operatorname{vol}(K)\alpha_{m_1,\ldots,m_t} + O(mN^{d-1}) \tag{C.11}$$

where $a_{m_1,\ldots,m_t}$ is as in (C.10).

Observe now that since χ_i's are compactly supported we have that $m_i \leqslant R^{O(1)}$ and thus $m \leqslant R^{O(1)}$. Therefore the contribution of the error term of (C.11) in (C.9) is $O(R^{O(1)}N^{d-1}\log^t R)$, which is $o(N^d)$ if γ in the definition of R is sufficiently small. Therefore it suffices to prove that

$$\log^t R \sum_{\substack{(m_{i,j})_{(i,j)\in\Omega}\in\mathbb{N}^\Omega \\ m_{i,j}\ \text{square-free}}} \Big(\prod_{(i,j)\in\Omega} \mu(m_{i,j})\chi_i\big(\frac{\log m_{i,j}}{\log R}\big)\Big)\alpha_{m_1,\ldots,m_t}$$
$$= \prod_{i\in[t]} c_{\chi_i,a_i} \prod_{p\in\mathbf{P}} \beta_{\Psi,p} + O(e^{O(X)}\log^{-1/20} R) \tag{C.12}$$

which is a genuinely simpler expression than (C.8) since the roles of K and N have been eliminated.

Step 2: Fourier expansion. In this step our goal is to replace the sum of the (LHS) of the previous expression by a product which is easier to cope with. To do this, at first we replace χ_is by integrals using the Fourier expansion we saw in (C.1). More precisely we have that for every square-free $m_{i,j}$

$$\chi_i\big(\frac{\log m_{i,j}}{\log R}\big) = \int_{\mathbb{R}} m_{i,j}^{-\frac{1+i\xi}{\log R}} \varphi_i(\xi)d\xi,$$

where φ_i's are as in (C.1). In order to simplify further the previous expression using the rapid decrease of the Fourier transform (Proposition C.1) and setting

$I = [-\log^{1/2} R, \log^{1/2} R]$ we have that for every $(i,j) \in \Omega$, every square-free $m_{i,j} \in \mathbb{N}$ and every $A > 0$, which we will choose later, we have

$$\chi_i\Big(\frac{\log m_{i,j}}{\log R}\Big) = \int_I m_{i,j}^{-\frac{1+\mathrm{i}\xi}{\log R}} \varphi_i(\xi) d\xi + O_A(m_{i,j}^{-1/\log R} \log^{-A} R). \tag{C.13}$$

Moreover, since every χ_i is Lipschitz continuous we have that for every $i \in [t]$ and every $m_{i,j}$, $\chi(\log m_{i,j}/\log R) = O(m_{i,j}^{-1/\log R})$. Therefore, by (C.13) we have

$$\prod_{(i,j)\in\Omega} \chi_i\Big(\frac{\log m_{i,j}}{\log R}\Big) = \int_I \cdots \int_I \prod_{(i,j)\in\Omega} m_{i,j}^{-z_{i,j}} \varphi_i(\xi_{i,j}) d\xi_{i,j} + O_A(\log^{-A} R \prod_{(i,j)\in\Omega} m_{i,j}^{-1/\log R}) \tag{C.14}$$

where $z_{i,j} = (1 + \mathrm{i}\,\xi_{i,j})/\log R$ for every $(i,j) \in \Omega$. For the error term of the previous expression we have the following lemma.

Lemma C.10. *There exists $A > 0$ such that*

$$\log^t R \sum_{\substack{(m_{i,j})_{(i,j)\in\Omega} \in \mathbb{N}^\Omega \\ m_{i,j} \text{ square-free}}} \Big(\prod_{(i,j)\in\Omega} \mu(m_{i,j}) \chi_i\Big(\frac{\log m_{i,j}}{\log R}\Big)\Big) \alpha_{m_1,\dots,m_t} O_A(\log^{-A} R \prod_{(i,j)\in\Omega} m_{i,j}^{-1/\log R})$$
$$= O(\log^{-1/20} R). \tag{C.15}$$

Thus we will have completed our proof as long as we show that

$$\log^t R \sum_{\substack{(m_{i,j})_{(i,j)\in\Omega} \in \mathbb{N}^\Omega \\ m_{i,j} \text{ square-free}}} \int_I \cdots \int_I \prod_{(i,j)\in\Omega} \mu(m_{i,j}) m_{i,j}^{-z_{i,j}} \alpha_{m_1,\dots,m_t} \varphi_i(\xi_{i,j}) d\xi_{i,j}$$
$$= \prod_{i\in[t]} c_{\chi_i,a_i} \prod_{p\in\mathbf{P}} \beta_{\Psi,p} + O(e^{O(X)} \log^{-1/20} R). \tag{C.16}$$

To this end, by exchanging sums and integrals, which can be done since I is compact and the summation is absolutely convergent[2], we see that the (LHS) of the previous equation equals

$$\log^t R \int_I \cdots \int_I \sum_{\substack{(m_{i,j})_{(i,j)\in\Omega} \in \mathbb{N}^\Omega \\ m_{i,j} \text{ square-free}}} \prod_{(i,j)\in\Omega} \mu(m_{i,j}) m_{i,j}^{-z_{i,j}} \alpha_{m_1,\dots,m_t} \varphi_i(\xi_{i,j}) d\xi_{i,j}$$

[2] One can use similar bounds with those that arise in the proof of Lemma C.10 (see Subsection C.2.2. below)

which in turn by the multiplicativity of α, which we saw in Claim C.8, may be written as an Euler product

$$\log^t R \int_t \cdots \int_t \prod_p E_{p,\xi} \cdot \prod_{(i,j)\in\Omega} \varphi_i(\xi_{i,j}) d\xi_{i,j},$$

where $\xi = (\xi_{i,j})_{(i,j)\in\Omega} \in I^\Omega$ and $E_{p,\xi}$ is the Euler factor

$$E_{p,\xi} = \sum_{(m_{i,j})_{(i,j)\in\Omega}\in\{1,p\}^\Omega} \Big(\prod_{(i,j)\in\Omega} \mu(m_{i,j}) m_{i,j}^{-z_{i,j}} \alpha_{m_1,\ldots,m_t} \Big). \tag{C.17}$$

These estimations along with (C.16) reduce further our task in to showing that

$$\log^t R \int_I \cdots \int_I \prod_{p\in\mathbf{P}} E_{p,\xi} \prod_{(i,j)\in\Omega} \varphi_i(\xi_{i,j}) d\xi_{i,j} = \prod_{i\in[t]} c_{\chi_i,a_i} \prod_{p\in\mathbf{P}} \beta_{\Psi,p} + O(e^{O(X)} \log^{-1/20} R). \tag{C.18}$$

In order to prove the previous equality we need to estimate the Euler product $\prod_p E_{p,\xi}$, which is the next step of the proof.

Step 3: The Euler product $\prod_p E_{p,\xi}$. We will "simplify" the Euler product $\prod_{p\in\mathbf{P}} E_{p,\xi}$ and to do so we first need to "simplify" the Euler factors $E_{p,\xi}$.

Lemma C.11 (Euler factor estimate). *Let $\xi = (\xi_{i,j})_{(i,j)\in\Omega} \in I^\Omega$, and let $p \in \mathbf{P}$. Set*

$$E'_{p,\xi} = \prod_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} \Big(1 - \frac{1}{p^{1+\sum_{(i,j)\in B} z_{i,j}}}\Big)^{|-1|^{|B|-1}}, \tag{C.19}$$

where for every $(i,j)\in\Omega$, $z_{i,j} = (1 + \mathrm{i}\,\xi_{i,j})/\log R$. Then, we have

$$E_{p,\xi} = \begin{cases} \big(1 + O(1/p^2)\big) E'_{p,\xi}, & \text{if } p > \log^{1/10} R \text{ and } p \notin P_\Psi, \\ \big(1 + O(1/p)\big) E'_{p,\xi}, & \text{if } p > \log^{1/10} R \text{ and } p \in P_\Psi \\ \big(\beta_{\Psi,p} + O(\frac{\log p}{\log^{1/2} R})\big) E'_{p,\xi}, & \text{if } p \leqslant \log^{1/10} R. \end{cases} \tag{C.20}$$

The previous lemma gives rise to the following one which completes *Step 3* of the proof of Theorem C.7.

Lemma C.12 (Euler product estimate). *For every $\xi \in I^\Omega$ we have*

$$\prod_{p\in\mathbf{P}} E_{p,\xi} = \Big(\prod_{p\in\mathbf{P}} \beta_{\Psi,p} + O(e^{O(X)} \log^{-1/20} R) \Big) \prod_{p\in\mathbf{P}} E'_{p,\xi},$$

where $E'_{p,\xi}$ is as in (C.19).

Step 4: Completion of the proof. We are ready now to prove (C.18). To do this we first have the following claim

Claim C.13. *For every $\xi = (\xi_{i,j})_{(i,j)\in\Omega} \in I^{\Omega}$ we have*

$$\prod_{p\in\mathbf{P}} E'_{p,\xi} = \big(1 + O(\log^{-1/2} R)\big) \prod_{\substack{B\subseteq\Omega \\ B \text{ vertical}}} \Big(\sum_{(i,j)\in B} z_{i,j}\Big)^{(-1)^{|B|-1}},$$

where $E'_{p,\xi}$ is as in (C.19) and for every $(i,j)\in\Omega$, $z_{i,j} = (1 + \mathrm{i}\,\xi_{i,j})/\log R$.

Moreover,

$$\prod_{p\in\mathbf{P}} |E'_{p,\xi}| \leqslant O\Big(\frac{1}{\log^t R} \prod_{(i,j)\in\Omega} (1 + |\xi_{i,j}|)^{O(1)}\Big). \tag{C.21}$$

We also have the following two lemmas.

Lemma C.14. *We have*

$$\log^t R \int_I \cdots \int_I \Big(\prod_{p\in\mathbf{P}} E'_{p,\xi}\Big) \prod_{(i,j)\in\Omega} \varphi_i(\xi_{i,j}) d\xi_{i,j} = \prod_{i\in[t]} c_{\chi_i,a_i} + O(\log^{-1/20} R), \tag{C.22}$$

where $E'_{p,\xi}$ is as in (C.19).

Lemma C.15. *We have*

$$\log^t R \int_I \cdots \int_I \prod_{p\in\mathbf{P}} |E'_{p,\xi}| \prod_{(i,j)\in\Omega} \big|\varphi_i(\xi_{i,j})\big| d\xi_{i,j} = O(1),$$

where $E'_{p,\xi}$ is as in (C.19).

Combining the two previous lemmas with Lemma C.12 we see that (C.18) holds true and thus the proof of Theorem C.7 is completed.

C.2.2. Proofs of the intermediate results.

As we have already stated this subsection is devoted to proving all the lemmas and claims that we presented in the previous section.

Proof of Claim C.8. Let $\{p_1, \ldots, p_s\}$ be the set of primes that divide m and observe that since m is square-free we have that $m = p_1 p_2 \ldots p_s$. Then, we need to show that

$$\mathbb{E}_{n\in\mathbb{Z}_m^d} \prod_{i\in[t]} \mathbf{1}_{m_i|\psi_i(n)} = \prod_{j=1}^{s} \mathbb{E}_{n\in\mathbb{Z}_{p_j}^d} \prod_{i\colon p_j|m_i} \mathbf{1}_{p_j|\psi_i(n)}. \tag{C.23}$$

To this end notice that $|\mathbb{Z}_m^d| = \prod_{p|m} |\mathbb{Z}_p^d|$ and by Lemma B.12 we have

$$\sum_{n\in\mathbb{Z}_m^d} \prod_{i\in[t]} \mathbf{1}_{m_i|\psi_i(n)} = \prod_{j=1}^{s} \sum_{n\in\mathbb{Z}_{p_j}^d} \prod_{i\colon p_j|m_i} \mathbf{1}_{p_j|\psi_i(n)}.$$

Thus (C.23) holds true and the proof of the claim is completed.

Proof of Lemma C.9. First observe that since ψ_is are affine linear forms the expression $\prod_{i\in[t]}\mathbf{1}_{m_i|\psi_i(n)}$ seen as a function of n is periodic with respect to the lattice $m\cdot\mathbb{Z}^d$. Having this in mind, the idea of the proof is to "fill" K with copies of $\mathbb{Z}_m^d$, the problem being that some of these copies may not lie entirely inside K. In order to quantify these "superfluous" copies we set ∂K to be the boundary of K, and also set

$$\mathcal{F}=\{A\subseteq\mathbb{Z}^d\colon A=x+\mathbb{Z}_m^d, \text{ for some } x\in m\cdot\mathbb{Z}^d \text{ and } A\cap K\neq\emptyset, A\cap(\mathbb{R}^d\setminus K)\neq\emptyset\}\subseteq\partial K.$$

Observe that the compactness of K implies that $\mathcal{F}$ is a finite family. Therefore, there exists some $0<l<1$ such that $\mathcal{F}\subseteq\partial K+[-ml,ml]^d$.

We have the following simple fact from convex geometry (see, e.g., [TV06, GT10]).

Fact C.16. *For any convex body $K\subseteq[-N,N]^d$ and for every $\varepsilon>0$ we have that*

$$\mathrm{vol}(\partial K+[-\varepsilon N,\varepsilon N])=O(\varepsilon N^d).$$

Applying the previous fact for $\varepsilon=ml/N$ and observing that $\mathrm{vol}(\bigcup\mathcal{F})=|\bigcup\mathcal{F}|$ we obtain that

$$|\bigcup\mathcal{F}|=O(m\cdot N^{d-1}).$$

Thus, using the last estimation and the periodicity of the expression $\prod_{i\in[t]}\mathbf{1}_{m_i|\psi_i(n)}$ we have that

$$\begin{aligned}\sum_{n\in K\cap\mathbb{Z}^d}\prod_{i\in[t]}\mathbf{1}_{m_i|\psi_i(n)} &= \frac{\mathrm{vol}(K)}{m^d}\sum_{n\in\mathbb{Z}_m^d}\prod_{i\in[t]}\mathbf{1}_{m_i|\psi_i(n)}+\sum_{n\in(\cup\mathcal{F})\cap K}\prod_{i\in[t]}\mathbf{1}_{m_i|\psi_i(n)}\\ &= \frac{\mathrm{vol}(K)}{m^d}\prod_{i\in[t]}\mathbf{1}_{m_i|\psi_i(n)}+O(1)|\bigcup\mathcal{F}|\\ &= \mathrm{vol}(K)\alpha_{m_1,\dots,m_t}+O(m\cdot N^{d-1})\end{aligned}$$

and the proof of the lemma is completed.

Proof of Lemma C.10. Taking absolute values one sees that the (LHS) of (C.15) is bounded up to a constant that depends on A by the quantity

$$(\log R)^{O(1)-A}\sum_{\substack{(m_{i,j})_{(i,j)\in\Omega}\in\mathbb{N}^\Omega\\ m_{i,j}\text{ square-free}}}\alpha_{m_1,\dots,m_t}\prod_{(i,j)\in\Omega}m_{i,j}^{-1/\log R}. \tag{C.24}$$

Then, by analyzing in prime factors and using Claim C.8 (see also Corollary B.10) we have that the previous expression can be rewritten as

$$(\log R)^{O(1)-A}\prod_p\sum_{(r_{i,j})_{(i,j)\in\Omega}\in\{0,1\}^\Omega}\alpha_{p^{r_1},\dots,p^{r_t}}p^{-(\sum_{(i,j)\in\Omega}r_{i,j})/\log R}, \tag{C.25}$$

where $r_i = \max(r_{i,1}, \ldots, r_{i,a_i})$. Thus, we need to eastimate $\alpha_{p^{r_1},\ldots,p^{r_t}}$ which is the exact purpose of the following claim.

CLAIM C.17. *For every prime p and for every $(r_1, \ldots, r_t) \in \{0,1\}^t \setminus \{0\}^t$ we have*

$$\alpha_{p^{r_1},\ldots,p^{r_t}} \leqslant \frac{1}{p}.$$

Furthermore, if $r_1 = \cdots = r_t = 0$, $\alpha_{p^{r_1},\ldots,p^{r_t}} = 1$.

PROOF OF CLAIM C.17. Let p be a prime number. First of all notice that $\alpha_{1,\ldots,1} = 1$, i.e. if $r_1 = \cdots = r_t = 0$ we have $\alpha_{p^{r_1},\ldots,p^{r_t}} = 1$.

Next, let $(r_1, \ldots, r_t) \in \{0,1\}^t \setminus \{0\}^t$. Let $n \in \mathbb{Z}_p^d$ be selected uniformly at random and observe that $\mathbf{1}_{p^{r_i} | \psi_i(n)}$ equals 1 with probability $1/p$, for every $i \in [t]$ such that $r_i \neq 0$. Thus,the product $\prod_{i\in[t]} \mathbf{1}_{p^{r_i}|\psi_i(n)}$ takes the value 1 with probability lower or equal that $1/p$, which completes the proof of the claim. □

By the previous claim we have that (C.25) is bounded by

$$(\log R)^{O(1)-A} \prod_{p\in\mathbf{P}} \Big(1 + \frac{1}{p}\big(\sum_{\substack{(r_{i,j})_{(i,j)\in\Omega}\in\{0,1\}^\Omega \\ \text{not all} 1'\text{s}}} p^{-(\sum_{(i,j)\in\Omega} r_{i,j})/\log R}\big)\Big) \tag{C.26}$$

and by the binomial theorem, applied for every $p \in \mathbf{P}$, we have

$$\begin{aligned}
1 \quad &+ \frac{1}{p}\big(\big(\frac{1}{p^{1/\log R}} + 1\big)^{|\Omega|} - 1\big) = 1 + \frac{1}{p}\sum_{k=1}^{|\Omega|} \binom{|\Omega|}{k} \frac{1}{p^{k/\log R}} \\
\leqslant &\quad 1 + \frac{1}{p^{1+1/\log R}} \sum_{k=0}^{|\Omega|} \binom{|\Omega|}{k} = 1 + \frac{2^{|\Omega|}}{p^{1+1/\log R}} \\
\leqslant &\quad \big(1 + \frac{1}{p^{1+1/\log R}}\big)^{2^{|\Omega|}} \leqslant \Big(\sum_{k=0}^{\infty} \frac{1}{p^{k(1+1/\log R)}}\Big)^{2^{|\Omega|}} \\
= &\quad \big(1 - \frac{1}{p^{1+1/\log R}}\big)^{-2^{|\Omega|}}.
\end{aligned}$$

Taking now product over all primes we have that (C.26) is bounded by

$$(\log R)^{O(1)-A} \prod_{p\in\mathbf{P}} \big(1 - \frac{1}{p^{1+1/\log R}}\big)^{-O(1)}. \tag{C.27}$$

But by Proposition B.14 we have

$$\prod_{p\in\mathbf{P}} \big(1 - \frac{1}{p^{1+1/\log R}}\big)^{-O(1)} = O(\log R^{O(1)})$$

and thus combining this estimation with (C.27) we see that the (LHS) of (C.15) is bounded by $O_A((\log R)^{O(1)-A})$. But for an adequate choice of A we have that $O_A\big((\log R)^{O(1)-A}\big) \leqslant O(\log^{-1/20} R)$ and thus the proof of the lemma is completed.

Proof of Lemma C.11. First of all notice that for every prime p and every $\xi = (\xi_{i,j})_{(i,j)\in\Omega} \in I^{\Omega}$ we may rewrite the Euler factor of (C.17) as follows

$$E_{p,\xi} = \sum_{B\subseteq\Omega} (-1)^{|B|} \frac{\alpha(p,B)}{p^{\sum_{(i,j)\in\Omega} z_{i,j}}},$$

where for every $(i,j) \in \Omega$, $z_{i,j} = (1 + \mathrm{i}\,\xi_{i,j})/\log R$. In the previous expression $\alpha(p,B) = \alpha_{p^{r_1},\dots,p^{r_t}}$, where $r_i = 1$ if $B \cap \mathcal{F}_i \neq \emptyset$ and $r_i = 0$ otherwise. We also have that $\alpha(p,\emptyset) = 1$. So, if we want to estimate $E_{p,\xi}$ the first thing to do is estimate $\alpha(p,B)$. To this end we split the family $\{B \subseteq \Omega, B \neq \emptyset\}$ in two main classes, the vertical sets and the rest. More precisely, call a set $\emptyset \neq B \subseteq \Omega$ *vertical* if there exists $i \in [t]$ such that $B \subseteq \mathcal{F}_i$, i.e. a set B is vertical if it is contained in a fibre $\mathcal{F}_i$, and non-vertical if there is no such fibre, i.e. if it intersects more than one fibres. Finally, notice that since N is large we may assume that

$$\log^{1/10} R \geqslant L. \tag{C.28}$$

We first have the following claim.

Claim C.18. *For every vertical set B and for every prime p with $p \geqslant \log^{1/10} R$ we have $\alpha(p,B) = \frac{1}{p}$.*

Proof. Let p be prime with $p \geqslant \log^{1/10} R$ and let B be a vertical set. Then there exists $i \in [t]$ such that $B \subseteq \mathcal{F}_i$ and therefore by definition $\alpha(p,B) = \sum_{n\in\mathbb{Z}_p^d} \mathbf{1}_{p|\psi_i(n)}$. The main ingredient of the proof is to show that since p is large enough we have that $\psi_i\colon \mathbb{Z}_p^d \to \mathbb{Z}_p$ uniformly covers $\mathbb{Z}_p^d$, i.e. it is a p^{d-1} to 1 mapping. To do so, since ψ_i is an affine linear form we only need to show that $\overline{\psi}_i(e_j) \neq 0 \mathrm{mod} p$, for some $1 \leqslant j \leqslant d$. Assume on the contrary that $p|\overline{\psi}_i(e_j)$ for every j. Since $p > L$ we have

$$|\overline{\psi}_i(e_j)| \leqslant \|\Psi\|_1 \leqslant L \overset{(C.28)}{\leqslant} \log^{1/10} R < p$$

and thus $\overline{\psi}_i(e_j) = 0$ for every j. But this is clearly a contradiction since ψ_i is not constant. Therefore ψ_i is a p^{d-1} to 1 mapping and hence

$$\alpha(p,B) = \frac{1}{p^d} \sum_{n\in\mathbb{Z}_p^d} \mathbf{1}_{p|\psi_i(n)} = \frac{p^{d-1}}{p^d} \sum_{n\in\mathbb{Z}_p} \mathbf{1}_{p|n} = \frac{1}{p}$$

which completes the proof of the claim. □

On the other hand, for the non-vertical sets we have the following claim.

CLAIM C.19. *For every non-vertical set $B \subseteq \Omega$, we have*

$$\alpha(p,B) = \begin{cases} O(1/p^2), & \text{when } p \notin P_\Psi \\ O(1/p), & \text{when } p \in P_\Psi. \end{cases}$$

PROOF. Let B be a non-vertical set and observe that there exist $1 \leqslant i < i' \leqslant t$ such that

$$\alpha(p,B) \leqslant \mathbb{E}_{n\in\mathbb{Z}_p^d} \mathbf{1}_{p|\psi_i(n)} \mathbf{1}_{p|\psi_{i'}(n)}.$$

We work as in Lemma C.6. If $p \notin P_\Psi$ let $n \in \mathbb{Z}_p^d$ be selected uniformly at random. Then, the expression $\mathbf{1}_{p|\psi_i(n)}\mathbf{1}_{p|\psi_{i'}(n)}$, seen as a function of n, takes the value 1 with probability $1/p^2$. Therefore $\alpha(p,B) = O(1/p^2)$. On the other hand, if $p \in P_\Psi$ then we have the following. If $p|\psi_i(n)$ we would have that $p|\psi_{i'}(n)$ also. Therefore the expression $\mathbf{1}_{p|\psi_i(n)}\mathbf{1}_{p|\psi_{i'}(n)}$ seen as a function of n takes the value 1 with probability $1/p$ and thus $\alpha(p,B) = O(1/p)$. Thus the proof of the claim is completed. □

Now, towards the proof of (C.20) assume first that $p \geqslant \log^{1/10} R$. If $p \notin P_\Psi$, then by claims C.18 and C.19 we have

$$\begin{aligned} E_{p,\xi} &= 1 - \sum_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} (-1)^{|B|-1} \frac{\alpha(p,B)}{p^{\sum_{(i,j)\in B} z_{i,j}}} + \sum_{\substack{B\subseteq\Omega,\\ B \text{ non-vertical}}} (-1)^{|B|} \frac{\alpha(p,B)}{p^{\sum_{(i,j)\in B} z_{i,j}}} \\ &= 1 - \sum_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} (-1)^{|B|-1} \frac{1}{p^{1+\sum_{(i,j)\in B} z_{i,j}}} + O(\frac{1}{p^2}) = \big(1 + O(\frac{1}{p^2})\big) E'_{p,\xi}, \end{aligned}$$

where the last equality derives from the binomial theorem. If on the other hand $p \in P_\Psi$ following the same steps as before we have $E_{p,\xi} = \big(1 + O(1/p)\big)E'_{p,\xi}$.

Assume now that $p \leqslant \log^{1/10} R$. First observe that since $\xi \in I^\Omega$ then for every $B \subseteq \Omega$ and every $(i,j) \in B$ we have that $|z_{i,j}| = O(\log^{-1/2} R)$ and thus

$$|p^{\sum_{(i,j)\in B} z_{i,j}}| = e^{\left|\sum_{(i,j)\in B} z_{i,j} \log p\right|} = 1 + O(\Big|\sum_{(i,j)\in B} z_{i,j}\Big| \log p) = 1 + O(\frac{\log p}{\log^{1/2} R}), \tag{C.29}$$

where we used the elementary inequality that for every $c < 1$ and for every $x \geqslant 0$ $e^{cx} \leqslant 1 + cx$. By Taylor expansion in $w = p^{\sum z_{i,j}}$ around $w = 1$, (C.29) and using once again the fact that $E'_{p,\xi} = O(1)$ for $p \leqslant \log^{1/10} R$ we have

$$\frac{E_{p,\xi}}{E'_{p,\xi}} = \frac{\tilde{E}_p}{\tilde{E}'_p} + O(\frac{\log p}{\log^{1/2} R}),$$

where $\widetilde{E}_p, \widetilde{E}'_p$ are defined setting all $z_{i,j} = 0$ in $E_{p,\xi}$ and $E'_{p,\xi}$ respectively, i.e.

$$\widetilde{E}_p = \sum_{B\subseteq\Omega} (-1)^{|B|} \alpha(p,B) \tag{C.30}$$

and

$$\widetilde{E'_p} = \sum_{\substack{B\subseteq\Omega \\ B \text{ vertical}}} (1-\frac{1}{p})^{(-1)^{|B|-1}}. \tag{C.31}$$

Therefore, in order to complete the proof of the lemma we need to show that

$$\sum_{B\subseteq\Omega}(-1)^{|B|}\alpha(p,B) = \beta_{\Psi,p} \sum_{\substack{B\subseteq\Omega \\ B \text{ vertical}}} (1-\frac{1}{p})^{(-1)^{|B|-1}}. \tag{C.32}$$

To this end, using the binomial theorem we see that for every $i \in [t]$ we have

$$\sum_{\emptyset\neq B\subseteq\{i\}\times[a_i]} (-1)^{|B|-1} = 1$$

and thus the (RHS) of (C.32) may be rewritten as $\beta_{\Psi,p}(1-p^{-1})^t$ which by (C.5) is equal to $\mathbb{E}_{n\in\mathbb{Z}_p^d}\prod_{i\in[t]}\mathbf{1}_{p\nmid\psi_i(n)}$ and and thus we reduced our task to showing

$$\sum_{B\subseteq\Omega}(-1)^{|B|}\alpha(p,B) = \mathbb{E}_{n\in\mathbb{Z}_p^d}\prod_{i\in[t]}\mathbf{1}_{p\nmid\psi_i(n)}. \tag{C.33}$$

By the inclusion-exclusion principle the (RHS) of the previous expression can be written as

$$\sum_{r_1,\dots,r_t\in\{0,1\}} (-1)^{r_1+\dots+r_t}\mathbb{E}_{n\in\mathbb{Z}_p^d}\prod_{i\colon r_i=1}\mathbf{1}_{p|\psi_i(n)}$$

which is equal to

$$\sum_{r_1,\dots,r_t\in\{0,1\}} (-1)^{r_1+\dots+r_t}\alpha_{p^{r_1},\dots,p^{r_t}}.$$

Thus we have to show that the (LHS) of (C.32) is equal to the previous expression. This will be done by comparing the coefficients of $\alpha_{p^{r_1},\dots,p^{r_t}}$ in these two expressions. For the (LHS) of (C.32) fix $a_{p^{r_1},\dots,p^{r_t}}$ and let $I = \{i \subseteq [t]\colon r_i = 1\}$. Then the coefficient of $a_{p^{r_1},\dots,p^{r_t}}$ equals

$$\sum_{\substack{B\subseteq\Omega, \\ B\cap\mathcal{F}_i\neq\emptyset, \text{ for every} i\in I}} (-1)^{|B|} = \prod_{i\in I}\sum_{B_i\subseteq[a_i]}(-1)^{|B_i|} = (-1)^{|I|},$$

where for the last equality we used the binomial theorem. With the previous estimation we showed that (C.33) holds true and thus the proof of Lemma C.11 is completed.

Proof of Lemma C.12. The main idea is to dispose at first the contribution of large primes ($p > \log^{1/10} R$) and then deal with the small ones.

Let $\xi \in I^{\Omega}$. From Lemma C.6 we have that $\beta_{\Psi,p} = 1 + O(1/p^2)$ if $p \notin P_{\Psi}$ and $\beta_{\Psi,p} = 1 + O(1/p)$ if $p \in P_{\Psi}$, which yields that

$$\prod_{p \in \mathbf{P}} \beta_{\Psi,p} \leqslant e^{O(X)}. \tag{C.34}$$

We also have

$$\prod_{\substack{p \in \mathbf{P}, \\ p \leqslant \log^{1/10} R}} \beta_{\Psi,p} \leqslant e^{O(X)}. \tag{C.35}$$

By these estimations we obtain that

$$\begin{aligned} \prod_{\substack{p \in \mathbf{P}, \\ p > \log^{1/10} R}} \beta_{\Psi,p} &\leqslant \exp\big(O(\sum_{\substack{p > \log^{1/10} R \\ p \in P_{\Psi}}} \frac{1}{p})\big) \leqslant \exp\big(O(\log^{-1/20} R \sum_{\substack{p > \log^{1/10} R \\ p \in P_{\Psi}}} \frac{1}{\sqrt{p}})\big) \\ &= \exp\big(O(X \log^{-1/20} R)\big) \leqslant 1 + O(e^{O(X)} \log^{-1/20} R), \end{aligned}$$

where for the last inequality we used the elementary inequality $e^{\lambda x} \leqslant 1 + \lambda e^{x}$, for every real numbers λ, x such that $\lambda \leqslant 1$ and $x \geqslant 0$. On the other hand using the estimations for the $\beta_{\Psi,p}$ once again and the inequalities $1 - x \leqslant e^{-x}$ for every $0 \leqslant x < 3/2$ and $e^{-\lambda x} \geqslant 1 - \lambda e^{x}$ for λ, x we also have

$$\prod_{\substack{p \in \mathbf{P}, \\ p > \log^{1/10} R}} \beta_{\Psi,p} \geqslant 1 + O(e^{O(X)} \log^{-1/20} R)$$

and thus

$$\prod_{\substack{p \in \mathbf{P}, \\ p > \log^{1/10} R}} \beta_{\Psi,p} = 1 + O(e^{O(X)} \log^{-1/20} R). \tag{C.36}$$

Thus by the previous estimation and by (C.35) we see that it suffices to show the following

$$\prod_{p \in \mathbf{P}} E_{p,\xi} = \Big(\prod_{\substack{p \in \mathbf{P}, \\ p \leqslant \log^{1/10} R}} \beta_{\Psi,p} + O(e^{O(X)} \log^{-1/20} R) \Big) \prod_{p \in \mathbf{P}} E'_{p,\xi}. \tag{C.37}$$

In order to do so, we use Lemma C.11 and obtain that

$$\begin{aligned}\prod_{\substack{p\in\mathbf{P},\\ p>\log^{1/10}R}} E_{p,\xi} &= \exp\Big(\sum_{\substack{p>\log^{1/10}R\\ p\notin P_\Psi}}\frac{1}{p^2}+\sum_{\substack{p>\log^{1/10}R\\ p\in P_\Psi}}\frac{1}{p}\Big)\prod_{\substack{p\in\mathbf{P},\\ p>\log^{1/10}R}} E'_{p,\xi}\\ &= \exp\big(O(1+X)\log^{-1/20}R\big)\prod_{\substack{p\in\mathbf{P},\\ p>\log^{1/10}R}} E'_{p,\xi}\\ &= \big(1+O(e^{O(X)}\log^{-1/20}R)\big)\prod_{\substack{p\in\mathbf{P},\\ p>\log^{1/10}R}} E'_{p,\xi},\end{aligned}$$

where for the last we equality we worked as in the proof of (C.36). But then we see that we have completed our first task, that is to discard the contribution of large primes, since now it suffices to prove

$$\prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R}} E_{p,\xi} = \Big(\prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R}}\beta_{\Psi,p}+O(e^{O(X)}\log^{-1/20}R)\Big)\prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R}} E'_{p,\xi}. \tag{C.38}$$

To this end , by Lemma C.11 it suffices to show that

$$\prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R}}\Big(\beta_{\Psi,p}+O(\frac{\log p}{\log^{1/2}R})\Big) = \prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R}}\beta_{\Psi,p}+O(e^{O(X)}\log^{-1/20}R).$$

Assume first that there exists some $p_0\leqslant\log^{1/10}R$ such that $\beta_{\Psi,p_0}=0$. Then, since $\beta_{\Psi,p}=1+O(1/p)$ (Lemma C.6) we have

$$\begin{aligned}\prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R}}\Big(\beta_{\Psi,p}+O(\frac{\log p}{\log^{1/2}R})\Big) &= O(\frac{\log p_0}{\log^{1/2}R})\prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R\\ p\neq p_0}}\Big(\beta_{\Psi,p}+O(\frac{\log p}{\log^{1/2}R})\Big)\\ &= O(\frac{\log p_0}{\log^{1/2}R})e^{O(X)} = O(e^{O(X)}\log^{-1/20}R).\end{aligned}$$

On the other hand, if we assume that no $\beta_{\Psi,p}$ vanishes we have the following. By Lemma C.6 we have that $\beta_{\Psi,p} = 1 + O(1/p)$ and thus

$$\begin{aligned}
\prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R}} \Big(\beta_{\Psi,p} + O(\frac{\log p}{\log^{1/2} R})\Big) &= \prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R}} \beta_{\Psi,p} \cdot \prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R}} (1 + O(\frac{\log p}{\log^{1/2} R})) \\
&= \Big(\prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R}} \beta_{\Psi,p}\Big)\big(1 + O(\log^{-1/3} R)\big) \\
&\overset{(C.35)}{=} \prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R}} \beta_{\Psi,p} + O(e^{O(X)}\log^{-1/3} R) \\
&= \prod_{\substack{p\in\mathbf{P},\\ p\leqslant\log^{1/10}R}} \beta_{\Psi,p} + O(e^{O(X)}\log^{-1/20} R).
\end{aligned}$$

This completes the proof of Lemma C.11.

Proof of Claim C.13. Let $\xi = (\xi_{i,j})_{(i,j)\in\Omega} \in I^{\Omega}$, and for every $(i,j)\in\Omega$ let $z_{i,j} = (1 + \mathrm{i}\,\xi_{i,j})/\log R$. Observe that $|z_{i,j}| = O(\log^{-1/2} R)$, for every $(i,j)\in\Omega$. By Lemma B.13 we see that

$$\begin{aligned}
\prod_{p\in\mathbf{P}} E'_{p,\xi} &= \prod_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} \Big(\frac{1}{\sum_{(i,j)\in B} z_{i,j}} + O(1)\Big)^{(-1)^{|B|}} \\
&= \prod_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} \big((1 + O(\log^{-1/2} R)\big)^{(-1)^{|B|}} \prod_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} \Big(\sum_{(i,j)\in B} z_{i,j}\Big)^{(-1)^{|B|-1}} \\
&= \big((1 + O(\log^{-1/2} R)\big) \prod_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} \Big(\sum_{(i,j)\in B} z_{i,j}\Big)^{(-1)^{|B|-1}},
\end{aligned}$$

where for the last equality we used the binomial theorem (see the proof of Lemma C.11 above). This completes the first part of the lemma. For the second part of the claim we work similarly. First, by the definition of $z_{i,j}$ we have

$$\prod_{p} E'_{p,\xi} = \prod_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} (\frac{1}{\log R})^{(-1)^{|B|-1}} \prod_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} (1 + \xi_{i,j})^{(-1)^{|B|-1}}. \tag{C.39}$$

By the binomial theorem the first factor of the (RHS) of the previous expression equals $\log^{-t} R$, while for the second factor we have

$$\Big|\prod_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} (1 + \xi_{i,j})^{(-1)^{|B|-1}}\Big| \leqslant O\Big(\prod_{(i,j)\in\Omega} (1 + |\xi_{i,j}|)^{O(1)}\Big)$$

Thus, combining the two previous estimations we see that (C.21) holds true. Therefore the proof of Claim C.13 is completed.

Proof of Lemma C.14. By the first part of Claim C.13 we have that the (LHS) of (C.22) equals

$$\log^t R \int_I \cdots \int_I \prod_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} \sum_{(i,j)\in B} (z_{i,j})^{(-1)^{|B|-1}} \prod_{(i,j)\in\Omega} \varphi_i(\xi_{i,j})d\xi_{i,j}$$
$$+ \quad O(\log^{-1/2} R)\log^t R \int_I \cdots \int_I \prod_{(i,j)\in\Omega} \varphi_i(\xi_{i,j})d\xi_{i,j}$$

By Proposition C.1 (the rapid decrease of φ) we have

$$\begin{aligned} &\log^t R \int_I \cdots \int_I \prod_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} \sum_{(i,j)\in B} (z_{i,j})^{(-1)^{|B|-1}} \prod_{(i,j)\in\Omega} \varphi_i(\xi_{i,j})d\xi_{i,j} \\ = \;&\log^t R \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \prod_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} \sum_{(i,j)\in B} (z_{i,j})^{(-1)^{|B|-1}} \prod_{(i,j)\in\Omega} \varphi_i(\xi_{i,j})d\xi_{i,j} + O(\log^{-1/20} R).\end{aligned}$$

and also

$$O(\log^{-1/2} R)\log^t R \int_I \cdots \int_I \prod_{(i,j)\in\Omega} \varphi_i(\xi_{i,j})d\xi_{i,j} = O(\log^{-1/20} R)$$

On the other hand,

$$\begin{aligned} &\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \prod_{\substack{B\subseteq\Omega,\\ B \text{ vertical}}} \sum_{(i,j)\in B} (z_{i,j})^{(-1)^{|B|-1}} \prod_{(i,j)\in\Omega} \varphi_i(\xi_{i,j})d\xi_{i,j} \\ = \;&\prod_{i\in[t]} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \prod_{\emptyset\neq B\subseteq\{i\}\times[a_i]} (\log R)^{(-1)^{|B|}} \Big(\sum_{(i,j)\in B} (1+i\xi_{i,j}) \Big)^{(-1)^{|B|-1}} \prod_{j=1}^{a_i} \varphi(\xi_{i,j})d\xi_{i,j} \\ = \;&\log^{-t} R \prod_{i\in[t]} c_{\chi_i,a_i}.\end{aligned}$$

Combining the three previous estimations the proof of Lemma C.14 follows.

Proof of Lemma C.15. By Proposition C.1 we have $|\phi_i(\xi_{i,j})| = O_A\big((1+\xi_{i,j})^{-A}\big)$, for every $i\in[t]$, $\xi_{i,j}\in I$ and $A>0$. Therefore if we choose A to be adequately large by (C.21) the proof of the Lemma C.15 follows.

Remark 4. By rerunning the proof we see that Theorem C.7 holds not only for convex bodies $K\subseteq[-N,N]^d$ but also for convex bodies that belong to translations of $[-N,N]^d$, i.e. for $K'\subseteq x+[-N,N]^d$ for some $x\in\mathbb{Z}^d$.

REMARK 5. If we assume that $\chi_1, \ldots, \chi_t$ are supported on $[-1,1]$ then we see that $\gamma \leqslant \frac{1}{10t}$.

C.3. The Goldston–Yildirim correlation estimates-A special case

Using theorem C.7 we will now prove a theorem also known as the Goldston–Yildirim correlation estimate, see [GT08, GT10, CFZ14]. This theorem gives as an immediate result Proposition 12.4. From now on let D be a positive integer, let $\chi \colon \mathbb{R} \to [0,1]$ be a smooth and supported on $[-1,1]$ function such that $\chi(0) = 1$ and $\int_0^1 |\chi'(x)|^2 dx = 1$, and let N be a large integer. Also, let $w = \log^{(4)} N$, $W = \prod_{p \in \mathbf{P}, p \leqslant w} p$, and $\widetilde{N} = \lfloor N/W \rfloor$. Finally, let $R = \widetilde{N}^{\gamma/2}$ for some small $\gamma = \gamma(\chi, D) > 0$.

THEOREM C.20 (Goldston–Yildirim correlation estimate). *Let $1 \leqslant d, t, L \leqslant D$, $b_1, \ldots, b_D \in \{0, \ldots, W-1\}$ be coprime to W and $\Psi = (\psi_1, \ldots, \psi_t)$ be a system of affine linear forms such that $\psi_i \colon \mathbb{Z}^d \to \mathbb{Z}$, $\|\Psi\|_1 = L$ and such that no two of the ψ_is are affinely related. Then, for any convex body $K \subseteq x + [-\widetilde{N}, \widetilde{N}]^d$, for some $x \in \mathbb{Z}^d$ we have that*

$$\Big(\frac{\phi(W)}{W}\Big)^t \sum_{K \cap \mathbb{Z}^d} \prod_{j \in [t]} \Lambda_{\chi,R,2}\big(W\psi_j(n) + b_{i_j}\big) = \mathrm{vol}(K) + o(\widetilde{N}^d), \tag{C.40}$$

for every $i_1, \ldots, i_t \in [t]$. In the previous expression $\Lambda_{\chi,R,2}$ is as in (12.8).

PROOF. Let $x \in \mathbb{Z}^d$, $K \subseteq x + [-\widetilde{N}, \widetilde{N}]^d$ be a convex body, let $i_1, \ldots, i_t \in [t]$, and let $\boldsymbol{b} = (b_{i_1}, \ldots, b_{i_t})$. Moreover, let $\beta_{W\Psi+\boldsymbol{b},p}$ be the p-local factor of $W\Psi + \boldsymbol{b}$, for every prime p, let $c_{\chi,2}$ be the sieve factor of χ with parameter 2, let $P_{W\Psi+\boldsymbol{b}}$ be as in (C.7) and $X = \sum_{p \in P_{W\Psi+\boldsymbol{b}}} p^{-1/2}$.

By Theorem C.7, Remark 4 and since by the choice of χ, $c_{\chi,2} = 1$ we have

$$\sum_{K \cap \mathbb{Z}^d} \prod_{j \in [t]} \Lambda_{\chi,R,2}\big(W\psi_j(n) + b_{i_j}\big) = \prod_{p \in \mathbf{P}} \beta_{W\Psi+\boldsymbol{b},p} \cdot \mathrm{vol}(K) + O\Big(e^{O(X)} \frac{\widetilde{N}^d}{\log^{1/20} R}\Big). \tag{C.41}$$

For the error term first we observe that no two of the linear forms $W\psi_i(n) + b_{i_j}$ are affinely related. Also we observe that if $p \in P_{W\Psi+\boldsymbol{b}}$ then $p \leqslant w$ which yields that $p = O(w) = O(\log^{(4)} N)$ and thus $X = O(\log\log\log^{1/2} N)$. Hence, $e^{O(X)} \log^{-1/20} R = o(1)$, and thus the error term of (C.41) becomes $o(\widetilde{N}^d)$.

It remains to show that $\prod_{p \in \mathbf{P}} \beta_{W\Psi+\boldsymbol{b},p} = (W/\phi(W))^t$. To this end, notice that if p is prime with $p \leqslant w$ we have $\beta_{W\Psi+\boldsymbol{b},p} = \big(p/(p-1)\big)^t$ and thus

$$\prod_{\substack{p \in \mathbf{P} \\ p \leqslant w}} \beta_{W\Psi+\boldsymbol{b},p} = \Big(\frac{W}{\phi(W)}\Big)^t.$$

Moreover, if p is prime with $p > w$ we have that the affine linear forms $W\psi_i(n) + b_{i_j}$ are not related modulo p. Thus by Claim C.19 we have that $\beta_{W\Psi+\boldsymbol{b},p} = 1 + O(1/p^2)$ and so $\prod_{p>w} \beta_{\Psi,p} = 1 + o(1)$. Combining the previous estimations we see that (C.40) holds true and thus the proof of Theorem C.20 is completed. $\square$

Bibliography

[AK06] F. Albiac and N. J. Kalton. *Topics in Banach space theory*, volume 233. Springer Science & Business Media, 2006.

[AN06] N. Alon and A. Naor. Approximating the cut-norm via Grothendieck's inequality. *SIAM Journal on Computing*, 35(4):787–803, 2006.

[Apo76] T. M. Apostol. *Introduction to analytic number theory*. Springer, 1976.

[AZHE10] M. Aigner, G. M. Ziegler, K. H. Hofmann, and P. Erdos. *Proofs from the Book*, volume 274. Springer, 2010.

[BCCZ14] 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. Available at https://arxiv.org/pdf/1401.2906v4.pdf, 2014.

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

[BCL+08] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs i: Subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219(6):1801–1851, 2008.

[Bil08] P. Billingsley. *Probability and measure*. John Wiley & Sons, 2008.

[BK17] S. Brazitikos and T. Karageorgos. An algorithmic regularity lemma for L_p regular sparse matrices. *Siam Journal on Discrete Mathematics*, 31(4):2301–2313, 2017.

[BN08] B. Bollobás and V. Nikiforov. An abstract Szemerédi regularity lemma. In *Building bridges*, pages 219–240. Springer, 2008.

[BR09] B. Bollobás and O. Riordan. Metrics for sparse graphs. *London Math. Soc. Lecture Note Ser*, pages 211–287, 2009.

[Bur82] D. L. Burkholder. A nonlinear partial differential equation and the unconditional constant of the haar system in L^p. *Bulletin of the American Mathematical Society*, 7(3):591–595, 1982.

[Can] P. Candela. Uniformity norms and nilsequences (an introduction). Available at http://www.renyi.hu/~candelap/unn-notes.pdf.

[CFZ13] D. Conlon, J. Fox, and Y. Zhao. Linear forms from the gowers uniformity norm. Available at http://arxiv.org/pdf/1305.5565.pdf, 2013.

[CFZ14] D. Conlon, J. Fox, and Y. Zhao. The Green–Tao theorem: an exposition. *EMS Surveys in Mathematical Sciences 1*, 1:249–282, 2014.

[CFZ15] D. Conlon, J. Fox, and Y. Zhao. A relative Szemerédi theorem. *Geometric and Functional Analysis*, 25(3):733–762, 2015.

[CM12] B. Cook and A. Magyar. Constellations in P^d. *International Mathematics Research Notices*, 2012(12):2794–2816, 2012.

[COCF10] A. Coja-Oghlan, C. Cooper, and A. Frieze. An efficient sparse regularity concept. *SIAM Journal on Discrete Mathematics*, 23(4):2000–2034, 2010.

[Dav00] H. Davenport. *Multiplicative number theory*. Springer, 2000.

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

[DKK15] P. Dodos, V. Kanellopoulos, and T. Karageorgos. L_p regular sparse hypergrahps:box norms. preprint, Available at https://arxiv.org/pdf/1510.07140v2.pdf, 2015.

[DKK16] P. Dodos, V. Kanellopoulos, and T. Karageorgos. Szemerédi's regularity lemma via martingales. *The Electronic Journal of Combinatorics*, 23(3), 2016.

[DKK18] P. Dodos, V Kanellopoulos, and T Karageorgos. L_p regular sparse hypergrahps. *Fundamenta Mathematicae*, (240):265–299, 2018.

[DKT13] P. Dodos, V. Kanellopoulos, and K. Tyros. A simple proof of the density Hales–Jewett theorem. *International Mathematics Research Notices*, 2014(12):3340–3352, 2013.

[DKT14] P. Dodos, V. Kanellopoulos, and K. Tyros. A density version of the Carlson-Simpson theorem. *Journal of the European Mathematical Society*, 2014.

[DKT16] P. Dodos, V. Kanellopoulos, and K. Tyros. A concentration inequality for product spaces. *Journal of Functional Analysis*, 270(2):609–620, 2016.

[Dur10] R. Durrett. *Probability: theory and examples*. Cambridge university press, 2010.

[FK99] A. Frieze and R. Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19(2):175–220, 1999.

[FZ15] J. Fox and Y. Zhao. A short proof of the multidimensional Szemerédi theorem in the primes. *American Journal of Mathematics*, 137(4):1139–1145, 2015.

[Gar07] D. J. H. Garling. *Inequalities: A journey into linear analysis.* Cambridge University Press, 2007.

[Gow01] W. T. Gowers. A new proof of Szemerédi's theorem. *Geometric and Functional Analysis*, 11(3):465–588, 2001.

[Gow06] W. T. Gowers. Quasirandomness, counting and regularity for 3-uniform hypergraphs. *Combinatorics, Probability and Computing*, 15(1-2):143–184, 2006.

[Gow07] W. T. Gowers. Hypergraph regularity and the multidimensional Szemerédi theorem. *Annals of Mathematics*, 166(3):897–946, 2007.

[Gow10] W. T. Gowers. Decompositions, approximate structure, transference, and the Hahn–Banach theorem. *Bulletin of the London Mathematical Society*, 42(4):573–606, 2010.

[Gre05] B. Green. Roth's theorem in the primes. *Annals of mathematics*, 161(3):1609–1636, 2005.

[GT08] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Annals of Mathematics*, 167(2):481–547, 2008.

[GT10] B. Green and T. Tao. Linear equations in primes. *Annals of Mathematics*, 171(3):1753–1850, 2010.

[GT12] B. Green and T. Tao. The Möbius function is strongly orthogonal to nilsequences. *Annals of Mathematics*, 175(2):541–566, 2012.

[GTZ12] B. Green, T. Tao, and T. Ziegler. An inverse theorem for the Gowers $U^{s+1}[N]$-norm. *Annals of Mathematics*, 176(2):1231–1372, 2012.

[GY] D. A. Goldston and C. Y. Yıldırım. Small gaps between primes, I. Available at https://front.math.ucdavis.edu/.

[GY03] D. A. Goldston and C. Y. Yıldırım. Higher correlations of divisor sums related to primes,I: Triple correlations. *Integers*, 3(A5):66, 2003.

[GY07] D. A. Goldston and C. Y. Yıldırım. Higher correlations of divisor sums related to primes III: Small gaps between primes. *Proceedings of the London Mathematical Society*, 95(3):653–686, 2007.

[Hås01] J. Håstad. Some optimal inapproximability results. *Journal of the ACM (JACM)*, 48(4):798–859, 2001.

[Hat09] H. Hatami. *On generalizations of Gowers norms.* PhD thesis, University of Toronto, 2009.

[IK04] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53. American Mathematical Society Providence, RI, 2004.

[KKMO07] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSP's. *SIAM Journal on Computing*, 37(1):319–357, 2007.

[Koh97] Y. Kohayakawa. Szemerédi' s regularity lemma for sparse graphs. In *Foundations of computational mathematics*, pages 216–230. Springer, 1997.

[KR03] Y. Kohayakawa and V. Rödl. Szemerédi' s regularity lemma and quasi-randomness. In *Recent advances in algorithms and combinatorics*, pages 289–351. Springer, 2003.

[LO94] R. Latała and K. Oleszkiewicz. On the best constant in the Khintchine–Kahane inequality. *Studia Mathematica*, 109(1):101–104, 1994.

[Lov12] L. Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[LS07] L. Lovász and B. Szegedy. Szemerédi' s lemma for the analyst. *GAFA Geometric And Functional Analysis*, 17(1):252–270, 2007.

[Nao04] A. Naor. Proof of the uniform convexity lemma. Available at https://web.math.princeton.edu/~naor/homepage%20files/inequality.pdf, 2004.

[NRS06] B. Nagle, V. Rödl, and M. Schacht. The counting lemma for regular k-uniform hypergraphs. *Random Structures & Algorithms*, 28(2):113–179, 2006.

[Pis12] G. Pisier. Grothendieck's theorem, past and present. *Bulletin of the American Mathematical Society*, 49:237–323, 2012.

[Pis16] G. Pisier. *Martingales in Banach spaces*, volume 155. Cambridge University Press, 2016.

[Pol12] DHJ Polymath. A new proof of the density Hales-Jewett theorem. *Annals of Mathematics*, 175:1283–1327, 2012.

[Ros84] H. E. Rose. *Subrecursion: functions and hierarchies*. Clarendon Press, 1984.

[RS78] I. Z. Ruzsa and E. Szemerédi. Triple systems with no six points carrying three triangles. volume 18, pages 939–945. 1978.

[RS04] V. Rödl and J. Skokan. Regularity lemma for k-uniform hypergraphs. *Random Structures & Algorithms*, 25(1):1–42, 2004.

[RTST06] V. Rödl, E Tengan, M Schacht, and N Tokushige. Density theorems and extremal hypergraph problems. *Israel Journal of Mathematics*, 152(1):371–380, 2006.

[RTTV08] O. Reingold, L. Trevisan, M. Tulsiani, and S. Vadhan. Dense subsets of pseudorandom sets. In *Foundations of Computer Science, 2008.*

FOCS'08. IEEE 49th Annual IEEE Symposium on, pages 76–85. IEEE, 2008.

[RX16] É. Ricard and Q. Xu. A noncommutative martingale convexity inequality. *Annals of Probability*, 44:867–882, 2016.

[Sha16] X. Shao. Narrow arithmetic progressions in the primes. *IMRN*, 2017(2):391–428, 2016.

[She88] S. Shelah. Primitive recursive bounds for van der waerden numbers. *Journal of the American Mathematical Society*, 1(3):683–697, 1988.

[Sol04] J. Solymosi. A note on a question of erdos and graham. *Combinatorics, Probability and Computing*, 13(2):263–267, 2004.

[SS03] E. M. Stein and R. Shakarchi. Fourier analysis: an introduction. *Princeton Press*, 2003.

[ST15] D. Saxton and A. Thomason. Hypergraph containers. *Inventiones mathematicae*, 201(3):925–992, 2015.

[Sze75] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith*, 27(2):199–245, 1975.

[Sze78] E. Szemerédi. Regular partitions of graphs. *Colloque Internationaux du CNRS*, 260:399–401, 1978.

[Tao06a] T. Tao. The gaussian primes contain arbitrarily shaped constellations. *Journal dAnalyse Mathématique*, 99(1):109–176, 2006.

[Tao06b] T. Tao. Szemerédi's regularity lemma revisited. *Contrib. Discrete Math*, 1:8–28, 2006.

[Tao06c] T. Tao. A variant of the hypergraph removal lemma. *Journal of combinatorial theory, Series A*, 113(7):1257–1280, 2006.

[Tao11] T. Tao. *An Epsilon of Room, II: Pages from Year Three of a Mathematical Blog*. American Mathematical Society, 2011.

[TSSW00] L. Trevisan, G. B. Sorkin, M. Sudan, and D. P. Williamson. Gadgets, approximation, and linear programming. *SIAM Journal on Computing*, 29(6):2074–2097, 2000.

[TV06] T. Tao and V. H. Vu. *Additive combinatorics*, volume 105. Cambridge University Press, 2006.

[TZ08] T. Tao and T. Ziegler. The primes contain arbitrarily long polynomial progressions. *Acta Mathematica*, 201(2):213–305, 2008.

[TZ15a] T. Tao and T. Ziegler. A multi-dimensional Szemerédi theorem for the primes via a correspondence principle. *Israel Journal of Mathematics*, 207(1):203–228, 2015.

[TZ15b] T. Tao and T. Ziegler. Narrow progressions in the primes. In *Analytic Number Theory*, pages 357–379. Springer, 2015.