



A Multidimensional Szemerédi Theorem in the Primes via Combinatorics

Brian Cook, Ákos Magyar and Tatchai Titichetrakun

Abstract. Let A be a subset of positive relative upper density of $\mathbb{P}^d$, the d -tuples of primes. We present an essentially self-contained, combinatorial argument to show that A contains infinitely many affine copies of any finite set $F \subseteq \mathbb{Z}^d$. This provides a natural multidimensional extension of the theorem of Green and Tao on the existence of long arithmetic progressions in the primes.

Mathematics Subject Classification. 05C65, 05C55, 11B30, 11N13.

Keywords. Hypergraph regularity, Primes, Progressions.

1. Introduction

1.1. Background

A celebrated theorem in additive combinatorics due to Green and Tao [7] establishes the existence of arbitrarily long arithmetic progressions in the primes. It is proved that if A is a subset of the primes of positive relative upper density, then A necessarily contains infinitely many affine copies of any finite set of integers. As such, it might be viewed as a relative version of Szemerédi's theorem [16] on the existence of long arithmetic progressions in dense subsets of the integers.

Another fundamental result in this area is the multidimensional extension of Szemerédi's theorem originally proved by Furstenberg and Katznelson [3]. It states that if $A \subseteq \mathbb{Z}^d$ is of positive upper density, then A contains an affine copy of any finite set $F \subseteq \mathbb{Z}^d$. The proof in [3] uses ergodic methods; however, a more recent combinatorial approach was developed by Gowers [5] and also independently by Nagel, Rödl and Schacht [13].

It is natural to ask if a multidimensional extension of the result of Green and Tao, or alternatively if a relative version of the Furstenberg–Katznelson

theorem can be established. In fact, this question was raised already in [17], where the existence of arbitrary constellations among the Gaussian primes was shown. A partial result was obtained earlier by the first two authors [2], where it was proved that relative dense subsets of $\mathbb{P}^d$ contain an affine copy of any finite set $F \subseteq \mathbb{Z}^d$ which is in *general position*, in the sense that each coordinate hyperplane contains at most one point of F . Recently, the existence of arbitrary constellations in relative dense subsets of $\mathbb{P}^d$ has been proved by Tao and Ziegler [20], based on an *infinite* number of linear form conditions to obtain a weighted version of the Furstenberg correspondence principle and a short proof by Fox and Zhao [4] has been obtained afterwards using sampling arguments.

Both the above proofs, however, rely on full force of the results of Green, Tao and Ziegler developed in [8–10] on the asymptotic number of prime solutions for systems linear equations, a monumental work encompassing more than 200 pages. Thus, it may be of interest to obtain an essentially self-contained, combinatorial proof using only the pseudo-randomness properties of the almost primes encapsulated by the so-called linear form condition of Green and Tao.

In this note, we provide such an argument, based on developing a hypergraph removal lemma for weighted hypergraphs. A crucial new feature is that, in our settings, there are weights attached possibly to any lower dimensional edge, and hence, known transference arguments [1, 7, 17] do not seem to apply. Our approach is different, and we are not trying to remove the weights and hence to reduce the problem to previously known results, but to extend the proof of the hypergraph regularity and removal lemmas directly to the weighted setting.

Finally, let us mention that the methods of Fox and Zhao [4] and Tao and Ziegler [20] do not provide bounds, while from our approach, one can extract quantitative statements. The bounds, though recursive, are rather weak, iterated tower-exponential type. In addition, as we rely only on sieve techniques, our approach is somewhat flexible, e.g., it can be modified to count the number of “small” copies of a finite set F , of size N^ε , in a set $A \subseteq [1, N]^d \cap \mathbb{P}^d$ of positive relative density, however, we do not pursue such extensions here.

1.2. Main Results

Let us recall that a set $A \subseteq \mathbb{P}^d$ is of positive relative upper density if

$$\limsup_{N \rightarrow \infty} \frac{|A \cap \mathbb{P}_N^d|}{|\mathbb{P}_N|^d} > 0,$$

where $\mathbb{P}_N$ denotes the set of primes up to N , and $|A|$ stands for the cardinality of a set A . If $F \subseteq \mathbb{Z}^d$ is a finite set, we say that a set F' is an affine copy of F , or alternatively that F' is a constellation defined by F , if

$$F' = x + t \cdot F = \{x + ty : y \in F\}.$$

We call F' *non-trivial* if $t \neq 0$. Our main result is the following.

Theorem 1.1. *If A is a subset of $\mathbb{P}^d$ of positive upper relative density, then A contains infinitely many non-trivial affine copies of any finite set $F \subseteq \mathbb{Z}^d$.*

Note that it is enough to show that the set A contains at least one non-trivial affine copy of F , as deleting the set F from A will not affect its relative density. In addition, by lifting the problem to a higher number of dimensions, it is easy to see that one can assume that F forms the vertices of a d -dimensional simplex. Indeed, let $F = \{0, x_1, \dots, x_k\}$, choose a set of k linearly independent vectors $\{y_1, \dots, y_k\} \subseteq \mathbb{Z}^k$, and define the set $\Delta := \{0, (x_1, y_1), \dots, (x_k, y_k), z_{k+1}, \dots, z_{k+d}\} \subseteq \mathbb{Z}^{k+d}$, such that the vectors of $\Delta \setminus \{0\}$ form a basis of $\mathbb{R}^{k+d}$. If the set $A' = A \times \mathbb{P}^k$ contains an affine copy of Δ , then clearly, A contains an affine copy of the set $\pi(\Delta) \supseteq F$, where $\pi: \mathbb{R}^d \times \mathbb{R}^k \rightarrow \mathbb{R}^d$ is the natural orthogonal projection.

In the case when $\Delta \subseteq \mathbb{Z}^d$ is a d -dimensional simplex, we prove a quantitative version of Theorem 1.1. To formulate it we define the quantity

$$l(\Delta) := \sum_{i=1}^d |\pi_i(\Delta)|, \quad (1.1)$$

$\pi_i: \mathbb{R}^d \rightarrow \mathbb{R}$ being the orthogonal projection to the i th coordinate axis.

Theorem 1.2. *Let $\alpha > 0$ and let $\Delta \subseteq \mathbb{Z}^d$ be a d -dimensional simplex. There exists a constant $\delta = \delta(\alpha, \Delta) > 0$, such that for any $N > 1$ and any set $A \subseteq \mathbb{P}_N^d$, such that $|A| \geq \alpha |\mathbb{P}_N|^d$, the set A contains at least $\delta N^{d+1} (\log N)^{-l(\Delta)}$ affine copies of the simplex Δ .*

Note that in Theorem 1.2, we do not require the copies of Δ to be non-trivial. Thus, without loss of generality, N can be assumed to be sufficiently large with respect to α and Δ . It is clear that Theorem 1.2 implies Theorem 1.1 as the number of trivial copies of Δ in A is at most $N^d (\log N)^{-d}$.

To see why the above lower bound is meaningful, note that there are $\approx N^{d+1}$ affine copies of Δ in $[1, N]^d$, and for a fixed i , the probability that all the i th coordinates of an affine copy Δ' are primes is roughly $(\log N)^{-|\pi_i(\Delta)|}$. Thus, if the prime tuples behave randomly, the probability that $\Delta' \subseteq \mathbb{P}_N^d$ is about $(\log N)^{-l(\Delta)}$.

In the contrapositive, Theorem 1.2 states that if a set $A \subseteq \mathbb{P}_N^d$ contains at most $\delta N^{d+1} (\log N)^{-l(\Delta)}$ affine copies of Δ , then its relative density is at most α . To prove Theorem 1.2, one formulates a statement involving a pseudo-random measure $\nu = \nu^{(n)}: [1, N] \rightarrow \mathbb{R}_+$, originally introduced in [7].

1.3. The Green–Tao Measure and the Linear Form Condition

As previously indicated, we use the pseudo-random measure ν defined in [7], and (a slight variant of) the so-called linear form condition, which we recall here briefly.

Let w be a sufficiently large number and let $W = \prod_{p \leq w} p$ be the product of primes up to w . For given b relatively prime to W , define the modified von Mangoldt function $\bar{\Lambda}_b: \mathbb{Z} \rightarrow \mathbb{R}_{\geq 0}$ by

$$\bar{\Lambda}_b(n) = \begin{cases} \frac{\phi(W)}{W} \log(Wn + b) & \text{if } Wn + b \text{ is a prime,} \\ 0 & \text{otherwise.} \end{cases}$$

Here, ϕ is the Euler function. Note that by Dirichlet's theorem on the distribution of primes in residue classes, one has that $\sum_{n \leq N} \bar{\Lambda}_b(n) = N(1 + o(1))$. A crucial fact is that the function $\bar{\Lambda}_b$ is majorized by the so-called Goldston–Yildirim divisor sum [7, 11]:

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

μ being the Mobius function and $R = N^{d^{-1}2^{-d-5}}$. For given small parameters $0 < \varepsilon_1 < \varepsilon_2 < 1$ (whose exact values will be specified later), recall the Green–Tao measure:

$$\nu_b(n) = \begin{cases} \frac{\phi(W)}{W} \frac{\Lambda_R(Wn+b)^2}{\log R} & \text{if } \varepsilon_1 N \leq n \leq \varepsilon_2 N, \\ 1 & \text{otherwise.} \end{cases}$$

Note that $\nu_b(n) \geq 0$ for all n , and in addition, it is easy to see that for N sufficiently large, one has that

$$\nu_b(n) \geq d^{-1}2^{-d-6} \bar{\Lambda}_b(n) \quad (1.2)$$

for all $\varepsilon_1 N \leq n \leq \varepsilon_2 N$. This is trivial unless $Wn + b$ is a prime, and in that case, since $\varepsilon_1 N > R$, $\Lambda_R(Wn + b) = \log R = d^{-1}2^{-d-5} \log N$.

Let us briefly recall its definition and the pseudo-randomness properties: the so-called linear form condition that we will need in the proof. This is not exactly the same formulation as given in [7]; however, the proof works without any changes.

Theorem A. (Linear form condition, [7]). *Let N , W , and the measures ν_b be as above, and let $m_0, t_0, k_0 \in \mathbb{N}$ be small parameters. Then, the following holds. For given $m \leq m_0$ and $t \leq t_0$, suppose that $\{l_{i,j}\}_{1 \leq i \leq m, 1 \leq j \leq t}$ are arbitrary rational numbers with numerator and denominator at most k_0 in absolute value, and that $\{b_i\}$ are arbitrary numbers relatively prime to W . If the linear forms*

$$L_i(x) = \sum_{j=1}^t l_{i,j} x_j,$$

are non-zero and pairwise linearly independent over the rationals then

$$\mathbb{E} \left(\prod_{i=1}^m \nu_{b_i}(L_i(x)); x \in \mathbb{Z}_N^t \right) = 1 + o_{N,W \rightarrow \infty; m_0, t_0, k_0}(1), \quad (1.3)$$

where the $o(1)$ term is independent of the choice of the b_i .

In the above formula, the linear forms $L_i(x)$ are considered as acting on $(\mathbb{Z}/N\mathbb{Z})^t$ and the error term $o_{N,W \rightarrow \infty; m_0, t_0, k_0}(1)$ denotes a quantity that tends to 0 as both $N \rightarrow \infty$ and $W \rightarrow \infty$ for any fixed choice of m_0, t_0, k_0 , i.e., the error term can be made smaller than any given $\varepsilon > 0$ by choosing both W and N sufficiently large with respect to m_0, t_0, k_0 , and ε . This refinement is important in obtaining the quantitative lower bound in Theorem 1.2, see also the remarks in [7, Sect. 11]. The error terms are independent of the choice of b_i , so we will write ν for ν_{b_i} for all $1 \leq i \leq d$ for simplicity of notations.

With the aid of this measure, we define the weight of a finite set $S \subseteq \mathbb{Z}^d$ as

$$w(S) := \prod_{i=1}^d \prod_{y \in \pi_i(\Delta)} \nu(y). \quad (1.4)$$

If $S = \{y\}$, we will write $w(y) := w(\{y\})$. The point is that if $Wy + b \in \mathbb{P}_N^d$, then

$$w(y) \approx (\log N)^d. \quad (1.5)$$

The implicit constant depends only on W which we will choose eventually sufficiently large but independent of N . Then, also

$$w(\Delta) \approx (\log N)^{l(\Delta)} \quad (1.6)$$

for a non-degenerate simplex Δ with $W\Delta + b \subseteq A \subseteq \mathbb{P}_N^d$. Thus, identifying $[1, N]$ with $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$, it is a straightforward argument to show (see Sect. 5) that Theorem 1.2 follows from

Theorem 1.3. *Let $\Delta = \{v_0, \dots, v_d\} \subseteq \mathbb{Z}^d$ be a d -dimensional simplex and let $\alpha > 0$. Then, there exists a constant $\delta = \delta(\alpha, \Delta) > 0$, such that the following holds. If N is a large prime and $A \subseteq \mathbb{Z}_N^d$ is a set for which*

$$\mathbb{E}_{x \in \mathbb{Z}_N^d, t \in \mathbb{Z}_N} \prod_{i=1}^d \mathbf{1}_A(x + tv_i) w(x + t\Delta) \leq \delta, \quad (1.7)$$

then

$$\mathbb{E}_{x \in \mathbb{Z}_N^d} \mathbf{1}_A(x) w(x) \leq \alpha + o_{N, W \rightarrow \infty, \Delta}(1).$$

Note that the error term $o_{N, W \rightarrow \infty, \Delta}(1)$ can be omitted if both N and W are chosen sufficiently large with respect to $\alpha > 0$.

1.4. A Removal Lemma for Weighted Hypergraph Systems

We will use the construction of a weighted hypergraph associated with a set $A \subseteq \mathbb{Z}_N^d$ and a simplex $\Delta = \{v_0, \dots, v_d\}$, given in [17]. This will enable us to reduce Theorem 1.3 to prove a removal lemma for weighted hypergraphs.

Let $J = \{0, 1, \dots, d\}$, $\overline{\mathcal{H}} := \{e : e \subseteq J\}$, $\mathcal{H}_d := \{e : e \subseteq J, |e| = d\}$. For a set $e \in \overline{\mathcal{H}}$, let $V_e = \mathbb{Z}_N^e = \prod_{j \in e} \mathbb{Z}_N$. Identify V_e as the subspace of elements $x = (x_0, \dots, x_d) \in V_J$, such that $x_j = 0$ for all $j \notin e$ and let $\pi_e : V_J \rightarrow V_e$ denote the natural projection. For $e = \{j\}$, we write $V_j := V_{\{j\}}$.

Remark 1.1. For a given $e \in \mathcal{H}$, $|e| = d$, we think of a point $x_e \in V_e$ as a d -simplex (or clique) with vertices $\{x_j : j \in e\}$, and then, a set $G_e \subseteq V_e$ may be viewed as a d -regular hypergraph with vertex sets V_j ($j \in e$). Similarly, a point $x \in V_J$ represents a $(d+1)$ -simplex with faces $x_e := \pi_e(x)$, for $e \in \mathcal{H}_d$.

For a given $e \subseteq J$, define the σ -algebra $\mathcal{A}_e = \{\pi_e^{-1}(F) : F \subseteq V_e\}$. We assign a $(d+1)$ -partite d -regular hypergraph to a given set $A \subseteq \mathbb{Z}_N^d$ as follows. For $e = J \setminus \{j\}$, define

$$E_e = \left\{ x \in V_J : \sum_{i=0}^d x_i(v_i - v_j) \in A \right\}. \quad (1.8)$$

Note that $E_e \in \mathcal{A}_e$ as the expression in (1.8) is independent of the coordinate x_j . We define $G := \bigcap_{e \in \mathcal{H}_d} E_e \subseteq V_J$ to be the hypergraph associated with A . From our point of view, the points $x \in G$ are $(d+1)$ -simplices (or cliques) with faces $x_e = \pi_e(x)$ belonging to the hypergraphs $G_e = \pi_e(E_e)$.

Example 1. Let $\Delta = \{v_0, v_1, v_2\}$ with $v_0 = (0, 0)$, $v_1 = (0, 2)$, $v_2 = (1, 1)$ and let $A \subseteq \mathbb{Z}_N^2$. Then, by (1.8), we have

$$\begin{aligned} G_{01} &= \{(x_0, x_1) : (-x_0 + x_1, -x_0 - x_1) \in A\}, \\ G_{02} &= \{(x_0, x_2) : (-2x_0 - x_2, x_2) \in A\}, \\ G_{12} &= \{(x_1, x_2) : (2x_1 + x_2, x_2) \in A\}, \quad \text{and} \\ G &= \{(x_0, x_1, x_2) : (-x_0 + x_1, -x_0 - x_1) \in A, (-2x_0 - x_2, x_2) \in A, \\ &\quad (2x_1 + x_2, x_2) \in A\} \end{aligned}$$

is the 3-partite graph associated with the set A .

To see how this hypergraph is related to detecting affine copies of the simplex Δ in a set $A \subseteq \mathbb{Z}_N^d$, consider the map $\Phi : \mathbb{Z}_N^{d+1} \rightarrow \mathbb{Z}_N^d \times \mathbb{Z}_N$, defined by

$$\Phi(x) = \left(\sum_{i=0}^d x_i v_i - \sum_{i=0}^d x_i \right) =: (y, t). \quad (1.9)$$

By (1.8) and (1.9), we have that $x \in E_e$ for $e = J \setminus \{j\}$ if and only if $y + tv_j \in A$, thus $x \in G = \bigcap_{e \in \mathcal{H}_d} E_e$ exactly when $y + t\Delta \subseteq A$. Since Φ is one to one (as $v_1 - v_0, \dots, v_d - v_0$ are linearly independent), this gives a parametrisation of all affine copies of Δ contained in A .

We now define a family of functions $\mu_e : V_e \rightarrow \mathbb{R}_+$, ($e \in \mathcal{H}_d$) and $\nu_e : V_e \rightarrow \mathbb{R}_+$, ($e \in \overline{\mathcal{H}}$). For $e \in \mathcal{H}_d$, $e = J \setminus \{j\}$ and $1 \leq k \leq d$, let

$$L_e^k(x) := \sum_{i=0}^d x_i (v_i^k - v_j^k), \quad (1.10)$$

where v_i^k denotes the k th coordinate of the vector v_i . A crucial property is that the forms $L_e^k(x)$ are pairwise linearly independent over $\mathbb{Z}_N$ for N being a sufficiently large prime. Indeed, for $e = J \setminus \{j\}$, the coefficients of these forms form the column vectors of the $d \times d$ matrix D_j with row vectors $v_i - v_j$ ($0 \leq i \leq d$, $i \neq j$). If any two columns would be linearly dependent over $\mathbb{Z}_N$, then N would divide the determinant of D_j . Since the row vectors are linearly independent and have integer entries, this is not possible for large enough N , which we will always assume.

We define the functions

$$\mu_e(x) := \prod_{k=1}^d \nu(L_e^k(x)). \quad (1.11)$$

We partition the family of forms

$$\mathcal{L} := \{L_e^k : |e| = d, 1 \leq k \leq d\}$$

according to which coordinates they depend on. For this, define the support of a linear form $L(x) = \sum_{k=1}^d a_k x_k$ as $\text{supp}(L) = \{k: a_k \neq 0\}$. For a given $e \in \bar{\mathcal{H}}$, let

$$\nu_e(x) = \prod_{\substack{\text{supp}(L)=e \\ L \in \mathcal{L}}} \nu(L(x)) \quad (1.12)$$

with the convention that $\nu_e \equiv 1$ if $\{L: \text{supp}(L) = e\} = \emptyset$.

Example 2. In the case of the hypergraph G constructed in Example 1, one obtains the following weights and measures on the edges and vertices:

$$\begin{aligned} \mu_{01}(x_0, x_1) &= \nu(-x_0 + x_1)\nu(-x_0 - x_1), \\ \mu_{02}(x_0, x_2) &= \nu(-2x_0 - x_2)\nu(x_2), \quad \mu_{12}(x_1, x_2) = \nu(2x_1 + x_2)\nu(x_2), \\ \nu_{01}(x_0, x_1) &= \nu(-x_0 + x_1)\nu(-x_0 - x_1), \\ \nu_{02}(x_0, x_2) &= \nu(-2x_0 - x_2), \quad \nu_{12}(x_1, x_2) = \nu(2x_1 + x_2), \end{aligned}$$

$\nu_2(x_2) = \nu(x_2)$, with the convention that $\nu_0(x_0) = \nu_1(x_1) \equiv 1$. The family of linear forms defining the weights is

$$\mathcal{L} = \{-x_0 + x_1, -x_0 - x_1, -2x_0 - x_2, 2x_1 + x_2, x_2\}.$$

If $\Delta = \{v_0, \dots, v_k\}$ is in general position, that is, if $v_i^k \neq v_j^k$ for all $i \neq j$ and for all k , then $\text{supp}(L_e^k) = e$ for all $e \in \mathcal{H}_d$, hence

$$\mu_e(x) = \nu_e(x) = \prod_{k=1}^d \nu(L_e^k(x)).$$

In general, we have

$$\mu_e(x) = \prod_{f \subseteq e} \nu_f(x). \quad (1.13)$$

This is because if for some $e' = J \setminus \{j'\}$, we have that $\text{supp}(L_{e'}^k) \subseteq e$, then by (1.10) $v_{j'}^k = v_j^k$ and $L_{e'}^k = L_e^k$. This ensures that the family $\mathcal{L}$ consists of pairwise linearly independent forms. In addition, for $e = J \setminus \{j\}$, $\mu_e(x)$ is well-defined on V_e as it depends only on the variables $\{x_i: i \neq j\}$. We will refer to the functions ν_e and μ_e as *weights* and *measures*. To emphasize this point of view, we will often use the integral notation and write

$$\int_{V_e} F(x) d\mu_e(x) := \mathbb{E}_{x \in V_e} F(x) \mu_e(x), \quad \text{and} \quad \mu_e(S) := \mathbb{E}_{x \in V_e} \mathbf{1}_S(x) \mu_e(x),$$

for functions $F: V_e \rightarrow \mathbb{R}$ and sets $S \subseteq V_e$.

The weight system $\{\nu_e\}_{e \subseteq J}$ has very strong pseudo-randomness properties. First, for any hypergraph $\mathcal{G} \subseteq \bar{\mathcal{H}}$, one has that $\mathbb{E}_{x \in V_j} \prod_{e \in \mathcal{G}} \nu_e(x_e) = 1 + o_{N,W \rightarrow \infty}$. This is immediate from (1.12) and the fact the family $\mathcal{L}$ is pairwise linearly independent. We will refer to this property as *order-1 pseudo-randomness* of the weight system $\{\nu_e\}_{e \subseteq J}$. However, much more is true, to formulate it recall the notion of a hypergraph bundle, see also Sect. 4.2. Let K be a finite set and let $\pi: K \rightarrow J$ be an onto map. Define $\bar{H}_\pi := \{g \subseteq K: |\pi(g)| = |g|\}$. If m is an integer such that $m_j := |\pi^{-1}(j)| \leq m$ for all $j \in J$,

then we say that the hypergraph bundle $\bar{H}_\pi$ is of order m . For given $k \in K$ set $V_k := V_{\pi(k)}$, intuitively, this means that we sample independently m_j points from each vertex set V_j . Thus, for $g \in \bar{H}_\pi$, $e = \pi(g)$, one has $V_g = V_e$ and can define the weights $\bar{\nu}_g(x_g) := \nu_e(x_g)$ on the points $x_g = (x_k)_{k \in g}$. Note that by (1.12), the weight $\bar{\nu}_g$ is defined via linear forms L for which $\text{supp}(L) = g$, and hence, the family of linear forms $\mathcal{L}_\pi$ defining the weight system $\{\bar{\nu}_g\}_{g \in \bar{H}_\pi}$ consists of pairwise linearly independent linear forms in the $x_K = (x_k)_{k \in K}$ variables. Thus, by the linear form condition, one has for any hypergraph $\mathcal{G} \subseteq \bar{H}_\pi$ of order m ,

$$\mathbb{E}_{x_K \in V_K} \prod_{g \in \mathcal{G}} \bar{\nu}_g(x_g) = \mathbb{E}_{x_K \in V_K} \prod_{g \in \mathcal{G}} \nu_{\pi(g)}(x_g) = 1 + o_{N, W \rightarrow \infty; d, m}(1).$$

We will refer to the above property as the *order- m pseudo-randomness* of the weight system $\{\nu_e\}_{e \subseteq J}$. For $m = 2$, this is essentially equivalent to the linear form condition on weighted hypergraphs, as given in [17, see def. 2.8]. What is needed in our proof is that the weight system $\{\nu_e\}$ is pseudo-random at order m , where $m = m(\alpha)$ is a fixed but sufficiently large number depending on the density α .

The measures μ_e are closely related to the weights w . Indeed, for $e = J \setminus \{j\}$ and $(y, t) = \Phi(x)$, one has

$$L_e^k(x) = \sum_{i=0}^d x_i(v_i^k - v_j^k) = \pi_k(y - tv_j), \quad (1.14)$$

where π_k is the orthogonal projection to the k -th coordinate axis. This implies that

$$\mu_e(x) = \prod_{k=1}^d \nu(L_e^k(x)) = w(y - tv_j) \quad (1.15)$$

and also

$$\mu_J(x) := \prod_{L \in \mathcal{L}} \nu(L(x)) = w(y - t\Delta). \quad (1.16)$$

Thus, assumption (1.7) in Theorem 1.3 translates to

$$\mathbb{E}_{x \in V_J} \prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}(x) \mu_J(x) = \mathbb{E}_{(y, t) \in \mathbb{Z}_N^{d+1}} w(y - t\Delta) \leq \delta. \quad (1.17)$$

On the other hand, if $M = \{x \in V_J : x_0 + \cdots + x_d = 0\}$, then $x \in M \cap_{e \in \mathcal{H}_d} E_e$ if and only if $\Phi(x) = (y, 0)$ with $y \in A$, thus by (1.9), (1.15)

$$\mathbb{E}_{y \in A} w(y) = \mathbb{E}_{x \in M} \prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}(x) \mu_{e'}(x), \quad (1.18)$$

for any fixed $e' \in \mathcal{H}_d$. Thus, it is easy to see that Theorem 1.3 follows from a *removal lemma* for weighted hypergraphs, which we first recall in the unweighted case (where $\nu_f \equiv 1$ for all f), see also [5, 13, 17].

Theorem B. (Simplex Removal Lemma [19]). *Let $E_e \in \mathcal{A}_e$ be given for $e \in \mathcal{H}_d$, and let $\varepsilon > 0$. In addition, let μ and μ_e denote the normalized counting measures on V and V_e . There exists $\delta = \delta(\varepsilon) > 0$, and for every index set $e \in \mathcal{H}_d$, there exists a set $E'_e \in \mathcal{A}_e$, such that the following holds. If*

$$\mathbb{E}_{x \in V_J} \prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}(x) \mu(x) \leq \delta,$$

then

$$\prod_{e \in \mathcal{H}_d} \mathbf{1}_{E'_e}(x) = 0 \quad \text{for all } x \in V, \quad \text{and} \\ \mu_e(E_e \setminus E'_e) \leq \varepsilon.$$

Now, let us observe the key properties of the family of linear forms $\mathcal{L}$ and introduce some terminology which we will use throughout the paper. Recall that for $e = J \setminus \{j\}$, $e' = J \setminus \{j'\}$, we have that $\text{supp}(L_{e'}^k) \subseteq e$ if and only if $v_j^k = v_{j'}^k$, and that is equivalent to $L_{e'}^k = L_e^k$. We call such a family $\mathcal{L}$ **well-defined**. Since for a given $e \in \mathcal{H}_d$, the forms $\{L_e^k : 1 \leq k \leq d\}$ are linearly independent, and $\mathcal{L}$ is a **pairwise linearly independent** family of linear forms. In addition, for $x \in M = \{x_0 + \cdots + x_d = 0\}$, we have $L_e^k(x) = L_{e'}^k(x)$ for all $e, e' \in \mathcal{H}_d$, $1 \leq k \leq d$. We call a family of linear forms $\mathcal{L} = \{L_e^k : e \in \mathcal{H}_d, 1 \leq k \leq d\}$ satisfying this property **symmetric**.

We now formulate a hypergraph removal lemma with respect to the weight system $\{\nu_e\}_{e \subseteq J}$ defined in (1.10)–(1.12), and show that it implies Theorem 1.3.

Theorem 1.4. (Weighted Simplex Removal Lemma). *Let $\{\nu_e\}_{e \subseteq J}, \{\mu_e\}_{e \subseteq J}$ be a system of weights and measures associated with a well-defined, pairwise linearly independent and symmetric family of linear forms $\mathcal{L}$, as defined in (1.10)–(1.12). Let $E_e \subseteq \mathcal{A}_e$, $g_e : V_e \rightarrow [0, 1]$ be given for $e \in \mathcal{H}_d$. Then, for a given $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon) > 0$, such that the following holds. If*

$$\mathbb{E}_{x \in V_J} \prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}(x) \mu_J(x) \leq \delta \tag{1.19}$$

then there exists a well-defined, symmetric family of linear forms $\tilde{\mathcal{L}} = \{\tilde{L}_e^k : e \in \mathcal{H}_d, 1 \leq k \leq d\}$, such that the associated system of weights and measures $\{\tilde{\nu}_e\}_{e \subseteq J}, \{\tilde{\mu}_e\}_{e \subseteq J}$ satisfies

$$\mathbb{E}_{x \in V_J} \prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}(x) \tilde{\mu}_J(x) = \mathbb{E}_{x \in V_J} \prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}(x) \mu_J(x) + o_{N, W \rightarrow \infty}(1) \tag{1.20}$$

and for all $e \in \mathcal{H}_d$,

$$\mathbb{E}_{x \in V_J} g_e(x) \tilde{\mu}_e(x) = \mathbb{E}_{x \in V_J} g_e(x) \mu_e(x) + o_{N, W \rightarrow \infty}(1). \tag{1.21}$$

In addition, there exist sets $E'_e \in \mathcal{A}_e$, such that

$$\bigcap_{e \in \mathcal{H}_d} (E_e \cap E'_e) = \emptyset \quad (1.22)$$

and for all $e \in \mathcal{H}_d$,

$$\mathbb{E}_{x \in V_J} \mathbf{1}_{E_e \setminus E'_e}(x) \tilde{\mu}_e(x) \leq \varepsilon + o_{N,W \rightarrow \infty}(1). \quad (1.23)$$

Naturally, one would like to establish (1.22)–(1.23) for $\{\tilde{\mu}_e\} = \{\mu_e\}$ as that would easily imply Theorem 1.3 and hence our main result Theorem 1.2. Formulas (1.20)–(1.21) mean that the measures $\tilde{\mu}_J, \tilde{\mu}_e$ are small perturbations of the measures μ_J, μ_e with respect to the functions $g = \prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}$ and g_e . The reason why we have to perturb the measures is the existence of weights μ_e on lower dimensional edges $1 < |e| < d$ when the configuration δ is not in general position. Removing these weights does not seem amenable to known “transference arguments”, see [1, 17].

Let us remark that if one drops the “symmetry” requirement $\mu_e(x) = \mu_{e'}(x)$, $e, e' \in \mathcal{H}_d$, $x \in M$, then one could formulate a simpler, purely combinatorial form of the weighted removal lemma, requiring only that the system of weights $\{\nu_e\}_{e \subseteq J}$ to be pseudo-random of an order $m = m(\varepsilon)$, without referring to the underlying linear forms. In that case, one could construct the perturbed weights from the weights on certain hypergraph bundles, emerging in our energy increment process. However, as we shall see below, we need to make sure that our weights and measures are invariant under certain linear transformations of the variables and have to go through the more tedious route of extending our underlying family of linear forms by introducing new variables, considered as parameters.

Proof. [Theorem 1.4 implies Theorem 1.3] Let $\alpha > 0$ and let A be a subset of $\mathbb{Z}_N^d$. Set $\varepsilon := \alpha/(d+2)$ and let $G = \bigcap_{e \in \mathcal{H}_d} E_e$ be the hypergraph associated with the set A as defined above. For a given $e' \in \mathcal{H}_d$, define the function $g_{e'}: V_{e'} \rightarrow [0, 1]$ as follows. Let $\phi_{e'}: V_{e'} \rightarrow M$ be the inverse of the projection map $\pi_{e'}: V_J \rightarrow V_{e'}$ restricted to M , and for $z \in V_{e'}$, let

$$g_{e'}(z) := \prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}(\phi_{e'}(z)).$$

Applying Theorem 1.4 to the system of weights $\{\nu_e\}$, the sets E_e and the functions $\{g_e\}$ give a constant $\delta = \delta(\varepsilon) > 0$, a system of measures $\tilde{\mu}_e$, and sets $E'_e \subseteq A_e$ satisfying (1.20)–(1.23).

Now, assume that, as in Theorem 1.3,

$$\mathbb{E}_{y,t} w(y - t\Delta) = \mathbb{E}_{x \in V_J} \prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}(x) \mu_J(x) \leq \delta.$$

By (1.9), we have that $x \in M \bigcap_{e \in \mathcal{H}_d} E_e$ if and only if $\Phi(x) = (y, 0)$ with $y \in A$. Moreover, in that case, $w(y) = \mu_e(x)$ for all $e \in \mathcal{H}_d$ by (1.15) and thus for any given $e' \in \mathcal{H}_d$

$$\begin{aligned}
\mathbb{E}_{y \in \mathbb{Z}_N^d} \mathbf{1}_A(y) w(y) &= \mathbb{E}_{x \in M} \prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}(x) \mu_{e'}(x) = \mathbb{E}_{z \in V_{e'}} g_{e'}(z) \mu_{e'}(z) \\
&= \mathbb{E}_{z \in V_{e'}} g_{e'}(z) \tilde{\mu}_{e'}(z) + o_{N,W \rightarrow \infty}(1) \\
&= \mathbb{E}_{x \in M} \prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}(x) \tilde{\mu}_{e'}(x) + o_{N,W \rightarrow \infty}(1).
\end{aligned}$$

By (1.22), clearly, $\prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}(x) \leq \sum_{e \in \mathcal{H}_d} \mathbf{1}_{E_e \setminus E'_e}(x)$. Then, the symmetry of the measures $\tilde{\mu}_e$ (i.e., $\tilde{\mu}_e(x) = \tilde{\mu}_{e'}(x)$ for $x \in M$), (1.23) and the fact that $\mathbf{1}_{E_e \setminus E'_e}$ is constant on the fibers $\pi_e^{-1}(x)$ implies

$$\begin{aligned}
\mathbb{E}_{x \in M} \prod_{e \in \mathcal{H}_d} \mathbf{1}_{E_e}(x) \tilde{\mu}_{e'}(x) &\leq \sum_{e \in \mathcal{H}_d} \mathbb{E}_{x \in M} \mathbf{1}_{E_e \setminus E'_e}(x) \tilde{\mu}_{e'}(x) \\
&= \sum_{e \in \mathcal{H}_d} \mathbb{E}_{x \in V_{\mathcal{J}}} \mathbf{1}_{E_e \setminus E'_e}(x) \tilde{\mu}_e(x) \\
&\leq (d+1) \epsilon(\delta) + o_{N,W \rightarrow \infty}(1) \leq \alpha,
\end{aligned}$$

choosing N, W sufficiently large with respect to α and Theorem 1.3 follows. $\square$

1.5. Weighted Box Norms and Hypergraph Regularity

The known proofs of the Simplex Removal Lemma rely on the so-called hypergraph regularity lemma and the associated counting lemma [5, 13, 19], and in particular the notion of a regular or pseudo-random hypergraph. This can be defined in different ways, we shall use a variant of Gowers' box norms [5] adapted to our settings.

Let $e \in \mathcal{H}_d$ be fixed. For a given $\omega \in \{0, 1\}^e$ (i.e., $\omega: e \rightarrow \{0, 1\}$), define the orthogonal projection $\omega_e: V_e \times V_e \rightarrow V_e$ by

$$\omega_e(x_e, q_e)_i = \begin{cases} x_i & \text{if } \omega_i = 0, \\ q_i & \text{if } \omega_i = 1, \end{cases} \quad (1.24)$$

for $i \in e$, and the weighted box norm of a function $F: V_e \rightarrow \mathbb{R}$, using the notation $x_f := \pi_f(x)$ for $f \subseteq J$, as

$$\|F\|_{\square_{\nu_e}}^{2^d} = \mathbb{E}_{x, q \in V_e} \prod_{\omega \in \{0, 1\}^e} F(\omega_e(x, q)) \prod_{f \subseteq e} \prod_{\omega \in \{0, 1\}^f} \nu_f(\omega_f(x_f, q_f)). \quad (1.25)$$

Note that if $\nu_f \equiv 1$ for all $f \subseteq e$, then $\|F\|_{\square_{\nu_e}} = \|F\|_{\square}$ is the usual box norm.

Example 3. Let $e = (1, 2)$ and $F: V_e = V_1 \times V_2 \rightarrow \mathbb{R}$. Then

$$\begin{aligned}
\|F\|_{\square_{\nu_e}}^4 &= \mathbb{E}_{x, q \in V_e} F(x_1, x_2) F(x_1, q_2) F(q_1, x_2) F(q_1, q_2) \\
&\quad \times \nu_e(x_1, x_2) \nu_e(x_1, q_2) \nu_e(q_1, x_2) \nu_e(q_1, q_2) \nu_1(x_1) \nu_2(x_2) \nu_1(q_1) \nu_2(q_2).
\end{aligned}$$

The points $\omega_e(x, q)$ and $\omega_f(x_f, q_f)$ may be viewed as the faces and edges of a d -dimensional octahedron $\mathcal{K}_d$ with vertices $\{x_j, q_j : j \in e\}$. The inner product in (1.25) represents the total weight of the octahedron obtained by multiplying the weights of all edges and vertices. The boxnorm itself is the weighted average of F over all embeddings of the hypergraph $\mathcal{K}_d$.

It is not hard to see that the $\square_\nu$ -norm is indeed a norm (for $d \geq 2$) and an appropriate version of the Gowers–Cauchy–Schwarz inequality holds (see the Appendix). The importance of this norm is that it controls weighted averages over $(d + 1)$ -dimensional simplices, which plays an important role in proving the counting lemma. More precisely one has the following.

Proposition 1.1 (Weighted von Neumann inequality). *Let $F_e: V_e \rightarrow \mathbb{R}$ be a given function, such that $|F_e| \leq 1$ for each $e \in \mathcal{H}_d$. Then, there is an absolute constant C such that*

$$\left| \mathbb{E}_{x \in V_J} \prod_{e \in \mathcal{H}_d} F_e(\pi_e(x)) \mu_J(x) \right| \leq C \min_{e \in \mathcal{H}_d} \|F_e\|_{\square_{\nu_e}} + o_{N,W \rightarrow \infty}(1). \quad (1.26)$$

The above inequality motivates the following

Definition 1.1. Let $e \in \mathcal{H}_d$ and $\varepsilon > 0$ be fixed and let $G_e \subseteq V_e$ be a d -regular hypergraph. We say that G_e is ε -regular with respect to the weight system $\{\nu_f\}_{f \subseteq e}$ if

$$\|\mathbf{1}_{G_e} - \mu_e(G_e) \mathbf{1}_{V_e}\|_{\square_{\nu_e}} \leq \varepsilon. \quad (1.27)$$

It is easy to see from Proposition 1.1 that if the sets $E_e \in \mathcal{A}_e$ are $c\varepsilon$ -regular for all $e \in \mathcal{H}_d$ ($0 < c \ll 1$), then Theorem 1.4 holds with $\tilde{\mu}_e = \mu_e$ and $\delta = c\varepsilon^{d+1}$. Indeed, writing $G_e = \pi_e(E_e)$, $\mathbf{1}_{G_e} = \mu_e(G_e) + F_e$, and substituting this decomposition into the left side of (1.19), we get $2^{d+1} - 1$ error terms each of which is bounded by $c'\varepsilon^{d+1}$ (with $0 < c' \ll 1$), and a main term of the form $\prod_{e \in \mathcal{H}_d} \mu_e(G_e) \mu_J(V_J)$. The forms defining the measure μ_J are pairwise linearly independent and hence $\mu_J(V_J) = 1 + o(1)$. Thus, by assumption (1.19) of Theorem 1.4 $\mu_J(E_e) = \mu_e(G_e)(1 + o_{N,W \rightarrow \infty}(1)) \leq \varepsilon$ (see Sect. 2), for at least one $e \in \mathcal{H}_d$ and the sets $E'_e := \emptyset$, $E'_{e'} := E_e$ ($e' \neq e$) satisfy the conclusion of Theorem 1.4.

In general, the hypergraphs G_e are not sufficiently regular and the bulk of our argument is to obtain a *regularity lemma* in our weighted setting. This roughly means to partition the sets G_e into sufficiently regular hypergraphs with respect to a system of measures $\tilde{\mu}_e$, which are small perturbations of the initial measures μ_e . Our proof is motivated by the iterative process described in [19]; however, we need to modify the entire argument because of the presence of weights on the lower dimensional edges. During the process, we construct parametric families of weight systems $\{\nu_{q,e}\}_{e \in \mathcal{H}, q \in \Omega}$ which for most values of the parameter, q will give rise to small perturbations of the initial system $\{\nu_e\}_{e \in \mathcal{H}}$.

1.6. Sketch of the Proof in Dimension 2

We describe in detail the main elements of the proof in dimension $d = 2$ (and illustrate some of the constructions of the configuration $\Delta = \{(0,0), (2,0), (1,1)\}$). We do this as in full generality both the notations and the arguments are quite complex, but the essential ideas can be illustrated in this case.

1. *The Energy Increment.* Let $\eta > 0$ and assume that for some edge e , say $e = (1, 2)$, the graph $G_e = \pi_e(E_e)$ is not η -regular. This means

$$\|F\|_{\square_{\nu_e}} \geq \eta, \quad (1.28)$$

where $F = \mathbf{1}_{G_e} - \mu_e(G_e) \mathbf{1}_{V_e}$. In view of definition (1.25), we may write

$$\|F\|_{\square_{\nu_e}}^4 = \int_{V_e} \int_{V_e} F(x) u_q^1(x_1) u_q^2(x_2) \nu_e(x_1, q_2) \nu_e(q_1, x_2) d\mu_e(x) d\mu_e(q) \geq \eta^4, \quad (1.29)$$

where $x = (x_1, x_2)$, $q = (q_1, q_2)$, $u_q^1(x_1) = F(x_1, q_2)$, and $u_q^2(x_2) = F(q_1, x_2)$ (q_1, q_2). We consider the variables q as parameters and define the measures $\mu_{q,e}$, depending on q , by

$$\mu_{q,e}(x) := \nu_e(x_1, q_2) \nu_e(q_1, x_2) \mu_e(x).$$

Example 4. If $\Delta = \{(0, 0), (2, 0), (1, 1)\}$, then $\nu_e(x) = \nu(2x_1 + x_2)$, $\mu_e(x) = \nu(2x_1 + x_2) \nu(x_2)$ and $\mu_{q,e}(x) = w(q, x) \mu_e(x)$ with $w(q, x) = \nu(2x_1 + q_2) \nu(2q_1 + x_2)$. Note that both forms $2x_1 + q_2$ and $2q_1 + x_2$ depend on both the x and q variables, which hold in general by (1.12).

The inner integral in (1.29) may be viewed as the inner product

$$\Gamma(q) := \langle F, u_q^1 \times u_q^2 \rangle_{\mu_{q,e}} = \int_{V_e} F(x) u_q^1(x_1) u_q^2(x_2) d\mu_{q,e}(x), \quad (1.30)$$

on the Hilbert space $L^2(V_e, \mu_{q,e})$. Thus, (1.29) translates to $\mathbb{E}_{q \in V_e} \Gamma(q) \mu_e(q) \geq \eta^4$, while using the linear form condition, it is easy to see that

$$\mathbb{E}_{q \in V_e} \Gamma(q)^2 \mu_e(q) \leq \mathbb{E}_{q,x,y} \mu_{q,e}(x) \mu_{q,e}(y) \mu_e(q) = 1 + o_{N,W \rightarrow \infty}(1),$$

as the underlying linear forms defining $\mu_{q,e}(x)$ are pairwise linearly independent and depend on at least one of the x variables. Alternatively, it also follows from the pseudo-randomness of the weight system $\{\nu_{12}, \nu_1, \nu_2\}$, applied to the hypergraph bundle obtained by sampling the variables x_1, y_1, q_1 from V_1 and x_2, y_2, q_2 from V_2 . Thus

$$\Gamma(q) \gtrsim \eta^4, \quad \text{for } q \in \Omega, \quad (1.31)$$

for a set $\Omega \subseteq V_e$ of measure $\mu_e(\Omega) \gtrsim \eta^8$.

Fix $q \in \Omega$. As the functions u_q^i ($i = 1, 2$) are bounded, without loss of generality, we may assume that they are indicator functions of the sets $U_q^i \subseteq V_i$. Let $\mathcal{B}_{q,e} := \mathcal{B}_q^1 \vee \mathcal{B}_q^2$ denote the σ -algebra generated by the sets U_q^i ($i = 1, 2$) on V_e , and let $\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e} | \mathcal{B}_q)$ be the conditional expectation function of $\mathbf{1}_{G_e}$ with respect to this σ -algebra and the measure $\mu_{q,e}$. Recall that if $(V, \mathcal{B}, \mu)$ is a finite atomic measure space, and $B \in \mathcal{B}_q$ is an atom, then $\mathbb{E}_{\mu}(f | \mathcal{B})(x) = \frac{1}{\mu(B)} \int_B f d\mu$ for all $x \in B$. It is easy to see that $\mathbb{E}_{\mu}(f | \mathcal{B})$ is the orthogonal projection of f to the subspace of $\mathcal{B}$ -measurable functions in $L^2(V, \mathcal{B}, \mu)$.

As $u_q^1 \times u_q^2$ is measurable with respect to $\mathcal{B}_q$, the inner product

$$\langle \mathbf{1}_{G_e} - \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e} | \mathcal{B}_{q,e}), u_q^1 \times u_q^2 \rangle_{\mu_{q,e}} = 0.$$

This together with (1.30) and (1.31) implies

$$\langle \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e} | \mathcal{B}_{q,e}) - \mathbb{E}_{\mu_e}(\mathbf{1}_{G_e} | \mathcal{B}_0), u_q^1 u_q^2 \rangle_{\mu_{q,e}} \gtrsim \eta^4,$$

where $\mathcal{B}_0 = \{V_e, \emptyset\}$ is the trivial σ -algebra, and hence, $\mathbb{E}_{\mu_e}(\mathbf{1}_{G_e} | \mathcal{B}_0) = \mu_e(G_e) \mathbf{1}_{V_e} + o_{N,W \rightarrow \infty}(1)$. Then, by the Cauchy–Schwarz inequality, we have

$$\|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e} | \mathcal{B}_{q,e}) - \mathbb{E}_{\mu_e}(\mathbf{1}_{G_e} | \mathcal{B}_0)\|_{L^2(\mu_{q,e})}^2 \gtrsim \eta^8. \quad (1.32)$$

By the Pythagoras theorem, if the second term on the left side would be a conditional expectation with respect to the measure $\mu_{q,e}$, then one would obtain an “energy increment”

$$\begin{aligned} & \|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e} | \mathcal{B}_{q,e}) - \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e} | \mathcal{B}_0)\|_{L^2(\mu_{q,e})}^2 \\ &= \|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e} | \mathcal{B}_{q,e})\|_{L^2(\mu_{q,e})}^2 - \|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e} | \mathcal{B}_0)\|_{L^2(\mu_{q,e})}^2 \gtrsim \eta^8. \end{aligned}$$

To overcome this “discrepancy”, a key observation of our approach is that the measures $\mu_{q,e}$ are negligible perturbations of the measure μ_e , for almost all values of the parameter q , which can be proved using the linear forms condition. To be more precise, for any given set $B \subseteq V_e$, one has for almost every $q \in V_e$

$$\mu_{q,e}(B) - \mu_e(B) = o_{N,W \rightarrow \infty}(1). \quad (1.33)$$

The detailed proof of this, and other similar statements will be given in Sect. 2, and we sketch the argument below. First, it is enough to show that

$$\Lambda := \mathbb{E}_{q \in V_e} |\mu_{q,e}(B) - \mu_e(B)|^2 \mu_e(q) = o_{N,W \rightarrow \infty}(1).$$

Since $\mu_{q,e}(x) = w(q, x) \mu_e(x)$, with $w(q, x) = \nu_e(x_1, q_2) \nu_e(q_1, x_2)$, one estimates

$$\begin{aligned} \Lambda &= \mathbb{E}_{q \in V_e} \mathbb{E}_{x, y \in V_e} \mathbf{1}_B(x) \mathbf{1}_B(y) (w(q, x) - 1)(w(q, y) - 1) \mu_e(x) \mu_e(y) \mu_e(q) \\ &\leq \mathbb{E}_{x, y \in V_e} \mu_e(x) \mu_e(y) |\mathbb{E}_{q \in V_e} (w(q, x) - 1)(w(q, y) - 1) \mu_e(q)|. \end{aligned}$$

By the linear form condition $\mathbb{E}_{x \in V_e} \mu_e(x) = 1 + o_{N,W \rightarrow \infty}(1) \leq 2$ and the Cauchy–Schwarz inequality,

$$\begin{aligned} \Lambda^2 &\lesssim \mathbb{E}_{x, y \in V_e} \mathbb{E}_{q, p \in V_e} (w(q, x) - 1)(w(q, y) - 1)(w(p, x) - 1)(w(p, y) - 1) \\ &\quad \times \mu_e(x) \mu_e(y) \mu_e(q) \mu_e(p). \end{aligned}$$

The above expression is an alternating sum of 16-terms, each being $1 + o_{N,W \rightarrow \infty}(1)$ by the linear form condition. Indeed, the linear forms defining the measures $\mu_e(x)$, $\mu_e(y)$, $\mu_e(q)$, $\mu_e(p)$ and the weight functions $w(q, x)$, $w(q, y)$, $w(p, x)$, $w(p, y)$ are pairwise linearly independent in the (x, y, q, p) variables, as either their support is different or by the fact that the measure $\mu_e(x)$ and the weight function $w(q, x)$ are defined via pairwise independent forms. Alternatively, it follows from order-4 pseudo-randomness of the weight system $\{\nu_{12}, \nu_1, \nu_2\}$ applied to the hypergraph bundle obtained by sampling the variables x_i, y_i, q_i, p_i from the vertex sets V_i , for $i = 1, 2$. This in turn implies that

$$\|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e} | \mathcal{B}_0) - \mathbb{E}_{\mu_e}(\mathbf{1}_{G_e} | \mathcal{B}_0)\|_{L^2(\mu_{q,e})} = o_{N,W \rightarrow \infty}(1) \quad (1.34)$$

and

$$\|\mathbb{E}_{\mu_e}(\mathbf{1}_{G_e} | \mathcal{B}_0)\|_{L^2(\mu_e)} = \|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e} | \mathcal{B}_0)\|_{L^2(\mu_{q,e})} + o_{N,W \rightarrow \infty}(1). \quad (1.35)$$

Then, from (1.32), assuming that N, W are sufficiently large with respect to η , we have that

$$\|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e}|\mathcal{B}_{q,e})\|_{L^2(\mu_{q,e})}^2 \geq \|\mathbb{E}_{\mu_e}(\mathbf{1}_{G_e}|\mathcal{B}_0)\|_{L^2(\mu_e)}^2 + c\eta^8, \quad (1.36)$$

uniformly for $q \in \Omega_1$, for a set $\Omega_1 \subseteq \Omega$ of measure $\nu_e(\Omega_1) \geq c\eta^8$.

If $F: V \rightarrow \mathbb{R}$ is a function and $(V, \mathcal{B}, \mu)$ is a measure space, the quantity $\|\mathbb{E}_\mu(F|\mathcal{B})\|_{L^2(\mu)}^2$ will be referred to as the “energy” of the function F with respect to the measure space $(V, \mathcal{B}, \mu)$, so (1.36) is telling that if G_e is not η -uniform with respect to the initial measure spaces $(V_e, \mathcal{B}_0, \mu_e)$, then its energy increases by $c\eta^8$, with an absolute constant $c > 0$, when passing to the measure spaces $(V_e, \mathcal{B}_{q,e}, \mu_{q,e})$ for $q \in \Omega_1$.

We symmetrize the forms $\mathcal{L}_{q,e}$ as follows. First pull back these forms to the subspace M via the map $\pi_e: M \rightarrow V_e$, i.e., define $\mathcal{L}_{q,M} := \{L \circ \pi_e: L \in \mathcal{L}_{q,e}\}$, and for an edge e' set $\mathcal{L}_{q,e'} := \{L \circ \pi_{e'}^{-1}: L \in \mathcal{L}_{q,M}\}$. If $e' = J \setminus \{j'\}$, this amounts to the change of variables $x_{j'} := -\sum_{i \neq j'} x_i$. The family of linear forms $\mathcal{L}_q := \bigcup_{e \in \mathcal{H}_2} \mathcal{L}_{q,e}$ is symmetric and well-defined in the x -variables, and also pairwise linearly independent in the (q, x) -variables. It is at this point, where we have to construct weights using linear transformations and not just extending our weights to hypergraph bundles.

Example 5. In the case $\Delta = \{(0, 0), (2, 0), (1, 1)\}$, the measure $\mu_{q,e}$ for $e = (1, 2)$ is defined via the family of linear forms $\mathcal{L}_{q,e} = \{2x_1 + x_2, 2q_1 + x_2, 2x_1 + q_2, x_2\}$, which are transferred to edges $e_1 = (0, 2)$ and $e_2 = (0, 1)$ to obtain $\mathcal{L}_{q,e_1} = \{-2x_0 - x_2, 2q_1 - x_2, -2x_0 - 2x_2 + q_2, x_2\}$ and $\mathcal{L}_{q,e_2} = \{x_1 - x_0, -x_0 - x_1, 2x_1 + q_2, 2q_1 - x_0 - x_1\}$.

For given q , we define the corresponding weights and measures as in (1.10) and (1.12), i.e., set

$$\nu_{q,e}(x) := \prod_{L \in \mathcal{L}_q, \text{supp}_x(L)=e} \nu(L(q, x)), \quad \mu_{q,e}(x) := \prod_{f \subseteq e} \nu_{q,f}(x),$$

where $\text{supp}_x(L)$ is the support of a linear form $L(q, x)$ in the x -variables. Finally, set $\mathcal{B}_{q,0} := \mathcal{B}_0 = \{V_0, \emptyset\}$, and note that for each $q \in \Omega_1$, we have a partition $\mathcal{B}_{q,i}$ (considered as a σ -algebra) on the vertex set V_i into at most 2 sets. This gives rise to a partition $\mathcal{B}_{q,e} = \mathcal{B}_{q,i} \vee \mathcal{B}_{q,j}$ for an edge $e = (i, j)$. By the stability property (1.32), we have

$$\begin{aligned} \|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e}|\mathcal{B}_{q,e})\|_{L^2(\mu_{q,e})}^2 &\geq \|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e}|\mathcal{B}_0)\|_{L^2(\mu_{q,e})}^2 \\ &= \|\mathbb{E}_{\mu_e}(\mathbf{1}_{G_e}|\mathcal{B}_0)\|_{L^2(\mu_e)}^2 + o_{N,W \rightarrow \infty}(1). \end{aligned}$$

Thus, one obtains a symmetric version of the energy increment (1.36), namely, for $q \in \Omega_1$

$$\sum_{e \in \mathcal{H}_2} \|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e}|\mathcal{B}_{q,e})\|_{L^2(\mu_{q,e})}^2 \geq \sum_{e \in \mathcal{H}_2} \|\mathbb{E}_{\mu_e}(\mathbf{1}_{G_e}|\mathcal{B}_0)\|_{L^2(\mu_e)}^2 + c\eta^8. \quad (1.37)$$

2. The Koopman–von Neumann Decomposition. The above energy increment argument applies with slight changes to the parametric system of weights and measures $\nu_{q,e}$ and $\mu_{q,e}$. To be more precise, assume that there is an edge, say

$e = (1, 2)$ and set $\Omega'_1 \subseteq \Omega_1$ of size $c_1 := \mu_e(\Omega'_1) \geq \mu_e(\Omega_1)/2$, such that for all $q \in \Omega'_1$, one has

$$\|\mathbf{1}_{G_e} - \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e} | \mathcal{B}_{q,e})\|_{\square_{\nu_{q,e}}} \geq \eta.$$

Writing $F_q := \mathbf{1}_{G_e} - \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e})$, we have as in (1.29)

$$\begin{aligned} \|F_q\|_{\square_{\nu_{q,e}}}^4 &= \int_{V_e} \int_{V_e} F_q(x) u_{q,p}^1(x_1) u_{q,p}^2(x_2) \nu_{q,e}(x_1, p_2) \nu_{q,e}(p_1, x_2) \\ &\quad \times d\mu_{q,e}(x) d\mu_{q,e}(p) \geq \eta^4, \end{aligned} \quad (1.38)$$

where $x = (x_1, x_2)$, $q = (q_1, q_2)$, $p = (p_1, p_2)$, and $u_{q,p}^1, u_{q,p}^2$ are functions bounded by 1. Define the measure

$$\mu_{q,p,e}(x) := \nu_{q,e}(x_1, p_2) \nu_{q,e}(p_1, x_2) \mu_{q,e}(x), \quad (1.39)$$

and note that it is defined by a parametric family of linear forms $\mathcal{L}_{q,p}$ which are pairwise independent in the (q, p, x) -variables. The space of parameters $Z := V_e \times V_e$ is equipped with the measure $\psi(q, p) = \mu_{q,e}(p) \mu_e(q)$. Then, the inner expression in (1.38) may be viewed as the inner product

$$\Gamma(q, p) := \langle F_q, u_{q,p}^1 \times u_{q,p}^2 \rangle_{\mu_{q,p,e}}, \quad (1.40)$$

on the Hilbert space $L^2(V_e, \mu_{q,p,e})$. Thus, (1.38) translates to $\mathbb{E}_{p \in V_e} \Gamma(q, p) \mu_{q,e}(p) \geq \eta^4$ for $q \in \Omega'_1$. By (1.32), one has that $\mu_{q,e}(V_e) = 1 + o_{N,W \rightarrow \infty}(1) \leq 2$ for “almost all” $q \in \Omega'_1$, thus

$$\int_{V_e} \mathbf{1}_{\{\Gamma(q,p) \geq \eta^4/4\}} \Gamma(q, p) d\mu_{q,e}(p) \geq \frac{\eta^4}{2}.$$

On the other hand using the linear form condition as in (1.31), one has

$$\int_{V_e} \int_{V_e} \Gamma(q, p)^2 d\mu_{q,e}(p) d\mu_e(q) = 1 + o_{N,W \rightarrow \infty}(1) \leq 2,$$

for N, W sufficiently large. If we set $\Omega'_2 := \{(q, p) \in \Omega'_1 \times V_e : \Gamma(q, p) \geq \eta^4/2\}$, then by the Cauchy–Schwarz inequality,

$$\frac{c_1^2 \eta^8}{4} \leq \left(\int_{\Omega'_2} \Gamma(q, p) d\mu_{q,e}(p) d\mu_e(q) \right)^2 \leq 2 \psi(\Omega'_2). \quad (1.41)$$

This means that $\Gamma(q, p) \geq \eta^4/4$ for $(q, p) \in \Omega'_2$, for a set $\Omega'_2 \subseteq \Omega'_1 \times V_e$ of measure $\psi(\Omega'_2) \geq 2^{-4} \eta^8 \mu_e(\Omega_1)^2$.

Since for $q' = (q, p) \in \Omega'_2$, the functions $u_{q'}^i$ ($i = 1, 2$) are bounded by 1, one may assume $u_{q'}^i = \mathbf{1}_{U_{q'}^i}$ for some sets $U_{q'}^i \subseteq V_i$, hence (1.40) becomes

$$\left\langle \mathbf{1}_{G_e} - \mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e} | \mathcal{B}_{q',e}), \mathbf{1}_{U_{q'}^1} \times \mathbf{1}_{U_{q'}^2} \right\rangle_{\mu_{q',e}} \geq \frac{1}{4} \eta^4. \quad (1.42)$$

Since $\mathbf{1}_{U_{q'}^1} \times \mathbf{1}_{U_{q'}^2}$ is constant on the atoms of $\mathcal{B}_{q',e}$, generated by $\mathcal{B}_{q,e}$ and the sets $\pi_i^{-1} \mathbf{1}_{U_{q'}^i}$ ($\pi_i : V_e \rightarrow V_i$ being the natural projection), one has

$$\left\langle \mathbf{1}_{G_e} - \mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e} | \mathcal{B}_{q',e}), \mathbf{1}_{U_{q'}^1} \times \mathbf{1}_{U_{q'}^2} \right\rangle_{\mu_{q',e}} = 0,$$

hence, uniformly for $q' = (q, p) \in \Omega'_2$

$$\left\langle \mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e}|\mathcal{B}_{q',e}) - \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e}|\mathcal{B}_{q,e}), \mathbf{1}_{U_{q'}^1} \times \mathbf{1}_{U_{q'}^2} \right\rangle_{\mu_{q',e}} \geq \frac{1}{4} \eta^4. \quad (1.43)$$

Then, by the Cauchy–Schwarz inequality

$$\|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e}|\mathcal{B}_{q',e}) - \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e}|\mathcal{B}_{q,e})\|_{L_{\mu_{q',e}}^2} \geq \frac{1}{8} \eta^4. \quad (1.44)$$

The measure $\mu_{q,p,e}(x)$ defined in (1.39) is of the form $\mu_{q,p,e}(x) = w(q, p, x)\mu_{q,e}(x)$ with a weight function $w(q, p, x)$ defined by pairwise linearly independent forms depending on *both* the x and p variables; hence, the stability properties (1.34)–(1.35) remain valid. Thus, for a set $\Omega_2 \subseteq \Omega'_2$, such that $\psi(\Omega'_2 \setminus \Omega_2) = o_{N,W \rightarrow \infty}(1)$, one has for all $q' = (q, p) \in \Omega_2$

$$\|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e}|\mathcal{B}_{q',e})\|_{L_{\mu_{q',e}}^2}^2 - \|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e}|\mathcal{B}_{q,e})\|_{L_{\mu_{q,e}}^2}^2 \geq 2^{-7} \eta^8. \quad (1.45)$$

Symmetrizing the forms $\mathcal{L}_{q',e}$, one obtains an energy increment of $2^{-8} \eta^8$ of the total energy:

$$\sum_{e \in \mathcal{H}_2} \|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e}|\mathcal{B}_{q',e})\|_{L_{\mu_{q',e}}^2}^2 \geq \sum_{e \in \mathcal{H}_2} \|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e}|\mathcal{B}_{q,e})\|_{L_{\mu_{q,e}}^2}^2 + 2^{-8} \eta^8, \quad (1.46)$$

uniformly for $q' = (q, p) \in \Omega_2$, for a set $\Omega_2 \subseteq Z$ of measure $\psi(\Omega_2) \geq 2^{-5} \eta^8 \mu_2(\Omega_1)^2$.

It is clear that the above procedure can be iterated leading to a family of measure spaces $(V_e, \mathcal{B}_{q_k,e}, \mu_{q_k,e})_{e \in \mathcal{H}_2}$ depending on a parameter $q_k \in \Omega_k$ with $\Omega_k \subseteq \mathbb{Z}_N^k$ of measure $\psi_k(\Omega_k) \geq (\eta/2)^{c_k}$ (for example, with $c_k = 2^{k+2}$), at the k th step. The procedure must stop at a step $k \leq 2^{10} \eta^{-8}$ as the total energy, given in (1.46), is increasing by at least $2^{-8} \eta^8$ in each step and, however, is bounded by $3(1 + o_{N,W \rightarrow \infty}(1)) \leq 4$, for sufficiently large W and N . At this step, there is a set $\Omega'_k \subseteq \Omega_k$, such that for all $q_k \in \Omega_k$ and $e \in \mathcal{H}_2$, we have

$$\|\mathbf{1}_{G_e} - \mathbb{E}_{\mu_{q_k,e}}(\mathbf{1}_{G_e}|\mathcal{B}_{q_k,e})\|_{\square_{\nu_{q_k,e}}} \leq \eta.$$

Given an edge $e = (i, j)$, this gives a *Koopman–von Neumann*-type decomposition $\mathbf{1}_{G_e} = f_{q_k,e} + g_{q_k,e}$ with $f_{q_k,e}$ being constant on the atoms of $\mathcal{B}_{q_k,e} = \mathcal{B}_{q_k}^i \vee \mathcal{B}_{q_k}^j$, while $g_{q_k,e}$ has small $\square_{\nu_{q_k,e}}$ -norm. This is not quite enough to prove a removal lemma as the size of the atoms are too small with respect to the regularity parameter η .

3. The Regularity Lemma. We will iterate the above decomposition to arrive at a partition of the vertex spaces, so that on most atoms the graphs G_e become sufficiently uniform. Originally, such partitions of graphs were obtained by Szemerédi [16], and we will follow the approach of Tao [18] as it is more amenable to extensions to hypergraphs.

Let $(V_e, \mathcal{B}_{q,e}, \mu_{q,e})_{e \in \mathcal{H}_2}$ be well-defined, symmetric system of measure spaces, depending on a parameter $q \in \mathbb{Z}_N^K$ and let $\Omega \subseteq \mathbb{Z}_N^K$ be a set of measure $\psi(\Omega) := c_0 > 0$. Assume that the σ -algebras $\mathcal{B}_{q,e} = \mathcal{B}_{q,i} \vee \mathcal{B}_{q,j}$ ($e = (i, j)$) have complexity (i.e., the least number of sets needed to generate them) bounded by M_0 uniformly in q .

Given $\varepsilon > 0$ and a function $F(M, \varepsilon) > 0$, what we would like to achieve is a symmetric, well-defined extension $(V_e, \mu_{q',e})_{e \in \mathcal{H}_2}$ and families of σ -algebras $\mathcal{B}_{q',i} \subseteq \mathcal{B}'_{q',i}$ ($i = 1, 2, 3$) for which the following holds. There is a set $\Omega' \subseteq \Omega \times \mathbb{Z}_N^l$ of measure $\psi'(\Omega') \geq c(F, M_0, c_0, \varepsilon) > 0$ such that uniformly for $q' = (q, p) \in \Omega'$, one has the following properties.

- (1) The complexity of the σ -algebras $\mathcal{B}_{q',i}$ is at most $M = O(M_0, F, \varepsilon)$.
- (2) For all $e \in \mathcal{H}_2$

$$\|\mathbf{1}_{G_e} - \mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e} | \mathcal{B}'_{q',e})\|_{\square_{\nu_{q',e}}} \leq \frac{1}{F(M, \varepsilon)}.$$

- (3) For all $e \in \mathcal{H}_2$

$$\|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e} | \mathcal{B}_{q',e}) - \mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e} | \mathcal{B}'_{q',e})\|_{L^2_{\mu_{q',e}}} \leq \varepsilon.$$

The second property means that the uniformity parameter can be taken arbitrary small with respect to the complexity M of the “coarse” σ -algebras $\mathcal{B}_{q',i}$. The third states that the approximations on the coarse and fine scales to the graphs G_e are sufficiently close in the L^2 sense, see [18].

Apply the Koopman–von Neumann decomposition with uniformity parameter $\eta := F(M_0, \varepsilon)^{-1}$ to the original system $(V_e, \mathcal{B}_{q,e}, \mu_{q,e})_{e \in \mathcal{H}_2}$. This generates a well-defined, symmetric extension $(V_e, \mathcal{B}'_{q',e}, \mu_{q',e})_{e \in \mathcal{H}_2}$. Set $\mathcal{B}_{q',e} := \mathcal{B}_{q,e}$ for $q' = (q, p)$. The new system satisfies (1) and (2) uniformly for $q' = (q, p) \in \Omega'$ for a set $\Omega' \subseteq \Omega \times \mathbb{Z}_N^l$ of measure $\psi'(\Omega_1) \geq \psi'(\Omega')/2 \geq c(F, \varepsilon, M_0, c_0) > 0$. There are two possibilities.

There exists a set $\Omega'_1 \subseteq \Omega'$ of measure $\psi'(\Omega'_1) \geq \psi'(\Omega')/2$, such that (3) also holds uniformly for $q' \in \Omega'_1$. In this case, all the three properties hold for this extension and the σ -algebras $\mathcal{B}_{q',e} \subseteq \mathcal{B}'_{q',e}$.

There exists a set $\Omega'_2 \subseteq \Omega'$ of measure $\psi'(\Omega'_1) \geq \psi'(\Omega')/2$, such that (3) fails for all $q' \in \Omega'_2$. Then, thanks to the stability condition (1.37) and the fact that $\mathcal{B}_{q',e} = \mathcal{B}_{q,e}$, we have for “almost every” $q' = (q, p) \in \Omega'_2$:

$$\begin{aligned} & \sum_{e \in \mathcal{H}_2} \|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e} | \mathcal{B}'_{q',e})\|_{L^2_{\mu_{q',e}}}^2 - \sum_{e \in \mathcal{H}_2} \|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e} | \mathcal{B}_{q,e})\|_{L^2_{\mu_{q,e}}}^2 \\ &= \sum_{e \in \mathcal{H}_2} (\|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e} | \mathcal{B}'_{q',e})\|_{L^2_{\mu_{q',e}}}^2 - \|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e} | \mathcal{B}_{q,e})\|_{L^2_{\mu_{q',e}}}^2) - o_{N,W \rightarrow \infty}(1) \\ &= \sum_{e \in \mathcal{H}_2} \|\mathbb{E}(\mathbf{1}_{G_e} | \mathcal{B}'_{q',e}) - \mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_e} | \mathcal{B}_{q',e})\|_{L^2_{\mu_{q',e}}}^2 - o_{N,W \rightarrow \infty}(1) \geq \frac{\varepsilon^2}{2}. \end{aligned}$$

This means that the total energy of the extended system $(V_e, \mu_{q',e}, \mathcal{B}'_{q',e})$ is at least $\frac{\varepsilon^2}{2}$ -larger than that of the system $(V_e, \mu_{q,e}, \mathcal{B}_{q,e})$ for all $q' = (q, p) \in \Omega'_2$. Repeating this step one generates systems $(V_e, \mu_{q_i,e})$ and σ -algebras $\mathcal{B}_{q_i,e} \subseteq \mathcal{B}'_{q_i,e}$, such that either (1)–(3) hold or the total energy of the system is increased by at least $\frac{\varepsilon^2}{2}$ at the i th step. Since the total energy is $O(1)$, the process must stop at a step $i = O(\varepsilon^{-2})$.

4. The Removal Lemma. We will sketch below how to use the decomposition obtained by the regularity lemma to prove Theorem 1.4. Recall that given

$\varepsilon > 0$, it is enough to show that there is a $\delta = \delta(\varepsilon) > 0$ and a $q \in \Omega$, such that if $\mu(E_{12} \cap E_{13} \cap E_{23}) \leq \delta$, then (1.20)–(1.23) hold for some sets $E'_{q,12}, E'_{q,13}, E'_{q,23}$ for the measures $\tilde{\mu}_e = \mu_{q,e}$, that is,

$$\bigcap_{e \in \mathcal{H}_2} (E_e \cap E'_{q,e}) = \emptyset, \quad \sum_{e \in \mathcal{H}_2} \mu_{q,e}(E_e \setminus E'_{q,e}) \lesssim \varepsilon. \quad (1.47)$$

Note that (1.20)–(1.21) hold for all $q \in \Omega$ outside an exceptional set $\mathcal{E}$ of measure $\psi(\mathcal{E}) = o_{N,W \rightarrow \infty}(1)$.

Therefore, let us assume that we have σ -algebras $\mathcal{B}_{q,i} \subseteq \mathcal{B}'_{q,i}$ on the vertex sets V_i obtained by the regularity lemma with respect to $\varepsilon_1 := \varepsilon^4 > 0$ and a function $F(M, \varepsilon)$; we choose $F(M, \varepsilon) := \varepsilon^{-8} 2^{6M+4}$. Note that the complexity of the σ -algebras $\mathcal{B}_{q,i} \leq M$ uniformly for q , where $M = M(F, \varepsilon) > 0$ is a constant depending only on F and ε .

We will work with the σ -algebras $\bar{\mathcal{B}}_{q,i} \subseteq \bar{\mathcal{B}}'_{q,i}$, $\bar{\mathcal{B}}_{q,i} := \pi_i^{-1}(\mathcal{B}_{q,i})$ together with the σ -algebras $\bar{\mathcal{B}}_{q,e} := \mathcal{B}_e$ generated by the single set E_e on $V = V_1 \times V_2 \times V_3$. The atoms of the σ -algebra $\bar{\mathcal{B}}_q = \bigvee_{e \in \bar{\mathcal{H}}} \bar{\mathcal{B}}_{q,e}$ are of the form $A_q = A_{12} \cap A_{13} \cap A_{23} \cap A_{q,1} \cap A_{q,2} \cap A_{q,3}$. The set $E_{12} \cap E_{13} \cap E_{23}$ is the disjoint union of all such atoms, where $A_{ij} = E_{ij}$.

It will be convenient to work with the single measure $\mu_q := \prod_{f \in \bar{\mathcal{H}}} \nu_{q,f}$ instead of the ensemble of measures $\mu_{q,e} = \prod_{f \subseteq e} \nu_{q,f}$; fortunately, this can be done at essentially no cost as $\mu_q(A_{q,e}) = \mu_q(\pi_e^{-1}(A_{q,e})) + o_{N,W \rightarrow \infty}(1)$ for all $e \in \bar{\mathcal{H}}$, uniformly for q outside a set $\mathcal{E} \subseteq \Omega$ of measure $\psi(\Omega) = o_{N,W \rightarrow \infty}(1)$. This is given in Lemma 2.2 in Sect. 2, and the argument is similar to the one given after (1.33).

The proof of Theorem 1.4 is roughly as follows. First, one shows that most atoms A_q are “large”, in fact,

$$\mu_q(A_q) \geq \frac{1}{F(M, \varepsilon)}, \quad (1.48)$$

in the sense that all “small” atoms A_q are contained in sets of “edges” $B_{q,e} \in \mathcal{A}_e$ of measure

$$\mu_q(B_{q,e}) = O(\varepsilon). \quad (1.49)$$

Let $\delta := \frac{1}{2}F(M, \varepsilon)^{-1}$. If $\mu_q(E_{12} \cap E_{13} \cap E_{23}) \leq \delta$, then it cannot contain any large atoms. If one defines $E'_{q,e} := V \setminus B_{q,e}$, then $\bigcap_{e \in \mathcal{H}_2} E'_{q,e}$ does not contain any small atoms; hence, $\bigcap_{e \in \mathcal{H}_2} (E_e \cap E'_{q,e})$ cannot contain any atoms at all and must be empty, and so (1.47) follows (1.49).

Let $A_q = \bigcap_{e \in \mathcal{H}_2} A_e \cap_i A_{q,i}$ be an atom of $\bar{\mathcal{B}}_q$. We say that A_q is (ε, M) -regular if the following holds for all $i = 1, 2, 3$ and $e = (i, j) \in \mathcal{H}_2$:

$$\mu_q(A_{q,i}) \geq \varepsilon 2^{-M}, \quad (1.50)$$

$$\mu_q(A_e \cap A_{q,i} \cap A_{q,j}) \geq \varepsilon \mu_q(A_{q,i} \cap A_{q,j}), \quad (1.51)$$

$$\begin{aligned} \int_V \left| \mathbb{E}_{\mu_q} \left(\mathbf{1}_{A_e} \mid \bar{\mathcal{B}}_{q,i} \vee \bar{\mathcal{B}}_{q,j} \right) - \mathbb{E}_{\mu_q} \left(\mathbf{1}_{A_e} \mid \bar{\mathcal{B}}_{q,i} \vee \bar{\mathcal{B}}_{q,j} \right) \right|^2 \mathbf{1}_{A_{q,i}} \mathbf{1}_{A_{q,j}} d\mu_q \\ \leq \varepsilon^4 \int_V \mathbf{1}_{A_{q,i}} \mathbf{1}_{A_{q,j}} d\mu_q. \end{aligned} \quad (1.52)$$

Let $B_{q,e}$ be the union of all atoms $A_e \cap A_{q,i} \cap A_{q,j}$ for which at least one of (1.50)–(1.52) fails. We show that $\mu_q(B_{q,e}) = O(\varepsilon)$. First, note that $\bar{\mathcal{B}}_{q,i}$ has at most 2^M atoms as its complexity is bounded by M ; thus, the total measure of atoms A_q when (1.50) fails is $O(\varepsilon)$. If (1.51) fails, then $\mu_q(A_q) \leq \varepsilon \mu_q(A_{q,i} \cap A_{q,j})$ for some $e = (i, j) \in \mathcal{H}_2$. Since the sets $A_{q,i} \cap A_{q,j}$ are disjoint, the total measure of such atoms is also $O(\varepsilon)$. Finally, summing over all pairs (i, j) when (1.52) fails, one estimates using property (3) of the regularity lemma:

$$\begin{aligned} & \sum_{i,j} \mu_q(A_{q,i} \cap A_{q,j}) \\ & \leq \varepsilon^{-1} \|\mathbb{E}_{\mu_q}(\mathbf{1}_{A_e} \mid \bar{\mathcal{B}}_{q,i} \vee \bar{\mathcal{B}}_{q,j}) - \mathbb{E}_{\mu_q}(\mathbf{1}_{A_e} \mid \bar{\mathcal{B}}'_{q,i} \vee \bar{\mathcal{B}}'_{q,j})\|_{L^2_{\mu_q}}^2 \\ & = \varepsilon^{-1} \|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{A_e} \mid \mathcal{B}_{q,i} \vee \mathcal{B}_{q,j}) - \mathbb{E}_{\mu_q}(\mathbf{1}_{A_e} \mid \mathcal{B}'_{q,i} \vee \mathcal{B}'_{q,j})\|_{L^2_{\mu_{q,e}}}^2 = O(\varepsilon^3). \end{aligned}$$

A crucial fact is that the measure of the regular atoms can be approximately determined which is the content of the so-called *Counting Lemma* [5, 19]. To state it let us introduce the relative densities $\delta_{q,e} := \mu_q(A_e \cap A_{q,i} \cap A_{q,j}) / \mu_q(A_{q,i} \cap A_{q,j})$ and set $\delta_{q,i} := \mu_q(A_{q,i})$. Then, for an (ε, M) -regular atom $A_q = \cap_{e \in \mathcal{H}_2} A_e \cap_i A_{q,i}$, we have

$$\begin{aligned} \mu_q(A_q) &= (1 + O(\varepsilon)) \delta_{q,12} \delta_{q,13} \delta_{q,23} \delta_{q,1} \delta_{q,2} \delta_{q,3} + O\left(\frac{1}{F(M, \varepsilon)}\right) \\ &\quad + o_{N, W \rightarrow \infty}(1). \end{aligned} \tag{1.53}$$

The above also holds for regular atoms of any subgraph $\mathcal{H}' \subseteq \mathcal{H}_2$. In fact, the proof proceeds by induction on the number of edges of $\mathcal{H}'$. If $\mathcal{H}' = \{1, 2, 3\}$, i.e., it has no edges, then one has the stronger approximation

$$\mu_q(A_{q,1} \cap A_{q,2} \cap A_{q,3}) = \mu_q(A_{q,1}) \mu_q(A_{q,2}) \mu_q(A_{q,3}) + o_{N, W \rightarrow \infty}(1). \tag{1.54}$$

This again follows, for q outside a negligible set, from the linear form condition by an argument similar to the one given after (1.33), see Sect. 2. Fix an edge, say $e = (1, 2)$ and assume that (1.53) holds for regular atoms of the graph $\mathcal{H}' = \mathcal{H}_2 \setminus \{e\}$. Write

$$\begin{aligned} \mathbf{1}_{A_e} &= \delta_{q,e} + b_{q,e} + c_{q,e}, \quad \text{where} \\ b_q &= \mathbb{E}_{\mu_q}(\mathbf{1}_{A_e} \mid \bar{\mathcal{B}}'_{q,1} \vee \bar{\mathcal{B}}'_{q,2}) - \mathbb{E}_{\mu_q}(\mathbf{1}_{A_e} \mid \bar{\mathcal{B}}_{q,1} \vee \bar{\mathcal{B}}_{q,2}), \\ c_q &= \mathbf{1}_{A_{q,e}} - \mathbb{E}_{\mu_q}(\mathbf{1}_{A_e} \mid \bar{\mathcal{B}}'_{q,1} \vee \bar{\mathcal{B}}'_{q,2}). \end{aligned}$$

Note that $\delta_{q,e} = \mathbb{E}_{\mu_q}(\mathbf{1}_{A_e} \mid \bar{\mathcal{B}}_{q,1} \vee \bar{\mathcal{B}}_{q,2})$ on the set $A_{q,1} \cap A_{q,2}$. Accordingly, integrating term by term

$$\begin{aligned} \int_V \prod_{f \in \mathcal{H}} \mathbf{1}_{A_{q,e}} d\mu_q &= \int_V (\delta_{q,e} + b_{q,e} + c_{q,e}) \prod_{f \neq e} \mathbf{1}_{A_{q,e}} d\mu_q \\ &= M_q + E_q^1 + E_q^2. \end{aligned}$$

For the main term, by induction for $\mathcal{H}' = \mathcal{H} \setminus \{e\}$, we have

$$M_q = \delta_{q,e}(1 + O(\varepsilon)) \prod_{f \neq e} \delta_{q,f} + O\left(\frac{1}{F(M, \varepsilon)}\right) + o_{N, W \rightarrow \infty}(1).$$

For the first error term, taking absolute values and discarding the factors $\mathbf{1}_{A_{13}}$, $\mathbf{1}_{A_{23}}$, we have

$$\begin{aligned} |E_q^1| &\leq \int_{x_1, x_2} |b_q(x_1, x_2)| \mathbf{1}_{A_{q,1}}(x_1) \mathbf{1}_{A_{q,2}}(x_2) \\ &\quad \times \left(\int_{x_3} \mathbf{1}_{A_{q,3}}(x_3) w_q(x_1, x_2, x_3) d\mu_{q,3} \right) d\mu_{q,12} \\ &= \int_{x_1, x_2} I_q(x_1, x_2) J_q(x_1, x_2) d\nu_{q,12}, \end{aligned}$$

where $w_q(x_1, x_2, x_3) = \nu_{q,13}(x_1, x_3) \nu_{q,23}(x_2, x_3)$. By (1.52),

$$\int I_q^2 d\mu_{q,12} \leq \varepsilon \int \mathbf{1}_{A_{q,1}} \mathbf{1}_{A_{q,2}} d\mu_{q,12} \lesssim \varepsilon^4 \delta_{q,1} \delta_{q,2}.$$

Expanding the square of the integral in x_3 ,

$$\begin{aligned} \int J_q^2 d\mu_{q,12} &= \int_{x_1, x_2, x_3, x_3'} \mathbf{1}_{A_{q,1}}(x_1) \mathbf{1}_{A_{q,2}}(x_2) \mathbf{1}_{A_{q,3}}(x_3) \mathbf{1}_{A_{q,3}}(x_3') d\mu'_q \\ &= (1 + o_{N, W \rightarrow \infty}(1)) \delta_{q,1} \delta_{q,2} \delta_{q,3}^2, \end{aligned}$$

where the second equality again holds outside an exceptional set $\mathcal{E} \subseteq \Omega$. Thus, by the Cauchy–Schwarz inequality, property (1.51) of regular atoms and the fact that $\mu_{q,12}(V_1 \times V_2) = 1 + o_{N, W \rightarrow \infty}(1)$, we have

$$|E_q^1| \lesssim \varepsilon^4 \delta_{q,1} \delta_{q,2} \delta_{q,3} \lesssim \varepsilon \delta_{q,12} \delta_{q,12} \delta_{q,23} \delta_{q,1} \delta_{q,2} \delta_{q,3}.$$

The error term E_q^2 can be estimated by Proposition 1.1, the weighted von Neumann inequality. Indeed, set $F_{q,12} := c_q$, $F_{q,13} := \mathbf{1}_{A_{13}} \mathbf{1}_{A_{q,1}} \mathbf{1}_{A_{q,3}}$ and $F_{q,23} := \mathbf{1}_{A_{23}} \mathbf{1}_{A_{q,2}}$. Then, by Proposition 1.1,

$$|E_q^2| \lesssim \|c_q\|_{\square_{\nu_{q,12}}} + o_{N, W \rightarrow \infty}(1) = O\left(\frac{1}{F(M, \varepsilon)}\right),$$

by the regularity condition (3) uniformly, for $q \in \Omega \setminus \mathcal{E}$ for a set of measure $\psi(\mathcal{E}) = o_{N, W \rightarrow \infty}(1)$. A detailed proof of the general case will be given in Sect. 4.2. Various versions of weighted von Neumann inequalities are obtained in [1, 8]. For the sake of completeness we include a proof of Proposition 1.1 in an Appendix. This concludes the proof of Theorem 1.4 for graphs, i.e., when $d = 2$.

1.7. Outline of the Paper

In Sect. 2, we describe the type of parametric weight systems $\{\nu_{q,f}\}_{f \in \mathcal{H}, q \in \mathbb{Z}}$ we shall encounter later on. Here, we also discuss their basic properties such as the stability and symmetry. Though the notations are more complex, most arguments are variants of those described in Sect. 1.6, see (1.32), based only on the Cauchy–Schwarz inequality and the linear form condition.

The main argument starts in Sect. 3, where we introduce the energy increment argument for parametric weight systems on hypergraphs, and prove a hypergraph regularity lemma. Section 4 is devoted to proving the counting and removal lemmas. Many of our arguments in Sects. 3 and 4 may be viewed as extensions of those in [19]. In the last section, we obtain our main results stated in the introduction. The basic properties of weighted box norms are discussed in an Appendix.

As for our notations, most of our variables are of vector type, although we do not emphasize this. We think of the initial data $\Delta = \{v_0, \dots, v_d\}$ being fixed throughout, and do not denote the dependence of various quantities on them. For example, we write $Y = O(X)$ or $Y \lesssim X$ if $Y \leq CX$ for some constant $C > 0$ depending only on the vectors v_i or the dimension d . If $y_1, \dots, y_s$ and X are additional parameters we write $O_{y_1, \dots, y_s}(X)$ for a quantity Y bounded by $C(y_1, \dots, y_s)X$ or equivalently $Y \lesssim_{y_1, \dots, y_s} X$.

Though most of our constructions depend on N , all other terms in our estimates are independent of N as well as the parameters appearing in them. We will utilize the linear form condition throughout the paper, giving rise to error terms which tend to 0 as both $N \rightarrow \infty$ and $W \rightarrow \infty$ for any fixed choice of the parameters $y_1, \dots, y_s$ on which they may depend. The standard notation for such terms would be $o_{N, W \rightarrow \infty; y_1, \dots, y_s}(1)$; for simplicity, we will write $o_{y_1, \dots, y_s}(1)$. Finally, as all estimates in the linear form condition involving the weights ν_b are independent of the choice of b , we write in certain places $\nu = \nu_b$ for the purpose of simplifying the notation.

2. Basic Properties of Parametric Weight Systems and Their Extensions

In this section, we define the type of parametric systems and associated families of measures that we encounter later and discuss their basic properties such as stability and symmetry. We also discuss the type of extensions of such systems which arise in our induction process. Although our arguments are increasing in complexity, they are in essence just variations of the argument described after (1.33) in the introduction, based on the Cauchy–Schwarz inequality and the linear form condition.

2.1. Parametric Weight Systems and Stability Properties

Recall the family of measures $\{\mu_e\}_{e \in \mathcal{H}}$ constructed in (1.8):

$$\mu_e(x) = \prod_{L \in \mathcal{L}, \text{supp}(L) \subseteq e} \nu(L(x)),$$

where the family $\mathcal{L}$ defined in (1.8) consists of pairwise linearly independent forms. The following statement is based on the linear form condition and is a prototype of many of the arguments in this section.

Lemma 2.1. *For all $e \in \mathcal{H}$, we have that*

$$\mu_e(V_e) = 1 + o(1), \tag{2.1}$$

moreover, if $g: V_e \rightarrow [-1, 1]$,

$$\mathbb{E}_{x_e \in V_e} g(x_e) \mu_e(x_e) = \mathbb{E}_{x \in V_J} g(\pi_e(x)) \mu_J(x) + o(1),$$

or equivalently,

$$\int_{V_e} g \, d\mu_e = \int_{V_J} (g \circ \pi_e) \, d\mu_J + o(1). \quad (2.2)$$

Proof. Note that the linear forms appearing on the right side of

$$\mu_e(V_e) = \mathbb{E}_{x \in V_e} \prod_{\text{supp}(L) \subseteq e} \nu(L(x))$$

are pairwise linearly independent, and as they are supported on e , they remain pairwise independent when restricted to V_e . Thus, (2.1) follows from the linear form condition.

To prove (2.2), let $e' = J \setminus e$ and write $x = (x_e, x_{e'})$ with $x_e = \pi_e(x)$, $x_{e'} = \pi_{e'}(x)$. Then

$$\begin{aligned} E &:= \mathbb{E}_{x \in V_J} (g \circ \pi_e)(x) \mu_J(x) - \mathbb{E}_{x_e \in V_e} g(x_e) \mu_e(x_e) \\ &= \mathbb{E}_{x_e \in V_e} g(x_e) \mu_e(x_e) \mathbb{E}_{x_{e'} \in V_{e'}} (w(x_e, x_{e'}) - 1), \end{aligned}$$

where $w(x_e, x_{e'}) = \prod_{f \not\subseteq e} \nu_f(x_{e \cap f}, x_{e' \cap f})$.

By (2.1), we have that $\mu_e(V_e) \lesssim 1$, and then, by the Cauchy–Schwartz inequality,

$$|E|^2 \lesssim \mathbb{E}_{x_e \in V_e} \mathbb{E}_{x_{e'}, y_{e'} \in V_{e'}} (w(x_e, x_{e'}) - 1)(w(x_e, y_{e'}) - 1) \mu_e(x_e).$$

The right-hand side of this expression is a combination of four terms and (2.2) follows from the fact that each term is $1 + o(1)$. Indeed, the linear forms appearing in the definition of the function $\mu_e(x_e)$ depend only on the variables x_j for $j \in e$ and are pairwise linearly independent. All linear forms involved in $w(x_e, x_{e'})$ depend also on some of the variables in x_j , $j \in e'$, while the ones in $w(x_e, y_{e'})$ depend on the variables in y_j , $j \in e'$, and hence, these forms depend on different sets of variables. Thus, the forms appearing in the expression $\mu_e(x_e)w(x_e, x_{e'})w(x_e, y_{e'})$ are pairwise linearly independent and (2.2) follows from the linear form condition. Note that the estimate is independent of the function g . $\square$

This will allow us to consider sets $G_e \subseteq V_e$ as sets $\overline{G}_e = \pi_e^{-1}(G_e) \subseteq V_J$, changing their measures only by a negligible amount:

$$\mu_J(\overline{G}_e) = \mu_e(G_e) + o(1). \quad (2.3)$$

Next, we define weight systems and associated families of measures depending on parameters. Let

$$\mathcal{L}_q := (L^1(q, x), \dots, L^s(q, x))$$

be a family of linear forms with integer coefficients depending on the parameters $q \in \mathbb{Q}^R$ and the variables $x \in \mathbb{Q}^D$. We call the family *pairwise linearly independent* if no two forms in the family are rational multiples of each other. If N is a sufficiently large prime with respect to the coefficients of the linear

forms $L^i(q, x)$, then the forms remain pairwise linearly independent when considered as forms over $Z \times V$, $Z = \mathbb{Z}_N^R$, $V = \mathbb{Z}_N^D$. We refer to the set $Z = \mathbb{Z}_N^R$ as the *parameter space* of the family $\mathcal{L}_q$. As our arguments will involve averaging over the parameter space Z , we call the family $\mathcal{L}_q$ *well-defined* if there is a measure on Z given by

$$\int_Z g(q) \, d\psi(q) = \mathbb{E}_{q \in Z} g(q) \psi(q), \quad \psi(q) = \prod_{i=1}^t \nu(Y_i(q)), \quad (2.4)$$

for a family of pairwise linearly independent forms Y_i defined over Z and all forms $L^i(q, x)$ depend on some of the x -variables.

If $V = V_J$, then we define an associated system of weights $\{\nu_{q,e}\}_{q \in Z, e \in \mathcal{H}}$ and measures $\{\mu_{q,e}\}_{q \in Z, e \in \mathcal{H}}$ as follows. For a form $L^k(q, x) = \sum_i b_i q_i + \sum_j a_j x_j$, define its x -support as $\text{supp}_x(L) = \{j \in J : a_j \neq 0\}$. For $e \subseteq J$ and $q \in Z$, let

$$\nu_{q,e}(x) := \prod_{\substack{L \in \mathcal{L}_q \\ \text{supp}_x(L) = e}} \nu(L(q, x)), \quad \mu_{q,e}(x) := \prod_{\substack{L \in \mathcal{L}_q \\ \text{supp}_x(L) \subseteq e}} \nu(L(q, x)). \quad (2.5)$$

We use the convention that $\nu_{q,e} \equiv 1$ if there is no form $L \subseteq \mathcal{L}_q$, such that $\text{supp}_x(L) = e$. Note that the x -support partitions the family of forms $\mathcal{L}_q$ independent of the parameters q , thus for given $e \in \mathcal{H}$

$$\mu_{q,e}(x) = \prod_{f \subseteq e} \nu_{q,e}(x), \quad \text{for all } q \in Z.$$

A crucial observation is that many of the properties of the measure system $\{\mu_e\}$ still hold for well-defined measure systems $\{\mu_{q,f}\}$ for almost every value of the parameter $q \in Z$. To formulate such statements, we say that the family $\mathcal{L}$ has *complexity* at most K if the dimension of the space Z , the number of linear forms $L^j(q, x)$, $Y_i(q)$, and the magnitude of their coefficients are all bounded by K . This quantity will control the dependence of the error terms in applications of the linear form condition. We have an analogue of Lemma 2.1.

Lemma 2.2. *Let $\{\mu_{q,e}\}_{e \in \mathcal{H}, q \in Z}$ be a well-defined parametric measure system of complexity at most K . For every $e \in \mathcal{H}$, there is a set $\mathcal{E}_e \subseteq Z$, such that $\psi(\mathcal{E}_e) = o_K(1)$, and for every $q \notin \mathcal{E}_e$*

$$\mu_{q,e}(V_e) = 1 + o_K(1). \quad (2.6)$$

Moreover, for every $e \in \mathcal{H}$, there is a set $\mathcal{E}_e \subseteq Z$ of measure $\psi(\mathcal{E}_e) = o(1)$, such that the following holds. For any function $g: Z \times V_e \rightarrow [-1, 1]$ and for every $q \notin \mathcal{E}_e$, one has the estimate

$$\int_{V_e} g(q, x_e) \, d\mu_{q,e}(x_e) = \int_{V_J} g(q, \pi_e(x)) \, d\mu_{q,J}(x) + o_K(1). \quad (2.7)$$

Proof. To prove (2.6), consider the quantity

$$\begin{aligned}\Lambda_e &:= \int_Z |\mu_{q,e}(V_e) - 1|^2 \, d\psi(q) \\ &= \int_Z \mathbb{E}_{x_e, y_e} \left(\prod_{\text{supp}_x(L) \subseteq e} \nu(L(q, x_e)) - 1 \right) \left(\prod_{\text{supp}_x(L) \subseteq e} \nu(L(q, y_e)) - 1 \right) \, d\psi(q).\end{aligned}$$

The above expression is a combination of four terms and note that the family of linear forms

$$\{Y_k(q), L^i(q, x_e), L^j(q, y_e)\}$$

is pairwise linearly independent in the (q, x_e, y_e) variables by our assumptions. Applying the linear form condition gives that each term is $1 + o_K(1)$ and so $\Lambda_e = o_K(1)$ and (2.6) follows.

Now, let $e' = J \setminus e$ and write $x = (x_e, x_{e'})$. Arguing as in Lemma 2.1, we have

$$\begin{aligned}\Lambda(q, e, g) &:= |\mathbb{E}_{x \in V_J} g(q, \pi_e(x)) \mu_{q,J}(x) - \mathbb{E}_{x_e \in V_e} g(q, x_e) \mu_{q,e}(x_e)| \\ &= |\mathbb{E}_{x_e \in V_e} g(q, x_e) \mu_{q,e}(x_e) \mathbb{E}_{x_{e'} \in V_{e'}} (w_q(x_e, x_{e'}) - 1)| \\ &\leq \mathbb{E}_{x_e \in V_e} \mu_{q,e}(x_e) |\mathbb{E}_{x_{e'} \in V_{e'}} (w_q(x_e, x_{e'}) - 1)|,\end{aligned}$$

where $w_q(x_e, x_{e'}) = \prod_{f \not\subseteq e} \nu_{q,f}(x_{e \cap f}, x_{e' \cap f})$.

Notice that the right-hand side of the above inequality is independent of the function g ; if we denote it by $\Lambda(q, e)$, then (2.7) would follow from the estimate $\mathbb{E}_{q \in Z} \Lambda(q, e) \, d\psi(q) = o_K(1)$. By the linear form condition $\mathbb{E}_{q, x_e} \, d\psi(q) \, d\mu_{q,e}(x_e) = 1 + o_K(1) \leq 2$, for N sufficiently large with respect to K . Then, by the Cauchy–Schwartz inequality, one has

$$\begin{aligned}(\mathbb{E}_{q \in Z} \Lambda(q, e) \, d\psi(q))^2 &\lesssim \mathbb{E}_{q \in Z, x_e \in V_e} \mathbb{E}_{x_{e'}, y_{e'} \in V_{e'}} (w_q(x_e, x_{e'}) - 1) \\ &\quad \times (w_q(x_e, y_{e'}) - 1) \, d\mu_{q,e}(x_e) \, d\psi(q).\end{aligned}$$

This is a combination of four terms; however, each term again is $1 + o_K(1)$ as the linear forms defining ψ depend on the variables q , while the ones defining $\mu_{q,e}$ depend also on the x_e variables. On the other hand, all linear forms appearing in the weight functions $w_q(x_e, x_{e'})$ (resp. $w_q(x_e, y_{e'})$) depend on the $x_{e'}$ (resp. $y_{e'}$) variables as well. Thus, the family of all linear forms in the above expressions is pairwise linearly independent in the $(q, x_e, x_{e'}, y_{e'})$ variables. $\square$

2.2. Extensions of Parametric Systems

During our iteration process, we will encounter extensions of parametric families of forms depending on more and more parameters. Roughly speaking, one extends a family by adding new parameters together with new forms depending also on the new parameters. More precisely, let $\mathcal{L}_{q_1}^1 = \{L_1^1(q_1, x), \dots, L_1^{s_1}(q_1, x)\}$ and $\mathcal{L}_{q_2}^2 = \{L_2^1(q_2, x), \dots, L_2^{s_2}(q_2, x)\}$ be two pairwise linearly independent families of linear forms defined on the parameter spaces $Z_1 = \mathbb{Z}_N^{k_1}$ and, respectively, $Z_2 = \mathbb{Z}_N^{k_2}$. Let ψ_1 and ψ_2 be measures on Z_1 and Z_2 defined by the families of linear forms $\{Y_1^1(q_1), \dots, Y_{s_1}^1(q_1)\}$ and $\{Y_1^2(q_2), \dots, Y_{s_2}^2(q_2)\}$.

Definition 2.1. We say that the family $\mathcal{L}_{q_2}^2$ is an extension of the family $\mathcal{L}_{q_1}^1$ if $Z_1 \leq Z_2$ and the following holds. The family of forms $L_2^i(q_2, x), Y_j^2(q_2)$ which depend only on the variables $q_1 = \pi(q_2)$ is exactly the family of forms $L_1^i(q_1, x), Y_j^1(q_1)$, where $\pi: Z_2 \rightarrow Z_1$ is the natural orthogonal projection.

If $V = V_J$, let $\mu^1 := \{\mu_{q_1, e}\}_{q_1 \in Z_1, e \in \mathcal{H}}$ and $\mu^2 := \{\mu_{q_2, f}\}_{q_2 \in Z_2, f \in \mathcal{H}}$ be the associated measure systems as defined in (2.5). We say that the measure system μ_2 is an *extension* of the system μ_1 . Let us make a few immediate observations. Writing $Z_2 = Z_1 \times \mathbb{Z}$, $Z = \mathbb{Z}_N^r$ and $q_2 = (q_1, q)$, we have

$$\psi_2(q_1, q) = \psi_1(q_1) \cdot \varphi(q_1, q), \quad (2.8)$$

where $\varphi(q, q_1) = \prod_{i=1}^t \nu(Y_i(q_1, q))$. The linear forms $Y_i(q_1, q)$ defining $\varphi(q, q_1)$ depend on some of the variables of $q = (q_i)_{1 \leq i \leq k}$ and are pairwise linearly independent. Similarly, one may write for any $e \in \mathcal{H}$

$$\mu_{(q_1, q), e}^2(x_e) = \mu_{q_1, e}^1(x_e) w_e(q_1, q, x_e), \quad (2.9)$$

where the linear forms $L_2^j(q_1, q, x_e)$ defining the function $w_e(q, q_1, x_e)$ depend on (some of) the variables q as well as on (all of) the variables x_e .

In the special case, when $\mathcal{L} = (L^1(x), \dots, L^s(x))$ is a family of linear forms, a parametric family $\mathcal{L}_q$ is called an extension of $\mathcal{L}$ if the set of forms in $\mathcal{L}_q$ which are independent of q is exactly the family $\mathcal{L}$. Similarly, the associated system of weights $\{\nu_{q, e}\}$ and measures $\{\mu_{q, e}\}$ is referred to as an extension of $\{\nu_e\}$ and $\{\mu_e\}$.

Lemma 2.3. Let $\{\mu_f\}_{f \in \mathcal{H}}$ be a well defined measure system, and let $\{\mu_{q, f}\}_{q \in Z, f \in \mathcal{H}}$ be a well-defined parametric extension of $\{\mu_f\}_{f \in \mathcal{H}}$ of complexity at most K . Then, for any $f \in \mathcal{H}$ and for any function $g: V_f \rightarrow [-1, 1]$, there is a set $\mathcal{E}_{g, f} \subseteq Z$ of measure $\psi(\mathcal{E}_{g, f}) = o_K(1)$, so that for all $q \notin \mathcal{E}_{g, f}$

$$\int_{V_f} g \, d\mu_{q, f} - \int_{V_f} g \, d\mu_f = o_K(1). \quad (2.10)$$

Similarly, if $\{\mu_{q_1, f}\}_{f \in \mathcal{H}, q_1 \in Z_1}$ is a well-defined parametric system and if $\{\mu_{q_2, f}\}_{f \in \mathcal{H}, q_2 \in Z_2}$ is an extension of complexity at most K_2 , then to any function $g: Z_1 \times V_f \rightarrow [-1, 1]$, there exists a set $\mathcal{E}_{g, f} \subseteq Z_2$ of measure $\psi_2(\mathcal{E}_{g, f}) = o_{K_2}(1)$, such that for all $q_2 = (q_1, q) \notin \mathcal{E}_{g, f}$,

$$\int_{V_f} g(q_1, x) \, d\mu_{q_2, f}(x) - \int_{V_f} g(q_1, x) \, d\mu_{q_1, f}(x) = o_{K_2}(1). \quad (2.11)$$

Proof. As $\mu_{q, f} = \mu_f(x_f) w_f(q, x_f)$, the left side of (2.10) may be written as

$$\Lambda_{f, g}(q) := \int_{V_f} g(x) (w_f(q, x) - 1) \, d\mu_f(x).$$

Consider

$$\Lambda_{f, g} := \int_Z |\Lambda_{f, g}(q)|^2 \, d\psi(q).$$

Using the Cauchy–Schwartz inequality, we estimate

$$\begin{aligned}\Lambda_{f,g} &= \int_Z \int_{V_f} \int_{V_f} (w_f(q, x) - 1)(w_f(q, y) - 1)g(x)g(y) \, d\mu_f(x)d\mu_f(y)d\psi(q) \\ &\leq \int_{V_f} \int_{V_f} \left| \int_Z (w_f(q, x) - 1)(w_f(q, y) - 1)d\psi(q) \right| d\mu_f(x)d\mu_f(y).\end{aligned}$$

Now, the Cauchy–Schwartz inequality and (2.1) give

$$\begin{aligned}|\Lambda_{f,g}|^2 &\lesssim \int_{V_f} \int_{V_f} \int_Z \int_Z (w_f(q, x) - 1)(w_f(q, y) - 1) \\ &\quad \times (w_f(p, x) - 1)(w_f(p, y) - 1) \, d\mu_f(x)d\mu_f(y)d\psi(q)d\psi(p).\end{aligned}$$

The last expression is a combination of 16 terms, where each term is $1 + o_K(1)$ by the linear forms condition. Indeed, the linear forms which can appear in any of these terms are $Y_{i_1}(q), Y_{i_2}(p), L^{i_3}(x), L^{i_4}(y), L^{i_5}(q, x), L^{i_6}(q, y), L^{i_7}(p, x)$, and $L^{i_8}(p, y)$. Note that the last four terms depend on both sets of variables (for example, $L^i(q, x)$ depends both on $q \in Z$ and on $x \in V_f$), and hence, the family of these forms is pairwise linearly independent in the (q, p, x, y) variables. This proves (2.10).

The proof of (2.11) is essentially the same. Set

$$\Lambda_{f,g}(q_2) := \int_{V_f} g(q_1, x)d\mu_{q_2,f}(x) - \int_{V_f} g(q_1, x)d\mu_{q_1,f}(x)$$

and

$$\Lambda_{f,g} := \int_Z |\Lambda_{f,g}(q_2)|^2 d\psi_2(q_2).$$

Write $Z_2 = Z_1 \times Z$, where $Z = \mathbb{Z}_N^k$, and $q_2 = (q_1, q)$ for $q_2 \in Z_2$. By (2.8), we estimate as above

$$\begin{aligned}\Lambda_{f,g} &\lesssim \int_{V_f} \int_{V_f} \int_{Z_1} d\psi_1(q_1)d\mu_{q_1,f}(x)d\mu_{q_1,f}(y) \\ &\quad \times |\mathbb{E}_{q \in Z} (w_f(q_1, q, x) - 1)(w_f(q_1, q, y) - 1)\varphi(q_1, q)|.\end{aligned}$$

The linear form condition gives

$$\int_{V_f} \int_{V_f} \int_{Z_1} d\psi_1(q_1)d\mu_{q_1,f}(x)d\mu_{q_1,f}(y) = 1 + o_{K_2}(1),$$

thus we have

$$\begin{aligned}|\Lambda_{f,g}|^2 &\lesssim \int_{V_f} \int_{V_f} \int_{Z_1} \mathbb{E}_{p,q \in Z} (w_f(q_1, q, x) - 1)(w_f(q_1, q, y) - 1) \\ &\quad \times (w_f(q_1, p, x) - 1)(w_f(q_1, p, y) - 1) \varphi(q_1, q)\varphi(q_1, p) d\psi_1(q_1)d\mu_{q_1,f}(x)d\mu_{q_1,f}(y).\end{aligned}$$

The point is that any linear form $L_f^i(q_1, q, x)$ depends on both the variables q and x . Thus, again, the left side is a combination of 16 terms, each being $1 + o_{K_2}(1)$ by the linear form condition as all the linear forms involved in any of these expressions are pairwise linearly independent in the (x, y, q_1, q, p) variables. $\square$

Lemma 2.3 is an example of what we refer to as the *stability* property. It means that the extension measures $\mu_{(q_1, q), f}$ are small perturbations of the measures $\mu_{q_1, f}$ with respect to quantities which are independent of q .

As a first application of this principle, we show that the weighted box norms, defined in (1.10), remain essentially unchanged under parametric extensions of the weight systems defining the norms. Let $\mathcal{L}_{q_1}$ be a pairwise linearly independent family of forms defined on the parameter space (Z_1, ψ_1) and let $\{\nu_{q_1, e}\}$ be the associated system of weights.

Let $g: Z_1 \times V_e \rightarrow \mathbb{R}$ be a function and let $e \in \mathcal{H}$, $|e| = d'$. For a given $q_1 \in Z_1$, recall the box norm of $g_{q_1}(x) = g(q_1, x)$:

$$\|g_{q_1}\|_{\square_{\nu_{q_1, e}}}^{2^{d'}} = \mathbb{E}_{p, x \in V_e} \prod_{\omega \in \{0, 1\}^e} g(q_1, \omega_e(p, x)) \prod_{f \subseteq e} \prod_{\omega_f \in \{0, 1\}^f} \nu_{q, f}(\omega_f(p_f, x_f)), \quad (2.12)$$

where $x_f = \pi_f(x)$, $p_f = \pi_f(p)$, $\pi_f: V_e \rightarrow V_f$ being the natural projection. The inner product on the right side of (2.12) is defined by the parametric family of forms:

$$\tilde{\mathcal{L}}_{q_1} = \bigcup_{f \subseteq e} \{L(q_1, \omega_f(p_f, x_f)): L \in \mathcal{L}_{q_1}, \text{supp}_x(L) = f, \omega \in \{0, 1\}^f\}. \quad (2.13)$$

It is easy to see that this is a pairwise linearly independent family of forms defined over $Z_1 \times V$, $V = V_e \times V_e$. Indeed, if we would have that

$$L'(q_1, \omega'_{f'}(p_{f'}, x_{f'})) = \lambda L(q_1, \omega_f(p_f, x_f)), \quad (2.14)$$

then restricting both forms to the subspace $\{p = x\}$ would imply that $L'(q_1, x_{f'}) = \lambda L(q_1, x_f)$, and hence, $f' = \text{supp}_x(L') = \text{supp}_x(L) = f$. Then, as L and L' depend exactly on the variables x_j for which $j \in f$, we have $\omega' = \omega$ and $L = L'$.

If $\{\tilde{\mu}_{q_1, f}\}_{q \in Z_1, f \subseteq e}$ denotes the associated system of measures and

$$G(q_1, p, x) := \prod_{\omega \in \{0, 1\}^e} g(q_1, \omega_e(p, x)), \quad (2.15)$$

then for given $q_1 \in Z_1$

$$\|g_{q_1}\|_{\square_{\nu_{q_1, e}}}^{2^{d'}} = \mathbb{E}_{p, x \in V_e} G_{q_1}(p, x) \tilde{\mu}_{q_1, e}(p, x). \quad (2.16)$$

Now, if $\mathcal{L}_{q_2}$ is a well-defined parametric extension of $\mathcal{L}_{q_1}$, then (2.13) yields a well-defined parametric extension $\tilde{\mathcal{L}}_{q_2}$ of the family $\tilde{\mathcal{L}}_{q_1}$. Then, by Lemma 2.3, and the simple observation that $|a^{2^{d'}} - b^{2^{d'}}| \leq \varepsilon$ implies $|a - b| \leq \varepsilon^{2^{-d'}}$ for $a, b \geq 0$, we obtain

Lemma 2.4. *Let $\{\nu_{q_1, f}\}_{f \in \mathcal{H}, q_1 \in Z_1}$ be a parametric weight system with a well-defined extension $\{\nu_{q_2, f}\}_{f \in \mathcal{H}, q_2 \in Z_2}$ of complexity at most K_2 . Then, to any $e \in \mathcal{H}$ and to any function $g: Z_1 \times V_e \rightarrow [-1, 1]$, there exists a set $\mathcal{E} = \mathcal{E}(g, e) \in Z_2$ of measure $\psi_2(\mathcal{E}) = o_{K_2}(1)$, such that for all $q_2 = (q_1, p) \notin \mathcal{E}$*

$$\|g_{q_1}\|_{\square_{\nu_{q_2, e}}} = \|g_{q_1}\|_{\square_{\nu_{q_1, e}}} + o_{K_2}(1). \quad (2.17)$$

Let $(V, \mathcal{B}, \mu)$ be a measure space and let $g: V \rightarrow \mathbb{R}$ be a function. An important construction, the so-called conditional expectation function is defined as

$$\mathbb{E}_\mu(g|\mathcal{B})(x) = \frac{1}{\mu(B(x))} \mathbb{E}_{y \in V} \mathbf{1}_{B(x)}(y) g(y) d\mu(y) = \frac{1}{\mu(B(x))} \int_{B(x)} g(y) d\mu(y),$$

where $B(x) \in \mathcal{B}$ is the atom containing x . If $\mu(B(x)) = 0$, then we set $\mathbb{E}_\mu(g|\mathcal{B})(x) = 1$.

The complexity of the σ -algebra $\mathcal{B}$, denoted by $\text{compl}(\mathcal{B})$, is defined as the minimum number of elements of $\mathcal{B}$ which generates $\mathcal{B}$. Note that the number of atoms of $\mathcal{B}$ is at most $2^{\text{compl}(\mathcal{B})}$. Next, we compare the conditional expectation functions of parametric systems.

Lemma 2.5. *Let $(\mu_{q_1, f})_{q_1 \in Z_1, f \in \mathcal{H}}$ be a well-defined parametric measure system with a well-defined extension $(\mu_{q_2, f})_{q_2 \in Z_2, f \in \mathcal{H}}$ of complexity at most K_2 . For $q_1 \in Z_1$ and $e \in \mathcal{H}$, let $\mathcal{B}_{q_1, e}$ be a σ -algebra on V_e , such that $\text{compl}(\mathcal{B}_{q_1, e}) \leq M$ for some fixed number M . For any function $g: Z_1 \times V_e \rightarrow [-1, 1]$, there exists a set $\mathcal{E} = \mathcal{E}(\mathcal{B}, g) \subseteq Z_2$ of measure $\psi_2(\mathcal{E}) = o_{M, K_2}(1)$, such that for any $q_2 = (q_1, q) \notin \mathcal{E}$*

(1) *we have*

$$\left\| \mathbb{E}_{\mu_{q_2, e}}(g_{q_1} | \mathcal{B}_{q_1, e}) - \mathbb{E}_{\mu_{q_1, e}}(g_{q_1} | \mathcal{B}_{q_1, e}) \right\|_{L^2(\mu_{q_2, e})}^2 = o_{M, K_2}(1) \quad (2.18)$$

(2) *and*

$$\left\| \mathbb{E}_{\mu_{q_2, e}}(g_{q_1} | \mathcal{B}_{q_1, e}) \right\|_{L^2(\mu_{q_2, e})}^2 = \left\| \mathbb{E}_{\mu_{q_1, e}}(g_{q_1} | \mathcal{B}_{q_1, e}) \right\|_{L^2(\mu_{q_1, e})}^2 + o_{M, K_2}(1). \quad (2.19)$$

Proof. Let $m = 2^M$ and enumerate the atoms of $\mathcal{B}_{q_1, e}$ as $B_{q_1}^1, \dots, B_{q_1}^m$, allowing some of them to be possibly empty. For a fixed $1 \leq i \leq m$, define the function

$$b_i(q_1, x) = \mathbf{1}_{B_{q_1}^i}(x) = \begin{cases} 1 & \text{if } x \in B_{q_1}^i, \\ 0 & \text{otherwise,} \end{cases}$$

and for $q_2 = (q_1, q) \in Z_2$, define the quantities

$$\begin{aligned} \mu_i(q_2, g) &:= \int_{V_e} g(q_1, x) b_i(q_1, x) d\mu_{q_2, e}(x), & \mu_i(q_2) &:= \mu_{q_2, e}(B_{q_1}^i), \\ \mu_i(q_1, g) &:= \int_{V_e} g(q_1, x) b_i(q_1, x) d\mu_{q_1, e}(x), & \mu_i(q_1) &:= \mu_{q_1, e}(B_{q_1}^i). \end{aligned}$$

By Lemma 2.3, we have that

$$\mu_i(q_2, g) = \mu_i(q_1, g) + o_{K_2}(1), \quad \mu_i(q_2) = \mu_i(q_1) + o_{K_2}(1) \quad (2.20)$$

for all $q_2 \notin \mathcal{E}_i$, where $\mathcal{E}_i \subseteq Z_2$ is a set of ψ_2 -measure $o_{K_2}(1)$. Let $\mathcal{E} = \bigcup_{i=1}^m \mathcal{E}_i$ then $\psi_2(\mathcal{E}) = o_{K_2, M}(1)$. The left-hand side of (2.18) takes the form:

$$\sum_{i=1}^m \left(\frac{\mu_i(q_2, g)}{\mu_i(q_2)} - \frac{\mu_i(q_1, g)}{\mu_i(q_1)} \right)^2 \mu_i(q_2), \quad (2.21)$$

with the convention that if $\mu_i(q_1) = 0$ or $\mu_i(q_2) = 0$, then $\mu_i(q_1, g)/\mu_i(q_1) := 1$ or $\mu_i(q_2, g)/\mu_i(q_2) := 1$.

If $q_2 = (q_1, q) \notin \mathcal{E}$, then by (2.20)

$$\varepsilon := \sum_{i=1}^m |\mu_i(q_2, g) - \mu_i(q_1, g)| + |\mu_i(q_2) - \mu_i(q_1)| = o_{K_2, M}(1). \quad (2.22)$$

Now, if $\mu_i(q_1) \leq 2\varepsilon^{1/4}$ then $\mu_i(q_2) \leq 3\varepsilon^{1/4}$ by (2.20); hence, the total contribution of such terms is bounded by $12m\varepsilon^{1/4} = o_{K_2, M}(1)$.

If $\mu_i(q_1) \geq 2\varepsilon^{1/4}$, then $\mu_i(q_2) \geq \varepsilon^{1/4}$ and we have the estimate

$$\left| \frac{\mu_i(q_2, g)}{\mu_i(q_2)} - \frac{\mu_i(q_1, g)}{\mu_i(q_1)} \right| \leq \frac{4\varepsilon(N)}{2\varepsilon(N)^{1/2}} \leq 2\varepsilon^{1/2} = o_{K_2, M}(1).$$

This proves (2.18). The proof of inequality (2.19) proceeds the same way, here, one needs to estimate the quantity

$$\begin{aligned} & \sum_{i=1}^m \left| \frac{\mu_i(q_2, g)^2}{\mu_i(q_2)} - \frac{\mu_i(q_1, g)^2}{\mu_i(q_1)} \right| \\ &= \sum_{i=1}^m \left| \left(\frac{\mu_i(q_2, g)}{\mu_i(q_2)} \right)^2 \mu_i(q_2) - \left(\frac{\mu_i(q_1, g)}{\mu_i(q_1)} \right)^2 \mu_i(q_1) \right|. \end{aligned} \quad (2.23)$$

If $\mu_i q_1 \leq 2\varepsilon^{1/4}$, then $\mu_i(q_2) \leq 3\varepsilon^{1/4}$ for $q_2 = (q_1, q) \notin \mathcal{E}$; thus, the contribution of such terms to the right side of (2.23) is trivially estimated by

$$3m\varepsilon^{1/4} = o_{M, K_2}(1).$$

The rest of the terms are bounded by $8\varepsilon^{1/2}$ and (2.19) follows. $\square$

We also need an analogue of the above result when the $\|\cdot\|_{L^2(\mu_{q,e})}$ norms are replaced by the more complicated $\|\cdot\|_{\square_{\nu_{q,e}}}$ norms.

Lemma 2.6. *Let $\{\nu_{q_2, f}\}_{f \in \mathcal{H}, q_2 \in Z_2}$ be a well-defined extension of the parametric weight system $\{\nu_{q_2, f}\}_{f \in \mathcal{H}, q_1 \in Z_1}$, of complexity at most K_2 . For $q_1 \in Z_1$ and $e \in \mathcal{H}$, let $\mathcal{B}_{q_1, e}$ be a σ -algebra of complexity at most M , for some fixed constant $M > 0$. Then*

$$\left\| \mathbb{E}_{\nu_{q_2, e}}(g_{q_1} | \mathcal{B}_{q_1, e}) - \mathbb{E}_{\nu_{q_1, e}}(g_{q_1} | \mathcal{B}_{q_1, e}) \right\|_{\square_{\nu_{q_2, e}}} = o_{M, K}(1), \quad (2.24)$$

for all $q_2 = (q_1, q) \notin \mathcal{E}$, where $\mathcal{E} = \mathcal{E}(g, \mathcal{B}) \subseteq Z_2$ is a set of measure $\psi_2(\mathcal{E}) = o_{M, K_2}(1)$.

Proof. First, we show that for any family of sets $\mathcal{A} = (A_{q_1})_{q_1 \in Z_1}$, $A_{q_1} \subseteq V_e$, there is a set $\mathcal{E}_1 = \mathcal{E}_1(g, \mathcal{A})$ of measure $\psi_2(\mathcal{E}_1) = o_{K_2}(1)$, such that for all $q_2 = (q_1, q) \notin \mathcal{E}_1$, we have

$$\|\mathbf{1}_{A_{q_1}}\|_{\square_{\nu_{q_2, e}}}^{2^d} \leq \mu_{q_2, e}(A_{q_1}) + o_{K_2}(1). \quad (2.25)$$

To see this, first note that for $q_2 = (q_1, q) \in Z_2$, one has

$$\begin{aligned} \|\mathbf{1}_{A_{q_1}}\|_{\square_{\nu_{q_2,e}}}^{2^d} &\leq \mathbb{E}_{x,p \in V_e} \mathbf{1}_{A_{q_1}}(x) \mu_{q_2,e}(x) \prod_{f \subseteq e} \prod_{\omega_f \neq 0} \nu_{q_2,f}(\omega_f(p_f, x_f)) \\ &= \mu_{q_2,e}(A_{q_1}) + E(q_2), \end{aligned}$$

with

$$E(q_2) \leq \mathbb{E}_{x \in V_e} \mu_{q_2,e}(x) |\mathbb{E}_{p \in V_e} (\mathcal{W}(q_2, p, x) - 1)|,$$

where

$$\mathcal{W}(q_2, p, x) = \prod_{f \subseteq e} \prod_{\omega_f \neq 0} \nu_{q_2,f}(\omega_f(p_f, x_f)).$$

Arguing as in Lemma 2.6, we see that

$$\mathbb{E}_{q_2 \in Z_2} \mathbb{E}_{x,p,p' \in V_e} \psi_2(q_2) d\mu_{q_2,e}(x) (\mathcal{W}(q_2, p, x) - 1)(\mathcal{W}(q_2, p', x) - 1) = o_{M,K}(1)$$

and (2.25) follows.

Now, let $\{B_{q_1}^i\}_{i=1}^m$ ($m = 2^M$) be the atoms of $\mathcal{B}_{q_1,e}$ and define the quantities $\mu_i(q_2, g)$, $\mu_i(q_2)$, $\mu_i(q_1, g)$, $\mu_i(q_1)$ as in Lemma 2.4. The expression in (2.18) is then estimated by

$$\begin{aligned} &\left\| \sum_{i=1}^m \left(\frac{\mu_i(q_2, g)}{\mu_i(q_2)} - \frac{\mu_i(q_1, g)}{\mu_i(q_1)} \right) \mathbf{1}_{B_{q_1}^i} \right\|_{\square_{\nu_{q_2,e}}} \\ &\leq \sum_{i=1}^m \left| \frac{\mu_i(q_2, g)}{\mu_i(q_2)} - \frac{\mu_i(q_1, g)}{\mu_i(q_1)} \right| \|\mathbf{1}_{B_{q_1}^i}\|_{\square_{\nu_{q_2,e}}} \\ &\lesssim \sum_{i=1}^m \left| \frac{\mu_i(q_2, g)}{\mu_i(q_2)} - \frac{\mu_i(q_1, g)}{\mu_i(q_1)} \right| \mu_{q_2,e}(B_{q_1}^i)^{2^{-d}} + o_{M,K}(1), \end{aligned}$$

for $q_2 = (q_1, q) \notin \mathcal{E}_1$, where $\mathcal{E}_1 = \mathcal{E}_1(\mathcal{B}_{q_1,e}, g)$ is a set of measure $o_{M,K}(1)$.

Using the facts that $\mu_i(q_2, g) = \mu_i(q_1, g) + o_{K_2}(1)$ and $\mu_i(q_2) = \mu_i(q_1) + o_{K_2}(1)$ outside a set of measure $o_{M,K_2}(1)$, and

$$\sum_{i=1}^m \mu_{q_2,e}(B_{q_1}^i) = \mu_{q_2,e}(V_e) = 1 + o_{K_2}(1),$$

it follows that the above expression is $o_{M,K_2}(1)$ by arguing as in Lemma 2.4. This completes the proof. $\square$

2.3. Symmetric Extensions

We will also need our parametric families of forms to be symmetric, to apply Theorem 1.3. Let for each $e \in \mathcal{H}_d$, $\mathcal{L}_{q,e} = \{L_e^1(q, x), \dots, L_e^s(q, x)\}$ be a pairwise linearly independent family of linear forms defined on V_J , depending on parameters $q \in Z$, such that $\text{supp}_x(L_e^j) \subseteq e$. We say that the family of forms $\mathcal{L}_q = \bigcup_{e \in \mathcal{H}_d} \mathcal{L}_{q,e}$ is *symmetric* if $L_e^j(q, x) = L_{e'}^j(q, x)$ for all $q \in Z$, $x \in M$, $e, e' \in \mathcal{H}_d$ and $1 \leq j \leq s$. Note that our initial family of forms defined in (1.10) has this property.

It is not hard to see that for a given family of forms $\mathcal{L}_{q,e}$ and for a fixed $e \in \mathcal{H}_d$, there is a unique symmetric family of forms $\mathcal{L}_q$, such that

$\mathcal{L}_{q,e} = \{L \in \mathcal{L}_q; \text{supp}_x(L) \subseteq e\}$. Indeed, if $\mathcal{L}_q$ is such a family, then for $e' \in \mathcal{H}_d$, $q \in Z$ and $x \in V_J$

$$L_{e'}^j(q, x) = L_{e'}^j(q, \pi_{e'}(x)) = L_{e'}^j(q, \phi_{e'} \circ \pi_{e'}(x)) = L_e^j(q, \phi_{e'} \circ \pi_{e'}(x)), \quad (2.26)$$

where $\phi_e: V_e \rightarrow M$ is the inverse of the projection $\pi_{e'}$ restricted to M . This shows the uniqueness of the family $\mathcal{L}_q$. Conversely, define $L_{e'}^j(q, x)$ by the above equality, then it is clear that $\text{supp}_x(L_{e'}^j) \subseteq e'$; moreover, if $x \in M$, then $x = \phi_{e'} \circ \pi_{e'}(x)$; hence, $L_{e'}^j(q, x) = L_e^j(q, x)$. In addition, if $\text{supp}_x L_{e'}^j \subseteq e$, then for all $q \in \mathbb{Z}_1$ and $x \in V_J$,

$$L_{e'}^j(q, x) = L_{e'}^j(q, \phi_e \circ \pi_e(x)) = L_e^j(q, \phi_e \circ \pi_e(x)) = L_e^j(q, x).$$

This shows that all the forms in $\mathcal{L}_q$ which depend only on the variables x_e are the forms of $\mathcal{L}_{q,e}$. Finally, if $\mathcal{L}_{q,e}$ is a pairwise linearly independent family, then so is $\mathcal{L}_q$, as linearly dependent forms must depend on the same set of variables. We will refer to the family $\mathcal{L}_q$ of forms as the *symmetrization* of the family $\mathcal{L}_{q,e}$. If $f \in \mathcal{H}$ for some edge $|f| = d' \leq d$ and $\mathcal{L}_{q,f}$ is a family of forms defined on V_f , then the above construction can be applied to obtain a symmetric family $\mathcal{L}_q$ simply by choosing an $e \in \mathcal{H}_d$, such that $f \subseteq e$ and considering $\mathcal{L}_{q,f}$ as a family of forms on V_e . Note that the construction is independent of the choice of $e \supseteq f$, as if $f \subseteq e'$ as well, then $L_e^j = L_{e'}^j$ for all $1 \leq j \leq s$.

In the next section, following [19], we will start an energy increment process to obtain a regularity lemma for weighted hypergraphs. At each stage, we will pass to an extension of a symmetric, well-defined and pairwise independent parametric family $\mathcal{L}_q$ defined for $q \in Z$ as follows. We choose an edge $e \in \mathcal{H}$ and consider the extension of the family $\mathcal{L}_{q,e}$, as given in (2.13), that is, replacing the forms $L^j(q, x_f)$ with the forms $L^j(q, \omega_f(p_f, x_f))$, $\omega \in \{0, 1\}^f$. This gives an extension $\tilde{\mathcal{L}}_{q,p,f}$ defined on the parameter space $(q, p) \in Z \times V_f$, which we symmetrize to obtain a new symmetric, well-defined and pairwise independent family $\tilde{\mathcal{L}}_{q,p}$. The first step of this process was described in the introduction in the special case $e = (1, 2)$.

3. Regularization of Parametric Systems

3.1. The Koopman-von Neumann Decomposition

Let $e \subseteq J$ and let $\mathcal{B}_f$ be a σ -algebra on V_f for $f \in \partial e$, where $\partial e = \{f \subseteq e: |f| = |e| - 1\}$ denotes the boundary of the edge e . Let $\mathcal{B} := \bigvee_{f \subseteq \partial e} \mathcal{B}_f$ be the σ -algebra generated by the sets $\pi_{ef}^{-1}(\mathcal{B}_f)$, where $\pi_{ef}: V_e \rightarrow V_f$ is the canonical projection. The atoms of $\mathcal{B}$ are the sets $G = \bigcap_{f \subseteq \partial e} \pi_{ef}^{-1}(G_f)$ with G_f being an atom of $\mathcal{B}_f$, which may be interpreted as the collection of simplices $x_e \in V_e$, whose faces x_f are in G_f for all $f \in \partial e$.

The starting point of the proof of the Regularity Lemma, given in [19], is to show that if a set $G_e \subseteq V_e$ is not sufficiently regular with respect to $\mathcal{B}$,

that is, if

$$\left\| \mathbf{1}_{G_e} - \mathbb{E} \left(\mathbf{1}_{G_e} \middle| \bigvee_{f \in \partial e} \mathcal{B}_f \right) \right\|_{\square} \geq \eta, \quad (3.1)$$

then there exists a σ -algebra $\mathcal{B}'_f \supseteq \mathcal{B}_f$ for each $f \in \partial e$, such that

$$\left\| \mathbb{E} \left(\mathbf{1}_{G_e} \middle| \bigvee_{f \in \partial e} \mathcal{B}'_f \right) \right\|_{L^2}^2 \geq \left\| \mathbb{E} \left(\mathbf{1}_{G_e} \middle| \bigvee_{f \in \partial e} \mathcal{B}_f \right) \right\|_{L^2}^2 + c\eta^2.$$

The quantity $\|\mathbb{E}(\mathbf{1}_G | \mathcal{B})\|_{L^2}^2$ is referred to as the *energy* (or *index*) of the set G with respect to the σ -algebra $\mathcal{B}$; thus, the above inequality means that the energy of the set G_e is increased by $c\eta^2$ by refining the σ -algebras $\mathcal{B}_f$. In addition, the complexity of the σ -algebras $\mathcal{B}'_f$, denoted by $\text{compl}(\mathcal{B}'_f)$ and defined as the minimal number of sets generating the σ -algebra, is at most 1 larger than that of $\mathcal{B}_f$.

In our settings, for a given $e \subseteq J$, we will have a parametric system of weights $\{\nu_{q,f}\}_{q \in Z, f \subseteq e}$ and measures $\{\mu_{q,f}\}_{q \in Z, f \subseteq e}$ associated with a well-defined, pairwise linearly independent family of forms $\mathcal{L}_q$ defined on $Z \times V_e$, as given in (2.5). For simplicity, we will refer to such systems of weights and measures as being *well-defined*.

Lemma 3.1. *For given $e \subseteq J$, $|e| = d'$, let $\{\mu_{q,f}\}_{q \in Z, f \subseteq e}$ be a well-defined family of measures of complexity at most K . For $q \in Z$, let $G_{q,e} \subseteq V_e$ and $\{\mathcal{B}_{q,f}\}_{f \in \partial e}$ be a σ -algebra on V_f . Assume*

$$\left\| \mathbf{1}_{G_{q,e}} - \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{G_{q,e}} \middle| \bigvee_{f \in \partial e} \mathcal{B}_{q,f} \right) \right\|_{\square_{\nu_{q,e}}}^{2^{d'}} \geq \eta, \quad (3.2)$$

for some $\eta > 0$ and for each $q \in \Omega$, where $\Omega \subseteq Z$ is a set of measure $\psi(\Omega) \geq c_0 > 0$. Then, for N, W sufficiently large with respect to the parameters c_0, η , there exists a well-defined extension $\{\mu_{q',f}\}_{q' \in Z', f \subseteq e}$ of the system $\{\mu_{q,f}\}$ of complexity $K' = O(K)$, and a set $\Omega' \subseteq \Omega \times V_e \subseteq Z'$, such that the following properties hold.

(1) We have

$$\psi'(\Omega') \geq 2^{-4} c_0^2 \eta^2, \quad (3.3)$$

where ψ' is the measure on the parameter space $Z' = Z \times V_e$.

(2) For all $q' = (q, p) \in Z'$ and $f \in \partial e$, there is a σ -algebra $\mathcal{B}_{q',f} \supseteq \mathcal{B}_{q,f}$ of complexity:

$$\text{compl}(\mathcal{B}'_{q,f}) \leq \text{compl}(\mathcal{B}_{q,f}) + 1. \quad (3.4)$$

(3) For all $q' = (q, p) \in \Omega'$, one has

$$\begin{aligned} & \left\| \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) \right\|_{L^2(\mu_{q',e})}^2 \\ & \geq \left\| \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q,f} \right) \right\|_{L^2(\mu_{q,e})}^2 + 2^{-8} \eta^2. \end{aligned} \quad (3.5)$$

and

$$\mu_{q',e}(V_e) \leq 2. \quad (3.6)$$

The meaning of the above lemma is that if there is a large “bad” set Ω of parameters q for which the set $G_{q,e}$ is not sufficiently uniform with respect to the σ -algebra $\bigvee_{f \in \partial e} \mathcal{B}_{q,f}$, then its energy will increase by a fixed amount when passing to a well-defined extension $\{\mathcal{B}_{q',f}\}$, $\{\mu_{q',e}\}$, for all $q' = (q, p) \in \Omega'$.

Proof. Let

$$g_q := \mathbf{1}_{G_{q,e}} - \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q,f} \right). \quad (3.7)$$

By (2.12), we have for each $q \in \Omega$

$$\|g_q\|_{\square_{\nu_{q,e}}}^{2^{d'}} = \int_{V_e} \left\langle g_q, \prod_{f \in \partial e} u_{q,p,f} \right\rangle_{\mu_{(q,p),e}} d\mu_{q,e}(p) \geq \eta, \quad (3.8)$$

where $u_{q,p,f}: V_e \rightarrow [-1, 1]$ are functions, and $\{\mu_{(q,p),e}\}_{(q,p) \in Z'}$ is the family of measures

$$\mu_{(q,p),e}(x) = \prod_{f \subseteq e} \prod_{\substack{\omega \in \{0,1\}^f \\ \omega \neq 0}} \nu_{q,f}(p_f, x_f).$$

As explained after (2.5), the measures $\mu_{(q,p),e}$ are defined by a pairwise independent family of forms $\mathcal{L}_{(q,p),e}$ depending on the parameters $(q, p) \in Z \times V_e$, which is a well-defined extension of the family $\mathcal{L}_{q,e}$ defining the measures $\mu_{q,e}$. It is clear from (3.8) that the measure ψ' on Z' has the form $\psi'(q, p) = \mu_{q,e}(p) \psi(q)$.

For $q' = (q, p)$, let

$$\Gamma(q, p) := \left\langle g_q, \prod_{f \in \partial e} u_{q,p,f} \right\rangle_{\mu_{q,p,f}}. \quad (3.9)$$

We show that there is a set $\Omega'_1 \subseteq \Omega \times V_e$ of measure:

$$\psi'(\Omega'_1) \geq 2^{-3} c_0^2 \eta^2, \quad (3.10)$$

such that for every $(q, p) \in \Omega'_1$ one has

$$\Gamma(q, p) \geq \frac{\eta}{4}. \quad (3.11)$$

By Lemma 2.2, we have that $\mu_{q,e}(V_e) = 1 + o_K(1) \leq 2$ for $q \notin \mathcal{E}_1$, where $\mathcal{E}_1 \subseteq \Omega$ is a set of measure $\psi(\mathcal{E}_1) = o_K(1)$. Thus, for $q \in \Omega \setminus \mathcal{E}_1 = \Omega_1$, we have by (3.8) that

$$\int_{V_e} \mathbf{1}_{\{\Gamma(q,p) \geq \eta/4\}} \Gamma(q,p) d\mu_{q,e}(p) \geq \frac{\eta}{2}, \quad (3.12)$$

where by (3.8) and (3.9), we have

$$\Gamma(q,p) = \int_{V_e} g_q(x) \prod_{f \in \partial e} u_{q,f} w_{q,p}(x) d\mu_{q,e}(x).$$

The function $w_{q,p}(x)$ is the product of weight functions of the form $\nu(L(q,p,x))$ depending on both variables p and x . Thus, using the bounds $|g_q| \leq 1$, $|u_{q,p,f}| \leq 1$, one has

$$\begin{aligned} \int_Z \int_{V_e} |\Gamma(q,p)|^2 d\mu_{q,e}(p) d\psi(q) &\leq \int_Z \int_{V_e} \int_{V_e} \int_{V_e} w_{q,p}(x) w_{q,p}(x') \\ &\quad \times d\mu_{q,e}(x) d\mu_{q,e}(x') d\mu_{q,e}(p) d\psi(q) \\ &= 1 + o_K(1) \leq 2 \end{aligned} \quad (3.13)$$

by the linear form condition, as the factors in the product depend on different sets of variables. Let $\Omega'_1 := \{(q,p) \in \Omega_1 \times V_e : \Gamma(q,p) \geq \eta/4\}$. Thus, by (3.12), (3.13) and the Cauchy–Schwartz inequality,

$$\begin{aligned} \frac{c_0^2 \eta^2}{4} &\leq \left(\int_{\Omega'_1} \Gamma(q,p) d\mu_{q,e}(p) d\psi(q) \right)^2 \leq \int_{\Omega'_1} \Gamma(q,p)^2 d\mu_{q,e}(p) d\psi(q) \psi'(\Omega'_1) \\ &\leq 2 \psi'(\Omega'_1). \end{aligned}$$

This shows $\psi'(\Omega'_1) \geq 2^{-3} c_0^2 \eta^2$ as claimed.

Since $|u_{q',f}| \leq 1$, decomposing each function $u_{q',f}$ into its positive and negative parts yields that

$$\left\langle g_q, \prod_{f \in \partial e} v_{q',f} \right\rangle_{\mu_{q',e}} \geq 2^{-2} \eta \quad (3.14)$$

for some functions $v_{q',f} : V_f \rightarrow [0, 1]$. For a given $f \in \partial e$ and some $0 \leq t_f \leq 1$; let

$$\mathcal{U}_{q',t_f} := \{x_f \in V_f : v_{q',f}(x_f) \geq t_f\}$$

be the level set of the functions $v_{q',f}$. Then, $v_{q',f}(x_f) = \int_0^1 \mathbf{1}_{\mathcal{U}_{q',t_f}}(x_f) dt_f$, and for each term in (3.14), we have

$$\int_0^1 \cdots \int_0^1 \left\langle g_q, \prod_{f \in \partial e} \mathbf{1}_{\mathcal{U}_{q',t_f}} \right\rangle_{\mu_{q',e}} dt \geq 2^{-2} \eta,$$

where $t = (t_f)_{f \in \partial e}$. Accordingly, the integrand must be at least $2^{-d-2} \eta$ for some value of the parameter t . Fix such a $t = (t_f)$ and write $\mathcal{U}_{q',f}$ for $\mathcal{U}_{q',t_f}$ for

the simplicity of notation. For $q' = (q, p) \in \Omega'_1$, define $\mathcal{B}_{q',f}$ to be the σ -algebra generated by $\mathcal{B}_{q,f}$, and $\mathcal{U}_{q',f}$. For $q' \notin \Omega'_1$, set $\mathcal{B}_{q',f} = \mathcal{B}_{q,f}$.

The function $\prod_{f \in \partial e} \mathbf{1}_{\mathcal{U}_{q',f}}$ is constant on the atoms of the σ -algebra $\bigvee_{f \in \partial e} \mathcal{B}_{q',f}$, and therefore, we have

$$\left\langle \mathbf{1}_{G_{q,e}} - \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right), \prod_{f \in \partial e} \mathbf{1}_{\mathcal{U}_{q',f}} \right\rangle_{\mu_{q',e}} = 0$$

for $q' \in \Omega'_1$. Hence, by (3.7) and (3.14), it follows that

$$\begin{aligned} & \left\langle \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) - \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q,f} \right), \prod_{f \in \partial e} \mathbf{1}_{\mathcal{U}_{q',f}} \right\rangle_{\mu_{q',e}} \\ & \geq 2^{-2}\eta. \end{aligned} \quad (3.15)$$

By Lemma 2.2, there is a set $\mathcal{E}_1 \subseteq Z'$, such that $\psi'(\mathcal{E}_1) = o_K(1)$ and

$$\left\| \prod_{f \in \partial e} \mathbf{1}_{\mathcal{U}_{q',f}} \right\|_{L^2(\mu_{q',e})} \leq \mu_{q',e}(V_e)^{1/2} = 1 + o_K(1) \leq 2$$

for $q' \in \Omega'_1 \setminus \mathcal{E}_1 =: \Omega'_2$. Then, by the Cauchy–Schwartz inequality,

$$\left\| \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) - \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q,f} \right) \right\|_{L^2(\mu_{q',e})} \geq 2^{-3}\eta,$$

for $q' \in \Omega'_2$. By Lemma 2.6, there is an exceptional set $\mathcal{E}_2 \subseteq Z'$ of measure $\psi'(\mathcal{E}_2) = o_{K,M}(1)$, such that for $q' = (q, p) \in \Omega'_3 := \Omega'_2 \setminus \mathcal{E}_2$, we have

$$\begin{aligned} & \left\| \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) - \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q,f} \right) \right\|_{L^2(\mu_{q',e})} \\ & \geq 2^{-3}\eta - o_{K,M}(1) \geq 2^{-4}\eta. \end{aligned} \quad (3.16)$$

Since $\mathcal{B}_{q,f} \subseteq \mathcal{B}_{q',f}$, for $q' = (q, p)$, (3.16) is equivalent to

$$\begin{aligned} & \left\| \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) \right\|_{L^2(\mu_{q',e})}^2 \\ & - \left\| \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q,f} \right) \right\|_{L^2(\mu_{q',e})}^2 \geq 2^{-8}\eta^2. \end{aligned} \quad (3.17)$$

Finally, by a further invocation of Lemma 2.6, there is a set $\mathcal{E}_3 \subseteq Z'$ of measure $\psi'(\mathcal{E}_3) = o_{K,M}(1)$, such that for $q' \in \Omega'_4 := \Omega'_3 \setminus \mathcal{E}_3$, we have (for N, W sufficiently large):

$$\begin{aligned} & \left\| \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) \right\|_{L^2(\mu_{q',e})}^2 - \left\| \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q,f} \right) \right\|_{L^2(\mu_{q,e})}^2 \\ & \geq 2^{-9} \eta^2. \end{aligned} \quad (3.18)$$

This proves the lemma by choosing $\Omega' = \Omega'_4$. $\square$

Iterating the above lemma leads to a parametric family of σ -algebras and measures, such that the sets $G_{q,e}$ become sufficiently uniform with respect to them. The associated decomposition of their indicator functions is sometimes referred to as the Koopman–von Neumann-type decomposition [19]. We will replace sets $G_e \subseteq V_e$ by σ -algebras $\mathcal{B}_e$ on V_e for $e \in \mathcal{H}_{d'}$, and for that, it is useful to define the *total energy* of the family $\{\mathcal{B}_e\}_{e \in \mathcal{H}_{d'}}$ with respect to a family of lower order σ -algebras $\{\mathcal{B}_f\}_{f \in \mathcal{H}_{d'-1}}$ and a family of measures $\{\mu_e\}_{e \in \mathcal{H}_{d'}}$ as

$$\sum_{e \in \mathcal{H}_{d'}, G_e \in \mathcal{B}_e} \left\| \mathbb{E}_{\mu_e} \left(\mathbf{1}_{G_e} \mid \bigvee_{f \in \partial e} \mathcal{B}_f \right) \right\|_{L^2(\mu_e)}^2. \quad (3.19)$$

Assuming that the measures μ_e are normalized, i.e., $\mu_e(V_e) = 1 + o(1) \leq 2$, a crude upper bound for the total energy is $2^{d+1} 2^{2^M} = O_M(1)$, where M is the complexity of the σ -algebras $\mathcal{B}_e$.

Lemma 3.2 (The Koopman–von Neumann decomposition). *Let $\{\mu_{q,f}\}_{q \in Z, f \in \mathcal{H}}$ be a well-defined, symmetric family of measures of complexity at most K . Let $1 \leq d' \leq d$, and let $\{\mathcal{B}_{q,e}\}_{q \in Z, e \in \mathcal{H}_{d'}}$ and $\{\mathcal{B}_{q,f}\}_{q \in Z, f \in \mathcal{H}_{d'-1}}$ be families of σ -algebras of complexity at most $M_{d'}$ and $M_{d'-1}$. Finally, let $\Omega \subseteq Z$ with $\psi(\Omega) \geq c_0 > 0$, and let $\delta > 0$ be a constant. Then, for N, W sufficiently large with respect to the constants $\delta, c_0, M_{d'}, M_{d'-1}$ and K , there exists a well-defined, symmetric extension $\{\mu_{q',f}\}_{q' \in Z', f \in \mathcal{H}}$ of the system $\{\mu_{q,f}\}$ of complexity at most $K' = O_{M_{d'}, K, \delta}(1)$ and a family of σ -algebras $\{\mathcal{B}_{q',f}\}_{q' \in Z', f \in \mathcal{H}_{d'-1}}$, such that the following properties hold.*

(1) *For all $q' = (q, p) \in Z'$ and $f \in \mathcal{H}_{d'-1}$, we have*

$$\mathcal{B}_{q,f} \subseteq \mathcal{B}_{q',f}, \quad \text{compl}(\mathcal{B}_{q',f}) \leq \text{compl}(\mathcal{B}_{q,f}) + O_{M_{d'}, \delta}(1). \quad (3.20)$$

(2) *There exists a set $\Omega' \subseteq \Omega \times V \subseteq Z'$ of measure $\psi'(\Omega') \geq c(c_0, \delta, M_{d'}) > 0$ such that for all $q' = (q, p) \in \Omega'$, and for all $G_{q,e} \in \mathcal{B}_{q,e}$, one has*

$$\left\| \mathbf{1}_{G_{q,e}} - \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) \right\|_{\square_{\nu_{q',e}}} \leq \delta, \quad (3.21)$$

and

$$\begin{aligned} \left\| \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q,f} \right) \right\|_{L^2(\mu_{q',e})}^2 &= \left\| \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q,f} \right) \right\|_{L^2(\mu_{q,e})}^2 \\ &\quad + o_{M_{d'}, K, \delta}(1). \end{aligned} \quad (3.22)$$

Proof. Initially, set $Z' = Z$, then (3.20) and (3.22) trivially hold for $q' = q$. If there is a set $\Omega_1 \subseteq \Omega$ of measure $\psi(\Omega_1) \geq \frac{c_0}{2}$, such that inequality (3.21) holds for all $q \in \Omega_1$ and $G_{q,e} \in \mathcal{B}_{q,e}$, then the conclusions of the lemma hold for the initial system of measures and σ -algebras $\{\mu_{q,f}\}$, $\{\mathcal{B}_{q,f}\}$ and the set Ω_1 . Otherwise, there is a set $\Omega_2 \subseteq \Omega$ of measure $\psi(\Omega_2) \geq \frac{c_0}{2}$, such that for each $q \in \Omega_2$, there is an $e \in \mathcal{H}_{d'}$ and a set $G_{q,e} \in \mathcal{B}_{q,e}$ for which the inequality (3.21) fails. By the pigeonholing, we may assume that $e \in \mathcal{H}_{d'}$ is independent of q . Then, by Lemma 3.1, with $\eta := \delta^{2^{d'}}$, there is a well-defined extension $\{\mu_{q',f}\}_{q' \in Z', f \subseteq e}$, a family of σ -algebras $\{\mathcal{B}_{q',f}\}_{q' \in Z', f \subseteq e}$ and a set $\Omega' \subseteq \Omega_2$ for which (3.3)–(3.5) hold. Let $\{\mu_{q',f}\}_{q' \in Z', f \in \mathcal{H}}$ be the symmetrization of the system $\{\mu_{q',f}\}_{q' \in Z', f \subseteq e}$, as described in Sect. 2.3, and set $\mathcal{B}_{q',f} := \mathcal{B}_{q,f}$ for $q' \notin \Omega'$ or $f \not\subseteq e$. By Lemma 2.5, one may remove a set $\mathcal{E}$ of measure $\psi'(\mathcal{E}) = o_{M_{d'},K}(1)$, such that for all $q' \in \Omega' \setminus \mathcal{E}$ (3.20) and (3.22) hold for the extended system, whose total energy is at least $2^{-10}\delta^{2^{d'}}$ larger than that of the initial system $\{\mu_{q,f}\}_{q \in Z, f \in \mathcal{H}}$.

Based on the above argument, we perform the following iteration. Let $\{\mu_{q',f}\}_{q' \in Z', f \in \mathcal{H}}$ be a well-defined, symmetric extension of the initial system $\{\mu_{q,f}\}_{q \in Z, f \in \mathcal{H}}$, $\{\mathcal{B}_{q',f}\}_{q' \in Z', f \in \mathcal{H}_{d'-1}}$ be a family of σ -algebras and let $\Omega' \subseteq \Omega \times V' \subseteq Z'$ for which (3.20) and (3.22) hold. If there is a set $\Omega'_1 \subseteq \Omega'$ of measure $\psi'(\Omega'_1) \geq \psi(\Omega')/2$, such that for all $q \in \Omega'_1$, $e \in \mathcal{H}_{d'}$ and $G_{q,e} \in \mathcal{B}_{q,e}$ inequality (3.22) holds, then the system $\{\mu_{q',f}\}$, $\{\mathcal{B}_{q',f}\}$ together with the set Ω'_1 satisfies the conclusions of the lemma.

Otherwise, there is a well-defined, symmetric extension $\{\mu_{q'',f}\}_{q'' \in Z'', f \in \mathcal{H}}$ together with a family of σ -algebras $\{\mathcal{B}_{q'',f}\}_{q'' \in Z'', f \in \mathcal{H}_{d'-1}}$ and a set $\Omega'' \subseteq \Omega' \times \mathbb{Z}_N^{d'}$ such that for all $q'' \in \Omega''$ inequalities (3.20) and (3.22) hold, and the total energy of the system $(\mu_{q'',f}, \mathcal{B}_{q,e}, \mathcal{B}_{q'',f})$ is at least $2^{-10}\delta^{2^{d'}}$ larger than that of the system $(\mu_{q',f}, \mathcal{B}_{q,e}, \mathcal{B}_{q',f})$. Set $Z' := Z$, $\mu_{q',f} := \mu_{q'',f}$ and $\mathcal{B}_{q',f} := \mathcal{B}_{q'',f}$.

By (3.19), the iteration process must stop in $O_{M_{d'},\delta}(1)$ steps and the system obtained satisfies (3.20)–(3.22). $\square$

3.2. Hypergraph Regularity Lemmas

The shortcoming of Lemma 3.2 is that the complexity of the σ -algebras $\mathcal{B}_{q,f}$ might be very large with respect to the parameter δ , which measures the uniformity of the graphs $G_{q,e}$. This issue can be taken care of by an iteration process using Lemma 3.2; along the lines, it was done in [19]. In the weighted settings, we have to pass to a new system of weights and measures at each iteration and have to exploit the stability properties of well-defined extensions to show that the iteration process terminates.

Lemma 3.3 (A preliminary regularity lemma). *Let $1 \leq d' \leq d$ and $M_{d'} > 0$ be a constant. Let $\{\mu_{q,f}\}_{q \in Z, f \in \mathcal{H}}$ be a well-defined, symmetric family of measures of complexity at most K , and $1 \leq d' \leq d$ and $\{\mathcal{B}_{q,e}\}_{q \in Z, e \in \mathcal{H}_{d'}}$ be a family of σ -algebras on V_e , so that for all $q \in Z$, $e \in \mathcal{H}_{d'}$*

$$\text{compl}(\mathcal{B}_{q,e}) \leq M_{d'}. \quad (3.23)$$

Let $\varepsilon > 0$ and $F : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a non-negative, increasing function, possibly depending on ε and $\Omega \subseteq Z$ be a set of measure $\psi(\Omega) \geq c_0 > 0$.

If N, W are sufficiently large with respect to the parameters $\varepsilon, c_0, M_{d'}, K$, and F , then there exists a well-defined, symmetric extension $\{\mu_{\bar{q},f}\}_{\bar{q} \in \bar{Z}}$ of complexity at most $O_{K,M_{d'},F,\varepsilon}(1)$, and families of σ -algebras $\mathcal{B}_{\bar{q},f} \subseteq \mathcal{B}'_{\bar{q},f}$ defined for $\bar{q} \in \bar{Z}$, $f \in \mathcal{H}_{d-1}$ and a set $\bar{\Omega} \subseteq \bar{Z}$, such that the following hold.

- (1) We have that $\bar{\Omega} \subseteq \Omega \times \bar{V} \subseteq \bar{Z} = Z \times \bar{V}$, where $\bar{V} = \mathbb{Z}_N^k$ is of dimension $k = O_{M_{d'},F,\varepsilon}(1)$. Moreover

$$\bar{\psi}(\bar{\Omega}) \geq c(c_0, F, M_{d'}, \varepsilon) > 0. \quad (3.24)$$

- (2) There is a constant $M_{d'-1} = O_{M_{d'},F,\varepsilon}(1)$, such that for all $\bar{q} \in \bar{Z}$ and $f \in \mathcal{H}_{d'-1}$, we have

$$\text{compl}(\mathcal{B}_{\bar{q},f}) \leq M_{d'-1}. \quad (3.25)$$

- (3) For all $\bar{q} = (q, p) \in \bar{\Omega}$, $e \in \mathcal{H}_{d'}$ and $G_{q,e} \in \mathcal{B}_{q,e}$, we have

$$\left\| \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{\bar{q},f} \right) - \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{\bar{q},f} \right) \right\|_{L^2(\mu_{\bar{q},e})} \leq \varepsilon \quad (3.26)$$

and

$$\left\| \mathbf{1}_{G_{q,e}} - \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{\bar{q},f} \right) \right\|_{\square_{\mu_{\bar{q},e}}} \leq \frac{1}{F(M_{d'-1})}. \quad (3.27)$$

Proof. Let $\{\mu_{q',f}\}_{q' \in Z', f \in \mathcal{H}}$ be a well-defined, symmetric extension of the initial system $\{\mu_{q,f}\}$ defined on a parameter space $Z' = Z \times V'$ of complexity at most K' . In addition, for $q' \in Z'$ and $f \in \mathcal{H}_{d'-1}$, let $\{\mathcal{B}_{q',f}\}_{q' \in Z', f \in \mathcal{H}_{d'-1}}$ be a family of σ -algebras of complexity at most $M_{d'-1}$. Set $\mathcal{B}_{q',e} := \mathcal{B}_{q,e}$ for $q' = (q, p) \in Z'$, $e \in \mathcal{H}_{d'}$, and apply Lemma 3.2 to the system $(\mu_{q',f}, \mathcal{B}_{q',e}, \mathcal{B}_{q',f})$, with $\delta = F(M_{d'-1})^{-1}$.

This generates a well-defined, symmetric extension $\{\mu_{\bar{q},f}\}_{\bar{q} \in \bar{Z}, f \in \mathcal{H}}$ and a family of σ -algebras $\{\mathcal{B}'_{\bar{q},f}\}_{\bar{q} \in \bar{Z}, f \in \mathcal{H}_{d'-1}}$ and a set $\bar{\Omega} \subseteq \bar{Z}$. Set $\mathcal{B}_{\bar{q},f} := \mathcal{B}'_{\bar{q},f}$ for $\bar{q} = (q', p) \in \bar{Z}$, $f \in \mathcal{H}_{d'-1}$. The new system $(\mu_{\bar{q},f}, \mathcal{B}_{\bar{q},f}, \mathcal{B}'_{\bar{q},f})$ satisfies (3.24)–(3.25) and (3.27) as long as the parameters $K', M_{d'-1}$ are of magnitude $O_{K,M_{d'},F,\varepsilon}(1)$. There are two possibilities.

- Case 1. There exists a set $\bar{\Omega}_1 \subseteq \bar{\Omega}$ of measure $\bar{\psi}(\bar{\Omega}_1) \geq \bar{\psi}(\bar{\Omega})/2$, such that (3.26) holds for all $\bar{q} \in \bar{\Omega}_1$. In this case, the conclusions of the lemma hold for the system $(\mu_{\bar{q},f}, \mathcal{B}_{\bar{q},f}, \mathcal{B}'_{\bar{q},f})$ and the set $\bar{\Omega}_1$.
- Case 2. There is a set $\bar{\Omega}_2 \subseteq \bar{\Omega}$ of measure $\bar{\psi}(\bar{\Omega}_2) \geq \frac{1}{2} \bar{\psi}(\bar{\Omega})$, so that inequality (3.26) fails for all $\bar{q} \in \bar{\Omega}_2$. Then, thanks to the stability condition (3.22) and the fact that $\mathcal{B}_{q',f} = \mathcal{B}_{\bar{q},f} \subseteq \mathcal{B}'_{\bar{q},f}$, we have for $\bar{q} \in \bar{\Omega}_2$, $q' = \pi'(\bar{q})$ and $q = \pi(\bar{q})$,

$$\begin{aligned}
& \sum_{e, G_{q,e}} \left\| \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{\bar{q},f} \right) \right\|_{L^2_{\mu_{\bar{q},e}}}^2 - \sum_{e, G_{q,e}} \left\| \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) \right\|_{L^2_{\mu_{q',e}}}^2 \\
& \geq \sum_{e, G_{q,e}} \left(\left\| \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{\bar{q},f} \right) \right\|_{L^2_{\mu_{\bar{q},e}}}^2 - \left\| \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) \right\|_{L^2_{\mu_{\bar{q},e}}}^2 \right) \\
& \quad - o_{M_{d'}, K', F}(1) \\
& = \sum_{e, G_{q,e}} \left\| \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{\bar{q},f} \right) - \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) \right\|_{L^2_{\mu_{\bar{q},e}}}^2 \\
& \geq \varepsilon^2 - o_{M_{d'}, K', F}(1), \tag{3.28}
\end{aligned}$$

where the summation is taken over all $e \in \mathcal{H}_{d'}$ and $G_{q,e} \in \mathcal{B}_{q,e}$.

Thus, for sufficiently large N, W , we have for all $\bar{q} = (q, p) \in \bar{\Omega}_2$ that the total energy of the system $(\mu_{\bar{q},f}, \mathcal{B}_{\bar{q},e}, \mathcal{B}'_{\bar{q},f})$ is at least $\frac{\varepsilon^2}{2}$ larger than that of the system $(\mu_{q',f}, \mathcal{B}_{q,e}, \mathcal{B}_{q',f})$. In this case, set $Z' := \bar{Z}$, $\Omega' := \bar{\Omega}_3$, $\mu_{q',f} := \mu_{\bar{q},f}$, and $\mathcal{B}_{q',f} := \mathcal{B}'_{\bar{q},f}$ and repeat the above argument. Starting with the original system $\mu_{q,f}, \mathcal{B}_{q,e}$ and σ -algebras, $\mathcal{B}_{q,f} = \{\emptyset, V_f\}$, the iteration process must stop in at most $\varepsilon^{-2} 2^{2M_{d'}+1} 2^{d+1} = O_{M_{d'}, \varepsilon}(1)$ steps, generating a system $(\mu_{\bar{q},f}, \mathcal{B}_{\bar{q},f}, \mathcal{B}'_{\bar{q},f})$ which satisfies the conclusions of the lemma. $\square$

To obtain a counting and a removal lemma starting from a given measure system $\{\mu_{q,e}\}$ and σ -algebras $\{\mathcal{B}_{q,e}\}$, we need to regularize the elements of the σ -algebras $\mathcal{B}_{\bar{q},e}$ for all $e \in \mathcal{H}$ with respect to the lower order σ -algebras $\bigvee_{f \in \partial e} \mathcal{B}_{q,f}$. This is done by applying Lemma 3.3 inductively, and provides the final form of the regularity lemma we need. Let us call a function $F: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ a *growth function* if it is continuous, increasing, and satisfies $F(x) \geq 1 + x$ for $x \geq 0$.

Theorem 3.1. (A Regularity Lemma). *Let $1 \leq d' \leq d$ and $M_{d'} > 0$ be a constant. Let $\{\mu_{q,f}\}_{q \in Z, f \in \mathcal{H}}$ be a well-defined, symmetric family of measures of complexity at most K , and $1 \leq d' \leq d$ and $\{\mathcal{B}_{q,e}\}_{q \in Z, e \in \mathcal{H}_{d'}}$ be a family of σ -algebras on V_e so that for all $q \in Z, e \in \mathcal{H}_{d'}$*

$$\text{compl}(\mathcal{B}_{q,e}) \leq M_{d'}. \tag{3.29}$$

Let $F: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a growth function, and $\Omega \subseteq Z$ be a set of measure $\psi(\Omega) \geq c_0 > 0$.

If N, W are sufficiently large with respect to the parameters $c_0, M_{d'}, K$, and F , then there exists a well-defined, symmetric extension $\{\mu_{\bar{q},f}\}_{\bar{q} \in \bar{Z}}$ of complexity at most $O_{K, M_{d'}, F}(1)$, and families of σ -algebras $\mathcal{B}_{\bar{q},f} \subseteq \mathcal{B}'_{\bar{q},f}$ defined for $\bar{q} \in \bar{Z}, f \in \mathcal{H}_{d-1}$ and a set $\bar{\Omega} \subseteq \bar{Z}$, such that the following hold.

- (1) *We have that $\bar{\Omega} \subseteq \Omega \times \bar{V} \subseteq \bar{Z} = Z \times \bar{V}$, where $\bar{V} = \mathbb{Z}_N^k$ is of dimension $k = O_{M_{d'}, F}(1)$. Moreover*

$$\bar{\psi}(\bar{\Omega}) \geq c(c_0, F, M_{d'}) > 0. \tag{3.30}$$

(2) *There exist numbers*

$$\begin{aligned} M_{d'} &< F(M_{d'}) \leq M_{d'-1} < F(M_{d'-1}) \\ &\leq \cdots \leq M_1 < F(M_1) \leq M_0 = O_{M_{d'}, F}(1) \end{aligned} \quad (3.31)$$

such that for all $1 \leq j < d'$, $f \in \mathcal{H}_j$, and $\bar{q} \in \bar{\mathbb{Z}}$,

$$\text{compl}(\mathcal{B}'_{\bar{q},f}) \leq M_j. \quad (3.32)$$

(3) *For all $1 \leq j \leq d'$, $e \in \mathcal{H}_j$, $\bar{q} = (q, p) \in \bar{\Omega}$, and $G_{\bar{q},e} \in \mathcal{B}_{\bar{q},e}$ (with $\mathcal{B}_{\bar{q},e} := \mathcal{B}_{q,e}$, if $j = d'$), one has*

$$\left\| \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{\bar{q},e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{\bar{q},f} \right) - \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{\bar{q},e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{\bar{q},f} \right) \right\|_{L^2(\mu_{\bar{q},e})} \leq \frac{1}{F(M_j)} \quad (3.33)$$

and

$$\left\| \mathbf{1}_{G_{\bar{q},e}} - \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{\bar{q},e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{\bar{q},f} \right) \right\|_{\square_{\nu_{\bar{q},e}}} \leq \frac{1}{F(M_1)}. \quad (3.34)$$

Proof. We proceed by induction on d' . If $d' = 1$, the statement follows from Lemma 3.3 with $\varepsilon = \frac{1}{F(M_1)}$, so assume that $d' \geq 2$ and the theorem holds for $d' - 1$. Apply Lemma 3.3 with a growth function $F^* \geq F$ (to be specified later) and with $\varepsilon = \frac{1}{2F^*(M_{d'})}$. This gives a well-defined, symmetric extension $\{\mu_{q',f}\}$ and a family of σ -algebras $\mathcal{B}'_{q',f} \subseteq \mathcal{B}'_{q',f}$ defined on a parameter space $Z' = Z \times V$, such that

$$\begin{aligned} &\left\| \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q',e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{q',f} \right) - \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q',e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) \right\|_{L^2(\mu_{q',e})} \\ &\leq \frac{1}{2F^*(M_{d'})} \end{aligned} \quad (3.35)$$

and

$$\left\| \mathbf{1}_{G_{q',e}} - \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q',e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{q',f} \right) \right\|_{\square_{\nu_{q',e}}} \leq \frac{1}{F^*(M_{d'-1})} \quad (3.36)$$

hold for all $q' = (q, p) \in \Omega'$, $e \in \mathcal{H}_{d'}$, and $G_{q',e} \in \mathcal{B}_{q',e} = \mathcal{B}_{q,e}$, where $\Omega' \subseteq \Omega \times V \subseteq Z'$ is a set of measure $\psi'(\Omega') \geq c(c_0, F, M_{d'}) > 0$.

Applying the induction hypothesis to the system $\{\mu_{q',f}\}_{q' \in Z'}$, $\{\mathcal{B}'_{q',f}\}_{q' \in Z', f \in \mathcal{H}_{d'-1}}$, the growth function F , and the set Ω' , one obtains an extension $\{\mu_{\bar{q},f}\}_{\bar{q} \in \bar{\mathbb{Z}}, f \in \mathcal{H}}$ and families of σ -algebras $\{\mathcal{B}_{\bar{q},f} \subseteq \mathcal{B}'_{\bar{q},f}\}_{\bar{q} \in \bar{\mathbb{Z}}, f \in \mathcal{H}_j}$, such that (3.32)–(3.34) hold for $j < d' - 1$, with constants

$$M_{d'-1} < F(M_{d'-1}) \leq \cdots \leq M_1 < F(M_1) = O_{M_{d'-1}, F}(1). \quad (3.37)$$

For $\bar{q} = (q', p) \in \bar{Z}$, $f \in \mathcal{H}_{d'-1}$ set $\mathcal{B}_{\bar{q},f} := \mathcal{B}_{q',f}$, and $\mathcal{B}'_{\bar{q},f} := \mathcal{B}'_{q',f}$. We show that inequalities (3.33) and (3.34) hold for $j = d'$. Indeed, by the stability property (2.19), one has

$$\begin{aligned}
 & \left\| \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{q',e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{q',f} \right) - \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{q',e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) \right\|_{L^2(\mu_{\bar{q},e})} \\
 &= \left\| \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q',e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{q',f} \right) - \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q',e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q',f} \right) \right\|_{L^2(\mu_{q',e})} \\
 &\quad + o_{K,M_{d'},F,F^*}(1) \\
 &\leq \frac{1}{2F^*(M_{d'})} + o_{K,M_{d'},F,F^*}(1), \tag{3.38}
 \end{aligned}$$

for all $\bar{q} = (q', p) \in \bar{\Omega} \setminus \bar{\mathcal{E}}_1$, $e \in \mathcal{H}_{d'}$, and $G_{q',e} \in \mathcal{B}_{q',e}$. Here, $\bar{\mathcal{E}}_1 \subseteq \bar{\Omega}$ is a set of measure $\bar{\psi}(\bar{\mathcal{E}}_1) = o_{K,M_{d'},F,F^*}(1)$.

Similarly, using the stability properties (2.17) and (2.24) of the box norms, we have

$$\begin{aligned}
 & \left\| \mathbf{1}_{G_{q',e}} - \mathbb{E}_{\mu_{\bar{q},e}} \left(\mathbf{1}_{G_{q',e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{q',f} \right) \right\|_{\square_{\nu_{\bar{q},e}}} \\
 &= \left\| \mathbf{1}_{G_{q',e}} - \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q',e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{q',f} \right) \right\|_{\square_{\nu_{\bar{q},e}}} + o_{K,M_{d'},F,F^*}(1) \\
 &= \left\| \mathbf{1}_{G_{q',e}} - \mathbb{E}_{\mu_{q',e}} \left(\mathbf{1}_{G_{q',e}} \mid \bigvee_{f \in \partial e} \bar{\mathcal{B}}_{q',f} \right) \right\|_{\square_{\nu_{q',e}}} + o_{K,M_{d'},F,F^*}(1) \\
 &\leq \frac{1}{2F^*(M_{d'-1})} + o_{K,M_{d'},F,F^*}(1), \tag{3.39}
 \end{aligned}$$

for all $\bar{q} = (q', p) \in \bar{\Omega} \setminus \bar{\mathcal{E}}_2$, $e \in \mathcal{H}_{d'}$ and $A_{q',e} \in \mathcal{B}_{q',e} = \mathcal{B}_{\bar{q},e}$, where $\bar{\mathcal{E}}_2 \subseteq \bar{\Omega}$ is a set of measure $\bar{\psi}(\bar{\mathcal{E}}_2) = o_{K,M_{d'},F,F^*}(1)$.

With $F(M_1) = O_{M_{d'-1},F}(1)$, we have that $F(M_1) \leq C(M_{d'-1}, F) =: \frac{1}{2}F^*(M_{d'-1})$ for a sufficiently rapidly growing function F^* depending only on F . Assuming that N , W are sufficiently large with respect to $M_{d'}$ and K , inequalities (3.33), (3.34) for $j = d'$ and $\bar{q} \in \bar{\Omega} \setminus (\bar{\mathcal{E}}_1 \cup \bar{\mathcal{E}}_2)$ follow from (3.35) and (3.36). The rest of the conclusions of the theorem are clear from the construction. $\square$

4. The Counting and the Removal Lemmas

4.1. The Removal Lemma

In this section, we formulate a so-called Counting Lemma and show how it implies Theorem 1.4. Our arguments will closely follow and are straightforward adaptations of those in [19] to the weighted settings; for the sake of completeness, we will include the details.

For $e \in \mathcal{H}_d$ let $G_e \subseteq V_e$ be a hypergraph, and let $\mathcal{B}_e = \{A_e, A_e^C, \emptyset, V_e\}$ be the σ -algebra generated by it. Let $\{\nu_e\}_{e \in \mathcal{H}}$ and $\{\mu_e\}_{e \in \mathcal{H}}$ be the weights and measures associated with a well-defined, symmetric family of forms $\mathcal{L} = \{L_e^k : e \in \mathcal{H}_d, 0 \leq k \leq d\}$. Take $M_d > 0$ and $F: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a growth function to be determined later and apply Theorem 3.1 with $d' = d$ to obtain a well-defined, symmetric parametric extension $\{\mu_{q,e}\}_{q \in Z, e \in \mathcal{H}}$ together with σ -algebras $\mathcal{B}_{q,e} \subseteq \mathcal{B}'_{q,e}$ and a set $\Omega \subseteq Z$, such that (3.30)–(3.34) hold.¹ Note that the complexity of the system as well as the σ -algebras is $O_{M_d, F}(1)$. We consider the system of measures $\mu_{q,e}$ and the σ -algebras $\mathcal{B}_{q,e}$, $\mathcal{B}'_{q,e}$ fixed for the rest of this section.

It will be convenient to define all our σ -algebras on the same space V_J and use only the measures $\mu_q := \mu_{q,J} = \prod_{f \in \mathcal{H}} \nu_{q,f}$ instead of using a whole ensemble of measures $\{\mu_{q,e}\}_{e \in \mathcal{H}}$. Thanks to the stability conditions (2.6)–(2.7), this can be done at essentially no cost. Indeed, for any $e \in \mathcal{H}$, there is an exceptional set $\mathcal{E}_e \subseteq \Omega$ of measure $\psi(\mathcal{E}_e) = o_{M_d, F}(1)$, such that for any family of sets $G_{q,e} \subseteq V_e$, we have that

$$\mu_q(\pi_e^{-1}(G_{q,e})) = \mu_{q,e}(G_{q,e}) + o_{M_d, F}(1), \quad (4.1)$$

uniformly for $q \in \Omega \setminus \mathcal{E}_e$. Let $\mathcal{E} = \bigcup_{e \in \mathcal{H}} \mathcal{E}_e$, $\Omega' := \Omega \setminus \mathcal{E}$ define the σ -algebras $\overline{\mathcal{B}}_{q,e} := \pi_e^{-1}(\mathcal{B}_{q,e})$ and $\overline{\mathcal{B}}'_{q,e} := \pi_e^{-1}(\mathcal{B}'_{q,e})$ on V_J . Note that $\overline{\mathcal{B}}_{q,e} = \overline{\mathcal{B}}_e$ for $e \in \mathcal{H}_d$ as the initial σ -algebras $\mathcal{B}_e$ are not altered in Theorem 3.1. Let $\overline{\mathcal{B}}_q := \bigvee_{e \in \mathcal{H}} \overline{\mathcal{B}}_{q,e}$ be the σ -algebra generated by the algebras $\overline{\mathcal{B}}_{q,e}$, and define similarly the σ -algebra $\overline{\mathcal{B}}'_q$. The atoms of $\overline{\mathcal{B}}_q$ are of the form $A_q = \bigcap_{f \in \mathcal{H}} A_{q,e}$, where $A_{q,e}$ is an atom of $\overline{\mathcal{B}}_{q,e}$. In particular, if $E_e \in \overline{\mathcal{B}}_e$, then $\bigcap_{e \in \mathcal{H}_d} E_e$ is the union of the atoms of $\overline{\mathcal{B}}_q$.

The so-called counting lemma [5, 13, 19] gives an approximate formula for the measure of “most” atoms A_q , and as a consequence, it shows that their measure is bounded below by a positive constant depending only on the initial data F and M_d . If, as in Theorem 1.4, one assumes that the measure of $\bigcap_{e \in \mathcal{H}_d} E_e$ is sufficiently small, then it cannot contain most of the atoms, thus removing the exceptional atoms from the sets E_e , the intersection of the remaining sets becomes empty, leading to a proof of Theorem 1.4.

To make this heuristic precise, let us start by defining the *relative density* $\delta_q(A_q|B_q) := \mu_q(A \cap B)/\mu_q(B)$ for $A, B \in \overline{\mathcal{B}}_q$, with the convention that $\delta_q(A_q|B_q) := 1$ if $\mu_q(B) = 0$.

¹The family $\{\nu_e\}$ can be considered as a parametric family of weights in a trivial way, setting $Z = \Omega = \{0\}$, and $\psi(0) = 1$.

Definition 4.1. Let $A_q = \bigcap_{e \in \mathcal{H}} A_{q,e}$ be an atom of $\mathcal{B}_q$. We say that the atom A_q is *regular* if the following hold.

(1) For all atoms $A_{q,e}$,

$$\delta_q \left(A_{q,e} \mid \bigcap_{f \in \partial e} A_{q,f} \right) \geq \frac{1}{\log F(M_j)}. \quad (4.2)$$

(2) Moreover,

$$\begin{aligned} & \int_{V_J} \left| \mathbb{E}_{\mu_q} \left(\mathbf{1}_{A_{q,e}} \mid \bigvee_{f \in \partial e} \overline{\mathcal{B}}'_{q,e} \right) - \mathbb{E}_{\mu_q} \left(\mathbf{1}_{A_{q,e}} \mid \bigvee_{f \in \partial e} \overline{\mathcal{B}}_{q,e} \right) \right|^2 \prod_{f \subsetneq e} \mathbf{1}_{A_{q,f}} d\mu_q \\ & \leq \frac{1}{F(M_j)} \int_{V_J} \prod_{f \subsetneq e} \mathbf{1}_{A_{q,f}} d\mu_q. \end{aligned} \quad (4.3)$$

This roughly means that all atoms $A_{q,e}$ are both somewhat large and regular on the intersection of the lower order atoms $A_{q,f}$, ($f \in \partial e$).

Proposition 4.1 (Counting lemma). *There is a set $\mathcal{E} \subseteq \Omega$ of measure $\psi(\mathcal{E}) = o_{N,W \rightarrow \infty; M_d, F}(1)$, such that if $q \in \Omega \setminus \mathcal{E}$ and $A_q = \bigcap_{e \in \mathcal{H}} A_{q,e} \in \bigvee_{e \in \mathcal{H}} \overline{\mathcal{B}}_{q,e}$ is a regular atom, then*

$$\begin{aligned} \mu_q(A_q) &= (1 + o_{M_d \rightarrow \infty}(1)) \prod_{e \in \mathcal{H}} \delta_q(A_{q,e}) + O_{M_1} \left(\frac{1}{F(M_1)} \right) \\ &+ o_{N,W \rightarrow \infty; M_d, F}(1). \end{aligned} \quad (4.4)$$

Next, following [19], we show that the total measure of irregular atoms is small. For any atom $A_{q,e} \in \overline{\mathcal{B}}_{q,e}$, let $B_{q,e,A_{q,e}}$ be the union of all sets of the form $\bigcap_{f \subsetneq e} A_{q,f}$ for which (4.2) or (4.3) fails. Note that if an atom $A_q = \bigcap_{e \in \mathcal{H}} A_{q,e}$ is irregular, then $A_q \subseteq A_{q,e} \cap B_{q,e,A_{q,e}}$ for some $e \in \mathcal{H}$. We claim that

$$\mu_q(A_{q,e} \cap B_{q,e,A_{q,e}}) \lesssim \frac{1}{\log F(M_j)} \quad (4.5)$$

for $q \notin \mathcal{E}_1$, where $\mathcal{E}_1 \subseteq \Omega$ is a set of measure $\psi(\mathcal{E}_1) = o_{M_d, F}(1)$. To see this, estimate first, the contribution of those sets $\bigcap_{f \subsetneq e} A_{q,f}$ to the left side of (4.11) for which (4.2) fails. This quantity is bounded by

$$\begin{aligned} & \sum_{\{A_{q,f}\}_{f \in \partial e}, \text{ (4.2) fails}} \mu_q \left(A_{q,e} \cap \bigcap_{f \in \partial e} A_{q,f} \right) \\ & \leq \frac{1}{\log F(M_j)} \sum_{\{A_{q,f}\}_{f \in \partial e}} \mu_q \left(\bigcap_{f \in \partial e} A_{q,f} \right) \\ & \leq \frac{1}{\log F(M_j)} \mu_q(V_J) \lesssim \frac{1}{\log F(M_j)}, \end{aligned}$$

as the summation is taken over the disjoint atoms of the σ -algebra $\bigvee_{f \in \partial e} \overline{\mathcal{B}}_{q,e}$.

Similarly, one estimates the total contribution of the disjoint atoms $\bigcap_{f \subsetneq e} A_{q,f}$ for which (4.3) fails as follows.

$$\begin{aligned} & \sum_{\{A_{q,f}\}_{f \subseteq e}, (4.3) \text{ fails}} \mu_q \left(\bigcap_{f \subsetneq e} A_{q,f} \right) \\ & \leq F(M_j) \int_{V_e} \left| \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{A_{q,e}} \bigvee_{f \in \partial e} \overline{\mathcal{B}}'_{q,f} \right) - \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{A_{q,e}} \bigvee_{f \in \partial e} \overline{\mathcal{B}}_{q,f} \right) \right|^2 d\mu_q \\ & \leq F(M_j) \frac{1}{F(M_j)^2} = \frac{1}{F(M_j)}. \end{aligned}$$

Since the sets $A_{q,e} \cap B_{q,e,A_{q,e}}$ contain all irregular atoms, and for given $e \in \mathcal{H}_j$, the number of all atoms of the σ -algebra $\overline{\mathcal{B}}_{q,e}$ is at most $2^{2^{M_j}}$, one estimates the total measure of all irregular atoms as

$$\begin{aligned} & \sum_{j=1}^d \sum_{e \in \mathcal{H}_j} \sum_{A_{q,e} \in \mathcal{B}_{q,e}} \mu_q(A_{q,e} \cap \mathcal{B}_{q,e,A_{q,e}}) \leq \sum_{j=1}^d \binom{d}{j} 2^{2^{M_j}} \frac{1}{\log F(M_j)} \\ & \leq \frac{1}{\sqrt{\log F(M_d)}} \leq 2^{-2^{M_d}} \end{aligned} \quad (4.6)$$

if, say $F(M) \geq 2^{2^{M_d+d}}$. This shows, choosing M_d sufficiently large, that most atoms are regular.

Another fact we need is that the measure of regular atoms is not too small. Indeed, by (4.2) and (4.4), we have that for $q \in \Omega$ and a regular atom $A_q = \bigcap_{e \in \mathcal{H}} A_{q,e}$,

$$\begin{aligned} \mu_q(A_q) & \geq \prod_{j \leq d} \prod_{e \in \mathcal{H}_j} \frac{1}{F(M_j)^{1/10}} - O_{d,M_1} \left(\frac{1}{F(M_1)} \right) + o_{M_d,F}(1) \\ & \geq \frac{1}{F(M_1)} > 0, \end{aligned} \quad (4.7)$$

as long as F is sufficiently rapid growing and M_d is sufficiently large with respect to d . It is clear from (3.31) that $F(M_1) \leq F^*(M_d)$ for a function F^* depending only on F and M_d .

After these preparations, assuming the validity of Proposition 4.1, it is easy to obtain the

Proof of Theorem 1.4. Let $\delta > 0$, $E_e \in \mathcal{A}_e$ and $g_e: V_e \rightarrow [0, 1]$ for $e \in \mathcal{H}_d$ be given. Let $\mathcal{E}_1 \subseteq \Omega$ be a set of measure $\psi(\mathcal{E}_1) = o_{M_d,F}(1)$, so that (4.1), (4.12) and (4.13) hold for $q \in \Omega/\mathcal{E}_1$. In addition, by the conditions in (2.11), we see that (1.20)–(1.21) hold for

$$\tilde{\mu}_J := \mu_{q,J} \quad \text{and} \quad \tilde{\mu}_e := \mu_{q,e} \quad (e \in \mathcal{H}_d), \quad (4.8)$$

for $q \notin \mathcal{E}_2$, for a set $\mathcal{E}_2 \subseteq \Omega$ is a set of measure $\psi(\mathcal{E}_2) = o_{M_d,F}(1)$.

Now, fix $q \notin \mathcal{E}_1 \cup \mathcal{E}_2$ and define $\tilde{\mu}_J$ and $\tilde{\mu}_e$ for $e \in \mathcal{H}_d$ as in (4.14). We claim that this system of measures satisfies the conclusions of the theorem. By

construction, the system is symmetric, so it remains to construct the sets E'_e and show that (1.22)–(1.23) hold. Given $e \in \mathcal{H}$, define the sets

$$E'_{q,e} = V_J \setminus \left(B_{q,e,A_e} \cup \bigcup_{f \not\supseteq e, A_{q,f}} (A_{q,f} \cap B_{q,f,A_{q,f}}) \right), \quad (4.9)$$

where $A_{q,f}$ ranges over the atoms of $\mathcal{B}_{q,f}$. As we have $\mathcal{B}_{q,e} = \mathcal{B}_e$, which is generated by a single set E_e , if $\bigcap_{e \in \mathcal{H}_d} E_e$ contains an atom $A_q = \bigcap_{f \in \mathcal{H}} A_{q,f}$ then $A_{q,e} = E_e$ for $e \in \mathcal{H}_d$. If such an atom would be regular, then by (1.19), its measure would satisfy

$$\frac{1}{F^*(M_d)} \leq \tilde{\mu}_J \left(\bigcap_{e \in \mathcal{H}_d} E_e \right) = \mu_J \left(\bigcap_{e \in \mathcal{H}_d} E_e \right) + o_{M_d, F}(1) < 2\delta.$$

Choosing M_d to be the largest positive integer, so that $F^*(M_d) \leq (2\delta)^{-1}$ we see that $\bigcap_{e \in \mathcal{H}_d} E_e$ contains only irregular atoms. From (4.15) and (4.12), we have

$$\tilde{\mu}_J(E_e \setminus E'_{q,e}) = \tilde{\mu}_J \left(\bigcup_{f \not\supseteq e, A_{q,f}} (A_{q,f} \cap B_{q,f,A_{q,f}}) \right) \leq 2^{-2^{M_d}}. \quad (4.10)$$

In addition, all irregular atoms $A_q = \bigcap_{f \in \mathcal{H}} A_{q,f} \subseteq \bigcap_{e \in \mathcal{H}_d} E_e$ are contained in one of the sets $E_e \setminus E'_{q,e}$, thus

$$\bigcap_{e \in \mathcal{H}_d} (E_e \cap E'_{q,e}) = \emptyset.$$

Finally, choosing $\varepsilon := 2^{-2^{M_d}}$, (1.23) holds by (4.16). Moreover, $\delta \rightarrow 0$ implies $M_d \rightarrow \infty$, and hence, $\varepsilon \rightarrow 0$. This proves Theorem 1.4. $\square$

4.2. Proof of Proposition 4.1

The proof proceeds by induction and uses the Cauchy–Schwartz inequality, causing to double certain sets of variables. As a consequence, we need a generalization of Proposition 4.1 which requires the following definition.

Definition 4.2 (*Weighted hypergraph bundles over $\mathcal{H}$*). Let K be a finite set together with a map $\pi: K \rightarrow J$, called the projection map of the bundle. Let $\mathcal{G}_K$ be the set of edges $g \subseteq K$, such that π is injective on any g and $\pi(g) \in \mathcal{H}$.

For any $g \in \mathcal{G}_K$, write

$$V_g := V_{\pi(g)} = \prod_{k \in g} V_{\pi(k)},$$

and define the weight functions $\bar{\nu}_{q,g}: V_g \rightarrow \mathbb{R}_+$ by

$$\bar{\nu}_{q,g}(x_g) := \nu_{q,\pi(g)}(x_g),$$

and the associated measures $\bar{\mu}_{q,g}$ with density functions:

$$d\bar{\mu}_{q,g}(x_g) = \prod_{g' \subseteq g} \bar{\nu}_{q,g'}(x_{g'}).$$

The total measure $\bar{\mu}_{q,K}$ on V_K is given by

$$d\bar{\mu}_{q,K}(x_g) = \prod_{g' \subseteq \mathcal{G}} \bar{\nu}_{q,g'}(x_{g'}).$$

A hypergraph $\mathcal{G} \subseteq \mathcal{G}_K$ which is closed in the sense that $\partial g \in \mathcal{G}$ for every $g \in \mathcal{G}$, together with the spaces V_g and the weight functions $\bar{\mu}_{q,g}$ for $g \in \mathcal{G}$ is called a weighted hypergraph bundle over $\mathcal{H}$. The quantity $d' = \sup_{g \in \mathcal{G}} |g|$ is called the order of $\mathcal{G}$.

Note that the underlying linear forms defining the weight system $\{\nu_{q,g}\}_{q \in \mathbb{Z}, g \in \mathcal{G}_K}$,

$$\bar{L}_g^s(q, x_g) = L_{\pi(g)}^s(q, x_g),$$

are pairwise linearly independent and the system of weights

$$\bar{\nu}_{q,g} := \prod_{s \in s(\pi(g))} \bar{L}_g^s(q, x_g)$$

is well-defined.

Lemma 4.1 (A Generalized Counting Lemma). *Let $\mathcal{G} \subseteq \mathcal{G}_K$ be a closed hypergraph bundle over $\mathcal{H}$ with the projection map $\pi: K \rightarrow J$, and $d' := \sup_{g \in \mathcal{G}} |g|$ be the order of $\mathcal{G}$. Then, for F growing sufficiently rapidly with respect to d and K , there exists a set $\mathcal{E} \subseteq \Omega$ of measure $\psi(\mathcal{E}) = o_{N \rightarrow \infty; M_d, K, F}(1)$, such that for $q \in \Omega \setminus \mathcal{E}$, we have*

$$\begin{aligned} \bar{\mu}_{q,K} \left(\bigcap_{g \in \mathcal{G}} \bar{A}_{q, \pi(g)} \right) &= (1 + o_{M_d \rightarrow \infty, K}(1)) \prod_{g \in \mathcal{G}} \delta_{q, f}(A_{q, \pi(g)}) \\ &\quad + O_{K, M_1} \left(\frac{1}{F(M_1)} \right) + o_{N \rightarrow \infty, K, M_d}(1). \end{aligned} \quad (4.11)$$

Note that Proposition 4.1 is the special case when $K = J$, π is the identity map and $\mathcal{G} = \mathcal{H}$.

Proof. We use a double induction. First, we induct on d' , the order of $\mathcal{G}$ (note that $d' \leq d$), and then, fixing K and π , we induct on the number of edges $r := |\{g \in \mathcal{G} : |g| = d'\}|$.

To start, assume that $d' = r = 1$, so that $\mathcal{G} = \{k\}$ and $j = \pi(k) \in J$. The left-hand side of (4.13) becomes

$$\begin{aligned} \bar{\mu}_{q,K}(\bar{A}_{q,j}) &= \int_{V_K} \mathbf{1}_{\bar{A}_{q,j}}(x_K) d\bar{\mu}_{q,K}(x_K) \\ &= \int_{V_k} \mathbf{1}_{A_{q,j}}(x_k) d\bar{\mu}_{q,k}(x_k) + o_{N \rightarrow \infty; M_d, K, F}(1) \\ &= \int_{V_j} \mathbf{1}_{A_{q,j}}(x_j) d\bar{\mu}_{q,j}(x_j) + o_{N \rightarrow \infty, K}(1) \\ &= \delta_{q,j}(A_{q,j}) + o_{N \rightarrow \infty; M_d, K, F}(1) \end{aligned}$$

for $q \notin \mathcal{E} \subseteq Z$, where $\mathcal{E}$ is a set of measure $\psi(\mathcal{E}) = o_{N \rightarrow \infty; M_d, K, F}(1)$. The second and fourth equalities follow from Lemma 2.2 using the fact that the family of measures $\{\bar{\mu}_{q,g}\}$ is well-defined. The third equality holds, because $V_K = V_j$ and

$$d\bar{\mu}_{q,K}(x_j) = \prod_{s \in s(j)} \bar{\nu}_{q,k}^s(x_j) = \prod_{s \in s(j)} \nu_{q,j}^s(x_j) = d\mu_{q,j}(x_j).$$

Let $\{A_{q,e}\}_{e \in \mathcal{H}}$ be a regular collection of atoms for $q \in \Omega$, and define the functions $b_{q,e}, c_{q,e}: V_e \rightarrow \mathbb{R}$ for $e \in \mathcal{H}$ by

$$b_{q,e} := \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{A_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{q,f} \right) - \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{A_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q,f} \right) \quad (4.12)$$

$$c_{q,e} := \mathbf{1}_{A_{q,e}} - \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{A_{q,e}} \mid \bigvee_{f \in \partial e} \mathcal{B}'_{q,f} \right). \quad (4.13)$$

Note that if $x_e \pi_e(\in \bigcap_{f \in \partial e} \bar{A}_{q,f}) \cap V_e$, then

$$\delta_{q,e} = \mathbb{E}_{\mu_{q,e}} \left(\mathbf{1}_{A_e} \mid \bigvee_{f \in \partial e} \mathcal{B}_{q,f} \right), \quad (4.14)$$

and thus, one has the decomposition

$$\mathbf{1}_{A_{q,e}}(x_e) = \delta_{q,e} + b_{q,e}(x_e) + c_{q,e}(x_e). \quad (4.15)$$

Let $g_0 \in \mathcal{G}$, such that $|g_0| = d'$ and use (4.17) to write

$$\prod_{g \in \mathcal{G}} \mathbf{1}_{A_{q,\pi(g)}}(x_g) = (\delta_{q,\pi(g_0)} + b_{q,\pi(g_0)}(x_{g_0}) + c_{q,\pi(g_0)}(x_{g_0})) \prod_{g \in \mathcal{G} \setminus \{g_0\}} \mathbf{1}_{A_{q,\pi(g)}}(x_g).$$

Now,

$$\begin{aligned} \bar{\mu}_q \left(\bigcap_{g \in \mathcal{G}} \bar{A}_{q,\pi(g)} \right) &= \int_{V_K} \prod_{g \in \mathcal{G}} \mathbf{1}_{\bar{A}_{q,\pi(g)}}(x) d\bar{\mu}_q(x) = \int_{V_K} \prod_{g \in \mathcal{G}} \mathbf{1}_{A_{q,\pi(g)}}(x_g) d\bar{\mu}_q(x) \\ &= \int_{V_K} (\delta_{q,\pi(g_0)} + b_{q,\pi(g_0)}(x_{g_0}) + c_{q,\pi(g_0)}(x_{g_0})) \\ &\quad \times \prod_{g \in \mathcal{G} \setminus \{g_0\}} \mathbf{1}_{A_{q,\pi(g)}}(x_g) d\bar{\mu}_q(x) \\ &= M_q + E_q^1 + E_q^2. \end{aligned} \quad (4.16)$$

Consider the first main term M_q . By the second induction hypothesis, we have

$$\begin{aligned} M_q &= \delta_{q,\pi(g_0)} \bar{\mu}_q \left(\bigcap_{g \in \mathcal{G} \setminus \{g_0\}} \bar{A}_{q,\pi(g)} \right) \\ &= \delta_{q,\pi(g_0)} (1 + o_{M_d \rightarrow \infty}(1)) \prod_{g \in \mathcal{G}} \delta_{q,\pi(g)} \\ &\quad + O_{K,M_1} \left(\frac{1}{F(M_1)} \right) + o_{N \rightarrow \infty, K, M_d}(1), \end{aligned}$$

and then, M_q agrees with the right side of (4.4). We continue to estimate the second error term by

$$\begin{aligned} E_q^2 &= \int_{V_K} c_{q,\pi(g_0)}(x_{g_0}) \prod_{g \in \mathcal{G} \setminus \{g_0\}} \mathbf{1}_{A_{q,\pi(g)}}(x_g) d\bar{\mu}_q(x) \\ &= \mathbb{E}_{x \in V_K} (c_{q,\pi(g_0)} \bar{\nu}_{q,g_0})(x_{g_0}) \prod_{g \in \mathcal{G} \setminus \{g_0\}} \mathbf{1}_{A_{q,\pi(g)}} \bar{\nu}_{q,g}(x_g) \\ &= \mathbb{E}_{x \in V_K} \prod_{|g|=d', g \in \mathcal{G}} f_{q,g}(x_g) \prod_{g' \in \mathcal{G}, |g'| < d'} \bar{\nu}_{q,g'}(x_{g'}), \end{aligned} \quad (4.17)$$

where $f_{q,g_0} = c_{q,\pi(g_0)} \bar{\nu}_{q,g_0}$ and $f_{q,g} = h_{q,g} \bar{\nu}_{q,g}$, for $g \in \mathcal{G}, g \neq g_0$ and $|g| = d'$ for a function $h_{q,g}$ of magnitude at most 1. Thus, we have $|f_{q,g}| \leq \bar{\nu}_{q,g}$ for all $g \in \mathcal{G}, |g| = d'$. Note that we are essentially in the situation of Proposition 1.1, the generalized von Neumann inequality. Indeed, applying the Cauchy–Schwartz inequality d' times successively in the variables x_j , $j \in g_0$, as in the Appendix, to clear all functions $f_{q,g}(x_g)$ which does not depend on at least one of these variables, we obtain

$$\begin{aligned} |E_q^2|^{2^{d'}} &\lesssim \|c_{q,\pi(g_0)}\|_{\square_{\bar{\nu}_{q,g_0}}}^{2^{d'}} + \mathbb{E}_{x_{g_0}, y_{g_0}} |\mathcal{W}_q(x_{g_0}, y_{g_0}) - 1| \\ &\quad \times \prod_{h \subseteq g_0} \prod_{\omega_h \in \{0,1\}^h} \bar{\nu}_{q,h}(\omega_h(x_h, y_h)), \end{aligned} \quad (4.18)$$

where setting $K' := K \setminus g_0$,

$$\mathcal{W}_q(x_{g_0}, y_{g_0}) = \mathbb{E}_{x \in V_{K'}} \prod_{g \in \mathcal{G}, g \not\subseteq g_0} \prod_{\omega_g \in \{0,1\}^{g \cap g_0}} \bar{\nu}_{q,g}(\omega_g(x_{g \cap g_0}, y_{g \cap g_0}), x_{g \setminus g_0}). \quad (4.19)$$

Note that the first term on the right-hand side of (4.18) is $O(F(M_1)^{2^{-d'}})$ by (3.12) and (4.13). To estimate the second term, we apply the Cauchy–Schwartz inequality one more time to see that it is $o_{N \rightarrow \infty; M_d, K, F}(1)$ for $q \notin \mathcal{E}_1$, $\mathcal{E}_1$ being a set of measure $o_{N \rightarrow \infty; M_d, K, F}(1)$ using the fact that the underlying linear forms are pairwise linearly independent in the variables $(q, x_{g_0}, y_{g_0}, x_{K'})$.

Finally, we estimate the error term E_q^1 defined as

$$E_q^1 = \int_{V_K} b_{q,\pi(g_0)}(x_{g_0}) \prod_{g \in \mathcal{G} \setminus \{g_0\}} \mathbf{1}_{A_{q,\pi(g)}}(x_g) d\bar{\mu}_q(x).$$

Taking absolute values and discarding all factors $\mathbf{1}_{A_{q,\pi(g)}}(x_g)$ for $|g| = d'$, $g \neq g_0$, one estimates

$$\begin{aligned} |E_q^1| &\leq \int_{V_{g_0}} \left(|b_{q,\pi(g_0)}(x_{g_0})| \prod_{g \subsetneq g_0} \mathbf{1}_{A_{q,\pi(g)}}(x_g) \right) \\ &\quad \times \left(\mathbb{E}_{x_{K'}} \prod_{g \in \mathcal{G}', |g| < d'} \mathbf{1}_{A_{q,\pi(g)}} \bar{\nu}_{q,g}(x_g) \prod_{h \in \mathcal{G}', |h|=d'} \bar{\nu}_{q,h}(x_h) \right) d\bar{\mu}_{q,g_0}(x_{g_0}), \end{aligned}$$

where $\mathcal{G}' = \{g \in \mathcal{G} : g \not\subseteq g_0\}$. Since $\bar{\nu}_{q,g_0}(V_{g_0}) = 1 + o_{N \rightarrow \infty; M_d, K, F}(1)$ outside a set $\mathcal{E}_2 \subseteq \Omega$ of measure $\psi(\mathcal{E}_2) = o_{N \rightarrow \infty; M_d, K, F}(1)$, we may write

$$E_q^1 = \int_{V_{g_0}} A(x_{g_0}) B(x_{g_0}) d\bar{\mu}_{q,g_0}(x_{g_0}),$$

and apply the Cauchy–Schwartz inequality to get

$$|E_q^1|^2 \lesssim \left(\int_{V_{g_0}} A(x_{g_0})^2 d\bar{\mu}_{q,g_0}(x_{g_0}) \right) \left(\int_{V_{g_0}} B(x_{g_0})^2 d\bar{\mu}_{q,g_0}(x_{g_0}) \right). \quad (4.20)$$

The first factor on the left side of (4.20) is estimated by

$$\mathbb{E}_{x_{g_0} \in V_{g_0}} b_{q,\pi(g_0)}(x_{g_0})^2 \prod_{g \subsetneq g_0} \mathbf{1}_{A_{q,\pi(g)}}(x_g) \prod_{g \subseteq g_0} \nu_{q,\pi(g)}(x_g). \quad (4.21)$$

Let $f_0 = \pi(g_0)$. Since $\pi : g_0 \rightarrow f_0$ is injective and $V_{g_0} = V_{f_0}$, we may write the expression in (4.21), by re-indexing the variables x_g to x_f for $g \subseteq g_0$, $f = \pi(g)$, as

$$\begin{aligned} &\int_{V_{f_0}} b_{q,f_0}^2(x_{f_0}) \prod_{f \subsetneq f_0} \mathbf{1}_{A_{q,f}}(x_f) d\mu_{q,f_0}(x_{f_0}) \\ &\lesssim \frac{1}{F(M_{d'})} \int_{V_{f_0}} \prod_{f \subsetneq f_0} \mathbf{1}_{A_{q,f}}(x_f) d\mu_{q,f_0}(x_{f_0}), \end{aligned} \quad (4.22)$$

where the inequality follows from assumption (4.2) on regular atoms. By the induction hypothesis, we further estimate the right side of (4.22) as

$$\begin{aligned} &\frac{1}{F(M_{d'})} (1 + o_{M_d \rightarrow \infty}(1)) \prod_{f \subsetneq f_0} \delta_{q,f}(A_{q,f}) \\ &\quad + O_{M_d} \left(\frac{1}{F(M_1)} \right) + o_{N \rightarrow \infty; M_d, K, F}(1). \end{aligned} \quad (4.23)$$

The second factor in (4.20) may be expressed in terms of a hypergraph bundle $\tilde{K}$ over K , involving the construction in [19]. Let $\tilde{K} = K_0 \oplus_{g_0} K$, the set $K \times \{0, 1\}$ with the elements $(k, 0)$ and $(k, 1)$ are identified for $k \in g_0$. Let $\phi : \tilde{K} \rightarrow K$ be the natural projection, and $\pi \circ \phi : \tilde{K} \rightarrow J$ be the associated

map down to J . Let $\mathcal{G}_0 = \{g \in \mathcal{G}, g \subseteq g_0\}$ and $\mathcal{G}' = \{g \in \mathcal{G}, g \not\subseteq g_0, |g| < d'\}$ and define the hypergraph bundle $\tilde{\mathcal{G}}$ on $\tilde{K}$ as $\tilde{\mathcal{G}} = \{g \times \{i\}; \mathcal{G}_0 \cup \mathcal{G}', i \in 1, 2\}$. Note that by our construction the edges $g \times \{0\}$ and $g \times \{1\}$ are identified for $g \in \mathcal{G}_0$. Define the weights

$$\tilde{\nu}_{q,g \times \{i\}}(x_{g \times \{i\}}) := \bar{\nu}_{q,g}(x_{g \times \{i\}}), \quad (4.24)$$

for $q \in Z$, $g \in \mathcal{G}_K, i = 0, 1$, that is; for all edges $\tilde{g} \in \tilde{\mathcal{G}}$, and let $\tilde{\mu}_{q,g \times \{i\}}$ be the associated family of measures. Then, we have for the second factor appearing in (4.20) (with $h_0 := K \setminus g_0$)

$$\begin{aligned} & \int_{V_{g_0}} B(x_{g_0})^2 d\bar{\mu}_{q,g_0}(x_{g_0}) \\ &= \int_{V_{g_0}} \left(\prod_{g \in g_0} \mathbf{1}_{A_{q,\pi(g)}(x_g)} \right) \\ & \quad \left(\mathbb{E}_{x_{h_0} \in V_{K \setminus g_0}} \prod_{g \in \mathcal{G} \setminus \{g_0\}} \mathbf{1}_{A_{q,\pi(g)}(x_g)} \bar{\nu}_{q,g}(x_g) \prod_{h \not\subseteq g_0, |h'|=d'} \bar{\nu}_{q,h}(x_h) \right)^2 d\bar{\mu}_{q,g_0}(x_{g_0}) \\ &= \int_{V_{\tilde{K}}} \prod_{\tilde{g} \in \tilde{\mathcal{G}}} \mathbf{1}_{A_{q,\pi \circ \phi(\tilde{g})}(x_{\tilde{g}})} d\tilde{\mu}_{q,\tilde{K}}(x_{\tilde{K}}). \end{aligned} \quad (4.25)$$

Indeed, when expanding the square of the inner sum in (4.23), we double all points in $K \setminus g_0$; thus, we eventually sum over $x_{\tilde{K}} \in V_{\tilde{K}}$, also double all edges $g \in \tilde{\mathcal{G}}$ to obtain the edges $g \times \{0\}, g \times \{1\}$. As for the weights, the procedure doubles all weights $\bar{\nu}_{q,g}(x_g)$ for $g \not\subseteq g_0, g \in \mathcal{G}_K$ to obtain the weights $\bar{\nu}_{q,g}(x_{g \times \{i\}})$ for $i = 0, 1$, while leaves the weights $\bar{\nu}_{q,g}(x_g)$ for $g \subseteq g_0$ unchanged. The order of $\tilde{g}$ is less than d' ; thus, by the first induction hypothesis, we have

$$\begin{aligned} & \tilde{\mu}_{q,\tilde{K}} \left(\bigcap_{\tilde{g} \in \tilde{\mathcal{G}}} \bar{A}_{q,\pi \circ \phi(\tilde{g})} \right) \\ &= (1 + o_{M_d \rightarrow \infty}(1)) \prod_{\tilde{g} \in \tilde{\mathcal{G}}} \delta_{q,\pi \circ \phi(\tilde{g})}(A_{q,\pi \circ \phi(\tilde{g})}) \\ & \quad + O_{K,M_1} \left(\frac{1}{F(M_1)} \right) + o_{N \rightarrow \infty; M_d, K, F}(1) \\ &= (1 + o_{M_d \rightarrow \infty}(1)) \prod_{g \in \mathcal{G}_0} \delta_{q,\pi(g)}(A_{q,\pi(g)}) \prod_{g \in \mathcal{G}'} \delta_{q,\pi(g)}^2(A_{q,\pi(g)}) \\ & \quad + O_{d,K,M_1} \left(\frac{1}{F(M_1)} \right) + o_{N \rightarrow \infty; M_d, K, F}(1), \end{aligned} \quad (4.26)$$

for $q \notin \mathcal{E}_{\tilde{K},\phi}$, where $\mathcal{E}_{\tilde{K},\phi} \subseteq \Omega$ is a set of measure $\psi(\mathcal{E}_{\tilde{K},\phi}) = o_{N \rightarrow \infty; M_d, K, F}(1)$. Note that there are only $O_K(1)$ choices for the set $\tilde{K}$ and the projection map $\phi: \tilde{K} \rightarrow K$; thus, taking the union of all possible exceptional sets $\mathcal{E}_{\tilde{K},\phi}$, we have that (4.26) holds for $q \notin \mathcal{E}'_K$ and $\psi(\mathcal{E}'_K) = o_{N \rightarrow \infty; M_d, K, F}(1)$. Combining the bounds (4.23) and (4.26), we obtain the error estimate $E_q^1 = o_{N \rightarrow \infty; M_d, K, F}(1)$

outside a set $\mathcal{E}'_K$ of measure $o_{N \rightarrow \infty; M_d, K, F}(1)$. This closes the induction and the proposition follows. $\square$

5. Proof of the Main Results

In this section, we finish the proof of our main result (Theorem 1.2). Since we have already proved the validity of Theorem 1.4 and, hence, by the argument in the introduction concerning Theorem 1.3, it remains to show that counting affine copies of Δ in a set $A \subseteq \mathbb{Z}_N^d$ with weights w translates to counting copies in $A \subseteq \mathbb{P}^d$ of positive relative density. This is standard, and we include the details for the sake of completeness, following the arguments given in [2].

First, let us identify $[1, N]^d$ with $\mathbb{Z}_N^d$ and let us recall that constellations in $\mathbb{Z}_N^d$ defined by the simplex Δ which are contained in a box $B \subseteq [1, N]^d$ of size εN are in fact genuine constellations contained in B . Note that we can assume that the simplex Δ is *primitive* in the sense that $t\Delta \not\subseteq \mathbb{Z}^d$ for any $0 < t < 1$, as any simplex is a dilation of a primitive one. To any simplex $\Delta \subseteq \mathbb{Z}^d$, there exists a constant $\tau(\Delta) > 0$ depending only on Δ , such that the following property holds.

Lemma 5.1 [2]. *Let $E \subseteq \mathbb{Z}^d$ be a primitive simplex. Then, there is a constant $0 < \varepsilon < \tau(\Delta)$, such that the following holds.*

Let N be sufficiently large, and let $B = I^d$ be a box of size εN contained in $[1, N]^d \simeq \mathbb{Z}_N^d$. If there exist $x \in \mathbb{Z}_N^d$ and $1 \leq t < N$, such that $x \in B$ and $x + tE \subseteq B$ as a subset on $\mathbb{Z}_N^d$, then either $x + t\Delta \subseteq B$ or $x + (t - N)\Delta \subseteq B$, also as a subset of $\mathbb{Z}^d$.

Proof. [Proof of Theorem 1.3 implies Theorem 1.2] Fix $\alpha > 0$ and let $\Delta = \{0, v_1, \dots, v_d\} \subseteq \mathbb{Z}^d$ be a non-degenerate simplex. Let $W = W(\alpha)$ be a sufficiently large integer and let $A \subseteq \mathbb{P}_N^d$ be a set with $|A| \geq \alpha |\mathbb{P}_N|^d$, where N is assumed to be sufficiently large with respect to both W and α . By the pigeonhole principle, choose $b = (b_j)_{1 \leq j \leq d}$, so that b_j is relatively prime to W for each j , and

$$|A \cap ((W\mathbb{Z})^d + b)| \geq \alpha \frac{N^d}{(\log N)^d \phi(W)^d}, \quad (5.1)$$

where ϕ is the Euler totient function. Set $N_1 := N/W$ and $A_1 := \{n \in [1, N_1]^d : Wn + b \in A\}$. Choose $\varepsilon_2 > 0$, such that $\varepsilon_2 < \tau(\Delta)$. By the prime number theorem, there is a prime N' , such that $\varepsilon_2 N' = N_1(1 + o_{N_1 \rightarrow \infty}(1))$; thus, we have

$$|A_1 \cap [1, \varepsilon_2 N']^d| \geq \frac{\alpha \varepsilon_2^d}{2} \frac{(N')^d W^d}{(\log N')^d \phi(W)^d}. \quad (5.2)$$

By Dirichlet's theorem on primes in arithmetic progressions, the number of $n \in [1, N']^d \setminus [\varepsilon_1 N', N']^d$ for which $Wn + b \in \mathbb{P}^d$ is of $O_d(\varepsilon_1 \frac{N'^d W^d}{(\log N')^d \phi(W)^d})$, thus (5.2) holds for the set $A' := A_1 \cap [\varepsilon_1 N', \varepsilon_2 N']^d$ as well, if $\varepsilon_1 \leq c_d \varepsilon_2^d \alpha$ for a small enough constant $c_d > 0$.

Fix $b = (b_j)_{1 \leq j \leq d}$, $0 < \varepsilon_1 < \varepsilon_2$ as above, and let $\nu_j = \nu_{b_j}$ be the Green–Tao measures on $\mathbb{Z}_{N'}$. For $n = (n_j)_{1 \leq j \leq d} \in A'$, one has that $Wn + b \in A \subseteq \mathbb{P}^d$ hence by (1.2) $w(x) = c_d(\frac{\phi(W)}{W} \log N)^d$. Thus

$$\mathbb{E}_{x \in \mathbb{Z}_{N'}^d} \mathbf{1}_{A'}(x) w(x) \geq c(d, \Delta) \alpha, \quad (5.3)$$

for some constant $c(d, \Delta) > 0$. Applying Theorem 1.3 for the set $A' \subseteq \mathbb{Z}_{N'}^d$, it follows that

$$\mathbb{E}_{(x,t) \in \mathbb{Z}_{N'}^{d+1}} \prod_{i=1}^d \mathbf{1}_{A'}(x + tv_i) w(x + t\Delta) \geq c(d, \Delta, \alpha) > 0. \quad (5.4)$$

If $x + t\Delta \subseteq A'$, then by (1.2) and (1.4), we have that $w(x + t\Delta) = c_\Delta(\frac{\phi(W)}{W} \log N)^{l(\Delta)}$; thus, the set $A' \subseteq \mathbb{Z}_{N'}^d$ must contain at least $c(d, \Delta, \alpha)(N')^{d+1}(\frac{\phi(W)}{W} \log N)^{-l(\Delta)}$ simplices of the form $\Delta' = x + t\Delta$. By Lemma 5.1, this remains true in $\mathbb{Z}^d$ as $A' \subseteq [\varepsilon_1 N, \varepsilon_2 N]^d$. Since $W\Delta' + b \subseteq WA' + b \subseteq A$ and $N' \approx N/\varepsilon_2 W$, it follows that the set A contains at least

$$\begin{aligned} c(d, \Delta, \alpha) W^{l(\Delta)-d-1} \phi(W)^{-l(\Delta)} N^{d+1} (\log N)^{-l(\Delta)} \\ = c'(d, \Delta, \alpha) N^{d+1} (\log N)^{-l(\Delta)} \end{aligned}$$

simplices of the form $\Delta' = y + s\Delta$, as we have chosen a sufficiently large but fixed integer $W = W(\alpha)$. This proves Theorem 1.2. $\square$

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix A: Basic Properties of Weighted Box Norms

In this appendix, we describe some basic facts about the weighted version of Gowers’s box norms defined in (1.29) for functions $F: V_e \rightarrow \mathbb{R}$. We will assume $e = \{1, \dots, d\} =: [d]$, and $V = V_{[d]} = \mathbb{Z}_N^d$ without loss of generality. To show that these are indeed norms (for $d \geq 2$) let us define a multilinear form referred to as the weighted Gowers’s inner product. Let $F_\omega: V_e \rightarrow \mathbb{R}$ for $\omega \in \{0, 1\}^e$, be a given family of functions and define

$$\begin{aligned} \langle F_\omega; \omega \in \{0, 1\}^d \rangle_{\square_\nu} \\ := \mathbb{E}_{x_{[d]}, y_{[d]} \in V} \prod_{\omega \in \{0, 1\}^d} F_\omega(\omega(x_{[d]}, y_{[d]})) \prod_{|I| < d} \prod_{\omega_I \in \{0, 1\}^I} \nu_I(\omega_I(x_I, y_I)). \end{aligned}$$

Therefore, $\langle F_\omega; \omega \in \{0, 1\}^d \rangle_{\square_\nu} = \|F\|_{\square_\nu}^{2^d}$, if $F_\omega = F$ for all $\omega \in \{0, 1\}^e$.

Lemma A.1. (The Gowers–Cauchy–Schwartz inequality). $|\langle F_\omega; \omega \in \{0, 1\}^d \rangle| \leq \prod_{\omega_{[d]}} \|F_\omega\|_{\square_\nu^d}.$

Proof. We will use the Cauchy–Schwartz inequality several times and the linear form condition. We have

$$\begin{aligned}
 & \langle F_\omega; \omega \in \{0, 1\}^d \rangle_{\square_\nu^d} \\
 &= \mathbb{E}_{x_{[2,d]}, y_{[2,d]}} \left(\left(\prod_{|I| < d, 1 \notin I} \prod_{\omega_I} \nu_I(\omega_I(x_I, y_I)) \right)^{1/2} \right. \\
 & \quad \times \left(\mathbb{E}_{x_1} \nu(x_1) \prod_{\omega_{[2,d]}} F_{\omega_{(0,[2,d])}}(x_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]})) \right) \\
 & \quad \times \prod_{|I| < d-1, 1 \notin I} \nu_{\{1\} \cup I}(x_1, \omega_I(x_I, y_I)) \left(\prod_{|I| < d, 1 \notin I} \prod_{\omega_I} \nu_I(\omega_I(x_I, y_I)) \right)^{1/2} \\
 & \quad \times \left(\mathbb{E}_{y_1} \nu(y_1) \prod_{\omega_{[2,d]}} F_{\omega_{(1,[2,d])}}(y_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]})) \right) \\
 & \quad \times \left. \prod_{|I| < d-1, 1 \notin I} \nu_{\{1\} \cup I}(y_1, \omega_I(x_I, y_I)) \right).
 \end{aligned}$$

Applying the Cauchy–Schwartz inequality in the x_1 variable, one has

$$\left| \langle F_\omega; \omega \in \{0, 1\}^d \rangle_{\square_\nu^d} \right|^2 \leq A \cdot B$$

here,

$$\begin{aligned}
 A &= \mathbb{E}_{x_{[2,d]}, y_{[2,d]}} \left(\prod_{|I| < d, 1 \notin I} \prod_{\omega_I} \nu_I(\omega_I(x_I, y_I)) \right. \\
 & \quad \times \mathbb{E}_{x_1, y_1} \nu(x_1) \nu(y_1) \prod_{\omega_{[2,d]}} F_{\omega_{(0,[2,d])}}(x_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]})) \\
 & \quad \times F_{\omega_{(0,[2,d])}}(y_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]})) \\
 & \quad \times \left. \prod_{|I| < d-1, 1 \notin I} \nu_{\{1\} \cup I}(x_1, \omega_I(x_I, y_I)) \right) = \left\langle F_\omega^{(0)}(\omega(x_{[d]}, y_{[d]})) \right\rangle_{\square_\nu^d},
 \end{aligned}$$

where

$$\begin{aligned}
 F_{(0, \omega_{[2,d]})}^{(0)}(x_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]})) &= F_{(1, \omega_{[2,d]})}^{(0)}(y_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]})) \\
 &:= F(x_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]}))
 \end{aligned}$$

for any $\omega_{[2,d]}$. Similarly

$$B = \left\langle F_\omega^{(1)}(\omega(x_{[d]}, y_{[d]})) \right\rangle_{\square_\nu^d},$$

where

$$\begin{aligned} F_{(0, \omega_{[2,d]})}^{(1)}(x_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]})) &= F_{(1, \omega_{[2,d]})}^{(1)}(y_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]})) \\ &:= F(y_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]})) \end{aligned}$$

for any $\omega_{[2,d]}$. In the same way, applying the Cauchy–Schwartz inequality in x_2 variable, we end up with

$$|\langle F_\omega; \omega \in \{0, 1\}^d \rangle_{\square_\nu^d}| \leq \prod_{\omega_{[0,1]}} \langle F_\omega^{[\omega_{[0,1]}]}; \omega \in \{0, 1\}^d \rangle_{\square_\nu^d}$$

and continuing this way with $x_3, \dots, x_d$ variables, we end up with

$$|\langle F_\omega; \omega \in \{0, 1\}^d \rangle_{\square_\nu^d}| \leq \prod_{\omega \in \{0, 1\}^d} \langle F_\omega, \dots, F_\omega \rangle_{\square_\nu^d} = \prod_{\omega \in \{0, 1\}^d} \|F_\omega\|_{\square_\nu^d}^{2^d}.$$

□

Corollary A.1. $\|\cdot\|_{\square_\nu^d}$ is a semi-norm for $d \geq 1$.

Proof. By the Gowers–Cauchy–Schwartz inequality, we have that $\|F\|_{\square_\nu} \geq 0$; moreover

$$\begin{aligned} \|F + G\|_{\square_\nu^d}^{2^d} &= \langle F + G, \dots, F + G \rangle_{\square_\nu^d} \\ &= \sum_{\omega \in \{0, 1\}^d} \langle h^{\omega_1}, \dots, h^{\omega_d} \rangle_{\square_\nu^d}, \quad h^\omega = \begin{cases} F, & \omega = 0 \\ G, & \omega = 1 \end{cases} \\ &\leq \sum_{\omega \in \{0, 1\}^d} \|h^{\omega_1}\|_{\square_\nu^d} \dots \|h^{\omega_d}\|_{\square_\nu^d} = (\|F\|_{\square_\nu^d} + \|G\|_{\square_\nu^d})^{2^d}. \end{aligned}$$

In addition, it follows directly from the definition that $\|\lambda F\|_{\square_\nu^d}^{2^d} = \lambda^{2^d} \|F\|_{\square_\nu^d}^{2^d}$; hence, $\|\lambda F\|_{\square_\nu^d} = |\lambda| \|F\|_{\square_\nu^d}$. □

Proof of Proposition 1.1. Let $\mathcal{H}' = \{f \in \mathcal{H}; |f| < d\}$, and write the left side of (1.26) as

$$\mathbb{E} = \mathbb{E}_{x \in V_J} \prod_{e \in \mathcal{H}_d} F_e(x_e) \prod_{f \in \mathcal{H}'} \nu_f(x_f).$$

Fix $e_0 = [d]$ and write $e_j := [d+1] \setminus \{j\}$ for the rest of the faces. The idea is to apply the Cauchy–Schwartz inequality successively in the $x_1, x_2, \dots, x_d$ variables to eliminate the functions $F_{e_1} \leq \nu_{e_1}, \dots, F_{e_d} \leq \nu_{e_d}$, using the linear form condition at each step. Using $F_{e_1} \leq \nu_{e_1}$, we have

$$|E| \leq \mathbb{E}_{x_2, \dots, x_{d+1}} \nu_{e_1}(x_1) \prod_{1 \notin f \in \mathcal{H}'} \nu_f(x_f) \left| \mathbb{E}_{x_1} \prod_{j \neq 2} F_{e_j}(x_j) \prod_{1 \in f \in \mathcal{H}'} \nu_f(x_f) \right|.$$

By the linear form condition $\mathbb{E}_{x_2, \dots, x_{d+1}} \nu_{e_1}(x_1) \prod_{1 \notin f \in \mathcal{H}'} \nu_f(x_f) = 1 + o_{N \rightarrow \infty}(1)$, thus by the Cauchy–Schwartz inequality,

$$\begin{aligned} E^2 &\lesssim \mathbb{E}_{x_2, \dots, x_{d+1}} \nu_{e_1}(x_1) \prod_{1 \notin f \in \mathcal{H}'} \nu_f(x_f) \mathbb{E}_{x_1, y_1} \prod_{j \neq 2} F_{e_j}(x_1, x_{e_j \setminus \{1\}}) F_{e_j}(y_1, x_{e_j \setminus \{1\}}) \\ &\quad \times \prod_{1 \in f \in \mathcal{H}'} \nu_f(y_1, x_{f \setminus \{1\}}) \nu_f(x_1, x_{f \setminus \{1\}}). \end{aligned} \quad (\text{A.1})$$

Note that, what happened is that we have replaced the function F_{e_1} by the measure ν_{e_1} , doubled the variable x_1 to the pair of variables (x_1, y_1) and also doubled each factor of the form $G_e(x_e)$ (which is either $F_e(x_e)$ or $\nu_e(x_e)$, for $e \in \mathcal{H}$) depending on the x_1 variable. To keep track of these changes as we continue with the rest of that variables, let us introduce some notations. Let $g \subseteq [d]$ and for a function $G_e(x_e)$ define

$$G_e^*(x_{e \cap g}, y_{e \cap g}, x_{e \setminus g}) := \prod_{\omega_e \in \{0,1\}^{e \cap g}} G_e(\omega_e(x_{e \cap g}, y_{e \cap g}), x_{e \setminus g}). \quad (\text{A.2})$$

We claim that after applying the Cauchy–Schwartz inequality in the $x_1, \dots, x_i$ variables, we have with $g = [i]$

$$\begin{aligned} E^{2^i} &\lesssim \mathbb{E}_{x_{[i]}, y_{[i]}, x_{J \setminus [i]}} \prod_{j \leq i} \nu_{e_j}^*(x_{[i] \cap e_j}, y_{[i] \cap e_j}, x_{e_j \setminus [d]}) \prod_{j > i} F_{e_j}^*(x_{[i] \cap e_j}, y_{[i] \cap e_j}, x_{e_j \setminus [d]}) \\ &\quad \times \prod_{f \in \mathcal{H}'} \nu_f^*(x_{f \cap [i]}, y_{f \cap [i]}, x_{f \setminus [i]}). \end{aligned} \quad (\text{A.3})$$

For $i = 1$, this can be seen from (A.1). Note that the linear forms appearing in any of these factors are pairwise linearly independent as our system is well-defined. Assuming it holds for i separating the factors independent of the x_{i+1} variable, replacing the function $F_{e_{i+1}}$ with $\nu_{e_{i+1}}$, and applying the Cauchy–Schwartz inequality we double the variable x_{i+1} to the pair (x_{i+1}, y_{i+1}) and each factor $G_e^*(x_{e \cap [i]}, y_{e \cap [i]}, x_{e \setminus [i]})$ depending on it, to obtain the factor $G_e^*(x_{e \cap [i+1]}, y_{e \cap [i+1]}, x_{e \setminus [i+1]})$; thus, the formula holds for $i + 1$. After finishing this process, we have by (A.2) and (A.3)

$$\begin{aligned} E^{2^d} &\lesssim \mathbb{E}_{x_{[d]}, y_{[d]}} \prod_{\omega \in \{0,1\}^d} F_{e_0}(\omega(x_{[d]}, y_{[d]})) \\ &\quad \times \prod_{f \subseteq [d], f \neq e_0} \prod_{\omega_f \in \{0,1\}^f} \nu_f(\omega_f(x_f, y_f)) \mathcal{W}(x_{[d]}, y_{[d]}), \end{aligned}$$

where

$$\mathcal{W}(x_{[d]}, y_{[d]}) = \mathbb{E}_{x_{d+1}} \prod_{d+1 \in e \in \mathcal{H}} \prod_{\omega_e \in \{0,1\}^{e \cap [d]}} \nu_e(\omega_e(x_{e \cap [d]}, y_{e \cap [d]}, x_{e \setminus [d]})).$$

Thus, as $F_{e_0} \leq \nu_{e_0}$, to prove (1.26), it is enough to show that

$$\mathbb{E}_{x_{[d]}, y_{[d]}} \prod_{f \subseteq [d]} \prod_{\omega_f \in \{0,1\}^f} \nu_f(\omega_f(x_f, y_f)) |\mathcal{W}(x_{[d]}, y_{[d]}) - 1| = o_{N \rightarrow \infty}(1).$$

This, similarly as in [7], can be done with one more application of the Cauchy–Schwartz inequality leading to four terms involving the “big” weight functions

$\mathcal{W}$ and $\mathcal{W}^2$. Each term is, however, $1 + o_{N \rightarrow \infty}(1)$ by the linear form condition, as the underlying linear forms are pairwise linearly independent. Indeed, the forms $L_f(\omega_f(x_f, y_f))$ are pairwise linearly independent for $f \subseteq [d]$, and depend on a different set of variables rather than the forms $L_e(\omega_e(x_{e \cap [d]}, y_{e \cap [d]}, x_{e \setminus [d]}))$ for $e \not\subseteq [d]$ defining the weight function $\mathcal{W}$. The new forms appearing in $\mathcal{W}^2$ are copies of the forms in $\mathcal{W}$ with the x_{d+1} variable replaced by a new variable y_{d+1} hence are independent of each other and the rest of the forms. This proves the proposition. $\square$

References

- [1] Conlon, D., Fox, J., Zhao, Y.: A relative Szemerédi theorem. *Geom. Funct. Anal.* 25(3), 733–762 (2015)
- [2] Cook, B., Magyar, A.: Constellations in $\mathbb{P}^d$. *Int. Math. Res. Not.* 2012(12), 2794–2816 (2012)
- [3] Furstenberg, H., Katznelson, Y.: An ergodic Szemerédi theorem for commuting transformations. *J. Analyse Math.* 34, 275–291 (1979)
- [4] Fox, J., Zhao, Y.: A short proof of the multidimensional Szemerédi theorem in the primes. *Amer. J. Math.* 137(4), 1139–1145 (2015)
- [5] Gowers, W.T.: Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. of Math.* 166(3), 897–946 (2007)
- [6] Gowers, W.T.: Decompositions, approximate structure, transference, and the Hahn-Banach theorem. *Bull. Lond. Math. Soc.* 42(4), 573–606 (2010)
- [7] Green, B., Tao, T.: The primes contain arbitrarily long arithmetic progressions. *Ann. of Math.* 167(2), 481–547 (2008)
- [8] Green, B., Tao, T.: Linear equations in primes. *Ann. of Math.* 171(3), 1753–1850 (2010)
- [9] Green, B., Tao, T.: The Möbius function is strongly orthogonal to nilsequences. *Ann. of Math.* 175(2), 541–566 (2012)
- [10] Green, B., Tao, T., Ziegler, T.: An inverse theorem for the Gowers $U^{s+1}[N]$ -norm. *Ann. of Math.* 176(2), 1231–1372 (2012)
- [11] Goldston, D.A., Yıldırım, C. Y.: Higher correlations of divisor sums related to primes I: Triple correlations. *Integers* 3, #A5 (2003)
- [12] Magyar, A., Titichetrakun, T.: Corners in dense subsets of P^d . Preprint
- [13] Nagle, B., Rödl, V., Schacht, M.: The counting lemma for regular k -uniform hypergraphs. *Random Structures Algorithms* 28(2), 113–179 (2006)
- [14] Reingold, O., Trevisan, L., Tulsiani, M., Vadham, S.: Dense subsets of pseudo-random sets. In: 49th Annual IEEE Symposium on Foundations of Computer Science, pp. 76–85. IEEE Computer Society, Philadelphia, PA (2008)

- [15] Solymosi, J.: Note on a generalization of Roth's theorem. In: Aronov, B., Basu, S., Pach, J., Shariv., M. (eds.) *Discrete and Computational Geometry, Algorithms Combin.* Vol. 25, pp. 825–827. Springer, Berlin (2003)
- [16] Szemerédi, E.: On sets of integers containing no k elements in arithmetic progression. *Acta Arith.* 27, 199–245 (1975)
- [17] Tao, T.: The Gaussian primes contain arbitrarily shaped constellations. *J. Anal. Math.* 99, 109–176 (2006)
- [18] Tao, T.: Szemerédi's regularity lemma revisited. *Contrib. Discrete Math.* 1(1), 8–28 (2006)
- [19] Tao, T.: A variant of the hypergraph removal lemma. *J. Combin. Theory Ser. A* 113(7), 1257–1280 (2006)
- [20] Tao, T., Ziegler, T.: A multi-dimensional Szemerédi theorem for the primes via a correspondence principle. *Israel J. Math.* 207(1), 203–228 (2015)

Ákos Magyar
University of Georgia
Athens, GA 30602
USA
e-mail: amagyar@uga.edu

Brian Cook
Kent State University
Kent, OH
USA
e-mail: bcook25@kent.edu

Tatchai Titichetrakun
University of British Columbia
Vancouver B.C.
Canada
e-mail: tatchai@math.ubc.ca

Received: 11 July 2018.