A MULTIDIMENSIONAL SZEMERÉDI THEOREM IN THE PRIMES

BRIAN COOK, ÁKOS MAGYAR, TATCHAI TITICHETRAKUN

ABSTRACT. Let A be a subset of positive relative upper density of $\mathbb{P}^d$, the d-tuples of primes. We prove that A contains an affine copy of any finite set $F \subseteq \mathbb{Z}^d$, which provides a natural multi-dimensional extension of the theorem of Green and Tao on the existence of long arithmetic progressions in the primes. The proof uses the hypergraph approach by assigning a pseudo-random weight system to the pattern F on a $d+1$-partite hypergraph; a novel feature being that the hypergraph is no longer uniform with weights attached to lower dimensional edges. Then, instead of using a transference principle, we proceed by extending the proof of the so-called hypergraph removal lemma to our settings, relying only on the linear forms condition of Green and Tao.

1. INTRODUCTION

1.1. **Background.** A celebrated theorem in additive combinatorics due to Green and Tao [7] establishes the existence of arbitrary 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 [17] on the existence of long arithmetic progressions in dense subsets of the integers.

Another fundamental result in this area is the multi-dimensional 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 [14].

It is natural to ask if a multi-dimensional 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 [18] 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.

A common feature of the above mentioned results is that they use an embedding of the underlying sets (the primes or the Gaussian primes) into a set which is sparse but sufficiently random with respect to the pattern F. In our case when the set F is not in general position (the simplest example being a 2-dimensional corner) this does not seem possible, due to the extra correlations arising from the direct product structure. For example if 3 vertices of a rectangle is in $\mathbb{P}^2$, then the fourth vertex is necessarily in $\mathbb{P}^2$, a type of self-correlation not present in the one dimensional case or the Gaussian primes.

Another approach, already partly used in [18], is to establish a hypergraph removal lemma [5], [14] for sparse uniform hypergraphs or alternatively with weights attached to the faces. This approach has been utilized by the second and third authors [13] to show the existence of d-dimensional corners (simplices with edges parallel to the coordinate axis) in dense subsets of $\mathbb{P}^d$. Recently a proof based on hypergraph theory

1991 Mathematics Subject Classification. 11B30, 05C55.

The second author is supported by NSERC grant 22R44824 and ERC-AdG. 321104.

using only the linear forms conditions, has been obtained in [1], covering both the original Green-Tao theorem and the case of the Gaussian primes.

In all of the above approaches the crucial point is to prove a removal lemma for a weighted (or sparse) *uniform* hypergraphs, using transference arguments to remove the weights from the hyperedges. As opposed, for a general constellation in $\mathbb{P}^d$ the hypergraph approach leads to a weighted *closed* hypergraph with weights attached possibly to any lower dimensional edge, and the usual transference principles do not apply. Our approach is different, 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 settings, which might be of independent interest. In this aspect our argument is essentially self-contained, relying only on results from sieve theory, namely on the so-called linear forms conditions [7].

Simultaneously with our original work on this problem, the existence of arbitrary constellations in relative dense subsets of $\mathbb{P}^d$ has also been shown by Tao and Ziegler [20], using an entirely different method based on an *infinite* number of linear forms conditions to obtain a weighted version of the Furstenberg correspondence principle, and a short, elegant proof by Fox and Zhao [4] has been obtained afterwards using sampling arguments. Both of the above proofs however rely on full force of the results of Green, Tao and Ziegler developed in [8],[9],[10] for the study of asymptotic number of prime solutions for systems linear equations. As such the methods of [20], and [4] are do not provide bounds, while from our approach one can extract quantitative statements. The bounds, though recursive, are rather poor (iterated tower-exponential type) and we do not pursue to explicitly calculate them here. Also, as we rely only on sieve-techniques our approach is somewhat flexible, i.e. it might not be hard to modify it 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.

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

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

where $\mathbb{P}_N$ denotes the set 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. Also, replacing the set F by $F' = F \cup (-F)$ one can require that the dilation parameter t is positive. 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 $F = \{0, x_1, \ldots, x_k\}$, choose a set of k linearly independent vectors $\{y_1, \ldots, y_k\} \subseteq \mathbb{Z}^k$, and define the set $\Delta := \{0, (x_1, y_1), \ldots, (x_k, y_k), z_{k+1}, \ldots, z_{k+d}\} \subseteq \mathbb{Z}^{k+d}$ such that the vectors of $\Delta \backslash \{\underline{0}\}$ 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 $\pi : \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}^d$ 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

$$l(\Delta) := \sum_{i=1}^{d} |\pi_i(\Delta)|, \tag{1.2.1}$$

$\pi_i : \mathbb{R}^d \to \mathbb{R}$ 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 $c(\alpha, \Delta) > 0$ 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 $c(\alpha, \Delta) N^{d+1} (\log N)^{-l(\Delta)}$ 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 2 implies Theorem 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 $\Delta' \subseteq \mathbb{P}^d$ 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 ϵ, where $\epsilon = \epsilon(\delta)$ is a quantity such that $\epsilon(\delta) \to 0$ as $\delta \to 0$. As for a number of similar results [7], [18], [2], [20], to prove this, one formulates a statement involving a pseudo-random measure $\nu = \nu^{(N)} : [1, N] \to \mathbb{R}_+$.

1.3. **The Green-Tao measure and the linear forms condition.** Let us recall the pseudo-random measure ν introduced by Green and Tao, and (a slight variant of) the so-called linear forms condition, see [7], Sec.9.

Let ω be a sufficiently large number and let $W = \prod_{p \leq \omega} p$ be the product of primes up to ω. For given b relative prime to W define the modified von Mangoldt function $\bar{\Lambda}_b : \mathbb{Z} \to \mathbb{R}_{\geq 0}$ 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 divisor sums closely related to 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 $R = N^{d^{-1} 2^{-d-5}}$. Indeed, for given small parameters $0 < \varepsilon_1 < \varepsilon_2 < 1$ (whose values will be specified later), recall the Green-Tao measure

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

Clearly $\nu_b(n) \geq 0$ for all n, and it is easy to see that

$$\nu_b(n) \geq d^{-1} 2^{-d-6} \bar{\Lambda}_b(n) \tag{1.3.1}$$

for all $\varepsilon_1 N \leq n \leq \varepsilon_2 N$, for N sufficiently large. Indeed, this is trivial unless $Wn + b$ is a prime, and in that case, since $\varepsilon_1 N > R$, $\Lambda_R(Wn + b) = \log R \geq d^{-1} 2^{-d-5} \log N$. Note that the measure ν is in fact dependent on N, however following [7] we do not explicitly indicate that.

Let us briefly recall the pseudo-randomness properties of the measures ν_b - the so called linear form condition - which we will need in the proof. This is a slight modification of the formulation given in [7], however the proof works without any changes.

Theorem A (Linear forms 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 $m \le m_0$ and $t \le t_0$, suppose that $\{l_{i,j}\}_{1\le i\le m, 1\le j\le t}$ are arbitrary integers at most k_0 in absolute value, and that $\{b_i\}$ are arbitrary numbers relative 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

$$\mathbb{E}\left(\prod_{i=1}^{m} \nu_{b_i}(L_i(x));\ x \in \mathbb{Z}_N^t\right) = 1 + o_{N,W\to\infty;m_0,t_0,k_0}(1), \tag{1.3.2}$$

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

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 $o_{N,W\to\infty;m_0,t_0,k_0}(1)$ denotes a quantity that tends to 0 as *both* $N \to \infty$ *and* $W \to \infty$, for any fixed choice of m_0, t_0, k_0. In our context it is important to let $W = \prod_{p\le\omega} p$ be independent of N to obtain the quantitative lower bound in Theorem 1.2, see also the remarks in [7] (Sec.11). As all error terms in (1.3.2) are independent of the choice of b_i's, we will write ν for ν_{b_i} for simplicity of notations.

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

$$w(S) := \prod_{i=1}^{d} \prod_{y\in\pi_i(S)} \nu(y) \tag{1.3.3}$$

where $\pi_i(S)$ is the canonical projection of S to the i-th coordinate axis. If $S = \{x\}$ we will write $w(x) := w(\{x\}) = \prod_{i=1}^{d} \nu(x_i)$. The point is that if $Wx + b \in \mathbb{P}_N^d$ (and $x \in [\varepsilon_1 N, \varepsilon_2 N]^d$), then

$$w(x) \approx (\log N)^d. \tag{1.3.4}$$

The implicit constant depends only on d and W - which we will choose sufficiently large but independent of N. Moreover for $\Delta \subseteq [\varepsilon_1 N, \varepsilon_2 N]^d$ such that $W\Delta + b \subseteq A \subseteq \mathbb{P}_N^d$ one has

$$w(\Delta) \approx (\log N)^{l(\Delta)}. \tag{1.3.5}$$

Thus identifying $[1, N]$ with $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$ it is easy to show that (see Sec.5) Theorem 1.2 follows from

Theorem 1.3. *Let $\Delta = \{v_0, \cdots, v_d\} \subseteq \mathbb{Z}^d$ be a d-dimensional simplex and let $\delta > 0$. Let N be a large prime and let $A \subseteq \mathbb{Z}_N^d$ such that*

$$\mathbb{E}_{x\in\mathbb{Z}_N^d, t\in\mathbb{Z}_N}\left(\prod_{i=0}^{d} \mathbf{1}_A(x + tv_i)\right) w(x + t\Delta) \le \delta \tag{1.3.6}$$

then there exists $\epsilon = \epsilon(\delta)$ such that

$$\mathbb{E}_{x\in\mathbb{Z}_N^d} \mathbf{1}_A(x) w(x) \le \epsilon(\delta) + o_{N,W\to\infty;\,\Delta}(1)$$

Moreover $\epsilon(\delta) \to 0$ as $\delta \to 0$.

We describe below some of the key elements of the proof. The details are given in the remaining sections.

1.4. **A Removal Lemma for weighted hypergraph systems.** We will use the construction of a weighted hypergraph associated to a set $A \subseteq \mathbb{Z}_N^d$ and a simplex $\Delta = \{v_0, \ldots, v_d\}$ given in [18].

Definition 1.1 (Hypergraph System.)**.** *Let* $J = \{0, 1, \ldots, d\}, \overline{\mathcal{H}} := \{e : e \subseteq J\}$*, and 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* $x = (x_0, \ldots, x_d) \in V_J$ *such that* $x_j = 0$ *for all* $j \notin e$ *and let* $\pi_e : V_J \to V_e$ *denote the natural projection. For* $e = \{j\}$ *we write* $V_j := V_{\{j\}}$ *and for a given* $\mathcal{H} \subseteq \overline{\mathcal{H}}$*, we will call the quadruplet* $(J, V_J, \mathcal{H}, d)$ *a hypergraph system.*

From a graph theoretical point of view we can think of a point x_e ($e \in \mathcal{H}$, $|e| = d$), as a d-simplex with vertices $\{x_j : j \in e\}$. A set $G_e \subseteq V_e$ then may be viewed as a d-regular d-partite hypergraph with vertex sets V_j ($j \in e$). Similarly a point $x \in V_J$ represents a $d+1$-simplex with faces $x_e := (x_j)_{j \in e}$ for $e \in \mathcal{H}_d := \{e \subseteq J, |e| = d\}$.

For a given $e \subseteq J$ define the $\sigma-$algebra $\mathcal{A}_e = \{\pi_e^{-1}(F) : F \subseteq V_e\}$, which will play an important role in the proof of the removal lemma. For a given set $A \subseteq \mathbb{Z}_N^d$ and for $e = J \backslash \{j\}$, let

$$E_e = \{x \in V_J : \sum_{i=0}^{d} x_i(v_i - v_j)\} \in A \tag{1.4.1}$$

Note that $E_e \in \mathcal{A}_e$ as the expression in (1.4.1) is independent of the coordinate x_j.

Definition 1.2 (Weighted system)**.** *We will define now a family of functions* $\nu_e : V_J \to \mathbb{R}_+, \mu_e : V_J \to \mathbb{R}_+$*. For* $e \in \mathcal{H}_d, e = J \backslash \{j\}$ *and* $1 \leq k \leq d$*. Define*

$$L_e^k(x) = \sum_{i=0}^{d} x_i(v_i^k - v_j^k) \tag{1.4.2}$$

where v_i^k *denotes the* $k^{th}-$*coordinate of the vector* v_i*. 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 we define the support of a linear form $L(x) = \sum_{k=0}^{d} a_k x_k$ *as supp*$(L) = \{k : a_k \neq 0\}$*. For a given* $e \subseteq J$*, define*

$$\nu_e(x) = \prod_{L \in \mathcal{L}, supp(L) = e} \nu(L(x)) \ , \quad \mu_e(x) = \prod_{L \in \mathcal{L}, supp(L) \subseteq e} \nu(L(x)), \tag{1.4.3}$$

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

Note that if $\Delta = \{v_0, \cdots, v_d\}$ is in general position, that is if $v_i^k \neq v_j^k$ for all $i \neq j$ and k then 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 $\mu_e(x) = \prod_{f \subseteq e} \nu_f(x)$ and also $\mu_e(x) = \mu_e(\pi_e(x))$, that is μ_e is constant along the fibers of the projection π_e. We will refer the functions ν_e and μ_e as *weights* and *measures* respectively. To emphasize this point of view we will often use the integral notation and write

$$\int_{V_J} F(x)\, d\mu_e(x) := \mathbb{E}_{x \in V_J} F(x)\, \mu_e(x), \ \ and \ \int_{V_e} F_e(x)\, d\mu_e(x) := \mathbb{E}_{x \in V_e} F_e(x)\, \mu_e(x),$$

for functions $F : V_J \to \mathbb{R}$ and $F_e : V_e \to \mathbb{R}$. Thus we could think of μ_e as a measure on V_J or on the subspace V_e, the exact interpretation will be clear from the context. Note that it follows easily from the linear forms condition that $\mu_e(V_e) = \int_{V_e} 1\, d\mu_e = 1 + o_{N,W \to \infty}(1)$ (similarly $\mu_e(V_J) = 1 + o_{N,W \to \infty}(1)$), see Lemma 2.1.

Let us observe now some properties of the family of linear forms $\mathcal{L}$ which will play a crucial role in the proof. If $e = J\backslash\{j\}$, $e' = J\backslash\{j'\}$ then $\mathrm{supp}(L^k_{e'}) \subseteq e$ if and only if $v^k_j = v^k_{j'}$ and that is equivalent to $L^k_{e'} = L^k_e$. We call such a family $\mathcal{L}$ *well-defined*. Since for a given $e \in \mathcal{H}_d$, the forms $\{L^k_e, 1 \leq k \leq d\}$ are linearly independent any two distinct forms of the family $\mathcal{L}$ are linearly independent. We will refer to such families of forms as being *pairwise linearly independent*. Also let $M = \{x \in V_J : x_0 + \ldots + x_d = 0\}$. Then for any $x \in M$, $L^k_e(x) = L^k_{e'}(x)$ for all $e, e' \in \mathcal{H}_d$ and k. We call a family of linear forms $\mathcal{L} = \{L^k_e; e \in \mathcal{H}_d, 1 \leq k \leq s\}$ satisfying this property *symmetric*.

To see how the weighted hypergraph $\{\nu_e\}_{e\in\overline{\mathcal{H}}}$ is related to our problem we follow [18] to parameterize affine copies of Δ. Define the map $\Phi : \mathbb{Z}_N^{d+1} \to \mathbb{Z}_N^{d+1}$ by

$$\Phi(x) = (\sum_{i=0}^{d} x_i v_i, -\sum_{i=0}^{d} x_i) := (y, t) \tag{1.4.4}$$

By (1.4.1) and (1.4.4) we have that $x \in E_e$ for $e = J\backslash\{j\}$ if and only if $y + tv_j \in A$ thus $x \in \bigcap_{e\in\mathcal{H}_d} E_e$ exactly when $y + t\Delta \subseteq A$. Since Φ is one to one, as we assume $\{v_1 - v_0, \ldots, v_d - v_0\}$ is a linearly independent family of vectors, this gives a parametrization of all affine copies of Δ contained in A $(\mathrm{mod}\ N)$. Also for $e = J\backslash\{j\}$

$$L^k_e(x) = \sum_{i=0}^{d} x_i(v^k_i - v^k_j) = \pi_k(y + tv_j) \tag{1.4.5}$$

where π_k is the orthogonal projection to the k^{th} coordinate axis. This implies that

$$\mu_e(x) = \prod_{\mathrm{supp}(L)\subseteq e} \nu(L(x)) = \prod_{k=1}^{d} \nu(L^k_e(x)) = w(y + tv_j), \tag{1.4.6}$$

and also

$$\mu_J(x) = \prod_{L\in\mathcal{L}} \nu(L(x)) = w(y + t\Delta). \tag{1.4.7}$$

Thus the assumption (1.3.6) in Theorem 1.3 translates to

$$\mathbb{E}_{x\in V_J} \prod_{e\in\mathcal{H}_d} \mathbf{1}_{E_e}(x)\mu_J(x) = \mathbb{E}_{(y,t)\in\mathbb{Z}_N^{d+1}}\, w(y + t\Delta) \leq \delta. \tag{1.4.8}$$

On the other hand, recall $M = \{x \in V_J : x_0 + \cdots + x_d = 0\}$ then $x \in M \cap \bigcap_{e\in\mathcal{H}_d} E_e$ if and only if $\Phi(x) = (y, 0)$ with $y \in A$, thus by (1.4.4), (1.4.6)

$$\mathbb{E}_{y\in A} w(y) = \mathbb{E}_{x\in M} \prod_{e\in\mathcal{H}_d} \mathbf{1}_{E_e}(x)\mu_{e'}(x) \tag{1.4.9}$$

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 [18], [5], [14].

Theorem B. *(Simplex Removal Lemma)*[19]. *Let $E_e \in \mathcal{A}_e$ be given for $e \in \mathcal{H}_d$, and let $\delta > 0$. Also let μ_J and μ_e denote the normalized counting measures on V_J and V_e. There exists $\varepsilon = \varepsilon(\delta) > 0$ 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

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

then

$$\prod_{e\in\mathcal{H}_d} \mathbf{1}_{E'_e}(x_e) = 0 \quad \textit{for all} \quad x \in V_J,$$

$$\mathbb{E}_{x\in V_e} \mathbf{1}_{E_e\setminus E'_e}(x)\mu_e(x) \leq \epsilon(\delta),$$

and

$$\varepsilon(\delta) \to 0 \quad \textit{as} \quad \delta \to 0.$$

Naturally one would like to extend Theorem 1.4 to the family of measures $\{\mu_e\}_{e\in\mathcal{H}_d}$ as that would easily imply Theorem 1.3 and hence our main result Theorem 1.2. The reason why this seems difficult is the existence of weights ν_e on lower dimensional edges $|e| < d$ when the configuration Δ is not in general position. Removing these weights does not seem amenable to known " transference arguments" developed in [18], [1], [6], [15]. What we prove instead is that the removal lemma extends to a family of measures $\tilde{\mu}_e$ which are sufficiently small perturbations of the measures μ_e with respect to a given family of functions $g_e : V_e \to \mathbb{R}$. [1]

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 to a well-defined, pairwise linearly independent and symmetric family of linear forms $\mathcal{L}$ as defined in* (1.4.3)*. Let $E_e \subseteq \mathcal{A}_e, g_e : V_e \to [0,1]$ be given for $e \in \mathcal{H}_d$. Then for a given $\delta > 0$ there exists an $\epsilon = \epsilon(\delta) > 0$ such that the following holds: If*

$$\mathbb{E}_{x\in V_J} \prod_{e\in\mathcal{H}_d} \mathbf{1}_{E_e}(x)\mu_J(x) \leq \delta \tag{1.4.10}$$

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

$$\mathbb{E}_{x\in V_J} \prod_{e\in\mathcal{H}_d} \mathbf{1}_{E_e}(x)\tilde{\mu}_J(x) = \mathbb{E}_{x\in V_J} \prod_{e\in\mathcal{H}_d} \mathbf{1}_{E_e}(x)\mu_J(x) + o_{N,W\to\infty}(1) \tag{1.4.11}$$

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

$$\mathbb{E}_{x\in V_e}\ g_e(x)\tilde{\mu}_e(x) = \mathbb{E}_{x\in V_e}\ g_e(x)\mu_e(x) + o_{N,W\to\infty}(1) \tag{1.4.12}$$

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

$$\bigcap_{e\in\mathcal{H}_d} (E_e \cap E'_e) = \emptyset \tag{1.4.13}$$

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

$$\mathbb{E}_{x\in V_e} \mathbf{1}_{E_e\setminus E'_e}(x)\tilde{\mu}_e(x) \leq \epsilon(\delta) + o_{N,W\to\infty}(1) \tag{1.4.14}$$

moreover

$$\epsilon(\delta) \to 0, \ \textit{as} \ \delta \to 0. \tag{1.4.15}$$

[1] It seems possible to formulate the properties of weight system $\{\nu_e\}_{e\subseteq J}$ so that Theorem 1.4 holds without referring to an underlying system of linear forms $\mathcal{L}$. For that one would need to formulate a 'linear forms' condition for weighted hypergraphs similar to [18] at an order depending on δ. We will not pursue this approach here.

Proof [Theorem 1.4 implies Theorem 1.3]

By assumption (1.3.6) in Theorem 1.3 and by (1.4.7),

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

For a given $e' \in \mathcal{H}_d$ define the function $g_{e'} : V_{e'} \to [0,1]$ as follows. Let $\phi_{e'} : V_{e'} \to M$ be the inverse of the projection map $\pi_{e'} : V_J \to V_{e'}$ restricted to M, and for $y \in V_{e'}$ let

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

Applying Theorem 1.4 to the system of weights $\{\nu_e\}$ and functions $\{g_e\}$ gives a system of measures $\tilde{\mu}_e$ and sets $E'_e \in A_e$ satisfying (1.4.11)-(1.4.15). By (1.4.4) we have that $x \in M \cap \bigcap_{e\in\mathcal{H}_d} E_e$ 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 (1.4.6), thus for any given $e' \in \mathcal{H}_d$

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

By (1.4.13), $\prod_{e\in\mathcal{H}_d} \mathbf{1}_{E_e} \leq \sum_{e\in\mathcal{H}_d} \mathbf{1}_{E_e\setminus E'_e}$. Then the symmetry of the measures $\tilde{\mu}_e$ (i.e. the fact that $\tilde{\mu}_e(x) = \tilde{\mu}_{e'}(x)$ for $x \in M$), (1.4.14) and the fact that $\mathbf{1}_{E_e\setminus E'_e}$ is constant on the fibers $\pi_e^{-1}(x)$ implies

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

Choosing N, W sufficiently large with respect to δ gives

$$\mathbb{E}_{y\in\mathbb{Z}_N^d} \mathbf{1}_A(y) w(y) \leq \epsilon'(\delta),$$

with, say $\epsilon'(\delta) := (d+2)\epsilon(\delta)$. $\square$

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 [19],[5],[14], and in particular the notion of a regular or pseudo-random hypergraph. This can be defined in different ways, we use a variant of Gowers's 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. $\omega : e \to \{0,1\}$), define the orthogonal projection $\omega_e : V_e \times V_e \to V_e$ by

$$\omega_e(x_e, q_e)_i = \begin{cases} x_i & \text{if} \quad \omega_i = 0 \\ q_i & \text{if} \quad \omega_i = 1 \end{cases} \tag{1.5.1}$$

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

$$\|F\|_{\square_{\nu_e}}^{2^d} = \mathbb{E}_{x,q\in V_e} \prod_{\omega\in\{0,1\}^e} F(\omega_e(x,q)) \prod_{f\subseteq e} \prod_{\omega\in\{0,1\}^f} \nu_f(\omega_f(x_f,q_f)) \tag{1.5.2}$$

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 1. *Let $e = (0,1)$ and $F: V_0 \times V_1 \to \mathbb{R}$. Then*

$$\begin{aligned}\|F\|_{\square_{\nu_e}}^4 = \mathbb{E}_{x_0,q_0\in V_0,\, x_1,q_1\in V_1}\ & F(x_0,x_1)F(x_0,q_1)F(q_0,x_1)F(q_0,q_1) \\ & \times \nu_e(x_0,x_1)\nu_e(x_0,q_1)\nu_e(q_0,x_1)\nu_e(q_0,q_1)\,\nu_0(x_0)\nu_0(q_0)\nu_1(x_1)\nu_1(q_1).\end{aligned}$$

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 (1.5.2) 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, something which plays an important role in proving the Counting Lemma. More precisely one has the following.

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

$$\Big|\mathbb{E}_{x\in V_J} \prod_{e\in\mathcal{H}_d} F_e(\pi_e(x))\, \mu_J(x)\Big| \leq C \min_{e\in\mathcal{H}_d} \|F_e\|_{\square_{\nu_e}} + o_{N,W\to\infty}(1) \tag{1.5.3}$$

The $\square_\nu$-norm has also been defined and studied in [8] see Appendix B-C there, where various forms of von Neumann type inequalities has been shown. In fact it is not hard to adapt the arguments given there to prove Proposition 1.1, however as our settings is somewhat different we will include a proof in an appendix. The above inequality motivates the following

Definition 1.3. *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*

$$\|\mathbf{1}_{G_e} - \mu_e(G_e)\,\mathbf{1}_{V_e}\|_{\square_{\nu_e}} \leq \varepsilon. \tag{1.5.4}$$

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 $e \in \mathcal{H}_d$ (with a sufficiently small constant $c > 0$), then Theorem 1.4 holds with $\{\tilde{\mu}_e\} = \{\mu_e\}$. Indeed, writing $G_e = \pi_e(E_e)$, $1_{G_e} = \mu_e(G_e)\,1_{V_e} + F_e$, and substituting this decomposition into the left side of (1.4.10) we get $2^{d+1} - 1$ error terms each of which is bounded by $c'\varepsilon$ (for some small absolute constant $c' > 0$ as long as N and W is sufficiently large with respect to ε), and a main term of the form $\prod_{e\in\mathcal{H}_d} \mu_e(G_e)$ which by the assumption of Theorem 1.4 should be less than, say 2ε. This implies that $\mathbb{E}_{x\in V_e} 1_{G_e}(x)\mu_e(x) = \mu_e(G_e) \leq \delta$ for $\delta = (2\varepsilon)^{\frac{1}{d+1}}$, for at least one $e \in \mathcal{H}_d$. Thus the sets $E'_e := \emptyset$, $E'_{e'} := E_e$ ($e' \neq e$) satisfy the conclusion of Theorem 1.4.

Of course in general the hypergraphs $G_e = \pi_e(E_e)$ are not sufficiently regular, the bulk of our argument is to obtain a "Regularity Lemma" in our weighted setting. This roughly says that one can 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 based on 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 increasing families of weight systems $\{\nu_{q,e}\}_{e\in\bar{\mathcal{H}},q\in\Omega}$ which for most values of the parameter q will give rise to small perturbations of the initial weight system $\{\nu_e\}_{e\in\bar{\mathcal{H}}}$.

Let us sketch below how the weights $\nu_{q,e}$ and the associated measures $\mu_{q,e}$ arise in the special case $d=2$, $\nu_1\equiv\nu_2\equiv 1$. [2] Assume that there is an edge e, say $e=(1,2)$, so that the graph $G_e=\pi_e(E_e)$ is not ε-regular. This means

$$\|F\|_{\square_{\nu_e}}\geq\varepsilon, \tag{1.5.5}$$

where $F=\mathbf{1}_{G_e}-\mu_e(G_e)\,\mathbf{1}_{V_e}$. In view of definition (1.5.2), we may write

$$\|F\|^4_{\square_{\nu_e}}=\int_{V_e}\int_{V_e}F(x)\,u^1_q(x_1)u^2_q(x_2)\nu_e(x_1,q_2)\nu_e(q_1,x_2)\,d\mu_e(x)\,d\mu_e(q)\geq\varepsilon^4, \tag{1.5.6}$$

where $x=(x_1,x_2)$, $q=(q_1,q_2)$, $u^1_q(x_1)=F(x_1,q_2)$, and $u^2_q(x_2)=F(q_1,x_2)F(q_1,q_2)$. If one defines the measures $\mu_{q,e}$, depending on the parameter q, by

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

then the inner expression in (1.5.6) can be viewed as the inner product

$$\Gamma(q):=\left\langle F,u^1_q\cdot u^2_q\right\rangle_{\mu_{q,e}}=\int_{V_e}F(x)\,u^1_q(x_1)u^2_q(x_2)\,d\mu_{q,e}(x), \tag{1.5.7}$$

on the Hilbert space $L^2(V_e,\mu_{q,e})$. Thus (1.5.6) translates to $\mathbb{E}_{q\in V_e}\,\Gamma(q)\,\mu_e(q)\geq\varepsilon^4$ while using the linear forms condition it is easy to see that $\mathbb{E}_{q\in V_e}\,\Gamma(q)^2\,\mu_e(q)\lesssim 1$ thus

$$\Gamma(q)\gtrsim\varepsilon^4,\ \ \textit{for}\ \ q\in\Omega, \tag{1.5.8}$$

for a set $\Omega\subseteq V_e$ of measure $\mu_e(\Omega)\gtrsim\varepsilon^8$. As the functions u^i_q are bounded, hence without loss of generality we may assume that they are indicator functions of sets $U^i_q\subseteq V_i$. Let $\mathcal{B}_q=\mathcal{B}^1_q\vee\mathcal{B}^2_q$ denote the σ-algebra generated by the sets $\pi_i^{-1}(U^i_q)$ $(i=1,2)$ 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 $\mu_{q,e}$. Then, as $u^1_q\,u^2_q$ is measurable with respect to $\mathcal{B}_q$, we have

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

This together with (1.5.7) and (1.5.8) implies for $q\in\Omega$

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

where $\mathcal{B}_0=\{V_e,\emptyset\}$ is the trivial σ-algebra, and $\mathbb{E}_{\mu_e}(\mathbf{1}_{G_e}|\mathcal{B}_0)=\mu_e(G_e)\,\mathbf{1}_{V_e}$. Then by the Cauchy-Schwartz inequality, we have

$$\|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{\mathcal{A}_e}|\mathcal{B}_q)-\mathbb{E}_{\mu_e}(\mathbf{1}_{\mathcal{A}_e}|\mathcal{B}_0)\|^2_{L^2(\mu_{q,e})}\gtrsim\varepsilon^8. \tag{1.5.9}$$

Note that 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"

$$\|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{\mathcal{A}_e}|\mathcal{B}_q)-\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{\mathcal{A}_e}|\mathcal{B}_0)\|^2_{L^2(\mu_{q,e})}=\|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{\mathcal{A}_e}|\mathcal{B}_q)\|^2_{L^2(\mu_{q,e})}-\|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{\mathcal{A}_e}|\mathcal{B}_0)\|^2_{L^2(\mu_{q,e})}\gtrsim\varepsilon^8.$$

To overcome this "discrepancy", using the linear forms condition, we show that for given $B\subseteq V_e$ one has for almost every $q\in V_e$

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

This in turn implies that

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

and

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

[2] Though our exposition later is self-contained, some familiarity with standard notions and arguments, such the conditional expectation, energy increment, discussed for example in [19], may be helpful here.

Then from (1.5.9) we have for almost every $q \in \Omega$, that

$$\|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_e}|\mathcal{B}_q)\|^2_{L^2(\mu_{q,e})} \geq \|\mathbb{E}_{\mu_e}(\mathbf{1}_{G_e}|\mathcal{B}_0)\|^2_{L^2(\mu_e)} + c\,\varepsilon^8. \tag{1.5.10}$$

If $F : V \to \mathbb{R}$ is a function and $(V, \mathcal{B}, \mu)$ is a measure space, the quantity $\|\mathbb{E}_\mu(F|\mathcal{B})\|^2_{L^2(\mu)}$ is sometimes referred to as the "energy" of the function F with respect to the measure space $(V, \mathcal{B}, \mu)$, so (1.5.10) is telling that if G_e is not ε-uniform with respect to the initial measure spaces $(V_e, \mathcal{B}_0, \mu_e)$ then its energy increases by a fixed amount when passing to the measure spaces $(V_e, \mathcal{B}_q, \mu_{q,e})$ for (almost) every $q \in \Omega$. One can iterate this argument to arrive to a family of measure spaces $(V_e, \mathcal{B}_{q,e}, \mu_{q,e})_{e\in\mathcal{H}_d,\, q\in\Omega}$ such that the atoms $G_{q,e} \in \mathcal{B}_{q,e}$ become sufficiently uniform, thus obtaining a parametric version of the so-called Koopman-von Neumann decomposition, see [19]. This can be further iterated to eventually obtain a regularity lemma. Note that the number of linear forms defining the measures $\mu_{q,e}$ is increasing at each step of the iteration, causing the linear forms condition to be used at a level depending eventually on the relative density of the set A and not just on the dimension d.

1.6. **Outline of the paper.** In Section 2 we describe the type of parametric weight systems $\{\nu_{q,f}\}_{f\in\mathcal{H},\, q\in Z}$ that we encounter later on. Here we also discuss their basic properties such as stability and symmetry. In Section 3 we introduce the energy increment argument for parametric systems, and prove a regularity lemma. Section 4 is devoted to proving the counting and removal lemmas. Many of our arguments in Section 3 and Section 4 may be viewed as an extension 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 vector type, although we do not emphasize this. We think of the initial data $\Delta = \{v_0, \ldots, v_d\}$ 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 $Y \leq C\,X$ for some constant $C > 0$ depending only on the vectors v_i or the dimension d. If $y_1, \ldots, y_s$ and X additional parameters we write $O_{y_1,\ldots,y_s}(X)$ for a quantity Y bounded by $C(y_1, \ldots, y_s)X$ or equivalently $Y \lesssim_{y_1,\ldots,y_s} X$.

Though most of our constructions depend on N and W, we will not indicate that to emphasize that all other terms in our estimates are independent of N as well as the parameters appearing in them. We'll utilize the linear forms condition throughout the paper, giving rise to error terms which tends to 0 as both $N \to \infty$ and $W \to \infty$ for any fixed choice of the parameters $y_1, \ldots, y_s$ on which they may depend. The standard notation for such terms would be $o_{N,W\to\infty;y_1,\ldots,y_s}(1)$, for simplicity we will write $o_{y_1,\ldots,y_s}(1)$. Finally as all estimates in the linear forms 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 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.

2.1. **Parametric weight systems and stability properties.** Recall the family of measures $\{\mu_e\}_{e\in\mathcal{H}}$ constructed in (1.4.1)

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

where the family $\mathcal{L}$ defined in (1.4.1) consists of pairwise linearly independent forms. The following statement is based on the linear forms 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*

$$\mu_e(V_e) = 1 + o(1), \tag{2.1.1}$$

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

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

or equivalently

$$\int_{V_e} g\, d\mu_e = \int_{V_J} (g \circ \pi_e)\, d\mu_J + o(1). \tag{2.1.2}$$

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

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

are pairwise linearly independent, and as they are supported on e they remain pairwise independent when restricted to V_e. Thus (2.1.1) follows from the linear forms condition.
To show (2.1.2), let $e' = J \backslash e$ and write $x = (x_e, x_{e'})$ with $x_e = \pi_e(x),\ x_{e'} = \pi_{e'}(x)$. Then

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

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

By (2.1.1) we have that $\mu_e(V_e) \lesssim 1$, and then by the Cauchy-Schwartz inequality

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

The right hand side of this expression is a combination of four terms and (2.1.2) 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 $\mathrm{w}(x_e, x_{e'})$ depend also on some of the variables in $x_j,\ j \in e'$, while the ones in $\mathrm{w}(x_e, y_{e'})$ depend on the variables in $y_j,\ j \in e'$, hence these forms depend on different sets of variables. Thus the forms appearing in the expression $\mu_e(x_e)\mathrm{w}(x_e, x_{e'})\mathrm{w}(x_e, y_{e'})$ are pairwise linearly independent and (2.1.2) follows from the linear forms condition. Note that the estimate is independent on the function g. □

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 measure only by a negligible amount

$$\mu_J(\overline{G}_e) = \mu_e(G_e) + o(1) \tag{2.1.3}$$

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

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

be a family of linear forms with integer coefficients depending on the parameters $q \in \mathbb{Z}^R$ and the variables $x \in \mathbb{Z}^D$. 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 measure on Z given by

$$\int_Z g(q)\, d\psi(q) = \mathbb{E}_{q \in Z}\, g(q)\, \psi(q), \quad \psi(q) = \prod_{i=1}^{t} \nu(Y_i(q)), \tag{2.1.4}$$

for a family of pairwise linearly independent linear forms Y_i defined over Z, and if 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 $supp_x(L) = \{j \in J;\ a_j \neq 0\}$. For $e \subseteq J$ and $q \in Z$, let

$$\nu_{q,e}(x) := \prod_{\substack{L \in \mathcal{L}_q \\ supp_x(L)=e}} \nu(L(q,x)), \qquad \mu_{q,e}(x) := \prod_{\substack{L \in \mathcal{L}_q \\ supp_x(L)\subseteq e}} \nu(L(q,x)) \tag{2.1.5}$$

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

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

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 $q \in Z$. In order 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 $L^j(q,x)$, $Y_l(q)$, 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 forms condition. We have the 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 $q \notin \mathcal{E}_e$*

$$\mu_{q,e}(V_e) = 1 + o_K(1). \tag{2.1.6}$$

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 the following holds. For any function $g : Z \times V_e \to [-1,1]$ and for every $q \notin \mathcal{E}_e$ one has the estimate

$$\int_{V_e} g(q,x_e)\, d\mu_{q,e}(x_e) = \int_{V_J} g(q,\pi_e(x))\, d\mu_{q,J}(x) + o_K(1). \tag{2.1.7}$$

Proof. To prove (2.1.6) consider the quantity

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

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 forms condition gives that each term is $1 + o_K(1)$ and so $\Lambda_e = o_K(1)$ and (2.1.6) follows.

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

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

where $\mathrm{w}_q(x_e,x_{e'}) = \prod_{f\nsubseteq e} \nu_{q,f}(x_{e\cap f}, x_{e'\cap f})$.

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 (2.1.7) would follow from the estimate $\mathbb{E}_{q\in Z}\ \Lambda(q,e)\, d\psi(q) = o_K(1)$. By the linear forms condition $\mathbb{E}_{q,x_e}\, d\psi(q)\, d\mu_{q,e}(x_e) = 1 + o_K(1) \le 2$, for N sufficiently large with respect to K. Then by the Cauchy-Schwartz inequality one has

$$(\mathbb{E}_{q\in Z}\ \Lambda(q,e)\, d\psi(q))^2 \lesssim \mathbb{E}_{q\in Z,\, x_e\in V_e}\ \mathbb{E}_{x_{e'},y_{e'}\in V_e}\ (\mathrm{w}_q(x_e,x_{e'})-1)(\mathrm{w}_q(x_e,y_{e'})-1)\, d\mu_{q,e}(x_e)\, d\psi(q).$$

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 $\mathrm{w}_q(x_e, x_{e'})$ (respectively, $\mathrm{w}_q(x_e, y_{e'})$) depend on the $x_{e'}$ (respectively, $y_{e'}$) 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. □

2.2. **Extension 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 $\mathcal{L}^1_{q_1} = \{L^1_1(q_1,x),...,L^{s_1}_1(q_1,x)\}$ and $\mathcal{L}^2_{q_2} = \{L^1_2(q_2,x),...,L^{s_2}_2(q_2,x)\}$ be two pairwise linearly indpendent families of linear forms defined on the parameter spaces $Z_1 = \mathbb{Z}_N^{k_1}$ and $Z_2 = \mathbb{Z}_N^{k_2}$. Let ψ_1 and ψ_2 be measures on Z_1 and Z_2 defined by the families of linear forms $\{Y^1_1(q_1),\dots Y^1_{s_1}(q_1)\}$ and $\{Y^2_1(q_2),\dots Y^2_{s_2}(q_2)\}$.

Definition 2.1. *We say that the family $\mathcal{L}^2_{q_2}$ is an extension of the family $\mathcal{L}^1_{q_1}$ if $Z_1 \le Z_2$ and the following holds. The family of forms $L^i_2(q_2,x), Y^2_j(q_2)$ which depend only on the variables $q_1 = \pi(q_2)$ is exactly the family of forms $L^i_1(q_1,x)$, $Y^1_j(q_1)$, where $\pi : Z_2 \to Z_1$ 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 (2.1.5). We say that the measure system μ^2 is an *extension* of the system μ^1.

Let us make a few immediate observations. Writing $Z_2 = Z_1 \times Z$, $Z = \mathbb{Z}_N^r$ and $q_2 = (q_1, q)$, we have

$$\psi_2(q_1,q) = \psi_1(q_1)\cdot\varphi(q_1,q) \tag{2.2.1}$$

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

$$\mu^2_{(q_1,q),e}(x_e) = \mu^1_{q_1,e}(x_e)\mathrm{w}_e(q_1,q,x_e) \tag{2.2.2}$$

where the linear forms $L^j_2(q_1,q,x_e)$ defining the function $\mathrm{w}_e(q,q_1,x_e)$ depend on (some of) the variables q as well as on (all of) the variables x_e.
In the special case when $\mathcal{L} = (L^1(x),..,L^s(x))$ 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 $g : V_f \to [-1,1]$ there is a set $\mathcal{E}_{g,f} \subseteq Z$ of measure $\psi(\mathcal{E}_{g,f}) = o_K(1)$, so that for all $q \notin \mathcal{E}_{g,f}$*

$$\int_{V_f} g\, d\mu_{q,f} - \int_{V_f} g\, d\mu_f = o_K(1). \tag{2.2.3}$$

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 $g : Z_1\times V_f \to [-1,1]$ 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 $q_2 = (q_1,q)\notin\mathcal{E}_{g,f}$

$$\int_{V_f} g(q_1,x)\, d\mu_{q_2,f}(x) - \int_{V_f} g(q_1,x)\, d\mu_{q_1,f}(x) = o_{K_2}(1). \tag{2.2.4}$$

Proof. As $\mu_{q,f} = \mu_f(x_f)\mathrm{w}_f(q,x_f)$, the left side of (2.2.3) may be written as

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

Consider

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

Using the Cauchy-Schwartz inequality we estimate

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

Now the Cauchy-Schwartz inequality and (2.1.1) gives

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

This last expression is a combination of 16 terms where each term is $1+o_K(1)$ by the linear form conditions. Indeed the linear forms which can appear in any of these terms are $Y_{i_1}(q)$,$Y_{i_2}(p)$,$L^{i_3}(x)$,$L^{i_4}(y)$, $L^{i_5}(q,x)$, $L^{i_6}(q,y)$, $L^{i_7}(p,x)$, $L^{i_8}(p,y)$. Note that the last 4 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 are pairwise linearly independent in the (q,p,x,y) variables. This Proves (2.2.3).

The proof of (2.2.4) 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 (2.2.1) we estimate as above

$$\Lambda_{f,g} \lesssim \int_{V_f}\int_{V_f}\int_{Z_1} d\psi_1(q_1)d\mu_{q_1,f}(x)d\mu_{q_1,f}(y)\ \left|\mathbb{E}_{q\in Z}\,(\mathrm{w}_f(q_1,q,x)-1)(\mathrm{w}_f(q_1,q,y)-1)\varphi(q_1,q)\right|.$$

The linear forms condition gives

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

so then we have

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

The point is that any linear form $L_f^i(q_1,q,x)$ depends both on 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 forms 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. □

Lemma 2.3 is an example of what we refer to as a *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 (1.4.2), 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 $g : Z_1 \times V_e \to \mathbb{R}$ be a function and let $e \in \mathcal{H}$, $|e| = d'$. For a given $q_1 \in Z_1$ recall the box norm of $g_{q_1}(x) = g(q_1, x)$

$$\|g_{q_1}\|_{\square_{\nu_{q_1,e}}}^{2^{d'}} = \mathbb{E}_{p,x\in V_e} \prod_{\omega_e\in\{0,1\}^e} g(q_1, \omega_e(p,x)) \prod_{f\subseteq e} \prod_{\omega_f\in\{0,1\}^f} \nu_{q,f}(\omega_f(p_f, x_f)), \tag{2.2.5}$$

where $x_f = \pi_f(x)$, $p_f = \pi_f(p)$, $\pi_f : V_e \to V_f$ being the natural projection. The inner product on the right side of (2.2.5) is defined by the parametric family of forms

$$\tilde{\mathcal{L}}_{q_1} = \bigcup_{f\subseteq e} \{L(q_1, \omega_f(p_f, x_f));\ L \in \mathcal{L}_{q_1},\ supp_x(L) = f,\ \omega_f \in \{0,1\}^f\}. \tag{2.2.6}$$

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'd have that

$$L'(q_1, \omega'_{f'}(p_{f'}, x_{f'})) = \lambda L(q_1, \omega_f(p_f, x_f)), \tag{2.2.7}$$

then restriction both forms to the subspace $\{p = x\}$ would imply that $L'(q_1, x_{f'}) = \lambda L(q_1, x_f)$ and hence $f' = supp_x(L') = supp_x(L) = f$. Then, as L and L' depend exactly variables x_j for $j \in f$, for (2.2.7) to hold, we should have $\omega'_f = \omega_f$ and $L = L'$.
If $\{\tilde{\mu}_{q_1,f}\}_{q\in Z_1, f\subseteq e}$ denotes the associated system of measures and

$$G(q_1, p, x) := \prod_{\omega\in\{0,1\}^e} g(q_1, \omega_e(p,x)), \tag{2.2.8}$$

then for given $q_1 \in Z_1$

$$\|g_{q_1}\|_{\square_{\nu_{q_1,e}}}^{2^{d'}} = \mathbb{E}_{p,x\in V_e} G_{q_1}(p,x)\, \tilde{\mu}_{q_1,e}(p,x). \tag{2.2.9}$$

Now, if $\mathcal{L}_{q_2}$ is a well-defined parametric extension of $\mathcal{L}_{q_1}$ then (2.2.6) yields to 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 $g : Z_1 \times V_e \to [-1,1]$ 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 $q_2 = (q_1, p) \notin \mathcal{E}$*

$$\|g_{q_1}\|_{\square_{\nu_{q_2,e}}} = \|g_{q_1}\|_{\square_{\nu_{q_1,e}}} + o_{K_2}(1). \tag{2.2.10}$$

Let $(V, \mathcal{B}, \mu)$ be a measure space and let $g : V \to \mathbb{R}$ 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 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 $\sigma-$algebra on V_e such that compl($\mathcal{B}_{q_1,e}$)$\leq M$ for some fixed number M. For any function $g : Z_1 \times V_e \to [-1,1]$ 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*

$$\left\|\mathbb{E}_{\mu_{q_2,e}}(g_{q_1}|\mathcal{B}_{q_1,e}) - \mathbb{E}_{\mu_{q_1,e}}(g_{q_1}|\mathcal{B}_{q_1,e})\right\|^2_{L^2(\mu_{q_2,e})} = o_{M,K_2}(1) \tag{2.2.11}$$

(2) *and*

$$\left\|\mathbb{E}_{\mu_{q_2,e}}(g_{q_1}|\mathcal{B}_{q_1,e})\right\|^2_{L^2(\mu_{q_2,e})} = \left\|\mathbb{E}_{\mu_{q_1,e}}(g_{q_1}|\mathcal{B}_{q_1,e})\right\|^2_{L^2(\mu_{q_1,e})} + o_{M,K_2}(1). \tag{2.2.12}$$

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

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

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

$$\mu_i(q_2, g) := \int_{V_e} g(q_1, x) b_i(q_1, x) d\mu_{q_2,e}(x), \qquad \mu_i(q_2) := \mu_{q_2,e}(B^i_{q_1}),$$

$$\mu_i(q_1, g) := \int_{V_e} g(q_1, x) b_i(q_1, x) d\mu_{q_1,e}(x), \qquad \mu_i(q_1) := \mu_{q_1,e}(B^i_{q_1})$$

By Lemma 2.3 we have that

$$\mu_i(q_2, g) = \mu_i(q_1, g) + o_{K_2}(1), \quad \mu_i(q_2) = \mu_i(q_1) + o_{K_2}(1) \tag{2.2.13}$$

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 (2.2.11) takes the form

$$\sum_{i=1}^m \left(\frac{\mu_i(q_2, g)}{\mu_i(q_2)} - \frac{\mu_i(q_1, g)}{\mu_i(q_1)} \right)^2 \mu_i(q_2), \tag{2.2.14}$$

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 (2.2.13)

$$\varepsilon := \sum_{i=1}^m \Big(|\mu_i(q_2, g) - \mu_i(q_1, g)| + |\mu_i(q_2) - \mu_i(q_1)| \Big) = o_{K_2,M}(1) \tag{2.2.15}$$

Now if $\mu_i(q_1) \leq 2\varepsilon^{1/4}$ then $\mu_i(q_2) \leq 3\varepsilon^{1/4}$ by (2.2.13), 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 $\mu_i(q_2) \geq \varepsilon^{1/4}$, 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 (2.2.11). The proof of inequality (2.2.12) proceeds the same way, here one needs to estimate the quantity

$$\sum_{i=1}^{m}\left|\frac{\mu_i(q_2,g)^2}{\mu_i(q_2)}-\frac{\mu_i(q_1,g)^2}{\mu_i(q_1)}\right|=\sum_{i=1}^{m}\left|\left(\frac{\mu_i(q_2,g)}{\mu_i(q_2)}\right)^2\mu_i(q_2)-\left(\frac{\mu_i(q_1,g)}{\mu_i(q_1)}\right)^2\mu_i(q_1)\right| \tag{2.2.16}$$

If $\mu_i(q_1)\le 2\varepsilon^{1/4}$ then $\mu_i(q_2)\le 3\varepsilon^{1/4}$ for $q_2=(q_1,q)\notin\mathcal{E}$, thus the contribution of such terms to the right side of (2.2.16) 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 (2.2.12) follows. □

We also need an analogue of the above result when the $\|\cdot\|_{L^2(\mu_{q,e})}$ norm is 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*

$$\|\mathbb{E}_{\nu_{q_2,e}}(g_{q_1}|\mathcal{B}_{q_1,e})-\mathbb{E}_{\nu_{q_1,e}}(g_{q_1}|\mathcal{B}_{q_1,e})\|_{\square_{\nu_{q_2,e}}}=o_{M,K}(1), \tag{2.2.17}$$

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

$$\|\mathbf{1}_{A_{q_1}}\|^{2^d}_{\square_{\nu_{q_2,e}}}\le\mu_{q_2,e}(A_{q_1})+o_{K_2}(1). \tag{2.2.18}$$

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

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

with

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

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 2.3, 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 (2.2.18) follows.

Now let $\{B^i_{q_1}\}_{i=1}^m$ $(m=2^M)$ 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 (2.2.11) is then estimated

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

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

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

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

2.3. **Symmetric extensions.** We will also need our parametric families of forms to be symmetric, to apply Theorem 1.3, which we define as follows. Let for each $e \in \mathcal{H}_d$, $\mathcal{L}_{q,e} = \{L^1_e(q,x), ..., L^s_e(q,x)\}$ be a pairwise linearly independent family of linear forms defined on $V = V_J$, depending on parameters $q \in Z$, such that $supp_x(L^j_e) \subseteq e$. We say that the family of forms $\mathcal{L}_q = \bigcup_{e\in\mathcal{H}_d} \mathcal{L}_{q,e}$ is *symmetric* if $L^j_e(q,x) = L^j_{e'}(q,x)$ for all $q \in Z$, $x \in M = \{x : x_0 + \cdots + x_d = 0\}$, e, $e' \in \mathcal{H}_d$ and $1 \le j \le s$. Note that that our initial family of forms defined in (1.4.2) has this property.

It is not hard to see that to a given family of forms $\mathcal{L}_{q,e}$, for a fixed $e \in \mathcal{H}_d$, there is a unique symmetric family of forms $\mathcal{L}_q$ such that $\mathcal{L}_{q,e} = \{L \in \mathcal{L}_q;\ supp_x(L) \subseteq e\}$. Indeed, if $\mathcal{L}_q$ is such a family, then for $e' \in \mathcal{H}_d$, $q \in Z$ and $x \in V_J$

$$L^j_{e'}(q,x) = L^j_{e'}(q, \pi_{e'}(x)) = L^j_{e'}(q, \phi_{e'} \circ \pi_{e'}(x)) = L^j_e(q, \phi_{e'} \circ \pi_{e'}(x)), \tag{2.3.1}$$

where $\phi_e : V_e \to M$ is the inverse of the projection $\pi_{e'}$ restricted to M. This shows the uniqueness of the family $\mathcal{L}_q$. Conversely, define $L^j_{e'}(q,x)$ by the above equality, then it is clear that $supp_x(L^j_{e'}) \subseteq e'$, moreover if $x \in M$ then $x = \phi_{e'} \circ \pi_{e'}(x)$ hence $L^j_{e'}(q,x) = L^j_e(q,x)$. Also, if $supp_x L^j_{e'} \subseteq e$ then for all $q \in Z_1$ and $x \in V_J$

$$L^j_{e'}(q,x) = L^j_{e'}(q, \phi_e \circ \pi_e(x)) = L^j_e(q, \phi_e \circ \pi_e(x)) = L^j_e(q,x),$$

This shows that all 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 of forms $\mathcal{L}_q$ as the *symmetrization* of the family $\mathcal{L}_{q,e}$. If $f \in \mathcal{H}$ for some edge $|f| = d' \le d$ 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 $L^j_e = L^j_{e'}$ for all $1 \le j \le s$.

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 (2.2.6), that is replacing the forms $L^j(q, x_f)$ with the forms $L^j(q, \omega_f(p_f, x_f))$, $\omega_f \in \{0,1\}^f$. 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. **A Koopman-von Neumann type 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 $\pi^{-1}_{ef}(\mathcal{B}_f)$ where $\pi_{ef} :\ V_e \to V_f$ is the canonical projection. The atoms of $\mathcal{B}$ are the sets $G = \bigcap_{f\subseteq\partial e} \pi^{-1}_{ef}(G_f)$ with G_f being an atom of $\mathcal{B}_f$, which may be interpreted as the collection of simplices $x \in V_e$ 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

$$\big\|\mathbf{1}_{G_e} - \mathbb{E}(\mathbf{1}_{G_e}|\bigvee_{f\in\partial e}\mathcal{B}_f)\big\|_\square \geq \eta, \tag{3.1.1}$$

then there exist σ-algebras $\mathcal{B}'_f \supseteq \mathcal{B}_f$ for $f \in \partial e$, such that

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

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 $compl(\mathcal{B}'_f)$ 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 to a well-defined, pairwise linearly independent family of of forms $\mathcal{L}_q$ defined on $Z \times V_e$, as given in (2.1.5). 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

$$\big\|\mathbf{1}_{G_{q,e}} - \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_{q,e}}|\bigvee_{f\in\partial e}\mathcal{B}_{q,f})\big\|_{\square_{\nu_{q,e}}}^{2^{d'}} \geq \eta, \tag{3.1.2}$$

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 $\Omega' \subseteq \Omega \times V_e \subseteq Z' = Z \times V_e$ such that the following hold.

(1) *We have*

$$\psi'(\Omega') \geq 2^{-4}c_0^2\eta^2, \tag{3.1.3}$$

where ψ' is the measure on the parameter space Z'.

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

$$compl(\mathcal{B}_{q',f}) \leq compl(\mathcal{B}_{q,f}) + 1. \tag{3.1.4}$$

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

$$\big\|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q,e}}|\bigvee_{f\in\partial e}\mathcal{B}_{q',f})\big\|_{L^2(\mu_{q',e})}^2 \geq \big\|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_{q,e}}|\bigvee_{f\in\partial e}\mathcal{B}_{q,f})\big\|_{L^2(\mu_{q,e})}^2 + 2^{-8}\eta^2, \tag{3.1.5}$$

(4) *and*

$$\mu_{q',e}(V_e) \leq 2. \tag{3.1.6}$$

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

$$g_q := \mathbf{1}_{G_{q,e}} - \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_{q,e}} | \bigvee_{f\in\partial e} \mathcal{B}_{q,f}). \tag{3.1.7}$$

Then by (2.2.5) we have for each $q \in \Omega$

$$\|g_q\|_{\Box_{\nu_{q,e}}}^{2^{d'}} = \int_{V_e} \langle g_q, \prod_{f\in\partial e} u_{q,p,f}\rangle_{\mu_{(q,p),e}} d\mu_{q,e}(p) \geq \eta, \tag{3.1.8}$$

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

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

As explained after (2.1.5) 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 (3.1.8) that the measure ψ' on Z' has the form $\psi'(q,p) = \mu_{q,e}(p)\,\psi(q)$.

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

$$\Gamma(q,p) := \langle g_q, \prod_{f\in\partial e} u_{q,p,f}\rangle_{\mu_{q,p,f}} \tag{3.1.9}$$

We show that there is a set $\Omega_1' \subseteq \Omega \times V_e$ of measure

$$\psi'(\Omega_1') \geq 2^{-3} c_0^2\, \eta^2, \tag{3.1.10}$$

such that for every $(q,p) \in \Omega_1'$ one has

$$\Gamma(q,p) \geq \frac{\eta}{4}. \tag{3.1.11}$$

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 $q \in \Omega \backslash \mathcal{E}_1 = \Omega_1$ we have by (3.1.8) that

$$\int_{V_e} \mathbf{1}_{\{\Gamma(q,p)\geq\eta/4\}} \Gamma(q,p) d\mu_{q,e}(p) \geq \frac{\eta}{2}, \tag{3.1.12}$$

where by (3.1.8) and (3.1.9) we have

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

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

$$\int_Z \int_{V_e} |\Gamma(q,p)|^2 d\mu_{q,e}(p) d\psi(q) \leq \int_Z \int_{V_e} \int_{V_e} \int_{V_e} \mathrm{w}_{q,p}(x)\mathrm{w}_{q,p}(x') d\mu_{q,e}(x) d\mu_{q,e}(x') d\mu_{q,e}(p) d\psi(q) \tag{3.1.13}$$

$$= 1 + o_K(1) \leq 2$$

by the linear forms 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 (3.1.12), (3.1.13) and the Cauchy-Schwartz inequality

$$\frac{c_0^2 \eta^2}{4} \leq \left(\int_{\Omega_1'} \Gamma(q,p) d\mu_{q,e}(p)\, d\psi(q)\right)^2 \leq \int_{\Omega_1'} \Gamma(q,p)^2 d\mu_{q,e}(p)\, d\psi(q)\ \psi'(\Omega_1') \leq 2\,\psi'(\Omega_1').$$

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

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

$$\langle g_q, \prod_{f\in\partial e} v_{q',f}\rangle_{\mu_{q',e}} \geq 2^{-2}\eta \tag{3.1.14}$$

for some functions $v_{q',f} : V_f \to [0,1]$. 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 (3.1.14) we have

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

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 simplicity of notation. For $q' = (q,p) \in \Omega'_1$, define $\mathcal{B}_{q',f}$ to be the $\sigma-$algebra generated by $\mathcal{B}_{q,f}$, and the $\mathcal{U}_{q',t_f}$. 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 $\sigma-$algebra $\bigvee_{f\in\partial e} \mathcal{B}_{q',f}$, and therefore we have for $q' \in \Omega'_1$

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

for $q' \in \Omega'_1$. Hence, by (3.1.7) and (3.1.14) it follows that

$$\langle \mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}_{q',f}) - \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}_{q,f}), \prod_{f\in\partial e} \mathbf{1}_{\mathcal{U}_{q',f}}\rangle_{\mu_{q',e}} \geq 2^{-2}\eta \tag{3.1.15}$$

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

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

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

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

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 $q' = (q,p) \in \Omega'_3 := \Omega'_2\backslash\mathcal{E}_2$ we have

$$\big\|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}_{q',f}) - \mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}_{q,f})\big\|_{L^2(\mu_{q',e})} \geq 2^{-3}\eta - o_{K,M}(1) \geq 2^{-4}\eta. \tag{3.1.16}$$

Since $\mathcal{B}_{q,f} \subseteq \mathcal{B}_{q',f}$, for $q' = (q,p)$, (3.1.16) is equivalent to

$$\big\|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}_{q',f})\big\|^2_{L^2(\mu_{q',e})} - \big\|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}_{q,f})\big\|^2_{L^2(\mu_{q',e})} \geq 2^{-8}\eta^2. \tag{3.1.17}$$

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 $q' \in \Omega'_4 := \Omega'_3\backslash\mathcal{E}_3$ we have (for N, W sufficiently large)

$$\|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q,e}}|\bigvee_{f\in\partial e}\mathcal{B}_{q',f})\|^2_{L^2(\mu_{q',e})}-\|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_{q,e}}|\bigvee_{f\in\partial e}\mathcal{B}_{q,f})\|^2_{L^2(\mu_{q,e})}\geq 2^{-9}\eta^2. \tag{3.1.18}$$

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

Iterating the above lemma leads to a parametric family of $\sigma-$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 a 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

$$\sum_{e\in\mathcal{H}_{d'},G_e\in\mathcal{B}_e}\|\mathbb{E}_{\mu_e}(\mathbf{1}_{G_e}|\bigvee_{f\in\partial e}\mathcal{B}_f)\|^2_{L^2(\mu_e)}. \tag{3.1.19}$$

Assuming 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 (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 hold.

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

$$\mathcal{B}_{q,f}\subseteq\mathcal{B}_{q',f},\quad \textit{compl}(\mathcal{B}_{q',f})\leq \textit{compl}(\mathcal{B}_{q,f})+O_{M_{d'},\delta}(1). \tag{3.1.20}$$

(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*

$$\|\mathbf{1}_{G_{q,e}}-\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q,e}}|\bigvee_{f\in\partial e}\mathcal{B}_{q',f})\|_{\square_{\nu_{q',e}}}\leq\delta. \tag{3.1.21}$$

and

$$\|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q,e}}|\bigvee_{f\in\partial e}\mathcal{B}_{q,f})\|^2_{L^2(\mu_{q',e})}=\|\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{G_{q,e}}|\bigvee_{f\in\partial e}\mathcal{B}_{q,f})\|^2_{L^2(\mu_{q,e})}+o_{M_{d'},K,\delta}(1), \tag{3.1.22}$$

Proof. Initially set $Z'=Z$, then (3.1.20) and (3.1.22) trivially holds 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 (3.1.21) 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 (3.1.21) 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 (3.1.3)-(3.1.5) 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 section 2.3, and set $\mathcal{B}_{q',f}:=\mathcal{B}_{q,f}$ for

$q' \notin \Omega'$ or $f \nsubseteq e$. 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 $q' \in \Omega' \backslash \mathcal{E}$, (3.1.20) and (3.1.22) 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 (3.1.20) and (3.1.22) 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 (3.1.21) 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 (3.1.20) and (3.1.22) hold, and total energy of the system $(\mu_{q'',e}, \mathcal{B}_{q,e}, \mathcal{B}_{q'',f})$ is at least $2^{-10}\delta^{2^{d'}}$ larger than that of the system $(\mu_{q',e}, \mathcal{B}_{q,e}, \mathcal{B}_{q',f})$. Set $Z' := Z$, $\mu_{q',e} := \mu_{q'',e}$ and $\mathcal{B}_{q',f} := \mathcal{B}_{q'',f}$.

By (3.1.19) the iteration process must stop in $O_{M_{d'},\delta}(1)$ steps and the system obtained satisfies (3.1.20)-(3.1.22).

□

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 with an iteration process using Lemma 3.2 repeatedly, 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 (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 $\sigma-$algebras on V_e so that for all $q \in Z,\ e \in \mathcal{H}_{d'}$*

$$compl\ (\mathcal{B}_{q,e}) \leq M_{d'}. \tag{3.2.1}$$

Let $\varepsilon > 0$ and $F : \mathbb{R}_+ \to \mathbb{R}_+$ 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 is sufficiently large with respect to the parameters $\varepsilon, c_0, M_{d'}, K,$ and F, then there exists a well-defined, symmetric extension $\{\mu_{\overline{q},f}\}_{\overline{q}\in\overline{Z}}$ of complexity at most $O_{K,M_{d'},F,\varepsilon}(1)$, and families of σ-algebras $\mathcal{B}_{\overline{q},f} \subseteq \mathcal{B}'_{\overline{q},f}$ defined for $\overline{q} \in \overline{Z},\ f \in \mathcal{H}_{d-1}$ and a set $\overline{\Omega} \subseteq \overline{Z}$ such that the following holds.

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

$$\overline{\psi}(\overline{\Omega}) \geq c(c_0, F, M_{d'}, \varepsilon) > 0. \tag{3.2.2}$$

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

$$compl(\mathcal{B}_{\overline{q},f}) \leq M_{d'-1}. \tag{3.2.3}$$

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

$$\big\|\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q,e}}|\bigvee_{f\in\partial e}\mathcal{B}'_{\overline{q},f}) - \mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q,e}}|\bigvee_{f\in\partial e}\mathcal{B}_{\overline{q},f})\big\|_{L^2(\mu_{\overline{q},e})} \leq \varepsilon \tag{3.2.4}$$

and

$$\|\mathbf{1}_{G_{q,e}} - \mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}'_{\overline{q},f})\|_{\square_{\nu_{\overline{q},e}}} \leq \frac{1}{F(M_{d'-1})}. \tag{3.2.5}$$

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'. Also 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 $(\mu_{q',e},\, \mathcal{B}_{q',e},\, \mathcal{B}_{q',f})$, with $\delta = F(M_{d'-1})^{-1}$.

This generates a well-defined, symmetric extension $\{\mu_{\overline{q},f}\}_{\overline{q}\in\overline{Z}, f\in\mathcal{H}}$ and a family of σ-algebras $\{\mathcal{B}'_{\overline{q},f}\}_{\overline{q}\in\overline{Z}, f\in\mathcal{H}_{d'-1}}$ and a set $\overline{\Omega} \subseteq \overline{Z}$. Set $\mathcal{B}_{\overline{q},f} := \mathcal{B}_{q',f}$ for $\overline{q} = (q',p) \in \overline{Z}$, $f \in \mathcal{H}_{d'-1}$. The new system $(\mu_{\overline{q},f}, \mathcal{B}_{\overline{q},e}, \mathcal{B}'_{\overline{q},f})$ satisfies (3.2.2)-(3.2.3) and (3.2.5) 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 $\overline{\Omega}_1 \subseteq \overline{\Omega}$ of measure $\overline{\psi}(\overline{\Omega}_1) \geq \overline{\psi}(\overline{\Omega})/2$ such that (3.2.5) holds for all $\overline{q} \in \overline{\Omega}_1$. In this case the conclusions of the lemma hold for the system $(\mu_{\overline{q},e}, \mathcal{B}_{\overline{q},e}, \mathcal{B}'_{\overline{q},f})$ and the set $\overline{\Omega}_1$.

- *Case 2:* There is a set $\overline{\Omega}_2 \subseteq \overline{\Omega}$ of measure $\overline{\psi}(\overline{\Omega}_2) \geq \frac{1}{2}\overline{\psi}(\overline{\Omega})$ so that inequality (3.2.5) fails for all $\overline{q} \in \overline{\Omega}_2$. Then, thanks to the stability condition (3.1.22) and the fact that $\mathcal{B}_{q',f} = \mathcal{B}_{\overline{q},f} \subseteq \mathcal{B}'_{\overline{q},f}$, we have for $\overline{q} \in \overline{\Omega}_2$, $q' = \pi'(\overline{q})$, and $q = \pi(\overline{q})$ that

$$\begin{aligned}
&\sum_{e,G_{q,e}} \|\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}'_{\overline{q},f})\|^2_{L^2_{\mu_{\overline{q},e}}} - \sum_{e,G_{q,e}} \|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}_{q',f})\|^2_{L^2_{\mu_{q',e}}}\\
&\geq \sum_{e,G_{q,e}} (\|\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}'_{\overline{q},f})\|^2_{L^2_{\mu_{\overline{q},e}}} - \|\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}_{q',f})\|^2_{L^2_{\mu_{\overline{q},e}}}) - o_{M_{d'},K',F}(1)\\
&= \sum_{e,G_{q,e}} \|\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}'_{\overline{q},f}) - \mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}_{q',f})\|^2_{L^2_{\mu_{\overline{q},e}}} - o_{M_{d'},K',F}(1)\\
&\geq \varepsilon^2 - o_{M_{d'},K',F}(1),
\end{aligned} \tag{3.2.6}$$

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 $\overline{q} = (q,p) \in \overline{\Omega}_2$ that the total energy of the system $(\mu_{\overline{q},f},\ \mathcal{B}_{\overline{q},e},\ \mathcal{B}'_{\overline{q},f})$ 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 $Z' := \overline{Z}$, $\Omega' := \overline{\Omega}_3$, $\mu_{q',f} := \mu_{\overline{q},f}$, and $\mathcal{B}_{q',f} := \mathcal{B}'_{\overline{q},f}$ 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 $\varepsilon^{-2}2^{2^{(M_{d'})+1}}2^{d+1} = O_{M_{d'},\varepsilon}(1)$ steps, generating a system $(\mu_{\overline{q},f}, \mathcal{B}_{\overline{q},e}, \mathcal{B}'_{\overline{q},f})$ which satisfies the conclusions of the lemma. □

This lemma is more widely applicable than Lemma 3.2 as the uniformity of the hypergraphs $G_{q,e}$ with respect to the (fine) $\sigma-$algebras $B'_{\overline{q},e}$ can be chosen to be arbitrarily small with respect to the complexity of the (coarse) $\sigma-$algebras $\mathcal{B}_{\overline{q},e}$, while the approximations $\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q,e}}|\bigvee \mathcal{B}'_{\overline{q},e})$ and $\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q,e}}|\bigvee \mathcal{B}_{\overline{q},e})$ stay very close in $L^2(\mu_{\overline{q},e})$.

In order 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 $\mathcal{B}_{\overline{q},e}$ 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 $F:\mathbb{R}_+\to\mathbb{R}_+$ a *growth function* if it is continuous, increasing, and satisfies $F(x)\geq 1+x$ for $x\geq 0$.

Theorem 3.1. *[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 $\sigma-$algebras on V_e so that for all $q\in Z,\ e\in\mathcal{H}_{d'}$*

$$compl\,(\mathcal{B}_{q,e})\leq M_{d'}. \tag{3.2.7}$$

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

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

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

$$\overline{\psi}(\overline{\Omega})\geq c(c_0,F,M_{d'})>0. \tag{3.2.8}$$

(2) *There exist numbers*

$$M_{d'}<F(M_{d'})\leq M_{d'-1}<F(M_{d'-1})\leq\cdots\leq M_1<F(M_1)\leq M_0=O_{M_{d'},F}(1) \tag{3.2.9}$$

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

$$compl(\mathcal{B}'_{\overline{q},f})\leq M_j. \tag{3.2.10}$$

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

$$\|\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{\overline{q},e}}|\bigvee_{f\in\partial e}\mathcal{B}'_{\overline{q},f})-\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{\overline{q},e}}|\bigvee_{f\in\partial e}\mathcal{B}_{\overline{q},f})\|_{L^2(\mu_{\overline{q},e})}\leq\frac{1}{F(M_j)} \tag{3.2.11}$$

and

$$\|\mathbf{1}_{G_{\overline{q},e}}-\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{\overline{q},e}}|\bigvee_{f\in\partial e}\mathcal{B}'_{\overline{q},f})\|_{\square_{\nu_{\overline{q}.e}}}\leq\frac{1}{F(M_1)}. \tag{3.2.12}$$

Proof. We proceed by an 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

$$\|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q',e}}|\bigvee_{f\in\partial e}\mathcal{B}'_{q',f})-\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q',e}}|\bigvee_{f\in\partial e}\mathcal{B}_{q',f})\|_{L^2(\mu_{q',e})}\leq\frac{1}{2F^*(M_{d'})} \tag{3.2.13}$$

and

$$\|\mathbf{1}_{G_{q',e}}-\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q',e}}|\bigvee_{f\in\partial e}\mathcal{B}'_{q',f})\|_{\square_{\nu_{q',e}}}\leq\frac{1}{F^*(M_{d'-1})}, \tag{3.2.14}$$

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 $\{\mu_{q',f}\}_{q'\in Z',f\in\mathcal{H}}$, $\{\mathcal{B}_{q',f}\}_{q'\in Z',f\in\mathcal{H}_{d'-1}}$, the growth function F, and the set Ω', one obtains an extension $\{\mu_{\overline{q},f}\}_{\overline{q}\in\overline{Z},f\in\mathcal{H}}$ and families of $\sigma-$algebras $\{\mathcal{B}_{\overline{q},f}\subseteq\mathcal{B}'_{\overline{q},f}\}_{\overline{q}\in\overline{Z},\ f\in\mathcal{H}_j}$ such that (3.2.10) - (3.2.12) hold for $j<d'-1$, with constants

$$M_{d'-1}<F(M_{d'-1})\leq\cdots\leq M_1<F(M_1)=O_{M_{d'-1},F}(1). \tag{3.2.15}$$

For $\overline{q}=(q',p)\in\overline{Z}$, $f\in\mathcal{H}_{d'-1}$ set $\mathcal{B}_{\overline{q},f}:=\mathcal{B}_{q',f}$, and $\mathcal{B}'_{\overline{q},f}:=\mathcal{B}'_{q',f}$. We show that inequalities (3.2.11) and (3.2.12) hold for $j=d'$. Indeed, by the stability property (2.2.12), one has

$$\begin{aligned}
&\big\|\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q',e}}|\bigvee_{f\in\partial e}\mathcal{B}'_{q',f})-\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q',e}}|\bigvee_{f\in\partial e}\mathcal{B}_{q',f})\big\|_{L^2(\mu_{\overline{q},e})}\\
&=\big\|\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q',e}}|\bigvee_{f\in\partial e}\mathcal{B}'_{q',f})-\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q',e}}|\bigvee_{f\in\partial e}\mathcal{B}_{q',f})\big\|_{L^2(\mu_{q',e})}+o_{K,M_{d'},F,F^*}(1)\\
&\leq\frac{1}{2F^*(M_{d'})}+o_{K,M_{d'},F,F^*}(1),
\end{aligned}\tag{3.2.16}$$

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

Similarly using the stability properties (2.2.10) and (2.2.17) of the box norms, we have

$$\begin{aligned}
&\big\|\mathbf{1}_{G_{q',e}}-\mathbb{E}_{\mu_{\overline{q},e}}(\mathbf{1}_{G_{q',e}}|\bigvee_{f\in\partial e}\mathcal{B}'_{q',f})\big\|_{\Box_{\nu_{\overline{q},e}}}=\big\|\mathbf{1}_{G_{q',e}}-\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q',e}}|\bigvee_{f\in\partial e}\mathcal{B}'_{q',f})\big\|_{\Box_{\nu_{\overline{q},e}}}+o_{K,M_{d'},F,F^*}(1)\\
&=\big\|\mathbf{1}_{G_{q',e}}-\mathbb{E}_{\mu_{q',e}}(\mathbf{1}_{G_{q,e}}|\bigvee_{f\in\partial e}\overline{\mathcal{B}}_{q',f})\big\|_{\Box_{\nu_{q',e}}}+o_{K,M_{d'},F,F^*}(1)\leq\frac{1}{2F^*(M_{d'-1})}+o_{K,M_{d'},F,F^*}(1),
\end{aligned}\tag{3.2.17}$$

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

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 N, W are sufficiently large with respect to $M_{d'}$ and K, inequalities (3.2.11), (3.2.12) for $j=d'$ and $\overline{q}\in\overline{\Omega}\backslash(\overline{\mathcal{E}}_1\cup\overline{\mathcal{E}}_2)$ follow from (3.2.13) and (3.2.14). The rest of the conclusions of the theorem are clear from the construction. □

4. 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 $\sigma-$algebra generated by it. Let $\{\nu_e\}_{e\in\mathcal{H}}$ and $\{\mu_e\}_{e\in\mathcal{H}}$ be the weights and measures associated to a well-defined, symmetric family forms $\mathcal{L}=\{L_e^k;\ e\in\mathcal{H}_d,\ 1\leq k\leq d\}$. Take $M_d>0$ and $F:\mathbb{R}_+\to\mathbb{R}_+$ 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 (3.2.8)-(3.2.12) hold.[3]

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

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 eventually replace the ensemble of measures $\{\mu_{q,e}\}_{e\in\mathcal{H}}$ with the measure $\mu_q := \mu_{q,J} = \prod_{f\in\mathcal{H}} \nu_{q,f}$. Thanks to the stability conditions (2.1.6)-(2.1.7) 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

$$\mu_q(\pi_e^{-1}(G_{q,e})) = \mu_{q,e}(G_{q,e}) + o_{M_d,F}(1), \tag{4.1.1}$$

uniformly for $q \in \Omega\backslash\mathcal{E}_e$. Let $\mathcal{E} = \bigcup_{e\in\mathcal{H}} \mathcal{E}_e$, $\Omega' := \Omega\backslash\mathcal{E}$, then (4.1.1) means that for any set $A_{q,e} \in \mathcal{A}_e$ one has that $\mu_q(A_{q,e}) = \mu_{q,e}(\pi_e(A_{q,e})) + o_{M_d,F}(1)$ uniformly for $q \in \Omega'$. We will write $\mu_{q,e}(A_{q,e}) = \int_{V_e} 1_{A_{q,e}}(x_e)\, d\mu_{q,e}(x_e)$ for simplicity of notations.

Define the σ-algebras $\overline{\mathcal{B}}_{q,e} := \pi_e^{-1}(\mathcal{B}_{q,e})$, $\overline{\mathcal{B}}'_{q,e} := \pi_e^{-1}(\mathcal{B}'_{q,e})$ on V_J, and 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 $A_q = \bigcap_{e\in\mathcal{H}} A_{q,e}$ 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 [19], [5], [14], gives an approximate formula for the measure of "most" atoms A_q and as 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* $\delta_{q,e}(A|B) := \mu_{q,e}(A\cap B)/\mu_q(B)$ for $A, B \in \overline{\mathcal{B}}_{q,e}$, with the convention that $\delta_{q,e}(A|B) := 1$ if $\mu_q(B) = 0$.

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

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

$$\delta_{q,e}(A_{q,e}| \bigcap_{f\in\partial e} A_{q,f}) \geq \frac{1}{\log F(M_j)}, \tag{4.1.2}$$

(2) *Moreover*

$$\int_{V_e} |\mathbb{E}_{\mu_q}(\mathbf{1}_{A_{q,e}}| \bigvee_{f\in\partial e} \overline{\mathcal{B}}'_{q,f}) - \mathbb{E}_{\mu_q}(\mathbf{1}_{A_{q,e}}| \bigvee_{f\in\partial e} \overline{\mathcal{B}}_{q,f})|^2 \prod_{f\subsetneq e} \mathbf{1}_{A_{q,f}}\, d\mu_{q,e} \leq \frac{1}{F(M_j)} \int_{V_J} \prod_{f\subsetneq e} \mathbf{1}_{A_{q,f}}\, d\mu_{q,e}. \tag{4.1.3}$$

This roughly means that all atoms $A_{q,e}$ are both somewhat large and regular on the intersection of the lower order atoms $A_{q,f}$, $(f \in \partial e)$. Note that if $|e| = 1$ then $\partial e = \emptyset$ and by convention we define $\bigcap_{f\in\partial e} A_{q,f} = V_J$, and the left side of (4.1.2) becomes $\mu_{q,e}(A_{q,e})$.

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

$$\mu_q(A_q) = (1 + o_{M_d\to\infty}(1)) \prod_{e\in\mathcal{H}} \delta_{q,e}(A_{q,e}| \bigcap_{f\in\partial e} A_{q,f}) + O_{M_1}\left(\frac{1}{F(M_1)}\right) + o_{N,W\to\infty;M_d,F}(1). \tag{4.1.4}$$

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 (4.1.2) or (4.1.3) 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

$$\mu_q(A_{q,e} \cap B_{q,e,A_{q,e}}) \lesssim \frac{1}{\log F(M_j)} \tag{4.1.5}$$

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, note that the measure μ_q can be replaced by the measure $\mu_{q,e}$ as they differ by a negligible quantity on sets which belong to $\mathcal{A}_e$. We estimate first the contribution of those sets $\bigcap_{f \subsetneq e} A_{q,f}$ to the left side of (4.1.5) for which (4.1.2) fails. This quantity is bounded by

$$\begin{aligned}\sum_{\{A_{q,f}\}_{f\in\partial e},\ \text{(4.1.2) fails}} \mu_{q,e}(\mathcal{A}_{q,e} \cap \bigcap_{f \in \partial e} A_{q,f}) &\le \frac{1}{\log F(M_j)} \sum_{\{A_{q,f}\}_{f\in\partial e}} \mu_{q,e}(\bigcap_{f\in\partial e} A_{q,f}) \\ &\le \frac{1}{\log F(M_j)}\, \mu_{q,e}(V_e) \lesssim \frac{1}{\log F(M_j)},\end{aligned}$$

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

Similarly, one estimates the total contribution of the disjoint atoms $\bigcap_{f \subsetneq e} A_{q,f}$ for which (4.1.3) fails as follows.

$$\begin{aligned}&\sum_{\{A_{q,f}\}_{f\subseteq e},\ \text{(4.1.3) fails}} \mu_{q,e}(\bigcap_{f \subsetneq e} A_{q,f}) \\ &\le F(M_j) \int_{V_e} |\mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{A_{q,e}} | \bigvee_{f\in\partial e} \overline{\mathcal{B}}'_{q,f}) - \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{A_{q,e}} | \bigvee_{f\in\partial e} \overline{\mathcal{B}}_{q,f})|^2 \, d\mu_{q,e} \\ &\le F(M_j)\, \frac{1}{F(M_j)^2} = \frac{1}{F(M_j)}.\end{aligned}$$

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

$$\sum_{j=1}^{d} \sum_{e \in \mathcal{H}_j} \sum_{A_{q,e \in \mathcal{B}_{q,e}}} \mu_q(A_{q,e} \cap \mathcal{B}_{q,e,A_{q,e}}) \le \sum_{j=1}^{d} \binom{d}{j} 2^{2^{M_j}} \frac{1}{\log F(M_j)} \le \frac{1}{\sqrt{\log F(M_d)}} \le 2^{-2^{M_d}} \tag{4.1.6}$$

if, say $F(M) \ge 2^{2^{2^{M_d+d}}}$. 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 (4.1.2), (4.1.4), we have that for $q \in \Omega$ and a regular atom $A_q = \cap_{e\in\mathcal{H}} A_{q,e}$,

$$\mu_q(A_q) \ \ge\ \prod_{j \le d} \prod_{e \in \mathcal{H}_j} \frac{1}{F(M_j)^{1/10}} - O_{d,M_1}\left(\frac{1}{F(M_1)}\right) + o_{M_d,F}(1) \ \ge\ \frac{1}{F(M_1)} > 0, \tag{4.1.7}$$

as long as F is sufficiently rapid growing and M_d is sufficiently large with respect to d. It is clear from (3.2.9) that $F(M_1) \le F^*(M_d)$ 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 $g_e : V_e \to [0,1]$ 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 (4.1.1), (4.1.6) and (4.1.7) hold for $q \in \Omega/\mathcal{E}_1$. Also by (2.2.4) conditions (1.4.11)-(1.4.12) hold for

$$\tilde{\mu}_J := \mu_{q,J} \quad \textit{and} \quad \tilde{\mu}_e := \mu_{q,e} \;\; (e \in \mathcal{H}_d), \tag{4.1.8}$$

for $q \notin \mathcal{E}_2$, for a set $\mathcal{E}_2 \subseteq \Omega$ be 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 is (4.1.8). We claim that this system of measures satisfy the conclusions of the theorem. By construction the system is symmetric so it remains to construct the sets E'_e and show (1.4.13)-(1.4.15) hold. For given $e \in \mathcal{H}$ define the sets

$$E'_{q,e} = V_J \backslash \, (B_{q,e,A_e} \cup \bigcup_{f \subsetneq e,\, A_{q,f}} (A_{q,f} \cap B_{q,f,A_{q,f}})), \tag{4.1.9}$$

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 (1.4.10) its measure would satisfy

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

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 (4.1.9) and (4.1.6) we have

$$\tilde{\mu}_J \, (E_e \backslash E'_{q,e}) \; = \; \tilde{\mu}_J \, (\bigcup_{f \subseteq e,\, A_{q,f}} (A_{q,f} \cap B_{q,f,A_{q,f}})) \; \leq \; 2^{-2^{M_d}}. \tag{4.1.10}$$

Also, 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 $E_e \backslash E'_{q,e,}$ thus

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

Finally, choosing $\varepsilon := 2^{-2^{M_d}}$, (1.4.14) holds by (4.1.10). Moreover $\delta \to 0$ implies $M_d \to \infty$ and hence $\varepsilon \to 0$ showing the validity of (1.4.15). This proves Theorem 1.4. □

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 $\pi : K \to J$, 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 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 weights and measures $\overline{\nu}_{q,g}$, $\overline{\mu}_{q,g} : V_g \to \mathbb{R}_+$ as

$$\overline{\nu}_{q,g}(x_g) := \nu_{q,\pi(g)}(x_g), \;\; \overline{\mu}_{q,g}(x_g) = \prod_{g' \subseteq g} \overline{\nu}_{q,g'}(x_{g'}).$$

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

$$\overline{\mu}_{q,K}(x) = \prod_{g\in\mathcal{G}} \overline{\nu}_{q,g}(x_g).$$

A hypergraph $\mathcal{G} \subseteq \mathcal{G}_K$ *which is closed in the sense that* $\partial g \subseteq \mathcal{G}$ *for every* $g \in \mathcal{G}$*, together with the spaces* V_g *and the weight functions* $\overline{\nu}_{q,g}$ *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 $\{\overline{\nu}_{q,g}\}_{q\in Z, g\in\mathcal{G}_K}$,

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

are pairwise linearly independent. Indeed, if $g \neq g'$ they depend on different sets of variables, and for a fixed sets of variables they are the same as the forms $L(q, x_g)$. What happens is that we sample a number variables from each space V_j and evaluate the forms $L(q, x)$ in the new variables. For example if we have $x_1, x_{1'} \in V_1$ and $x_2, x_{2'} \in V_2$ then to the edge $(1, 2) \in \mathcal{H}$ there correspond the edges $(1, 2), (1, 2'), (1', 2)$ and $(1', 2')$ in $\mathcal{G}$, and to every linear form $L(q, x_1, x_2)$ there also correspond the forms $L(q, x_1, x_{2'})$, $L(q, x_{1'}, x_2)$ and $L(q, x_{1'}, x_{2'})$ defining the weights on the appropriate edges.

Proposition 4.2. *[Generalized Counting Lemma] Let* $\mathcal{G} \subseteq \mathcal{G}_K$ *be a closed hypergraph bundle over* $\mathcal{H}$ *with the projection map* $\pi : K \to J$*, 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* $\psi(\mathcal{E}) = o_{N\to\infty;M_d,K,F}(1)$ *such that for* $q \in \Omega\backslash\mathcal{E}$ *we have*

$$\int_{V_K} \prod_{g\in\mathcal{G}} \mathbf{1}_{A_{q,\pi(g)}}(x_g)\, d\overline{\mu}_{q,K}(x) \tag{4.2.1}$$
$$= (1 + o_{M_d\to\infty,K}(1)) \prod_{g\in\mathcal{G}} \delta_{q,\pi(g)}(A_{q,\pi(g)}| \bigcap_{f\in\partial\pi(g)} A_{q,f}) + O_{K,M_1}(\frac{1}{F(M_1)}) + o_{N\to\infty,K,M_d}(1).$$

Note that Proposition 3 is the special case when $\mathcal{G} = \mathcal{H}$ and π is the identity map.

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 (4.2.1) becomes

$$\int_{V_k} \mathbf{1}_{A_{q,j}}(x_k)\, d\overline{\mu}_{q,k}(x_k) = \int_{V_j} \mathbf{1}_{A_{q,j}}(x_j)\, d\mu_{q,j}(x_j) = \delta_{q,j}(A_{q,j}).$$

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

$$b_{q,e} := \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{A_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}'_{q,f}) - \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{A_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}_{q,f}) \tag{4.2.2}$$

$$c_{q,e} := \mathbf{1}_{A_{q,e}} - \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{A_{q,e}}| \bigvee_{f\in\partial e} \mathcal{B}'_{q,f}) \tag{4.2.3}$$

and introduce the shorthand notation

$$\delta_{q,e} = \delta_{q,e}(A_{q,e}| \bigcap_{f\in\partial e} A_{q,f}).$$

Note that if $x \in A_{q,e} \bigcap_{f\in\partial e} \mathcal{A}_{q,f}$ then

$$\delta_{q,e} = \mathbb{E}_{\mu_{q,e}}(\mathbf{1}_{A_e}| \bigvee_{f\in\partial e} \mathcal{B}_{q,f})(x_e), \tag{4.2.4}$$

and thus one has the decomposition

$$\mathbf{1}_{A_{q,e}}(x_e) = \delta_{q,e} + b_{q,e}(x_e) + c_{q,e}(x_e) \tag{4.2.5}$$

on the set $\bigcap_{f\in\partial e} \mathcal{A}_{q,f}$. Let $g_0 \in \mathcal{G}$ such that $|g_0| = d'$ and use (4.2.5) to write

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

Consider the contribution of the terms separately

$$\begin{aligned}
&\int_{V_K} \prod_{g\in\mathcal{G}} \mathbf{1}_{A_{q,\pi(g)}}(x_g) d\overline{\mu}_{q,K}(x) \\
&= \int_{V_K} (\delta_{q,\pi(g_0)} + b_{q,\pi(g_0)}(x_{g_0}) + c_{q,\pi(g_0)}(x_{g_0})) \prod_{g\in\mathcal{G}\setminus\{g_0\}} \mathbf{1}_{A_{q,\pi(g)}}(x_g) d\overline{\mu}_{q,K}(x) \\
&= M_q + E_q^1 + E_q^2
\end{aligned} \tag{4.2.6}$$

For main term M_q, by the second induction hypothesis we have

$$\begin{aligned}
M_q &= \delta_{q,\pi(g_0)} \int_{V_K} \prod_{g\in\mathcal{G}\setminus\{g_0\}} \mathbf{1}_{A_{q,\pi(g)}}(x_g) d\overline{\mu}_{q,K}(x) \\
&= \delta_{q,\pi(g_0)} (1 + o_{M_d\to\infty}(1)) \prod_{g\in\mathcal{G}\setminus g_0} \delta_{q,\pi(g)} + \ O_{K,M_1}(\frac{1}{F(M_1)}) + \ o_{N,W\to\infty;K,M_d}(1),
\end{aligned}$$

and hence M_q agrees with the right side of (4.2.1). We continue to estimate the second error term by

$$\begin{aligned}
E_q^2 &= \int_{V_K} c_{q,\pi(g_0)}(x_{g_0}) \prod_{g\in\mathcal{G}\setminus\{g_0\}} \mathbf{1}_{A_{q,\pi(g)}}(x_g) d\overline{\mu}_q(x) = \mathbb{E}_{x\in V_K}(c_{q,\pi(g_0)}\overline{\nu}_{q,g_0})(x_{g_0}) \prod_{g\in\mathcal{G}\setminus\{g_0\}} \mathbf{1}_{A_{q,\pi(g)}}\overline{\nu}_{q,g}(x_g) \\
&= \mathbb{E}_{x\in V_K} \prod_{|g|=d',g\in\mathcal{G}} f_{q,g}(x_g) \prod_{g'\in\mathcal{G},|g'|<d'} \overline{\nu}_{q,g'}(x_{g'}),
\end{aligned} \tag{4.2.7}$$

where $f_{q,g_0} = c_{q,\pi(g_0)}\overline{\nu}_{q,g_0}$ and $f_{q,g} = h_{q,g}\overline{\nu}_{q,g}$, 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 $|f_{q,g}| \leq \overline{\nu}_{q,g}$ 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 $f_{q,g}(x_g)$, $g \neq g_0$, which does not depend on at least one of these variables, we obtain

$$|E_q^2|^{2^{d'}} \lesssim \|c_{q,\pi(g_0)}\|_{\square_{\overline{\nu}_{q,g_0}}}^{2^{d'}} + \mathbb{E}_{x_{g_0},y_{g_0}} |\mathcal{W}_q(x_{g_0}, y_{g_0}) - 1| \prod_{h\subseteq g_0} \prod_{\omega\in\{0,1\}^h} \overline{\nu}_{q,h}(\omega_h(x_h, y_h)), \tag{4.2.8}$$

where setting $K' := K\backslash g_0$

$$\mathcal{W}_q(x_{g_0}, y_{g_0}) = \mathbb{E}_{x\in V_{K'}} \prod_{g\in\mathcal{G}, g\not\subseteq g_0} \prod_{\omega\in\{0,1\}^{g\cap g_0}} \overline{\nu}_{q,g}(\omega_{g\cap g_0}(x_{g,\cap g_0}, y_{g\cap g_0}), x_{g\backslash g_0}). \tag{4.2.9}$$

Note that the first term on the right hand side of (4.2.8) is $O(F(M_1)^{2^{-d'}})$ by (3.1.12) and (4.2.3). To estimate the second term we apply the Cauchy-Schwartz inequality one more time to see that it is $o_{N,W\to\infty;M_d,K,F}(1)$ for $q \notin \mathcal{E}_1$, where $\mathcal{E}_1$ is a set of measure $o_{N,W\to\infty;M_d,K,F}(1)$ 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

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

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

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

where $\mathcal{G}' = \{g \in \mathcal{G};\ g \not\subseteq g_0\}$. Writing $A(x_{g_0})$ for the expression in the first parenthesis, and $B(x_{g_0})$ for the expression in the first parenthesis. Thus we have

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

thus by the Cauchy-Schwartz inequality we get

$$|E_q^1|^2 \lesssim \left(\int_{V_{g_0}} A(x_{g_0})^2\, d\overline{\mu}_{q,g_0}(x_{g_0}) \right) \left(\int_{V_{g_0}} B(x_{g_0})^2\, d\overline{\mu}_{q,g_0}(x_{g_0}) \right). \tag{4.2.10}$$

Since $\overline{\nu}_{q,g_0}(V_{g_0}) = 1 + o_{M_d,K,F}(1)$ outside a set $\mathcal{E}_2 \subseteq \Omega$ of measure $\psi(\mathcal{E}_2) = o_{M_d,K,F}(1)$, the first factor on the left side of (4.2.10) is estimated by

$$\mathbb{E}_{x_{g_0}\in V_{g_0}} b_{q,\pi(g_0)}(x_{g_0})^2 \prod_{g\subsetneq g_0} \mathbf{1}_{A_{q,\pi(g)}}(x_g) \prod_{g\subseteq g_0} \nu_{q,\pi(g)}(x_g). \tag{4.2.11}$$

Let $f_0 = \pi(g_0)$, since $\pi : g_0 \to f_0$ is injective and $V_{g_0} = V_{f_0}$, we may write the expression in (4.2.11), by re-indexing the variables x_g to x_f, $f = \pi(g)$ for $g \subseteq g_0$, as

$$\int_{V_{f_0}} b_{q,f_0}^2(x_{f_0}) \prod_{f\subsetneq f_0} \mathbf{1}_{A_{q,f}}(x_f)\, d\mu_{q,f_0}(x_{f_0}) \lesssim \frac{1}{F(M_{d'})} \int_{V_{f_0}} \prod_{f\subsetneq f_0} \mathbf{1}_{A_{q,f}}(x_f) d\mu_{q,f_0}(x_{f_0}), \tag{4.2.12}$$

where the inequality follows from by assumption (4.1.3) on regular atoms. By the induction hypothesis we further estimate the right side (4.2.12) as

$$\frac{1}{F(M_{d'})}\,(1 + o_{M_d\to\infty}(1)) \prod_{f\subsetneq f_0} \delta_{q,f} \,+\, O_{M_d}(\frac{1}{F(M_1)}) + o_{N,W\to\infty;M_d,K,F}(1). \tag{4.2.13}$$

The second factor in (4.2.10) may be expressed in terms of a hypergraph bundle $\tilde{K}$ over K, by using the construction given 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 $k \subseteq g_0$. Let $\phi : \tilde{K} \to K$ be the natural projection, and $\pi \circ \phi : \tilde{K} \to J$ 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 $\tilde{K}$ to consist of the edges $g \times \{0\}$ and $g \times \{1\}$ for $g \in \mathcal{G}_0 \cup \mathcal{G}'$, two edges bing identified for $g \in \mathcal{G}_0$. Define the weights

$$\tilde{\nu}_{q,g\times\{i\}}(x_{g\times\{i\}}) := \overline{\nu}_{q,g}(x_{g\times\{i\}}), \tag{4.2.14}$$

for $q \in Z,\ g \in \mathcal{G}_K, i = 0,1$, that is for all edges $\tilde{g} \in \mathcal{G}_{\tilde{K}}$, and let $\tilde{\mu}_{q,g\times\{i\}}$ be the associated family of measures. Then we have for the second factor appearing in (4.2.10)

$$\begin{aligned}
&\int_{V_{g_0}} B(x_{g_0})^2 d\overline{\mu}_{q,g_0}(x_{g_0}) \\
&= \int_{V_{g_0}} \Big[\prod_{g\in g_0} \mathbf{1}_{A_{q,\pi(g)}}(x_g)\Big]\Big[\mathbb{E}_{x\in V_{K\setminus g_0}} \prod_{g\in\mathcal{G}\setminus\{g_0\}} \mathbf{1}_{A_{q,\pi(g)}}\overline{\nu}_{q,g}(x_g) \prod_{h\not\subseteq g_0,|h|=d'} \overline{\nu}_{q,h}(x_h)\Big]^2 d\overline{\mu}_{q,g_0}(x_{g_0}) \\
&= \int_{V_{\tilde{K}}} \prod_{\tilde{g}\in\tilde{\mathcal{G}}} \mathbf{1}_{A_{q,\pi\circ\phi(\tilde{g})}}(x_{\tilde{g}})\, d\tilde{\mu}_{q,\tilde{K}}(x_{\tilde{K}}).
\end{aligned} \tag{4.2.15}$$

Indeed, when expanding the square of inner sum in (4.2.13) 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 $g \in \tilde{G}$ to obtain the edges $g\times\{0\}, g\times\{1\}$. As for the weights, the procedure doubles all weights $\overline{\nu}_{q,g}(x_g)$ for $g \not\subset g_0,\ g \in \mathcal{G}_K$ to obtain the weights $\overline{\nu}_{q,g}(x_{g\times\{i\}})$ for $i = 0,1$ while leaves the weights $\overline{\nu}_{q,g}(x_g)$ for $g \subseteq g_0$ unchanged. The order of $\tilde{g}$ is less than d' thus by the first induction hypothesis, we have

$$\begin{aligned}
&\int_{V_{\tilde{K}}} \prod_{\tilde{g}\in\tilde{\mathcal{G}}} \mathbf{1}_{A_{q,\pi\circ\phi(\tilde{g})}}(x_{\tilde{g}})\, d\tilde{\mu}_{q,\tilde{K}}(x_{\tilde{K}}) = \\
&= (1 + o_{M_d\to\infty}(1)) \prod_{\tilde{g}\in\tilde{\mathcal{G}}} \delta_{q,\pi\circ\phi(\tilde{g})} + O_{K,M_1}(\frac{1}{F(M_1)}) + o_{N,W\to\infty;,M_d,K,F}(1) \\
&= (1 + o_{M_d\to\infty}(1)) \prod_{g\in\mathcal{G}_0} \delta_{q,\pi(g)} \prod_{g\in\mathcal{G}'} \delta^2_{q,\pi(g)} + O_{K,M_1}(\frac{1}{F(M_1)}) + o_{N,W\to\infty;M_d,K,F}(1),
\end{aligned} \tag{4.2.16}$$

for $q \notin \mathcal{E}_{\tilde{K},\phi}$ where $\mathcal{E}_{\tilde{K},\phi} \subseteq \Omega$ is a set of measure $\psi(\mathcal{E}_{\tilde{K},\phi}) = o_{N,W\to\infty;M_d,K,F}(1)$. Note that there are only $O_K(1)$ choices for choosing the set $\tilde{K}$ and the projection map $\phi : \tilde{K} \to K$ thus taking the union of all possible exceptional sets $\mathcal{E}_{\tilde{K},\phi}$ we have that (4.2.16) holds for $q \notin \mathcal{E}'_K$ if measure $\psi(\mathcal{E}'_K) = o_{N,W\to\infty;M_d,K,F}(1)$. Combining the bounds (4.2.13) and (4.2.16) we obtain the error estimate

$$|E^1_q|^2 = (o_{M_d\to\infty}(1)) \prod_{g\in\mathcal{G}} \delta^2_{q,\pi(g)} + O_{K,M_1}(\frac{1}{F(M_1)}) + o_{N,W\to\infty;M_d,K,F}(1),$$

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

5. Proof of The Main Results

In this section we finish the proof of our main result Theorem 1.2. Since we have already shown the validity of Theorem 1.4 and hence that of Theorem 1.3 by the argument in the introduction, it remains to show that counting affine copies of Δ in a set $A \subseteq \mathbb{Z}^d_N$ with weights w translates to counting copies in $A \subseteq \mathbb{P}^d$ of relative density $\alpha > 0$. This is standard, we include the details for the sake of completeness, using the arguments given in [2].

First, let us identify $[1,N]^d$ with $\mathbb{Z}_N^d$ and 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 $t\Delta \nsubseteq \mathbb{Z}^d$ for any $0<t<1$, as any simplex is a dilate 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 holds.

Lemma 5.1. [2] *Let $\Delta \subseteq \mathbb{Z}^d$ be a primitive simplex. Then there is constant $0<\varepsilon<\tau(\Delta)$ so 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 $1 \le t < N$ such that $x \in B$ and $x + t\Delta \subseteq B$ 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 [Theorem 1.3 implies Theorem 1.2]

Let N, W be sufficiently large positive integers and assume that $|A| \ge \alpha\, |\mathbb{P}_N|^d$ for a set $A \subseteq \mathbb{P}_N^d$. By the pigeonhole principle choose $b = (b_j)_{1\le j\le d}$ so that b_j is relative prime to W for each j, and

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

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 $\varepsilon_2 > 0$ so that $2\varepsilon_2 < \tau(\Delta)$. By the Prime Number Theorem there is a prime N' so that $\varepsilon_2 N' = N_1(1 + o_{N_1\to\infty}(1))$, thus we have

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

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

If $x \in A'$ then $\varepsilon_1 N' \le x_i \le \varepsilon_2 N'$ and $Wx_i + b_i \in \mathbb{P}$ for $1 \le i \le d$, thus by the definition of the Green-Tao measure $\nu_b : [1,N'] \to \mathbb{R}_+$ given in Section 1.3, we have

$$w(x) = \prod_{i=1}^d \nu_{b_i}(x_i) \ge c_d \left(\frac{\phi(W)\log N}{W}\right)^d. \tag{5.3}$$

as $\log N' - \log N$ assuming N sufficiently large with respect to W. Thus

$$\mathbb{E}_{x\in\mathbb{Z}_{N'}^d} \mathbf{1}_{A'}(x)\, w(x) \ \ge\ c_d \varepsilon_2^d \alpha \tag{5.4}$$

for some constant $c_d > 0$. Applying the contrapositive of Theorem (1.3) for the set A' with $\varepsilon := c_d \varepsilon_2^d \alpha$ gives

$$\mathbb{E}_{x\in\mathbb{Z}_{N'}^d,\, t\in\mathbb{Z}_{N'}} \left(\prod_{j=0}^d \mathbf{1}_{A'}(x + tv_j)\right) w(x + t\Delta) \ \ge\ \delta \tag{5.5}$$

with a constant $\delta = \delta(\alpha, \Delta) > 0$ depending only on α and the simplex $\Delta = \{v_0, \ldots, v_d\}$. Similarly as in (5.3)

$$w(x + t\Delta) \le C_d \left(\frac{\phi(W)\ \log\ N}{W}\right)^{l(\Delta)}, \tag{5.6}$$

since all coordinates of $x+t\Delta$ are primes, bigger then R. Thus the number of copies $\Delta' = x+t\Delta$ which are contained in A' as a subset of $\mathbb{Z}_{N'}^d$ is at least $c\, N^{d+1}\, (\log\ N)^{-l(\Delta)}$, for some constant $c = c(\alpha, \Delta, W) > 0$

depending only on the initial data α, Δ and the number W. Since $A' \subseteq [\varepsilon_1 N', \varepsilon_2 N']^d$, by Lemma 5.1 at least half of the simplices Δ' are contained in A' as a subset of $\mathbb{Z}^d$, and then the simplices $\Delta'' := W\Delta' + b$ are contained in A.

Now choose $W = W(\alpha, \Delta)$ large enough so that Theorem 1.3 holds for all sufficiently large N, and then A contain at least $c'(\alpha, \Delta)\, N^{d+1}\, (\log\, N)^{-l(\Delta)}$ similar copies of Δ for some constant $c'(\alpha, \Delta) > 0$ depending only on α and the simplex Δ. This proves Theorem 1.2 $\square$

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 (1.5.2) for functions $F : V_e \to \mathbb{R}$. These norms have also been defined in [8], Appendix B, and in fact all the properties we prove here, including Proposition 1.1, can be deduced from the arguments given there. However as our settings is slightly different, we include the proofs below.

We will assume $e = \{1, \ldots, d\} =: [d]$, 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 $F_\omega : V_e \to \mathbb{R}$ for $\omega \in \{0,1\}^e$, be a given family of functions and define

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

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

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

Proof. We will use Cauchy-Schwartz inequality several times and the linear forms condition.

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

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

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

here,

$$A = \mathbb{E}_{x_{[2,d]},y_{[2,d]}}\Bigg[\prod_{|I|<d, 1\notin I} \prod_{\omega_I} \nu_I(\omega_I(x_I, y_I))$$

$$\times \Bigg(\mathbb{E}_{x_1,y_1} \nu(x_1)\nu(y_1) \prod_{\omega_{[2,d]}} F_{\omega(0,[2,d])}(x_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]})) F_{\omega(0,[2,d])}(y_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]}))$$

$$\times \prod_{|I|<d-1, 1\notin I} \nu_{\{1\}\cup I}(x_1, \omega_I(x_I, y_I)) \Bigg] = \left\langle F_\omega^{(0)}(\omega(x_{[d]}, y_{[d]})) \right\rangle_{\square_\nu^d},$$

where

$$F^{(0)}_{(0,\omega_{[2,d]})}(x_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]})) = F^{(0)}_{(1,\omega_{[2,d]})}(y_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]}))$$
$$:= F(x_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]}))$$

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

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

where

$$F^{(1)}_{(0,\omega_{[2,d]})}(x_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]})) = F^{(1)}_{(1,\omega_{[2,d]})}(y_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]}))$$
$$:= F(y_1, \omega_{[2,d]}(x_{[2,d]}, y_{[2,d]}))$$

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

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

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

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

□

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

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

Also 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. Let $\mathcal{H}' = \{f \in \mathcal{H};\ |f| < d$, and write the left side of (1.5.3) 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 $e_j := [d+1]\backslash\{j\}$ for the rest of the faces. The idea is to apply the Cauchy-Schwartz inequality successively in the $x_1, x_2, \ldots, x_d$ variables to eliminate the functions $F_{e_1} \leq \nu_{e_1}, \ldots, F_{e_d} \leq \nu_{e_d}$, using the linear forms condition at each step. Using $F_{e_1} \leq \nu_{e_1}$ we have

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

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

$$E^2 \lesssim \mathbb{E}_{x_2,\ldots,x_{d+1}} \nu_{e_1}(x_1) \prod_{1\notin f\in\mathcal{H}'} \nu_f(x_f)\, \mathbb{E}_{x_1,y_1} \prod_{j\neq 2} F_{e_j}(x_1, x_{e_j\backslash\{1\}}) F_{e_j}(y_1, x_{e_j\backslash\{1\}}) \tag{A.1}$$

$$\times \prod_{1\in f\in\mathcal{H}'} \nu_f(y_1, x_{f\backslash\{1\}})\, \nu_f(x_1, x_{f\backslash\{1\}})$$

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

$$G^*_e(x_{e\cap g}, y_{e\cap g}, x_{e\backslash g}) := \prod_{\omega_e\in\{0,1\}^{e\cap g}} G_e(\omega_e(x_{e\cap g}, y_{e\cap g}), x_{e\backslash g}). \tag{A.2}$$

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

$$E^{2^i} \lesssim \mathbb{E}_{x_{[i]},y_{[i]},x_{J\backslash[i]}} \prod_{j\leq i} \nu^*_{e_j}(x_{[i]\cap e_j}, y_{[i]\cap e_j}, x_{e_j\backslash[d]}) \prod_{j>i} F^*_{e_j}(x_{[i]\cap e_j}, y_{[i]\cap e_j}, x_{e_j\backslash[d]}) \tag{A.3}$$

$$\times \prod_{f\in\mathcal{H}'} \nu^*_f(x_{f\cap[i]}, y_{f\cap[i]}, x_{f\backslash[i]}). \tag{A.4}$$

For $i = 1$ this can be seem 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 $G^*_e(x_{e\cap[i]}, y_{e\cap[i]}, x_{e\backslash[i]})$ depending on it, to obtain the factor $G^*_e(x_{e\cap[i+1]}, y_{e\cap[i+1]}, x_{e\backslash[i+1]})$, thus the formula holds for $i+1$. After finishing this process we have by (A.2) and (A.3)

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

where

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

Thus, as $F_{e_0} \leq \nu_{e_0}$, to prove (1.5.3) it is enough to show that

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

This, similarly as in [7], can be done with one more application of the Cauchy-Schwartz inequality leading to 4 terms involving the "big" weight functions $\mathcal{W}$ and $\mathcal{W}^2$. Each terms is however $1 + o_{N\to\infty}(1)$ by the linear forms condition, as the underlying linear forms are pairwise linearly independent. Indeed the forms $L_f(\omega_f(x_f, y_f)$ are pairwise independent for $f \subseteq [d]$, and depend on a different set of variables then the forms $L_e(\omega_e(x_{e\cap[d]}, y_{e\cap[d]}, x_{e\backslash[d]}))$ for $e \nsubseteq [d]$ 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. □

Acknowledgements. We would like to thank Terence Tao for some helpful correspondence during this research.

REFERENCES

[1] D. CONLON, J. FOX, Y. ZHAO, *A relative Szemeredi theorem*,Geometric and Functional Analysis , to appear.
[2] B. COOK, A. MAGYAR *Constellations in Pd* International Mathematics Research Notices 2012.12 (2012), 2794-2816.
[3] H. FURSTENBERG, Y. KATZNELSON, *An ergodic Szemerédi theorem for commuting trnasformations*, J. Analyse Math. 31 (1978), 275-291
[4] J. FOX, Y. ZHAO, *A short proof of the multidimensional Szemerédi theorem in the primes.*, American Journal of Mathematics, ti appear.
[5] W.T. GOWERS, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Annals of Math. 166/3 (2007), 897-946
[6] W.T. GOWERS, *Decompositions, approximate structure, transference, and the Hahn-Banach theorem*, Bull. London Math. Soc. 42 (4) (2010), 573-606
[7] B. GREEN AND T. TAO, *The primes contain arbitrary long arithmetic progressions*, Annals of Math. 167 (2008), 481-547
[8] B. GREEN AND T. TAO, *Linear equations in primes.* Annals. of Math.(2) 171.3 (2010), 1753-1850.
[9] B. GREEN, T. TAO, *The Mobius function is asymptotically orthogonal to nilsequences*, Annals of Math., 175 (2012), 541-566
[10] B. GREEN, T. TAO, T. ZIEGLER, *An inverse theorem for the Gowers $U^{s+1}[N]$ norm*, Annals of Math., 176 (2012), no. 2, 1231-1372
[11] D. GOLDSTON, C. YILDIRIM, *Higher correlations of divisor sums related to primes I: triple correlations*, Integers: Electronic Journal of Combinatorial Number theory, 3 (2003), 1-66
[12] D. GOLDSTON, C. YILDIRIM, *Higher correlations of divisor sums related to primes III: small gaps between primes* , Proc. London Math. Soc. 95 (2007), 653-686
[13] A. MAGYAR, T. TITICHETRAKUN *Corners in dense subsets of P^d*, preprint
[14] B. NAGLE, V. RDL, M. SCHACHT, *The counting lemma for regular k-uniform hypergraphs*, Random Structures and Algorithms, 28(2), (2006), 113-179
[15] O. REINGOLD, L. TREVISAN, M. TULSIANI, S. VADHAM, *Dense subsets of pseudorandom sets* Electronic Colloquium of Computational Complexity, Report TR08-045 (2008)
[16] J. SOLYMOSI, *Note on a generalization of Roths theorem*, Discrete and Computational Geometry, Algorithms Combin. 25, (2003), 825-827
[17] E. SZEMERÉDI, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27 (1975), 299-345
[18] T. TAO, *The Gaussian primes contain arbitrarily shaped constellations*, J. Analyse Math., 99/1 (2006), 109-176
[19] T. TAO, *A variant of the hypergraph removal lemma* Journal of Combinatorial Theory, Series A 113.7 (2006): 1257-1280
[20] T. TAO AND T. ZIEGLER, *A multi-dimensional Szemeredi theorem for the primes*, Israel Journal of Mathematics, to appear.