Forbidden vector-valued intersections

Peter Keevash[*] Eoin Long[†]

September 5, 2019

Abstract

We solve a generalised form of a conjecture of Kalai motivated by attempts to improve the bounds for Borsuk's problem. The conjecture can be roughly understood as asking for an analogue of the Frankl-Rödl forbidden intersection theorem in which set intersections are vector-valued. We discover that the vector world is richer in surprising ways: in particular, Kalai's conjecture is false, but we prove a corrected statement that is essentially best possible, and applies to a considerably more general setting. Our methods include the use of maximum entropy measures, VC-dimension, Dependent Random Choice and a new correlation inequality for product measures.

1 Introduction

Intersection theorems have been a central topic of Extremal Combinatorics since the seminal paper of Erdős, Ko and Rado [9], and the area has grown into a vast body of research (see [2], [4] or [19] for an overview). The Frankl-Rödl forbidden intersection theorem is a fundamental result of this type, which has had a wide range of applications to different areas of mathematics, including discrete geometry [12], communication complexity [28] and quantum computing [6].

To state their result we introduce the following notation. Let $[n] = \{1, \ldots, n\}$ and let $\binom{[n]}{k} = \{A \subset [n] : |A| = k\}$. For $\mathcal{A} \subset \binom{[n]}{k}$ and $t \in [n]$ let $\mathcal{A} \times_t \mathcal{A}$ be the set of all $(A, B) \in \mathcal{A} \times \mathcal{A}$ with $|A \cap B| = t$. Note that $\binom{[n]}{k} \times_t \binom{[n]}{k}$ is non-empty if and only if $\max(2k - n, 0) \leq t \leq k$. Frankl and Rödl proved the following 'supersaturation theorem', showing that if t is bounded away from these extremes and $\mathcal{A}$ is 'exponentially dense' in $\binom{[n]}{k}$ then $\mathcal{A} \times_t \mathcal{A}$ is 'exponentially dense' in $\binom{[n]}{k} \times_t \binom{[n]}{k}$.

Theorem 1.1 (Frankl–Rödl [11])**.** *Let*[1] $0 < n^{-1} \ll \delta \ll \varepsilon < 1$ *and* $\max(2k - n, 0) + \varepsilon n \leq t \leq k - \varepsilon n$. *Suppose* $\mathcal{A} \subset \binom{[n]}{k}$ *with* $|\mathcal{A}| \geq (1-\delta)^n \binom{n}{k}$. *Then* $|\mathcal{A} \times_t \mathcal{A}| \geq (1-\varepsilon)^n \left| \binom{[n]}{k} \times_t \binom{[n]}{k} \right|$.

In a recent survey on the Borsuk problem, Kalai [21] remarked that the Frankl–Rödl theorem can be used to give a counterexample to the Borsuk conjecture (the Frankl–Wilson intersection theorem [13] was used in Kahn and Kalai's celebrated counterexample [20]), and suggested that improved bounds might follow from a suitably generalised Frankl–Rödl theorem. He proposed the following

[*]Mathematical Institute, University of Oxford, Oxford, UK. E-mail: keevash@maths.ox.ac.uk.
Research supported in part by ERC Consolidator Grant 647678.
[†]Mathematical Institute, University of Oxford, Oxford, UK. E-mail: long@maths.ox.ac.uk.
Supported in part by ERC Starter Grant 633509.

[1]Throughout the paper we will write $\alpha \ll \beta$ to denote that the statement is valid, provided α is sufficiently small as a function of β, i.e. $\alpha \leq \alpha(\beta)$. Thus, in Theorem 1.1 '$n^{-1} \ll \delta \ll \varepsilon$' means that for any $0 < \varepsilon < 1$ there is $\delta_0 > 0$ such that for any $0 < \delta < \delta_0$ there is n_0 such that for $n \geq n_0$ the theorem holds.

supersaturation conjecture as a possible step in this direction, in which one measures a set by its size $|A| = \sum_{i\in A} 1$ and its sum $\sum A = \sum_{i\in A} i$. Let $[n]_{k,s}$ be the set of $A \subset [n]$ with $|A| = k$ and $\sum A = s$. For $\mathcal{A} \subset [n]_{k,s}$ write

$$\mathcal{A} \times_{(t,w)} \mathcal{A} = \Big\{(A, B) \in \mathcal{A} \times \mathcal{A} : A \neq B \text{ with } |A \cap B| = t \text{ and } \sum(A \cap B) = w\Big\}.$$

Conjecture 1.2 (Kalai)**.** *Let* $0 < n^{-1} \ll \delta \ll \varepsilon, \alpha_1, \alpha_2, \beta_1, \beta_2 < 1$, $k = \lfloor \alpha_1 n \rfloor$, $s = \lfloor \alpha_2 \binom{n}{2} \rfloor$, $t = \lfloor \beta_1 n \rfloor$ *and* $w = \lfloor \beta_2 \binom{n}{2} \rfloor$. *Suppose* $\mathcal{A} \subset [n]_{k,s}$ *with* $|\mathcal{A}| \geq (1-\delta)^n |[n]_{k,s}|$. *Then* $|\mathcal{A} \times_{(t,w)} \mathcal{A}| \geq \big\lfloor (1-\varepsilon)^n \big| [n]_{k,s} \times_{(t,w)} [n]_{k,s} \big| \big\rfloor$.

Somewhat surprisingly, this conjecture is false! In fact, although the conjecture holds in a number of natural special cases, it fails quite dramatically in general; for most pairs (α_1, α_2) there is exactly one choice of (β_1, β_2) for which Conjecture 1.2 holds. Before stating this result, we first remark that Conjecture 1.2 is only non-trivial when $[n]_{k,s}$ is exponentially large in n (when $|[n]_{k,s} \times_{(t,w)} [n]_{k,s}| \geq (1-\varepsilon)^{-n}$, requiring $|[n]_{k,s}| \geq (1-\varepsilon)^{-n/2}$). Defining α_1, α_2 as in Conjecture 1.2, we can therefore assume that (α_1, α_2) belongs to

$$\Lambda := \{(x, y) : 0 < x < 1, x^2 < y < 2x - x^2\}.$$

We say that $\mathbf{g} = (\alpha_1, \alpha_2, \beta_1, \beta_2)$ is (n, δ, ε)-*Kalai* if Conjecture 1.2 holds for $\mathbf{g}$, i.e. any $\mathcal{A} \subset [n]_{k,s}$ with $|\mathcal{A}| \geq (1-\delta)^n |[n]_{k,s}|$ satisfies $|\mathcal{A} \times_{(t,w)} \mathcal{A}| \geq \big\lfloor (1-\varepsilon)^n \big| [n]_{k,s} \times_{(t,w)} [n]_{k,s} \big| \big\rfloor$. We will classify the Kalai parameters $\mathbf{g}$ in terms of the following set Γ; note that the definition of Γ_1 uses two functions $\beta_1, \beta_2 : \Lambda \to \mathbb{R}$ that will be defined in Section 10, see (9). Let $\Gamma = \bigcup_{i\in[3]} \Gamma_i$, where

(Popular intersections) $\quad \Gamma_1 = \Big\{(\alpha_1, \alpha_2, \beta_1, \beta_2) : (\alpha_1, \alpha_2) \in \Lambda,\ \alpha_1 \neq \alpha_2 \text{ and } \beta_i = \beta_i(\alpha_1, \alpha_2)\Big\};$

(Doubly random) $\quad \Gamma_2 = \Big\{(\alpha, \alpha, \beta, \beta) : 0 < \alpha < 1 \text{ and } \max(2\alpha - 1, 0) < \beta < \alpha\Big\};$

(Uniformly random sets) $\quad \Gamma_3 = \Big\{(1/2, 1/2, \beta_1, \beta_2) : (2\beta_1, 2\beta_2) \in \Lambda\Big\}.$

Theorem 1.3. *Supppose* $\mathbf{g} = (\alpha_1, \alpha_2, \beta_1, \beta_2) \in [0,1]^4$ *with* $(\alpha_1, \alpha_2) \in \Lambda$ *and* $n^{-1} \ll \delta \ll \varepsilon \ll \varepsilon' \ll \mathbf{g}$. *Let* $d(\mathbf{g}, \Gamma) = \inf\{\|\mathbf{g} - \mathbf{g}'\|_1 : \mathbf{g}' \in \Gamma\}$.

i. If $d(\mathbf{g}, \Gamma) \leq \delta$ *then* $\mathbf{g}$ *is* (n, δ, ε)*-Kalai.*

ii. If $d(\mathbf{g}, \Gamma) \geq \varepsilon'$ *and* $|[n]_{k,s} \times_{(t,w)} [n]_{k,s}| \geq (1-\varepsilon)^{-n}$ *then* $\mathbf{g}$ *is not* (n, δ, ε)*-Kalai.*

Remark 1.4. The condition $|[n]_{k,s} \times_{(t,w)} [n]_{k,s}| \geq (1-\varepsilon)^{-n}$ in Theorem 1.3 just appears to ensure that the lower bound on $|\mathcal{A} \times_{(t,w)} \mathcal{A}|$ in Conjecture 1.2 is non-trivial.

The labels assigned to the parts of Γ correspond to the following interpretations:

- **Popular intersections:** For $(\alpha_1, \alpha_2) \in \Lambda$ with $\alpha_1 \neq \alpha_2$ there is exactly one $(\alpha_1, \alpha_2, \beta_1, \beta_2) \in \Gamma_1$. For n large, the value $(\beta_1 n, \beta_2 \binom{n}{2})$ is essentially the most popular intersection between sets in $[n]_{k,s}$, where $(k, s) = (\alpha_1 n, \alpha_2 \binom{n}{2})$.

- **Doubly random:** If $A \subset [n]$ is a uniformly random set of size $k = \alpha n$, the expected size of $\sum A$ is $s = \alpha \binom{n}{2} + o(n^2)$. Similarly, if two sets A and B in $[n]_{k,s}$ with $|A \cap B| = \beta n$ are randomly selected then the expected value of $\sum(A \cap B)$ is $\beta \binom{n}{2} + o(n^2)$. Theorem 1.3 for Γ_2 shows Conjecture 1.2 holds for 'random-like βn intersections' between 'random-like αn sets', provided α and β satisfy the Frankl-Rödl conditions.

- **Uniformly random sets:** Most sets $A \subset [n]$ have $k = \frac{1}{2}n + o(n)$, $s = \frac{1}{2}\binom{n}{2} + o(n^2)$ and $|A \cap [2L]| = L \pm o(n)$ for all $L \leq n/2$. Intersections of type $(t, w) = (\beta_1 n, \beta_2 \binom{n}{2})$ can only occur between such sets if $(2\beta_1 + o(1), 2\beta_2 + o(1)) \in \Lambda$. Theorem 1.3 for Γ_3 shows that Conjecture 1.2 is true for $(\alpha_1, \alpha_2) = (1/2, 1/2)$ provided this necessary condition is fulfilled.

Remark 1.5. While Theorem 1.3 shows that Conjecture 1.2 is often false, we will also give two concrete counterexamples, which illustrate two different reasons why the conjecture fails. We defer these examples to Section 5 as their analysis uses some results from Sections 2–4.

Although the bounds from Conjecture 1.2 in general do not hold, it is still natural to ask whether we can find *any* (t, w)-intersection in such 'exponentially dense' subsets $\mathcal{A} \subset [n]_{k,s}$. If so, what is the optimal lower bound on $|\mathcal{A} \times_{(t,w)} \mathcal{A}|$? This paper investigates these questions; in particular, we give a natural correction to Conjecture 1.2.

Our results will apply to the following more general setting of vector-valued set 'sizes': given vectors $\mathcal{V} = (\mathbf{v}_i : i \in [n])$ in $\mathbb{R}^D$, we define the $\mathcal{V}$-size of $A \subset [n]$ by

$$|A|_{\mathcal{V}} = \sum_{i \in A} \mathbf{v}_i.$$

We note that the Frankl-Rödl theorem concerns $\mathcal{V}$-sizes where $D = 1$ and all $\mathbf{v}_i = 1$, and the Kalai conjecture concerns $\mathcal{V}$-sizes where $D = 2$ and $\mathbf{v}_i = (1, i)$.

1.1 Vector-valued intersections

In order to prove our forbidden $\mathcal{V}$-intersection theorem, we need to work over a general alphabet, where we associate a vector with each possible value of each coordinate, as follows.

Definition 1.6. Suppose $\mathbf{v}^i_j \in \mathbb{Z}^D$ for all $i \in [n]$ and $j \in J$. We call $\mathcal{V} = (\mathbf{v}^i_j)$ an (n, J)-array in $\mathbb{Z}^D$. For $\mathbf{a} \in J^n$ we define

$$\mathcal{V}(\mathbf{a}) = \sum_{i \in [n]} \mathbf{v}^i_{a_i}.$$

For $\mathcal{A} \subset J^n$ and $\mathbf{w} \in \mathbb{Z}^D$ we define $\mathcal{A}^{\mathcal{V}}_{\mathbf{w}} = \{\mathbf{a} \in \mathcal{A} : \mathcal{V}(\mathbf{a}) = \mathbf{w}\}$.

By identifying subsets of $[n]$ with their characteristic vectors in $\{0, 1\}^n$, and pairs of subsets of $[n]$ with vectors in $(\{0, 1\} \times \{0, 1\})^n$, this definition extends the definition of $\mathcal{V}$-size and $\mathcal{V}$-intersection via the following specialisation (note that $\mathcal{V}(A) = |A|_{\mathcal{V}}$ and $\mathcal{V}_\cap(A, B) = |A \cap B|_{\mathcal{V}}$).

Definition 1.7. Suppose $\mathcal{V} = (\mathbf{v}_i : i \in [n])$, where $\mathbf{v}_i \in \mathbb{Z}^D$ for all $i \in [n]$. We also let $\mathcal{V} = (\mathbf{v}^i_j)$ denote the $(n, \{0, 1\})$-array in $\mathbb{Z}^D$, where $\mathbf{v}^i_1 = \mathbf{v}_i$ and $\mathbf{v}^i_0 = 0$. We let $\mathcal{V}_\cap = ((\mathbf{v}_\cap)^i_{j,j'})$ denote the $(n, \{0, 1\} \times \{0, 1\})$-array in $\mathbb{Z}^D$, where $(\mathbf{v}_\cap)^i_{1,1} = \mathbf{v}_i$ and $(\mathbf{v}_\cap)^i_{j,j'} = 0$ otherwise.

We also introduce a class of norms on $\mathbb{R}^D$ to account for the possibility that different coordinates of vectors in $\mathcal{V}$ may operate at different scales. In the following definition we think of $\mathbf{R}$ as a scaling; e.g. for the Kalai vectors $(1, i)$, we take $\mathbf{R} = (1, n)$.

Definition 1.8. Suppose $\mathbf{R} = (R_1, \ldots, R_D) \in \mathbb{R}^D$. We define the $\mathbf{R}$-norm on $\mathbb{R}^D$ by $\|\mathbf{v}\|_{\mathbf{R}} = \max_{d \in [D]} |v_d|/R_d$. We say that $\mathcal{V} = (\mathbf{v}^i_j)$ is $\mathbf{R}$-bounded if all $\|\mathbf{v}^i_j\|_{\mathbf{R}} \leq 1$.

Our $\mathcal{V}$-intersection theorem requires two properties of the set of vectors $\mathcal{V}$. The first property, roughly speaking, says that any vector in $\mathbb{Z}^D$ can be efficiently generated by changing the values of coordinates, and that furthermore this holds even if a small set of coordinates are frozen, so that no coordinate is overly significant. To see why such a condition is necessary, suppose that $D = 1$ and almost all coordinates have only even values: then there are large families where all intersections have a fixed parity.

Definition 1.9. Let $\mathcal{V} = (\mathbf{v}^i_j)$ be an (n, J)-array in $\mathbb{Z}^D$. We say that $\mathcal{V}$ is γ-robustly $(\mathbf{R}, k)$-generating in $\mathbb{Z}^D$ if for any $\mathbf{v} \in \mathbb{Z}^D$ with $\|\mathbf{v}\|_{\mathbf{R}} \leq 1$ and $T \subset [n]$ with $|T| \leq \gamma n$ there is $S \subset [n] \setminus T$ with $|S| \leq k$ and $j_i, j'_i \in J$ for all $i \in S$ such that $\mathbf{v} = \sum_{i \in S}(\mathbf{v}^i_{j_i} - \mathbf{v}^i_{j'_i})$.

Note that if $\mathcal{V} = (\mathbf{v}_i : i \in [n])$, considered as an $(n, \{0, 1\})$-array, then Definition 1.9 says that for all such $\mathbf{v}$ and T there are disjoint $S, S' \subset [n] \setminus T$ with $|S| + |S'| \leq k$ such that $\mathbf{v} = \sum_{i \in S} \mathbf{v}_i - \sum_{i \in S'} \mathbf{v}_i$. To illustrate the definition, and for future reference, we note that

$$\text{the Kalai vectors are 0.1-robustly } ((1, n), 7)\text{-generating in } \mathbb{Z}^2. \tag{1}$$

To see this, first note that for any vector $(0, b)$ with $b \in [n/2]$, there are $n/3$ disjoint pairs $\{i_1, i_2\}$ with $(0, b) = (1, i_1) - (1, i_2)$. This implies that for any vector $(0, b)$ with $|b| \leq n$ there are $n/9$ disjoint sets $\{i_1, i_2, j_1, j_2\}$ with $(0, b) = (1, i_1) + (1, i_2) - (1, j_1) - (1, j_1)$. Also, there are $n/6$ disjoint triples $\{i_1, i_2, i_3\}$ with $(1, 0) = (1, i_1) + (1, i_2) - (1, i_3)$. Combined, given $T \subset [n]$ with $|T| \leq n/10 < n/9 - 3$ and $(a, b) \in \mathbb{Z}^2$ with $\|(a, b)\|_{\mathbf{R}} \leq 1$, there are disjoint $S, S' \subset [n] \setminus T$ with $|S| + |S'| \leq 7$ with $(a, b) = \sum_{i \in S} \mathbf{v}_i - \sum_{i \in S'} \mathbf{v}_i$. Thus (1) holds.

We also make the following 'general position' assumption for $\mathcal{V}$.

Definition 1.10. Suppose $\mathcal{V} = (\mathbf{v}_i)$ is an $(n, \{0, 1\})$-array in $\mathbb{Z}^D$. For $I \in \binom{[n]}{D}$, let $\mathcal{V}_I = \{\mathbf{v}_i : i \in I\}$ and say that I is $(\gamma, \mathbf{R})$-generic if $|\det(\mathcal{V}_I)| \geq \gamma \prod_{d \in [D]} R_d$. We say that $\mathcal{V}$ is γ'-robustly $(\gamma, \mathbf{R})$-generic if for any $X \subset [n]$ with $|X| > \gamma' n$, some $I \subset X$ is $(\gamma, \mathbf{R})$-generic for $\mathcal{V}$.

We note for future reference that

$$\text{the Kalai vectors are } \gamma\text{-robustly } (\gamma/2, \mathbf{R})\text{-generic for any } \gamma > 0. \tag{2}$$

Indeed, if $X \subset [n]$ with $|X| \geq \gamma n$ then we can choose $i, i' \in X$ with $|i - i'| \geq \gamma n - 1 \geq \gamma n/2$, and then $(1, i)$ and $(1, i')$ span a parallelogram of area $|i - i'| \geq (\gamma/2) \cdot 1 \cdot n$.

We are now in a position to state our main theorem. It shows that, under the above assumptions on $\mathcal{V}$, there are only two obstructions to a set $\mathcal{X} = (\{0, 1\}^n)^{\mathcal{V}}_{\mathbf{z}}$ satisfying a supersaturation result as in Kalai's conjecture (case *i*): either (case *ii*) there is a small set $\mathcal{B}_{full} \subset \mathcal{X}$ responsible for almost all $\mathbf{w}$-intersections in $\mathcal{X}$, or (case *iii*) there is a large set $\mathcal{B}_{empty} \subset \mathcal{X}$ containing no $\mathbf{w}$-intersections. Furthermore, in case *ii* we obtain optimal supersaturation relative to $\mathcal{B}_{full}$.

Theorem 1.11. *Let $n^{-1} \ll \delta \ll \gamma_1, \gamma'_1 \ll \gamma_2, \gamma'_2 \ll \varepsilon, D^{-1}, C^{-1}, k^{-1}$ and $\mathbf{R} \in \mathbb{R}^D$ with $\max_d R_d \leq n^C$. Suppose $\mathcal{V} = (\mathbf{v}_i : i \in [n])$ where each $\mathbf{v}_i \in \mathbb{Z}^D$ is $\mathbf{R}$-bounded and $\mathcal{V}$ is γ'_j-robustly $(\gamma_j, \mathbf{R})$-generic and γ_j-robustly $(\mathbf{R}, k)$-generating for $j = 1, 2$. Let $\mathbf{z}, \mathbf{w} \in \mathbb{Z}^D$ with $\mathbf{z} \neq \mathbf{w}$ and let $\mathcal{X} = (\{0, 1\}^n)^{\mathcal{V}}_{\mathbf{z}}$. Then one of the following holds:*

i. All $\mathcal{A} \subset \mathcal{X}$ with $|\mathcal{A}| \geq (1 - \delta)^n |\mathcal{X}|$ satisfy $|(\mathcal{A} \times \mathcal{A})^{\mathcal{V}_\cap}_{\mathbf{w}}| \geq (1 - \varepsilon)^n |(\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}|$.

ii. There exists $\mathcal{B}_{full} \subset \mathcal{X}$ with $|\mathcal{B}_{full}| \leq (1 - \delta)^n |\mathcal{X}|$ satisfying

$$|(\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}} \setminus (\mathcal{B}_{full} \times \mathcal{B}_{full})^{\mathcal{V}_\cap}_{\mathbf{w}}| \leq (1 - \delta)^n |(\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}|.$$

iii. There is $\mathcal{B}_{empty} \subset \mathcal{X}$ with $|\mathcal{B}_{empty}| \ge \lfloor (1-\varepsilon)^n |\mathcal{X}| \rfloor$ satisfying $(\mathcal{B}_{empty} \times \mathcal{B}_{empty})^{\mathcal{V}_\cap}_{\mathbf{w}} = \emptyset$. Furthermore, if ii holds and iii does not then any $\mathcal{B} \subset \mathcal{B}_{full}$ with $|\mathcal{B}| \ge (1-\delta)^n |\mathcal{B}_{full}|$ satisfies $|(\mathcal{B} \times \mathcal{B})^{\mathcal{V}_\cap}_{\mathbf{w}}| \ge (1-\varepsilon)^n |(\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}|$.

Remark 1.12.

i. Theorem 1.11 applies to (t, w)-intersections in $[n]_{k,s}$, as we have shown above that its hypotheses hold for the Kalai vectors.
ii. As indicated above, cases *ii* and *iii* of Theorem 1.11 may simultaneously hold (see counterexample 1 of Section 5).
iii. The assumption that $\mathcal{V}$ is γ_1-robustly $(\mathbf{R}, k)$-generating is redundant, as it is implied by γ_2-robustly $(\mathbf{R}, k)$-generating, but the assumptions of γ'_j-robustly $(\gamma_j, \mathbf{R})$-generic for $j = 1, 2$ are incomparable, and our proof seems to require this 'multiscale general position'.

We have highlighted Theorem 1.11 as our main result for the sake of giving a clean combinatorial statement. However, we will in fact obtain considerably more general results in two directions, whose precise statements are postponed until later in the paper.

- Our most general result, Theorem 6.3, implies cross-intersection theorems for two or more families and applies to families of vectors over any finite alphabet.
- Theorem 1.11 leaves open the question of how many $\mathbf{w}$-intersections are guaranteed in large subsets of $\mathcal{X}$ when case (ii) holds; this is answered by Theorem 11.1.

It is natural to ask under which conditions the alternate cases of Theorem 1.11 hold. These conditions are best understood in relation to our proof framework, so we postpone this discussion to section 1.4, after we have introduced the two principal components of the proof.

1.2 A probabilistic forbidden intersection theorem

A key paradigm of our approach is that $\mathcal{V}$-intersection theorems often have equivalent formulations in terms of certain product measures (the maximum entropy measures described in the next subsection), and that the necessary condition for these theorems appears naturally as a condition on the product measures. (A similar idea arose in the new proof of the density Hales-Jewett theorem developed by the first Polymath project [26], although in this case the natural 'equal slices' distribution was not a product measure.)

To illustrate this point, we recast the Frankl-Rödl theorem in such terms. Again we identify subsets of $[n]$ with their characteristic vectors in $\{0,1\}^n$, on which we introduce the product measure $\mu_p(\mathbf{x}) = \prod_{i \in [n]} p_{x_i}$, where $p_1 = k/n$ and $p_0 = 1 - p_1$. Pairs of subsets are identified with $\{0,1\}^n \times \{0,1\}^n$, which we can identify with $(\{0,1\} \times \{0,1\})^n$, on which we introduce the product measure $\mu_q(\mathbf{x}, \mathbf{x}') = \prod_{i \in [n]} q_{x_i, x'_i}$, where $q_{1,1} = t/n$, $q_{0,1} = q_{1,0} = (k-t)/n$ and $q_{0,0} = (n - 2k + t)/n$. It follows from our general large deviation principle in the next subsection (or is easy to see directly in this case) that the hypothesis of Theorem 1.1 is essentially equivalent to $\mu_p(\mathcal{A}) > (1-\delta)^n$ and the conclusion to $\mu_q(\mathcal{A} \times_t \mathcal{A}) > (1-\varepsilon)^n$. Furthermore, the assumption on t can be rephrased as $q_{j,j'} \ge \varepsilon$ for all $j, j' \in \{0,1\}$, and this indicates the condition that we need in general.

Let us formalise the above discussion of product measures in a general context. Although we only considered the cases when the 'alphabet' J is $\{0,1\}$ or $\{0,1\} \times \{0,1\}$, we remark that it is essential for our arguments to work with general alphabets, as the proofs of our results even in the binary case rely on reductions that increase the alphabet size.

Definition 1.13. Suppose $\mathbf{p} = (p^i_j : i \in [n], j \in J)$ with all $p^i_j \in [0,1]$ and $\sum_{j\in J} p^i_j = 1$ for all $i \in [n]$. The product measure $\mu_{\mathbf{p}}$ on J^n is given, for $\mathbf{a} \in J^n$, by $\mu_{\mathbf{p}}(\mathbf{a}) = \prod_{i\in[n]} p^i_{a_i}$.

Given an (n, J)-array $\mathcal{V}$ and a measure μ on J^n, we write $\mathcal{V}(\mu) = \mathbb{E}_{\mathbf{a}\sim\mu}\mathcal{V}(\mathbf{a})$.

Suppose $\mu_{\mathbf{q}}$ is a product measure on $(\prod_{s\in S} J_s)^n$, with $\mathbf{q} = (q^i_{\mathbf{j}} : i \in [n], \mathbf{j} \in \prod_{s\in S} J_s)$. For $s \in S$ the s-marginal of $\mu_{\mathbf{q}}$ is the product measure $\mu_{\mathbf{p}_s}$ on J^n_s with $(p_s)^i_j = \sum q^i_{\mathbf{j}}$ for all $i \in [n]$, $j \in J_s$, where the sum is over all $\mathbf{j}$ with $j_s = j$.

We say that $\mu_{\mathbf{q}}$ has marginals $(\mu_{\mathbf{p}_s} : s \in S)$. We say that $\mu_{\mathbf{q}}$ is κ-bounded if $q^i_{\mathbf{j}} \in [\kappa, 1-\kappa]$ for all $\mathbf{j} \in \prod_{s\in S} J_s$. Note that if $\mu_{\mathbf{q}}$ is κ-bounded then so are its marginals.

Remark 1.14. We will often simply write $q^i_{j_1,j_2}$ for $q^i_{(j_1,j_2)}$, etc.

A rough statement of our probabilistic forbidden intersection theorem (Theorem 1.17 below) is that if $\mathcal{A}$ has 'large measure' then the set of $\mathbf{w}$-intersections in $\mathcal{A}$ has 'large measure'. We will combine this with an equivalence of measures discussed in the next subsection to deduce our main theorem. First will highlight two special cases of Theorem 1.17 that have independent interest. The first is the following result, which ignores the intersection conditions, and is only concerned with the relationship between the measures of $\mathcal{A}$ and $\mathcal{A} \times \mathcal{A}$; it is a new form of correlation inequality (see Theorem 7.1 for a more general statement that applies to several families defined over general alphabets).

Theorem 1.15. *Let $0 < n^{-1}, \delta \ll \kappa, \varepsilon < 1$ and $\mu_{\mathbf{q}}$ be a κ-bounded product measure on $(\{0,1\} \times \{0,1\})^n$ with both marginals $\mu_{\mathbf{p}}$. Suppose $\mathcal{A} \subset \{0,1\}^n$ with $\mu_{\mathbf{p}}(\mathcal{A}) > (1-\delta)^n$. Then $\mu_{\mathbf{q}}(\mathcal{A} \times \mathcal{A}) > (1-\varepsilon)^n$.*

Next we consider the problem of finding $\mathcal{V}$-intersections that are *close* to $\mathbf{w}$, which is also natural, and somewhat easier than finding $\mathcal{V}$-intersections that are (exactly) $\mathbf{w}$. We require some notation. For $r > 0$ let $B_{\mathbf{R}}(\mathbf{w}, r) = \{\mathbf{w}' \in \mathbb{Z}^D : \|\mathbf{w} - \mathbf{w}'\|_{\mathbf{R}} \le r\}$. For $\mathcal{A} \subset \mathcal{P}[n]$ and $L \subset \mathbb{Z}^D$ let $(\mathcal{A} \times \mathcal{A})^{\mathcal{V}_\cap}_L = \{(A, B) \in \mathcal{A} \times \mathcal{A} : \mathcal{V}_\cap(A, B) \in L\}$.

Theorem 1.16. *Let $0 < n^{-1}, \delta \ll \zeta \ll \kappa, \varepsilon \ll D^{-1}$ and $\mathbf{R} \in \mathbb{Z}^D$. Suppose that*

i. $\mu_{\mathbf{q}}$ is a κ-bounded product measure on $(\{0,1\} \times \{0,1\})^n$ with both marginals $\mu_{\mathbf{p}}$,

ii. $\mathcal{V} = (\mathbf{v}_i : i \in [n])$ is an $\mathbf{R}$-bounded array in $\mathbb{Z}^D$,

iii. $\mathcal{V}_\cap(\mu_{\mathbf{q}}) = \mathbf{w} \in \mathbb{R}^D$ and $L := B_{\mathbf{R}}(\mathbf{w}, \zeta n)$.

Then any $\mathcal{A} \subset \{0,1\}^n$ with $\mu_{\mathbf{p}}(\mathcal{A}) > (1-\delta)^n$ satisfies $\mu_{\mathbf{q}}((\mathcal{A} \times \mathcal{A})^{\mathcal{V}_\cap}_L) > (1-\varepsilon)^n$.

Theorem 1.16 naturally fits into the wide literature on forbidden L-intersections in extremal set theory (see [2], [4] or [19]). Here one aims to understand how large certain families of sets can be if all intersections between elements of $\mathcal{A}$ are restricted to lie in some set L. For example, the Erdős-Ko-Rado theorem [9] can be viewed as an L_0-intersection theorem for families $\mathcal{A} \subset \binom{n}{k}$, where $L_0 = \{l \in \mathbb{N} : 1 \le l \le k\}$. Similarly, Katona's t-intersection theorem [22] can be viewed as an $L_{\ge t}$-intersection theorem for families $\mathcal{A} \subset \mathcal{P}[n]$, where $L_{\ge t} = \{l \in \mathbb{N} : l \ge t\}$. We also note that when $D = 1$ the set $L = B_{\mathbf{R}}(\mathbf{w}, \zeta n)$ in Theorem 1.16 is simply an interval, and this case naturally arose in Frankl and Rödl's original proof of Theorem 1.1.

Now we state our probabilistic forbidden intersection theorem: if $\mathcal{V}$ is robustly generating then Theorem 1.16 can be upgraded to find fixed $\mathcal{V}$-intersections.

Theorem 1.17. *Let $0 < n^{-1}, \delta \ll \zeta \ll \kappa, \gamma, \varepsilon \ll D^{-1}, C^{-1}, k^{-1}$ and $\mathbf{R} \in \mathbb{Z}^D$ with $\max_d R_d < n^C$. Suppose that* *i.* $\mu_{\mathbf{q}}$ *is a* κ*-bounded product measure on* $(\{0,1\} \times \{0,1\})^n$ *with both marginals* $\mu_{\mathbf{p}}$*,*

ii. $\mathcal{V} = (\mathbf{v}_i : i \in [n])$ *is* $\mathbf{R}$*-bounded and* γ*-robustly* $(\mathbf{R}, k)$*-generating in* $\mathbb{Z}^D$*,*

iii. $\mathbf{w} \in \mathbb{Z}^D$ *with* $\|\mathbf{w} - \mathcal{V}_\cap(\mu_{\mathbf{q}})\|_{\mathbf{R}} < \zeta n$*.*

Then any $\mathcal{A} \subset \{0,1\}^n$ *with* $\mu_{\mathbf{p}}(\mathcal{A}) > (1-\delta)^n$ *satisfies* $\mu_{\mathbf{q}}((\mathcal{A} \times \mathcal{A})^{\mathcal{V}_\cap}_{\mathbf{w}}) > (1-\varepsilon)^n$*.*

1.3 Maximum entropy and large deviations

Next we will discuss an equivalence of measures that will later combine with Theorem 1.17 to yield Theorem 1.11. Here we are guided by the maximum entropy principle (proposed by Jaynes [18] in the context of Statistical Mechanics) which suggests considering the distribution with maximum entropy[2] subject to the constraints of our problem, as defined in the following lemma (the proof is easy and appears in Lemma 2.6 in Section 2).

Lemma 1.18. *Suppose* $\mathcal{V} = (\mathbf{v}^i_j)$ *is an* (n, J)*-array in* $\mathbb{Z}^D$ *and* $\mathbf{w} \in \mathbb{Z}^D$*. Let* $\mathcal{M}^{\mathcal{V}}_{\mathbf{w}}$ *be the set of probability measures* μ *on* J^n *such that* $\mathcal{V}(\mu) = \mathbf{w}$*. Then, provided* $\mathcal{M}^{\mathcal{V}}_{\mathbf{w}}$ *is non-empty, there is a unique distribution* $\mu^{\mathcal{V}}_{\mathbf{w}} \in \mathcal{M}^{\mathcal{V}}_{\mathbf{w}}$ *with* $H(\mu^{\mathcal{V}}_{\mathbf{w}}) = \max_{\mu \in \mathcal{M}^{\mathcal{V}}_{\mathbf{w}}} H(\mu)$*, and* $\mu^{\mathcal{V}}_{\mathbf{w}}$ *is a product measure* $\mu_{\mathbf{p}^{\mathcal{V}}_{\mathbf{w}}}$ *on* J^n*, where* $\sum_{i \in [n], j \in J} (p^{\mathcal{V}}_{\mathbf{w}})^i_j \mathbf{v}^i_j = \mathbf{w}$*.*

We will show that $\mu^{\mathcal{V}}_{\mathbf{w}}$ is equivalent to the uniform measure on $(J^n)^{\mathcal{V}}_{\mathbf{w}}$, in the sense of exponential contiguity, defined as follows. (It is reminiscent of, but distinct from, the more well-known theory of contiguity, see [17, Section 9.6].)

Definition 1.19. Let $\mu = (\mu_n)_{n \in \mathbb{N}}$ and $\mu' = (\mu'_n)_{n \in \mathbb{N}}$, where μ_n and μ'_n are probability measures on a finite set Ω_n for all $n \in \mathbb{N}$. Let $\mathcal{F} = (\mathcal{F}_n)_{n \in \mathbb{N}}$ where each $\mathcal{F}_n$ is a set of subsets of Ω_n.

We say that μ' exponentially dominates μ relative to $\mathcal{F}$, and write $\mu \lesssim_{\mathcal{F}} \mu'$, if for $n^{-1} \ll \delta \ll \varepsilon \ll 1$ and $A_n \in \mathcal{F}_n$ with $\mu_n(A_n) > (1-\delta)^n$ we have $\mu'_n(A_n) > (1-\varepsilon)^n$. We say that μ and μ' are exponentially contiguous relative to $\mathcal{F}$, and write $\mu \approx_{\mathcal{F}} \mu'$ if $\mu \lesssim_{\mathcal{F}} \mu'$ and $\mu' \lesssim_{\mathcal{F}} \mu$.

If $\Delta = (\Delta_n)_{n \in \mathbb{N}}$ with each $\Delta_n \subset \Omega_n$ then we write $\mu \lesssim_\Delta \mu'$ if $\mu \lesssim_{\mathcal{F}} \mu'$, where $\mathcal{F}_n$ is the set of all subsets of Δ_n; we define $\mu \approx_\Delta \mu'$ similarly.

Note that $\lesssim_{\mathcal{F}}$ is a partial order and $\approx_{\mathcal{F}}$ is an equivalence relation.

The following result establishes the required equivalence of measures under the same hypotheses as in the previous subsection. It can be regarded as a large deviation principle for conditioning $\mathbf{x} \in J^n$ on the event $\mathcal{V}(\mathbf{x}) = \mathbf{w}$ (see [8] for an overview of this area).

Theorem 1.20. *Let* $0 < n^{-1} \ll \gamma, \kappa, k^{-1}, D^{-1}, C^{-1} < 1$ *and* $\mathbf{R} \in \mathbb{R}^D$ *with* $\max_d R_d < n^C$*. Suppose* $\mathcal{V} = (\mathbf{v}^i_j)$ *is an* $\mathbf{R}$*-bounded* γ*-robustly* $(\mathbf{R}, k)$*-generating* (n, J)*-array in* $\mathbb{Z}^D$*, and* $\mathbf{w} \in \mathbb{Z}^D$ *such that* $\mu_{\mathbf{p}^{\mathcal{V}}_{\mathbf{w}}}$ *is* κ*-bounded. Let* ν *be the uniform distribution on* $\Delta_n := (J^n)^{\mathcal{V}}_{\mathbf{w}}$*. Then* $\mu_{\mathbf{p}^{\mathcal{V}}_{\mathbf{w}}} \approx_\Delta \nu$*.*

To apply Theorem 1.20 under combinatorial conditions, we will use the following lemma which shows that $\mu_{\mathbf{p}^{\mathcal{V}}_{\mathbf{w}}}$ is κ-bounded under our general position condition on $\mathcal{V}$. (See also Section 4 for a more general result based on VC-dimension that applies to larger alphabets.)

Lemma 1.21. *Let* $0 < n^{-1} \ll \kappa \ll \gamma, \gamma' \ll \alpha, D^{-1}$*. Suppose* $\mathcal{V} = (\mathbf{v}^i)$ *is an* $\mathbf{R}$*-bounded* γ'*-robustly* $(\gamma, \mathbf{R})$*-generic* $(n, \{0,1\})$*-array in* $\mathbb{Z}^D$ *and* $|(\{0,1\}^n)^{\mathcal{V}}_{\mathbf{w}}| \geq (1+\alpha)^n$*. Then* $\mu^{\mathcal{V}}_{\mathbf{w}}$ *is* κ*-bounded.*

Alexander Barvinok remarked (personal communication) that similar results to Theorem 1.20 and Lemma 1.21 were obtained by Barvinok and Hartigan in [3]. Theorem 3 of [3] gives stronger

[2] The entropy of a distribution μ defined on a finite set X is $H(\mu) = \sum_{x \in X} -\mu(x) \log_2 \mu(x)$.

bounds on $|(\{0,1\}^n)^{\mathcal{V}}_{\mathbf{w}}|$ where applicable, but their assumptions are very different to ours (they assume bounds for quadratic forms of certain inertia tensors), and they also require that the vectors all operate at the 'same scale', so their results do not apply to the Kalai vectors. Although our bounds are weaker, our proofs are considerably shorter, and furthermore, stronger bounds here would not give any improvements elsewhere in our paper, as they account for a term subexponential in n, while our working tolerance is up to a term exponential in n.

1.4 Supersaturation

We now give a brief overview of the strategy for combining the results of the previous two subsections to prove supersaturation, and also indicate the conditions that determine which case of Theorem 1.11 holds. Under the set up of Theorem 1.11, a telegraphic summary of the argument is:

$$|\mathcal{A}| \geq (1-\delta)^n|\mathcal{X}| \overset{\text{Theorem 1.20}}{\Longrightarrow} \mu_{\mathbf{p}^{\mathcal{V}}_{\mathbf{z}}}(\mathcal{A}) \geq (1-\delta')^n \overset{\text{Theorem 1.17}}{\Longrightarrow} \mu_{\mathbf{q}}((\mathcal{A}\times\mathcal{A})^{\mathcal{V}_\cap}_{\mathbf{w}}) \geq (1-\varepsilon)^n,$$

where $\mu_{\mathbf{q}}$ is chosen to optimise the lower bound on $|(\mathcal{A}\times\mathcal{A})^{\mathcal{V}_\cap}_{\mathbf{w}}|$ implied by the final inequality.

The best possible supersaturation bound (case i of Theorem 1.11) arises when Theorem 1.17 is applicable with $\mu_{\mathbf{q}}$ equal to the maximum entropy measure $\mu_{\widetilde{\mathbf{q}}}$ that represents $(\mathcal{X}\times\mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}$: this case holds when $\mu_{\widetilde{\mathbf{q}}}$ is κ-bounded and has marginals $\mu_{\widetilde{\mathbf{p}}}$ close to $\mu_{\mathbf{p}} := \mu_{\mathbf{p}^{\mathcal{V}}_{\mathbf{z}}}$.

Case ii of Theorem 1.11 holds if $\mu_{\widetilde{\mathbf{q}}}$ is κ-bounded but $\mu_{\widetilde{\mathbf{p}}}$ is not close to $\mu_{\mathbf{p}}$: then $\mu_{\widetilde{\mathbf{p}}}$ is concentrated on a small subset $\mathcal{B}_{full}$ of $\mathcal{X}$, which is responsible for almost all $\mathbf{w}$-intersections in $\mathcal{X}$.

Lastly, case iii of Theorem 1.11 holds if $\mu_{\widetilde{\mathbf{q}}}$ is not κ-bounded. The key to understanding this case is the well-known [31] Vapnik-Chervonenkis dimension, defined as follows.

Definition 1.22. We say that $\mathcal{A}\subset J^n$ shatters $X\subset[n]$ if for any $(j_x : x\in X)\in J^X$ there is $\mathbf{a}\in\mathcal{A}$ with $a_x = j_x$ for all $x\in X$. The VC-dimension $\dim_{VC}(\mathcal{A})$ of $\mathcal{A}$ is the largest size of a subset of $[n]$ shattered by $\mathcal{A}$.

To see why it is natural to consider the VC-dimension, consider the problem of finding an intersection of size $n/3$ among subsets of $[n]$ of size $2n/3$. The conditions of the Frankl-Rödl theorem are not satisfied with these parameters, and indeed the conclusion that every family from $\binom{[n]}{2n/3}$ of density $(1-o(1))^n$ contains an $n/3$ intersection is not true: take $\mathcal{A} = \{A\in\binom{[n]}{2n/3} : 1\notin A\}$. Considering $\binom{[n]}{2n/3}\times_{n/3}\binom{[n]}{2n/3}$ as a subset of $(\{0,1\}\times\{0,1\})^n$, we see that no coordinate can take the value $(0,0)$, so there is not even a shattered set of size 1! Modifying this example in the obvious way we see that it is natural to assume a bound that is linear in n. We also note that this example shows that the 'Frankl-Rödl analogue' of Conjecture 1.2 is not true (i.e. although $\binom{[n]}{k}\times_t\binom{[n]}{k}$ is exponentially large, there are exponentially dense subsets of $\binom{[n]}{k}$ with no t intersections) and hints towards a counterexample for Kalai's conjecture. More generally, we will prove that κ-boundedness of $\mu_{\mathbf{q}}$ is roughly equivalent to the VC-dimension of $(\mathcal{X}\times\mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}$ being large as a subset of $(\{0,1\}\times\{0,1\})^n$ (see Lemma 4.8). Case iii of Theorem 1.11 will apply when $(\mathcal{X}\times\mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}$ has low VC-dimension.

The above outline also gives some indication of how the values in Theorem 1.3 arise. As described above, the supersaturation conclusion desired by Conjecture 1.2 (case i of Theorem 1.11) needs $\mu_{\widetilde{\mathbf{q}}}$ to have marginals $\mu_{\widetilde{\mathbf{p}}}$ close to $\mu_{\mathbf{p}} := \mu_{\mathbf{p}^{\mathcal{V}}_{\mathbf{z}}}$. We can describe $\mu_{\widetilde{\mathbf{q}}}$ and $\mu_{\mathbf{p}}$ explicitly using Lagrange multipliers: they are Boltzmann distributions (see Lemma 10.1). In general, it is not possible for one Boltzmann distribution to be a marginal of another, which explains why Conjecture 1.2 is generally false. An analysis of the special conditions under which it is possible gives rise to the characterisation of Γ in Theorem 1.3.

The outline also suggests a possible characterisation of the optimal level of supersaturation in all cases (i.e. including those for which Kalai's conjecture fails). Any choice of $\mu_{\mathbf{q}}$ satisfying the hypotheses of Theorem 1.17 with marginal distributions $\mu^{\mathcal{V}}_{\mathbf{z}}$ gives a lower bound on $|(\mathcal{A} \times \mathcal{A})^{\mathcal{V}_\cap}_{\mathbf{w}}|$, and the optimal such lower bound is obtained by taking such a measure with maximum entropy. Is this essentially tight? We wil give a positive answer to this question by proving a matching upper bound in Section 11.

Finally, we remark that our method allows different vectors defining the sizes of intersections from those defining the sizes of sets in the family, i.e. $\mathcal{V}'$-intersections in $(\{0,1\}^n)^{\mathcal{V}}_{\mathbf{z}}$; in Section 6.3 we show such an application to give a new proof of a theorem of Frankl and Rödl [11, Theorem 1.15] on intersection patterns in sequence spaces.

1.5 Organisation of the paper

In the next section we collect some probabilistic methods that will be used throughout the paper. We prove the large deviation principle (Theorem 1.20) in Section 3. In Section 4 we establish the connection between VC-dimension and boundedness of maximum entropy measures. Section 5 is expository: we give two concrete counterexamples to Kalai's Conjecture 1.2. As mentioned above, the reader may wish to first read this section and then return to fill in the details. Next we introduce a more general setting in Section 6, state our most general result (Theorem 6.3), and show that it implies our probabilistic intersection theorem (Theorem 1.17). In Section 7 we prove a correlation inequality needed for the proof of Theorem 6.3; as far as we are aware, the inequality is quite unlike other such inequalities in the literature. We prove Theorem 6.3 in Section 8, and then deduce our main theorem (1.11) in Section 9. Our corrected form of Kalai's conjecture (Theorem 1.3) is proved in Section 10; we also show here in much more generality that supersaturation of the form conjectured by Kalai is rare. In Section 11 we give a complete characterisation of the optimal level of supersaturation in terms of a certain optimisation problem for measures. Lastly, in section 12 we recast our results in terms of 'exponential continuity': a notion that arises naturally when comparing distributions according to exponential contiguity, and may be interpreted in terms of robust statistics for social choice: this point and several potential directions for future research are addressed in the concluding remarks.

1.6 Notation

We identify subsets of a set with their characteristic vectors: $A \subset X$ corresponds to $\mathbf{a} \in \{0,1\}^X$, where $a_i = 1 \Leftrightarrow i \in A$. The Hamming distance between vectors $\mathbf{a}$ and $\mathbf{a}'$ in a product space J^n is $d(\mathbf{a}, \mathbf{a}') = |\{i \in [n] : a_i \neq a'_i\}|$. Given a set X, we write $\binom{X}{k} = \{A \subset X : |A| = k\}$. We write $\delta \ll \varepsilon$ to mean for any $\varepsilon > 0$ there exists $\delta_0 > 0$ such that for any $\delta \leq \delta_0$ the following statement holds. Statements with more constants are defined similarly. We write $a = b \pm c$ to mean $b - c \leq a \leq b + c$. Throughout the paper we omit floor and ceiling symbols where they do not affect the argument. All vectors appear in boldface.

2 Probabilistic methods

In this section we gather several probabilistic methods that will be used throughout the paper: concentration inequalities, entropy, an application of Dependent Random Choice to the independence number of product graphs, and an alternative characterisation of exponential contiguity.

2.1 Concentration inequalities

We start with the well-known Chernoff bound (see e.g. [1, Appendix A]).

Lemma 2.1 (Chernoff's inequality)**.** *Suppose $t \geq 0$ and $X := \sum_{i\in[n]} X_i$, where $X_1, \dots, X_n$ are independent random variables with $|X_i - \mathbb{E}X_i| \leq a_i$ for all $i \in [n]$. Then $\mathbb{P}(|X - \mathbb{E}X| \geq t) \leq 2e^{-t^2/(2\sum_{i=1}^n a_i^2)}$.*

An easy consequence is the following concentration inequality for random sums of vectors.

Lemma 2.2. *Suppose $\mu_{\mathbf{p}}$ is a product measure on J^n, and $\mathcal{V} = (\mathbf{v}_j^i)$ is an $\mathbf{R}$-bounded (n, J)-array in $\mathbb{Z}^D$. Let $X = \mathcal{V}(\mathbf{a})$ with $\mathbf{a} \sim \mu_{\mathbf{p}}$ and $t \geq 0$. Then $\mathbb{P}(\|X - \mathbb{E}X\|_{\mathbf{R}} \geq t) \leq 2De^{-t^2/8n}$.*

Proof. For each $d \in D$, we have $X_d = \sum_{i\in[n]} X_{d,i}$, where $X_{d,i}(\mathbf{a}) = v^i_{a_i,d}$ are independent random variables with $|X_{d,i}| \leq R_d$ for all $i \in [n]$. By Chernoff's inequality we have $\mathbb{P}(|X_d - \mathbb{E}X_d| \geq tR_d) \leq 2e^{-t^2/8n}$, so the lemma follows from a union bound. □

We will also use the following consequence of Azuma's martingale concentration inequality (see e.g. [25]). We say that $f : J^n \to \mathbb{R}$ is *b-Lipschitz* if for any $\mathbf{a}, \mathbf{a}' \in J^n$ differing only in a single coordinate we have $|f(\mathbf{a}) - f(\mathbf{a}')| \leq b$.

Lemma 2.3. *Suppose $Z = (Z_1, \dots, Z_n)$ is a sequence of independent random variables, and $X = f(Z)$, where f is b-Lipschitz. Then $\mathbb{P}(|X - \mathbb{E}X| > a) \leq 2e^{-a^2/2nb^2}$.*

2.2 Entropy

In this subsection we record some basic properties of entropy (see [7] for an introduction to information theory). The entropy of a probability distribution $\mathbf{p} = (p_1, \dots, p_n)$ is $H(\mathbf{p}) = -\sum_{i\in[n]} p_i \log_2 p_i$. The entropy of a random variable X taking values in a finite set S is $H(X) = H(\mathbf{p})$, where $\mathbf{p} = (p_s : s \in S)$ is the law of X, i.e. $p_s = \mathbb{P}(X = s)$. When $\mathbf{p} = (p, 1-p)$ takes only two values we write $H(p) = H(\mathbf{p}) = -p \log_2 p - (1-p)\log_2(1-p)$.

Entropy is subadditive: if $X = (X_1, \dots, X_n)$ then $H(X) \leq \sum_{i=1}^n H(X_i)$, with equality if and only if the X_i are independent. An equivalent reformulation is the following lemma.

Lemma 2.4. *Suppose μ is a probability measure on $\prod_{s\in[S]} J_s$ with marginals $(\mu_s : s \in [S])$. Then $H(\mu) \leq \sum_{s\in[S]} H(\mu_s)$, with equality if and only if $\mu = \prod_{s\in[S]} \mu_s$.*

It is easy to deduce Lemma 1.18 from Lemma 2.4. Indeed, consider $\mu \in \mathcal{M}^{\mathcal{V}}_{\mathbf{w}}$ with maximum entropy. Let $p^i_j = \mathbb{P}_{\mathbf{x}\sim\mu}(x_i = j)$. Then $\mathbf{w} = \mathcal{V}(\mu) = \sum_{i\in[n], j\in J} p^i_j \mathbf{v}^i_j$, so the product measure $\mu_{\mathbf{p}}$ is in $\mathcal{M}^{\mathcal{V}}_{\mathbf{w}}$, and $H(\mu) \leq H(\mu_{\mathbf{p}})$, with equality if and only if $\mu = \mu_{\mathbf{p}}$. As $\mathcal{M}^{\mathcal{V}}_{\mathbf{w}}$ is convex, uniqueness follows from strict concavity of the entropy function, which we will now prove in a stronger form (see Lemma 2.6 below). It is often convenient to use the notation $H(\mathbf{p}) = \sum_{i\in[n]} L(p_i)$, where $L(p) = -p\log_2 p = -p\frac{\log p}{\log 2}$. Note that $L'(p) = -\frac{1+\log p}{\log 2}$ and $L''(p) = -\frac{1}{p \log 2} < 0$, so L is strictly concave. The following lemma is immediate from these formulae and the mean value form of Taylor's theorem: $f(a + t) = f(a) + f'(a)t + f''(a + t')t^2/2$ for some $0 < t' < t$.

Lemma 2.5. *If $|t| < \min(p, 1-p)$ then*

i. $L(p + t) - L(p) = -\big(\frac{1+\log p}{\log 2}\big)t \pm (p - |t|)^{-1}t^2$,

ii. $L(p + t) + L(p - t) - 2L(p) \leq -\frac{t^2}{\log 2}$.

We deduce the following 'stability version' of the uniqueness of the maximum entropy measure, which quantifies the decrease in entropy in terms of distance from the maximiser.

Lemma 2.6. *Suppose* $\mu_{\mathbf{p}} = \mu_{\mathbf{z}}^{\mathcal{Y}}$ *and* $\mu_{\widetilde{\mathbf{p}}} \in \mathcal{M}_{\mathbf{z}}^{\mathcal{Y}}$. *If* $\|\mathbf{p} - \widetilde{\mathbf{p}}\|_1 > \delta n$ *then* $H(\widetilde{\mathbf{p}}) < H(\mathbf{p}) - \delta^2 n$.

Proof. Let $\mathbf{p}' = (\mathbf{p} + \widetilde{\mathbf{p}})/2$ and note that $\mu_{\mathbf{p}'} \in \mathcal{M}_{\mathbf{z}}^{\mathcal{Y}}$. By definition of $\mu_{\mathbf{z}}^{\mathcal{Y}}$ we have $H(\mu_{\mathbf{p}'}) \le H(\mu_{\mathbf{p}})$, so $H(\mu_{\mathbf{p}}) - H(\mu_{\widetilde{\mathbf{p}}}) \ge 2H(\mu_{\mathbf{p}'}) - H(\mu_{\mathbf{p}}) - H(\mu_{\widetilde{\mathbf{p}}}) \ge \sum_{i \in [n]} (p_i - \widetilde{p}_i)^2 \ge \delta^2 n$, by Lemma 2.5 *ii* and then Cauchy-Schwarz. □

We conclude this subsection with a perturbation lemma.

Lemma 2.7. *Suppose* μ *is a probability distribution on* X *and* $-\log_2 \mu(x) > H(\mu)$ *for some* $x \in X$. *Then there is* $t > 0$ *such that* $\nu = (1-t)\mu + t1_x$ *has* $H(\nu) > H(\mu)$.

Proof. For $\mu(x) \neq 0$, for small enough t by Lemma 2.5 *i* we have

$$\begin{aligned} L(\nu(x)) - L(\mu(x)) &= (\nu(x) - \mu(x))L'(\mu(x)) \pm 2\mu(x)^{-1}(\nu(x) - \mu(x))^2 \\ &= -\Big(\frac{1 + \log_2 \mu(x)}{2}\Big)(1 - \mu(x))t - O(t^2). \end{aligned}$$

If $y \neq x$ and $\mu(y) = 0$ then $L(\nu(y)) - L(\mu(y)) = 0$. Therefore, by Lemma 2.5(i), for all $y \neq x$ we have

$$L(\nu(y)) - L(\mu(y)) = \Big(\frac{1 + \log_2 \mu(y)}{2}\Big)\mu(y)t - O(t^2). \tag{3}$$

All combined, this gives

$$\begin{aligned} H(\nu) - H(\mu) = \sum_{y \in X} L(\nu(y)) - L(\mu(y)) &= \sum_{y \in X} \Big(\frac{1 + \log_2 \mu(y)}{2}\Big)\mu(y)t - \Big(\frac{1 + \log_2 \mu(x)}{2}\Big)t - O(t^2) \\ &= t\Big(-\frac{\log_2 \mu(x)}{2} - \frac{H(\mu)}{2} - O(t)\Big) > 0, \end{aligned}$$

for small $t > 0$. The case $\mu(x) = 0$ is similar, using $L(\nu(x)) - L(\mu(x)) = -t\log_2 t$ with (3). □

2.3 Dependent Random Choice

We will use the following version of Dependent Random Choice (see [23, Lemma 11] for a proof and [10] for a comprehensive survey of the method). We write $N_G(u, u') := \{v \in V(G) : uv, u'v \in E(G)\}$ for the set of common neighbours of u and u' in a graph G.

Lemma 2.8. *Let* $t \in \mathbb{N}$ *and* $G = (V_1, V_2, E)$ *be a bipartite graph with* $|V_i| = N_i$ *and* $|E| = \alpha N_1 N_2$. *Then there is* $U \subset V_1$ *with* $|U| \ge \alpha^t N_1/2$ *such that* $|N_G(u, u')| \ge \alpha N_1^{-1/t} N_2$ *for all* $u, u' \in U$.

The following is an immediate consequence of Lemma 2.8, applied with $t = \lceil 2/c\varepsilon \rceil$.

Lemma 2.9. *Let* $0 < N^{-1} \ll \delta \ll \varepsilon, c < 1$. *Suppose* $G = (V_1, V_2, E)$ *is a bipartite graph with each* $|V_i| = N_i$, *where* $N \le N_1^c \le N_2 \le N_1^{1/c}$ *and* $e(G) > (N_1 N_2)^{1-\delta}$. *Then there is* $U \subset V_1$ *with* $|U| > N_1^{1-\varepsilon}$ *such that* $|N_G(u, u')| > N_2^{1-\varepsilon}$ *for all* $u, u' \in U$.

We say that $S \subset V(G)$ is independent if it contains no edges of G. The independence number $\alpha(G)$ of G is the maximum size of an independent set in G. Given graphs $G_1, \dots, G_k$, we write $G_1 \times \cdots \times G_k$ for the graph on vertex set $V(G_1) \times \cdots \times V(G_k)$, in which vertices $(u_1, \cdots, u_k)$ and $(v_1, \cdots, v_k)$ are joined by an edge if $u_i v_i \in E(G_i)$ for all $i \in [k]$.

Lemma 2.10. *Let $0 < N^{-1} \ll \delta \ll \varepsilon, c < 1$ and $N \le N_1^c \le N_2 \le N_1^{1/c}$. Suppose for $i = 1, 2$ we have graphs G_i on V_i with $|V_i| = N_i$ and $\alpha(G_i) \le N_i^{1-\varepsilon}$. Then $\alpha(G_1 \times G_2) \le (N_1 N_2)^{1-\delta}$.*

Proof. Suppose $E \subset V_1 \times V_2$ with $|E| > (N_1 N_2)^{1-\delta}$. Consider the bipartite graph $G = (V_1, V_2, E)$. Let U be as in Lemma 2.9. As $|U| > \alpha(G_1)$, there is an edge $u_1 u_2$ of G_1 in U. As $|N_G(u_1, u_2)| > \alpha(G_2)$, there is an edge $v_1 v_2$ of G_2 in $N_G(u_1, u_2)$. Then $(u_1, v_1)(u_2, v_2) \in E$, so E is not independent in $G_1 \times G_2$. □

By repeated application of the previous lemma, we obtain the following corollary.

Lemma 2.11. *Let $0 < N^{-1} \ll \delta \ll \varepsilon, c, k^{-1} < 1$ and $N_1, \dots, N_k \in \mathbb{N}$ with $N \le N_i^c \le N_j \le N_i^{1/c}$ for all $i, j \in [k]$. Suppose for $i \in [k]$ we have graphs G_i on V_i with $|V_i| = N_i$ and $\alpha(G_i) \le N_i^{1-\varepsilon}$. Then $\alpha(G_1 \times \cdots \times G_k) \le (N_1 \cdots N_k)^{1-\delta}$.*

2.4 Exponential Contiguity

We conclude this section with an alternative characterisation of exponential contiguity.

Lemma 2.12. *$\mu \lesssim_\Delta \mu'$ if and only if for $n^{-1} \ll \delta \ll \varepsilon \ll 1$ and $B_n = \{x \in \Delta_n : \mu'_n(x) < (1-\varepsilon)^n \mu_n(x)\}$ we have $\mu_n(B_n) \le (1-\delta)^n$.*

Proof. Let $n^{-1} \ll \delta \ll \varepsilon \ll 1$. Suppose first that if $A_n \subset \Delta_n$ with $\mu_n(A_n) > (1-\delta)^n$ then $\mu'_n(A_n) > (1-\varepsilon)^n$. As $\mu'_n(B_n) \le (1-\varepsilon)^n \mu_n(B_n) \le (1-\varepsilon)^n$, we cannot have $\mu_n(B_n) > (1-\delta)^n$, so we have $\mu_n(B_n) \le (1-\delta)^n$. Conversely, suppose $\mu_n(B_n) \le (1-\delta)^n$ and $A_n \subset \Delta_n$ with $\mu_n(A_n) > (1-\delta/2)^n$. Then $\mu'_n(A_n) \ge \mu'_n(A_n \setminus B_n) \ge (1-\varepsilon)^n \mu_n(A_n \setminus B_n) > (1-2\varepsilon)^n$. □

3 Large deviations of fixed sums

In this section we prove Theorem 1.20. Our first lemma will be used to show that the maximum entropy measure is exponentially dominated by the uniform measure.

Lemma 3.1. *Let $0 < n^{-1} \ll \eta \ll \kappa, |J|^{-1} \ll 1$. Suppose $\mu_{\mathbf{p}}$ is a κ-bounded product measure on J^n. Let $\mathcal{B} = \{\mathbf{x} \in J^n : \log_2 \mu_{\mathbf{p}}(\mathbf{x}) \notin -H(\mu_{\mathbf{p}}) \pm \eta n\}$. Then $\mu_{\mathbf{p}}(\mathcal{B}) \le (1-\eta^3)^n$.*

Proof. Consider $\mathbf{x} \sim \mu_{\mathbf{p}}$ and $X := \log_2 \mu_{\mathbf{p}}(\mathbf{x}) = \sum_{i \in [n]} X_i$, where $X_i = \sum_{j \in J} \mathbf{1}_{x_i = j} \log_2(p_j^i)$. As $\mu_{\mathbf{p}}$ is κ-bounded, the X_i satisfy $|X_i - \mathbb{E}(X_i)| \le \log_2(\kappa^{-1})$ for all $i \in [n]$. As these random variables are independent and $\mathbb{E}X = -H(\mu_{\mathbf{p}})$, the bound on $\mu_{\mathbf{p}}(\mathcal{B})$ follows from Chernoff's inequality. □

Our next lemma gives a lower bound for point probabilities of maximum entropy measures, which implies an upper bound on the number of solutions of $\mathcal{V}(\mathbf{x}) = \mathbf{w}$.

Lemma 3.2. *Suppose $\mathcal{V} = (\mathbf{v}_j^i)$ is an (n, J)-array in $\mathbb{Z}^D$ and $\mathbf{w} \in \mathbb{Z}^D$ and $\mu_{\mathbf{p}} = \mu_{\mathbf{p}_{\mathbf{w}}^{\mathcal{V}}}$. Then for all $\mathbf{x} \in (J^n)_{\mathbf{w}}^{\mathcal{V}}$ we have $-\log_2 \mu_{\mathbf{p}}(\mathbf{x}) \le H(\mu_{\mathbf{p}})$. In particular, $\log_2 |(J^n)_{\mathbf{w}}^{\mathcal{V}}| \le H(\mu_{\mathbf{p}})$.*

Proof. If some $\mathbf{x} \in (J^n)_{\mathbf{w}}^{\mathcal{V}}$ satisfies $-\log_2 \mu_{\mathbf{p}}(\mathbf{x}) > H(\mu_{\mathbf{p}})$ then Lemma 2.7(i) shows that $\nu = (1-t)\mu_{\mathbf{p}} + t 1_{\mathbf{x}}$ satisfies $H(\nu) > H(\mu_{\mathbf{p}})$ for some $t > 0$. However as $\nu \in \mathcal{M}_{\mathbf{w}}^{\mathcal{V}}$ this would contradict the choice of $\mathbf{p}_{\mathbf{w}}^{\mathcal{V}}$. The second statement now follows as $|(J^n)_{\mathbf{w}}^{\mathcal{V}}| 2^{-H(\mu_{\mathbf{p}})} \le \sum_{\mathbf{x}} \mu_{\mathbf{p}}(\mathbf{x}) \le 1$. □

Our final lemma will give an approximate formula for the number of solutions of $\mathcal{V}(\mathbf{x}) = \mathbf{w}$ (as mentioned in the introduction, [3, Theorem 3] gives stronger bounds under different hypotheses).

First we require a small set that efficiently generates $\mathbb{Z}^D$, as described by the following definition and associated lemma, which shows that such a set exists under the mild assumption of polynomial growth for the coordinate scale vector $\mathbf{R}$ (this will also be used later in Theorem 6.3).

Definition 3.3. We say that $\mathcal{U} \subset \mathbb{Z}^D$ is $(k, B, \mathbf{R})$-generating if for any $\mathbf{v} \in \mathbb{Z}^D$ we have $\mathbf{v} = \sum_{\mathbf{u}\in\mathcal{U}} c_\mathbf{u}\mathbf{u}$, with each $c_\mathbf{u} \in \mathbb{Z}$ with $|c_\mathbf{u}| \le k\|\mathbf{v}\|_\mathbf{R} + B$.

Lemma 3.4. *If $n^{-1} \ll \beta, C^{-1}, D^{-1}$ and $\max_d R_d < n^C$ then there is a $(1, \beta n, \mathbf{R})$-generating $\mathbf{R}$-bounded $\mathcal{U} \subset \mathbb{Z}^D$ with $|\mathcal{U}| \le D(C+2)$.*

Proof. Let $\mathcal{U}$ be the set of all $\mathbf{u} = (u_1, \dots, u_D)$ such that for some $d \in [D]$ we have $u_{d'} = 0$ for all $d' \neq d$ and $u_d = R_d$ or $u_d = \lfloor \beta n \rfloor^{a_d} \in [R_d]$ for some integer $a_d \ge 0$. □

Lemma 3.5. *Let $0 < n^{-1} \ll \lambda \ll \delta \ll \gamma, \kappa, k^{-1}, D^{-1}, C^{-1} < 1$ and $\mathbf{R} \in \mathbb{R}^D$ with $\max_d R_d < n^C$.*

Suppose

i. $\mathcal{V} = (\mathbf{v}^i_j)$ is an $\mathbf{R}$-bounded γ-robustly $(\mathbf{R}, k)$-generating (n, J)-array in $\mathbb{Z}^D$,

ii. $\mu_\mathbf{p}$ is a κ-bounded product measure on J^n with $\|\mathbb{E}\mathcal{V}(\mathbf{x}) - \mathbf{w}\|_\mathbf{R} \le \lambda n$, where $\mathbf{w} \in \mathbb{Z}^D$.

Then $\log_2 |(J^n)^\mathcal{V}_\mathbf{w}| \ge H(\mu_\mathbf{p}) - \delta n$; in particular, $(J^n)^\mathcal{V}_\mathbf{w} \neq \emptyset$.

Thus, if the maximum entropy measure $\mu_{\mathbf{p}^\mathcal{V}_\mathbf{w}}$ is κ-bounded then $\log_2 |(J^n)^\mathcal{V}_\mathbf{w}| = H(\mu_{\mathbf{p}^\mathcal{V}_\mathbf{w}}) \pm \delta n$.

Proof. We first note that the final statement of the lemma follows from the first: the latter gives the lower bound, as $\mathbb{E}\mathcal{V}(\mathbf{x}) = \mathbf{w}$ when $\mathbf{x} \sim \mu_{\mathbf{p}^\mathcal{V}_\mathbf{w}}$, and the upper bound follows from Lemma 3.2.

It remains to prove the first statement of the lemma. Let $\mathcal{F}$ be the set of $\mathbf{x} \in J^n$ such that there is $\mathbf{x}' \in (J^n)^\mathcal{V}_\mathbf{w}$ with Hamming distance $d(\mathbf{x}, \mathbf{x}') < \delta^2 n$. We claim that $\mu_\mathbf{p}(\mathcal{F}) > 1/2$.

First we assume the claim and deduce the lower bound. Given $\mathbf{x} \in \mathcal{F}$ there is $\mathbf{x}' \in (J^n)^\mathcal{V}_\mathbf{w}$ with Hamming distance $d(\mathbf{x}, \mathbf{x}') < \delta^2 n$ and so $|\mathcal{F}| \le |\{(\mathbf{x}, \mathbf{x}') \in J^n \times (J^n)^\mathcal{V}_\mathbf{w} : d(\mathbf{x}, \mathbf{x}') < \delta^2 n\}| \le |(J^n)^\mathcal{V}_\mathbf{w}|\binom{n}{\delta^2 n}|J|^{\delta^2 n}$, and as $|J| \le \kappa^{-1} \ll \delta^{-1}$ this gives $\log_2 |(J^n)^\mathcal{V}_\mathbf{w}| \ge \log_2 |\mathcal{F}| - \delta^{3/2} n$. Now let $\mathcal{B} = \{\mathbf{x} \in J^n : \log_2 \mu_\mathbf{p}(\mathbf{x}) > -H(\mu_\mathbf{p}) + \delta^2 n\}$. We have $\mu_\mathbf{p}(\mathcal{B}) \ll 1/4$ by Lemma 3.1, so $\mu_\mathbf{p}(\mathcal{F} \setminus \mathcal{B}) \ge \mu_\mathbf{p}(\mathcal{F}) - \mu_\mathbf{p}(\mathcal{B}) \ge 1/4$ by the claim. The definition of $\mathcal{B}$ gives $\log_2 |\mathcal{F} \setminus \mathcal{B}| - H(\mu_\mathbf{p}) + \delta^2 n \ge \log_2 \mu_\mathbf{p}(\mathcal{F} \setminus \mathcal{B}) \ge -2$, so $\log_2 |(J^n)^\mathcal{V}_\mathbf{w}| \ge \log_2 |\mathcal{F}| - \delta^{3/2} n \ge \log_2 |\mathcal{F} \setminus \mathcal{B}| - \delta^{3/2} n \ge H(\mu_\mathbf{p}) - \delta n$, as required.

To prove the claim, we consider $\mathbf{x} \sim \mu_\mathbf{p}$ and show that with probability at least $1/2$ there is $\mathbf{x}' \in (J^n)^\mathcal{V}_\mathbf{w}$ with $d(\mathbf{x}, \mathbf{x}') < \delta^2 n$. Let $\mathcal{B}_1$ be the event that $\|\mathcal{V}(\mathbf{x}) - \mathbf{w}\|_\mathbf{R} \ge \delta^3 n$. If $\mathcal{B}_1$ holds, by the triangle inequality $\|\mathcal{V}(\mathbf{x}) - \mathbb{E}\mathcal{V}(\mathbf{x})\|_\mathbf{R} \ge \delta^3 n - \lambda n \ge \delta^3 n/2$ and so $\mathbb{P}(\mathcal{B}_1) \le 2De^{-\delta^6 n/2^7}$ by Lemma 2.2. Next, by Lemma 3.4 we can fix some $(1, \delta^2 n, \mathbf{R})$-generating $\mathbf{R}$-bounded $\mathcal{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_M\} \subset \mathbb{Z}^D$ with $M \le D(C+2)$. Since $\mathcal{V}$ is γ-robustly $(\mathbf{R}, k)$-generating, by repeatedly applying Definition 1.9, we can choose pairwise disjoint $S_{mt} \subset [n]$ for each $m \in [M]$ and $t \in [\gamma n/kM]$ with each $|S_{mj}| \le k$ and $j_i, j'_i \in J$ for all $i \in S_{mt}$ such that $\mathbf{u}_m = \sum_{i\in S_{mt}} (\mathbf{v}^i_{j_i} - \mathbf{v}^i_{j'_i})$. Let $\mathcal{B}_2$ be the event that for some m we have $|\{t : x_i = j_i \ \forall i \in S_{mt}\}| < \kappa^k \gamma n/2kM$ or $|\{t : x_i = j'_i \ \forall i \in S_{mt}\}| < \kappa^k \gamma n/2kM$. Then $\mathbb{P}(\mathcal{B}_2) < e^{-\delta n}$ by Chernoff's inequality.

Thus with probability at least $1 - e^{-\delta^7 n} > 1/2$ neither $\mathcal{B}_1$ or $\mathcal{B}_2$ holds for $\mathbf{x}$. As $\mathcal{U}$ is $(1, \delta^3 n, \mathbf{R})$-generating and $\|\mathcal{V}(\mathbf{x}) - \mathbf{w}\|_\mathbf{R} < \delta^3 n$, we have $\mathcal{V}(\mathbf{x}) - \mathbf{w} = \sum_{m\in[M]} c_m \mathbf{u}_m$, with each $c_m \in \mathbb{Z}$ with $|c_m| \le 1 \cdot \|\mathcal{V}(\mathbf{x}) - \mathbf{w}\|_\mathbf{R} + \delta^3 n \le 2\delta^3 n$. Now we modify $\mathbf{x}$ to obtain $\mathbf{x}'$, where for each $m \in [M]$, if $c_m > 0$ we fix c_m values of t such that $x_i = j_i$ for all $i \in S_{mt}$ and let $x'_i = j'_i$ for all such i, and if $c_m < 0$ we fix c_m values of t such that $x_i = j'_i$ for all $i \in S_{mt}$ and let $x'_i = j_i$ for all such i. Then $\mathcal{V}(\mathbf{x}') = \mathbf{w}$, i.e. $\mathbf{x}' \in (J^n)^\mathcal{V}_\mathbf{w}$, and $d(\mathbf{x}, \mathbf{x}') < k \sum_{m\in[M]} |c_m| < \delta^2 n$. This completes the proof of the claim, and so of the lemma. □

We deduce Theorem 1.20, which states that under the hypotheses of the above lemmas, we have $\mu_\mathbf{p} \approx_\Delta \nu$, where $\mu_\mathbf{p} = \mu_{\mathbf{p}^\mathcal{V}_\mathbf{w}}$, and ν is the uniform distribution on $\Delta_n := (J^n)^\mathcal{V}_\mathbf{w}$.

Proof of Theorem 1.20. Let $0 < n^{-1} \ll \delta \ll \varepsilon \ll \gamma, \kappa, k^{-1}, D^{-1}, C^{-1} < 1$. Note that $\nu(\mathbf{x}) = |(J^n)^{\mathcal{V}}_{\mathbf{w}}|^{-1}$ for all $\mathbf{x} \in (J^n)^{\mathcal{V}}_{\mathbf{w}}$, and $\log_2 |(J^n)^{\mathcal{V}}_{\mathbf{w}}| = H(\mu_{\mathbf{p}}) \pm \delta n$ by Lemma 3.5.

Consider $\mathcal{C} = \{\mathbf{x} \in \Delta_n : \mu_{\mathbf{p}}(\mathbf{x}) < (1-\varepsilon)^n \nu(\mathbf{x})\}$. If there is $\mathbf{x} \in \mathcal{C}$ then $\log_2 \mu_{\mathbf{p}}(\mathbf{x}) < -\log_2 |(J^n)^{\mathcal{V}}_{\mathbf{w}}| - \varepsilon n < H(\mu_{\mathbf{p}})$, which contradicts Lemma 3.2, so $\mathcal{C} = \emptyset$. Thus $\nu \lesssim_\Delta \mu_{\mathbf{p}}$ by Lemma 2.12.

Now consider $\mathcal{C}' = \{\mathbf{x} \in \Delta_n : \nu(\mathbf{x}) < (1-\varepsilon)^n \mu_{\mathbf{p}}(\mathbf{x})\}$. For any $\mathbf{x} \in \mathcal{C}'$ we have $\log_2 \mu_{\mathbf{p}}(\mathbf{x}) > -\log_2 |(J^n)^{\mathcal{V}}_{\mathbf{w}}| + \varepsilon n$, so $\log_2 \mu_{\mathbf{p}}(\mathcal{C}') < -\varepsilon^3 n < -\delta n$ by Lemma 3.1, i.e. $\mu_{\mathbf{p}} \lesssim_\Delta \nu$. □

4 Boundedness, feasibility and universal VC-dimension

In this section we will give several combinatorial characterisations of the boundedness condition on maximum entropy measures required in our probabilistic intersection theorem. The characterisations hold under the following 'multiscale general position' assumption, which extends Definition 1.10 to all finite alphabets (by 'multiscale' we mean that the parameter γ can be arbitrary, which is true of the Kalai vectors).

Definition 4.1. (robustly generic) Suppose $\mathcal{V} = (\mathbf{v}^i_j)$ is an (n, J)-array in $\mathbb{Z}^D$. Let $I \in \binom{[n]}{D}$ and $\mathbf{c} = (\mathbf{c}^i : i \in I)$ with $\mathbf{c}^i \in \mathbb{R}^J$ and $\sum_{j \in J} c^i_j = 0$ for all $i \in I$.

We say that $(I, \mathbf{c})$ is $(\gamma, \mathbf{R})$-generic for $\mathcal{V}$ if all $|c^i_j| \le \gamma^{-1}$, and writing $\mathbf{w}^i = \sum_{j \in J} c^i_j \mathbf{v}^i_j$ and $W = (w^i_d : i \in I, d \in [D])$, we have $|\det(W)| \ge \gamma \prod_{d \in [D]} R_d$.

We say that $\mathcal{V}$ is γ'-robustly $(\gamma, \mathbf{R})$-generic if for any $X \subset [n]$ with $|X| > \gamma' n$ there is some $(I, \mathbf{c})$ with $I \subset X$ that is $(\gamma, \mathbf{R})$-generic for $\mathcal{V}$.

We say that a sequence $(\mathcal{V}_n, \mathbf{R}_n)$ of (n, J)-arrays and scalings is robustly generic if $\mathcal{V}_n$ is γ'-robustly $(\gamma, \mathbf{R}_n)$-generic whenever $n^{-1} \ll \gamma \ll \gamma'$.

It will also be convenient to use the following sequence formulation of Definition 1.9.

Definition 4.2. We say that $(\mathcal{V}_n, \mathbf{R}_n)$ is robustly generating if there are $\gamma > 0$ and $k, n_0 \in \mathbb{N}$ such that $\mathcal{V}_n$ is γ-robustly $(\mathbf{R}_n, k)$-generating for all $n > n_0$.

Next we will define the combinatorial conditions that appear in our characterisation. We recall the definition of VC-dimension and also define a universal variant that will be important in the proof of Theorem 1.11 in section 9.

Definition 4.3. We say that $\mathcal{A} \subset J^n$ shatters $X \subset [n]$ if for any $(j_x : x \in X) \in J^X$ there is $\mathbf{a} \in \mathcal{A}$ with $a_x = j_x$ for all $x \in X$.

The VC-dimension $\dim_{VC}(\mathcal{A})$ of $\mathcal{A}$ is the largest natural k such that $\mathcal{A}$ shatters *some* subset of $[n]$ of size k.

The universal VC-dimension $\dim_{UVC}(\mathcal{A})$ of $\mathcal{A}$ is the largest natural k such that $\mathcal{A}$ shatters *every* subset of $[n]$ of size k.

Next we give a feasibility condition, which can be informally understood as saying that we can solve any small perturbation of the equation $\mathcal{V}_n(\mathbf{x}) = \mathbf{z}_n$.

Definition 4.4. Let $(\mathcal{V}_n, \mathbf{R}_n, \mathbf{z}_n)$ be a sequence of (n, J)-arrays, scalings and vectors in $\mathbb{Z}^D$. We say $(\mathcal{V}_n, \mathbf{R}_n, \mathbf{z}_n)$ is λ-feasible if there is n_0 such that for any $n > n_0$, any $\mathbf{z}'_n \in \mathbb{Z}^D$ with $\|\mathbf{z}'_n - \mathbf{z}_n\|_{\mathbf{R}_n} \le \lambda n$, and any (n', J)-array $\mathcal{V}'_{n'}$ obtained from $\mathcal{V}_n$ by deleting at most λn co-ordinates, we have $(J^{n'})^{\mathcal{V}'_{n'}}_{\mathbf{z}'_n} \neq 0$.

Our final property appears to be a substantial weakening of our κ-boundedness condition, so it is quite surprising that it also gives a characterisation.

Definition 4.5. Suppose $\mu_{\mathbf{p}}$ is a product measure on J^n. We say that $\mu_{\mathbf{p}}$ is κ-dense if there are at least κn coordinates $i \in [n]$ such that $p^i_j \geq \kappa$ for all $j \in J$.

Now we can state the main theorem of this section. The sense of the equivalences in the statement is that the implied constants are bounded away from zero together. For example, the implication $ii \Rightarrow i$ means that for any $\delta > 0$ there is $\varepsilon > 0$ such that if $\mu^{\mathcal{V}_n}_{\mathbf{z}_n}$ is δ-dense then $\mu^{\mathcal{V}_n}_{\mathbf{z}_n}$ is ε-bounded.

Theorem 4.6. *Let $(\mathcal{V}_n, \mathbf{R}_n)$ be a robustly generic and robustly generating sequence of (n, J)-arrays and scalings in $\mathbb{Z}^D$, and $(\mathbf{z}_n)$ a sequence of vectors in $\mathbb{Z}^D$. The following are equivalent:*

i. $\mu^{\mathcal{V}_n}_{\mathbf{z}_n}$ *is* $\Omega(1)$*-bounded.*
ii. $\mu^{\mathcal{V}_n}_{\mathbf{z}_n}$ *is* $\Omega(1)$*-dense.*
iii. $\dim_{VC}((J^n)^{\mathcal{V}_n}_{\mathbf{z}_n}) = \Omega(n)$.
iv. $\dim_{UVC}((J^n)^{\mathcal{V}_n}_{\mathbf{z}_n}) = \Omega(n)$.
v. $(\mathcal{V}_n, \mathbf{R}_n, \mathbf{z}_n)$ *is* $\Omega(1)$*-feasible.*

The main step in the proof of Theorem 4.6 is Lemma 4.8, which provides the implication $iii \Rightarrow i$. It also implies Lemma 1.21, as for binary vectors the following coarse version of the Sauer-Shelah theorem shows that linear VC-dimension is equivalent to exponential growth.

Lemma 4.7. *[27, 29] For $\mathcal{A} \subset \{0,1\}^n$ we have $\dim_{VC}(\mathcal{A}) = \Omega(n) \Leftrightarrow \log_2 |\mathcal{A}| = \Omega(n)$.*

Lemma 4.8. *Let $0 < n^{-1} \ll \kappa \ll \gamma, \gamma' \ll \lambda \ll D^{-1}, |J|^{-1}$. Suppose $\mathcal{V} = (\mathbf{v}^i_j)$ is an $\mathbf{R}$-bounded γ'-robustly $(\gamma, \mathbf{R})$-generic (n, J)-array in $\mathbb{Z}^D$. If $\dim_{VC}((J^n)^{\mathcal{V}}_{\mathbf{w}}) \geq \lambda n$ then $\mu^{\mathcal{V}}_{\mathbf{w}}$ is κ-bounded.*

The proof of Lemma 4.8 is immediate from the next two lemmas, which give the implications $iii \Rightarrow ii$ and $ii \Rightarrow i$ of Theorem 4.6.

Lemma 4.9. *Let $0 < n^{-1} \ll \kappa \ll \lambda \ll D^{-1}, |J|^{-1}$. Suppose $\mathcal{V}$ is an (n, J)-array in $\mathbb{Z}^D$. Let $\mathbf{w} \in \mathbb{Z}^D$ and $\mathbf{p} = \mathbf{p}^{\mathcal{V}}_{\mathbf{w}}$. If $\dim_{VC}((J^n)^{\mathcal{V}}_{\mathbf{w}}) \geq \lambda n$ then $\mu_{\mathbf{p}}$ is κ-dense.*

Proof. Fix Z with $|Z| > \lambda n$ such that $(J^n)^{\mathcal{V}}_{\mathbf{w}}$ shatters Z. Suppose for a contradiction that $\mu_{\mathbf{p}}$ is not κ-dense. Then we have $Y \subset Z$ with $|Y| \geq |Z|/2$ and $(j'_y : y \in Y) \in J^Y$ such that $p^y_{j'_y} < \kappa$ for all $y \in Y$. As $(J^n)^{\mathcal{V}}_{\mathbf{w}}$ shatters Z, we can choose $\mathbf{j} \in (J^n)^{\mathcal{V}}_{\mathbf{w}}$ with $j_y = j'_y$ for all $y \in Y$. Note that $\mu_{\mathbf{p}}(\mathbf{j}) \leq \kappa^{|Y|} \leq \kappa^{\lambda n/2}$, so $-\log_2 \mu_{\mathbf{p}}(\mathbf{j}) \geq -(\log \kappa)\lambda n/2 > |J|n > H(\mu_{\mathbf{p}})$. By Lemma 2.7 we can find $\nu = (1-t)\mu_{\mathbf{p}} + t 1_{\mathbf{j}} \in \mathcal{M}^{\mathcal{V}}_{\mathbf{w}}$ with $H(\nu) > H(\mu_{\mathbf{p}})$. This contradicts the definition of $\mu^{\mathcal{V}}_{\mathbf{w}}$. □

Lemma 4.10. *Let $0 < n^{-1} \ll \kappa \ll \gamma, \gamma' \ll \lambda \ll D^{-1}, |J|^{-1}$. Suppose $\mathcal{V} = (\mathbf{v}^i_j)$ is an $\mathbf{R}$-bounded γ'-robustly $(\gamma, \mathbf{R})$-generic (n, J)-array in $\mathbb{Z}^D$. Let $\mathbf{p} = \mathbf{p}^{\mathcal{V}}_{\mathbf{w}}$. If $\mu_{\mathbf{p}}$ is λ-dense then $\mu_{\mathbf{p}}$ is κ-bounded.*

Proof. As $\mu_{\mathbf{p}}$ is λ-dense, we can fix $Y \subset [n]$ with $|Y| \geq \lambda n$ such that $p^i_j \geq \lambda$ for all $i \in Y$ and $j \in J$. As $\mathcal{V}$ is γ'-robustly $(\gamma, \mathbf{R})$-generic, we can fix $(I, \mathbf{c})$ with $I \subset Y$ that is $(\gamma, \mathbf{R})$-generic for $\mathcal{V}$, i.e. all $|c^i_j| \leq \gamma^{-1}$, $\sum_{j \in J} c^i_j = 0$ for all $i \in I$, and writing $\mathbf{w}^i = \sum_{j \in J} c^i_j \mathbf{v}^i_j$ and $W = (w^i_d : i \in I, d \in [D])$, we have $|\det(W)| \geq \gamma \prod_{d \in [D]} R_d$.

Now we show that $\mu_{\mathbf{p}}$ is κ-bounded. For suppose on the contrary that $p^{i'}_{j'} < \kappa$ for some $i' \in [n]$ and $j' \in J$. Fix $j'' \in J$ such that $p^{i'}_{j''} \geq |J|^{-1}$. As $(\mathbf{w}^i : i \in I)$ are linearly independent, we can write $\mathbf{v}^{i'}_{j''} - \mathbf{v}^{i'}_{j'} = \sum_{i \in I} b_i \mathbf{w}^i$. By Cramer's rule we have $b_i = \det(W)^{-1} \det(W_i)$, where W_i is the matrix obtained from W by replacing $\mathbf{w}^i$ with $\mathbf{v}^{i'}_{j''} - \mathbf{v}^{i'}_{j'}$. As $\mathcal{V}$ is $\mathbf{R}$-bounded, we can write $\det(W_i) = \det(A_i)\prod_{d\in D} R_d$, where all entries of A_i have modulus at most 1, so $|\det(A_i)| \le D!$ (or $D^{D/2}$ by Hadamard's inequality). Therefore $|b_i| \le D!\gamma^{-1}$ for all $i \in I$.

Consider a product measure $\mu_{\mathbf{p}'}$ where for some $t > 0$ we have $p'^i_j = p^i_j + tb_i c^i_j$ for all $i \in I$ and $j \in J$, $p'^{i'}_{j'} = p^{i'}_{j'} + t$, $p'^{i'}_{j''} = p^{i'}_{j''} - t$, and $p'^i_j = p^i_j$ otherwise. Note that $\mathbb{E}_{\mathbf{x}\sim\mu_{\mathbf{p}'}} \sum_{i\in[n]} \mathbf{v}^i_{x_i} = \mathbb{E}_{\mathbf{x}\sim\mu_{\mathbf{p}}} \sum_{i\in[n]} \mathbf{v}^i_{x_i} + t\sum_{i\in I, j\in J} b_i c^i_j \mathbf{v}^i_j + t\mathbf{v}^{i'}_{j'} - t\mathbf{v}^{i'}_{j''} = \mathbb{E}_{\mathbf{x}\sim\mu_{\mathbf{p}}} \sum_{i\in[n]} \mathbf{v}^i_{x_i} = \mathbf{w}$, so $\mu_{\mathbf{p}'} \in \mathcal{M}^{\mathcal{V}}_{\mathbf{w}}$.

We claim that we can choose $t > 0$ such that $H(\mu_{\mathbf{p}'}) > H(\mu_{\mathbf{p}})$. This will contradict the definition of $\mu^{\mathcal{V}}_{\mathbf{w}}$, showing that $\mu_{\mathbf{p}}$ is κ-bounded. To see this, note that $H(\mu_{\mathbf{p}'}) - H(\mu_{\mathbf{p}}) = \sum_{i\in I\cup\{i'\}}(H(p'^i) - H(p^i))$ and $H(p'^i) - H(p^i) = -\sum_{j\in J}\left[p'^i_j \log_2(p'^i_j) - p^i_j \log_2(p^i_j)\right]$ for each $i \in [n]$. Also, by Lemma 2.5 i, we have $-(p+t)\log_2(p+t) + p\log_2 p = -\left(\frac{1+\log p}{\log 2}\right)t \pm (p-|t|)^{-1}t^2 = \log_2(p^{-1})t \pm 4t$ provided $|t| < \frac{1}{2}\min(p, 1-p)$. If $p^{i'}_{j'} > 0$ then for $t < \frac{1}{2}\min\left(p^{i'}_{j'}, 1-p^{i'}_{j'}, (2|J|)^{-1}\right)$ this gives

$$H(p'^{i'}) - H(p^{i'}) \ge -t\log_2\kappa - 4t - 3t\log_2|J| - 2|J|t^2,$$

using $p^{i'}_{j'} < \kappa$, $p^{i'}_{j''} \ge |J|^{-1}$ and $|J| \ge 2$. If instead $p^{i'}_{j'} = 0$ then for $t < \min(\kappa, (2|J|)^{-1})$ we have

$$H(p'^{i'}) - H(p^{i'}) \ge -t\log_2\kappa - 3t\log_2|J| - 2|J|t^2,$$

Lastly, we have

$$H(p'^i) - H(p^i) \ge 2|J|D!\gamma^{-2}t\log\lambda - 2\lambda^{-1}|J|(tD!\gamma^{-2})^2$$

for $i \in I$, as all $|b_i c^i_j| \le D!\gamma^{-2}$ and $p^i_j \in (\lambda, 1-\lambda)$. Thus the dominant term in $H(\mu_{\mathbf{p}'}) - H(\mu_{\mathbf{p}})$ as $t \to 0$ is $-t\log_2\kappa$, and so we can choose $t > 0$ so that $H(\mu_{\mathbf{p}'}) - H(\mu_{\mathbf{p}}) > 0$, as required. □

Proof of Theorem 4.6. It remains to prove the implications $i \Rightarrow v$ and $v \Rightarrow iv$ (note that $iv \Rightarrow iii$ is trivial).

For $i \Rightarrow v$, let $n^{-1} \ll \lambda \ll \kappa \ll \gamma, k^{-1}$, suppose $\mathcal{V}_n$ is γ-robustly $(\mathbf{R}_n, k)$-generating, $\mu^{\mathcal{V}_n}_{\mathbf{z}_n}$ is κ-bounded, $\mathbf{z}'_n \in \mathbb{Z}^D$ with $\|\mathbf{z}'_n - \mathbf{z}_n\|_{\mathbf{R}_n} \le \lambda n$, and $\mathcal{V}'$ is obtained from $\mathcal{V}_n$ by deleting $S \subset [n]$ with $|S| \le \lambda n$. Then $\mathcal{V}'$ is $\mathbf{R}_n$-bounded and $(\gamma/2)$-robustly $(\mathbf{R}_n, k)$-generating. Also, the restriction $\mu_{\mathbf{p}'}$ of $\mu^{\mathcal{V}_n}_{\mathbf{z}_n}$ to $J^{[n]\setminus S}$ is κ-bounded, and $\mathbb{E}_{\mathbf{x}\sim\mu_{\mathbf{p}'}}\mathcal{V}'(\mathbf{x}) = \mathbf{z}^*$, where $\|\mathbf{z}^* - \mathbf{z}'_n\|_{\mathbf{R}_n} \le 2\lambda n$. Therefore $(J^{[n]\setminus S})^{\mathcal{V}'_{n'}}_{\mathbf{z}'_n} \ne \emptyset$ by Lemma 3.5, as required.

For $v \Rightarrow iv$, let $n^{-1} \ll \kappa$, and suppose $(\mathcal{V}_n, \mathbf{R}_n, \mathbf{z}_n)$ is κ-feasible. Fix $S \subset [n]$ with $|S| = \kappa n$ and $\mathbf{y} \in J^S$. We need to show that there is $\mathbf{x} \in (J^n)^{\mathcal{V}_n}_{\mathbf{z}_n}$ with $\mathbf{x}|_S = \mathbf{y}$. Let $\mathcal{V}'$, $\mathcal{V}^0$ be obtained from $\mathcal{V}_n$ by respectively deleting, retaining the coordinates of S. Let $\mathbf{z}'_n = \mathbf{z}_n - \mathcal{V}^0(\mathbf{y})$. Then $\|\mathbf{z}'_n - \mathbf{z}_n\|_{\mathbf{R}_n} \le \kappa n$, so by definition of κ-feasibility we can find $\mathbf{x}' \in (J^{[n]\setminus S})^{\mathcal{V}'_{n'}}_{\mathbf{z}'_n}$. Then $\mathbf{x} = (\mathbf{x}', \mathbf{y})$ is as required. □

We conclude this section by noting the following lemma which is immediate from the preceding proof and Lemma 4.8.

Lemma 4.11. *Let $0 < n^{-1} \ll \lambda \ll \gamma, \gamma' \ll \alpha, D^{-1}, J^{-1}$. Suppose $\mathcal{V} = (\mathbf{v}^i_j)$ is an $\mathbf{R}$-bounded, γ-robustly $(\mathbf{R}, k)$-generating, γ'-robustly $(\gamma, \mathbf{R})$-generic (n, J)-array in $\mathbb{Z}^D$.*

Let $\mathbf{w} \in \mathbb{Z}^D$ with $\dim_{VC}((J^n)^{\mathcal{V}}_{\mathbf{w}}) \ge \alpha n$. Then $\dim_{UVC}((J^n)^{\mathcal{V}}_{\mathbf{w}}) \ge \lambda n$.

5 Counterexamples to Conjecture 1.2

Theorem 1.3 will precisely describe the conditions under which the conclusion of Conjecture 1.2 is valid, and so show the existence of counterexamples in most cases. However, our proof is not constructive, so for expository purposes, in this section we will present two concrete counterexamples, each illustrating a different 'breaking point' of Theorem 1.11. The first will illustrate case *ii* by showing that we may have $[n]_{k,s} \times_{(t,w)} [n]_{k,s}$ large, but very few sets in $[n]_{k,s}$ are involved in any (t,w)-intersection. The second will illustrate case *iii* by showing that we may have almost all sets in $[n]_{k,s}$ involved in some (t,w)-intersection, but a large subset of $[n]_{k,s}$ containing no (t,w)-intersections. The first example also shows that cases *ii* and *iii* can hold simultaneously.

Counterexample 1: Set $\alpha_1 = 1/2$ and $\alpha_2 = 7/16$ and $\beta_1 = 1/4$ and $\beta_2 = 1/16 + \zeta$, where $\zeta > 0$ is small to be selected. Given $n \in \mathbb{N}$, take k, s, t and w as in Conjecture 1.2. We will show that $|[n]_{k,s} \times_{(t,w)} [n]_{k,s}| > (1+c)^n$ for some constant $c > 0$ but that there is a set $\mathcal{A} \subset [n]_{k,s}$ with $|\mathcal{A}| \geq (1-o(1))\big|[n]_{k,s}\big|$ satisfying $\mathcal{A} \times_{(t,w)} \mathcal{A} = \emptyset$.

First we show that $|[n]_{k,s} \times_{(t,w)} [n]_{k,s}|$ is large. To see this, we start by finding $C \subset [(1/4+\zeta^{1/2})n]$ with $(|C|, \sum(C)) = (t,w)$. Let $C_0 = [t]$ and note that $\sum(C_0) = \binom{t+1}{2} < w$. By a sequence of moves, each removing some i and adding $i+1$, we can obtain $C_1 = [(1/4+\zeta^{1/2})n - t + 1, (1/4+\zeta^{1/2})n]$, with $\sum(C_1) > w$. Clearly some intermediate set has $(C, \sum(C)) = (t,w)$. A similar argument gives a set $S \subset [(1/4+\zeta^{1/2})n + 1, n]$ with $\big(|S|, \sum(S)\big) = 2(k-t, s-w)$. Next we note that the maximum entropy measure $\mu_{\widetilde{\mathbf{p}}}$ on $\{0,1\}^S$ with $\sum_{i \in S} \widetilde{p}_i(1,i) = (k-t, s-w)$ is the constant vector $(1/2)_{i \in S}$. By Theorem 1.20 we deduce $|\{D \subset S : (|D|, \sum(D)) = (k-t, s-w)\}| = 2^{(1-o(1))|S|} = 2^{(1/2-o(1))n}$. However, for any $D \subset S$ with $(D, \sum(D)) = (k-t, s-w)$ we have $(E, \sum(E)) = (k-t, s-w)$, where $E = S \setminus D$. Taking $A = C \cup D$ and $B = C \cup E$ we find $(A,B) \in [n]_{k,s} \times_{(t,w)} [n]_{k,s}$. We have at least as many (t,w)-intersections as choices of D, so $|[n]_{k,s} \times_{(t,w)} [n]_{k,s}| \geq 2^{(1/2-o(1))n}$.

Next we show that sets in $[n]_{k,s}$ involved in any (t,w)-intersection are very restricted. Let $A, B \in [n]_{k,s}$ with $(|A|, \sum(A)) = (|B|, \sum(B)) = (\alpha_1 n, \alpha_2 \binom{n}{2}) = (n/2, \frac{7}{16}\binom{n}{2})$. Suppose A and B are (t,w)-intersecting. Let $\ell := |A \cap B \cap [n/4]|$. Then

$$\Big(\frac{1}{16} + \zeta\Big)\binom{n}{2} = w = \sum_{i \in A \cap B} i = \sum_{\substack{i \in A \cap B: \\ i < n/4}} i + \sum_{\substack{i \in A \cap B: \\ i \geq n/4}} i \geq \binom{\ell+1}{2} + \Big(\frac{n}{4} - \ell\Big)\frac{n}{4}.$$

Rearranging gives $(\ell - n/4)^2 - \zeta n^2 + (1/16+\zeta)n + \ell \leq 0$, so $\ell \geq n/4 - \sqrt{\zeta}n$. In particular, $|A \cap [n/4]| \geq n/4 - \sqrt{\zeta}n$.

We now show that almost all elements of $[n]_{k,s}$ do not have this restricted form. Fix constants $\delta_2 \ll \delta_1 \ll \kappa \ll 1$. Let $\mu_{\mathbf{p}}$ be the maximum entropy measure with $\sum p_i(1,i) = (k,s)$. Then $\mu_{\mathbf{p}}$ is κ-bounded by Lemma 1.21. Let $\mathcal{E} := \{A \subset [n] : \big|A \cap [\frac{n}{4}]\big| \geq \big(1 - \frac{\kappa}{2}\big)\frac{n}{4}\}$. Then $\mu_{\mathbf{p}}(\mathcal{E}) \leq (1-\delta_1)^n$ by Chernoff's inequality, so $\big|[n]_{k,s} \cap \mathcal{E}\big| \leq (1-\delta_2)^n |[n]_{k,s}|$ by Theorem 1.20. Choosing $\zeta < (\kappa/2)^2$, all (t,w)-intersecting pairs from $[n]_{k,s}$ lie within $\mathcal{E}$, which illustrates case *ii* of Theorem 1.11. Furthermore, $[n]_{k,s} \setminus \mathcal{E}$ is a set of size $(1-o(1))|[n]_{k,s}|$ containing no (t,w)-intersections, which illustrates case *iii* of Theorem 1.11.

Counterexample 2: This counterexample is a modification of the family in Section 1.4, related to VC-dimension. Let $\alpha_1 = \alpha_2 = 2/3$ and $\beta_1 = \beta_2 = 1/3 + \zeta$, where $\zeta > 0$ is small. Let n, k, s, t and w be as in Conjecture 1.2. It is not hard to see that $|[n]_{k,s} \times_{(t,w)} [n]_{k,s}| = (1 - o_\zeta(1))^n \binom{n}{t, k-t, k-t, n-2k+t} = (1-o_\zeta(1))^n 3^n$ and almost all elements of $[n]_{k,s}$ are involved in a (t,w)-intersection. However, for any set $U \subset [n]$ with $|U| = 2\zeta n + 1$, taking $\mathcal{A}_U = \{A \in [n]_{k,s} : A \cap U = \emptyset\}$, we have $|A \cap B| \geq t+1$ for all $A, B \in \mathcal{A}_U$ and so $\mathcal{A}_U \times_{(t,w)} \mathcal{A}_U = \emptyset$. On the other hand, if we select such a set U uniformly at random we find $\mathbb{E}_U(|\mathcal{A}_U|) \geq (1-o_\zeta(1))^n |[n]_{k,s}|$. Thus for some U we have $|\mathcal{A}_U| \geq (1-o_\zeta(1))^n |[n]_{k,s}|$ and $\mathcal{A}_U \times_{(t,w)} \mathcal{A}_U = \emptyset$.

6 The general setting

In this section we state our most general result, Theorem 6.3; we will defer the proof to section 8. This is in fact the main result of the paper in some sense, as we will show in this section that it implies Theorem 1.17 (in a more general cross-intersection form). However, the hypothesis of 'transfers' in Theorem 6.3 appears to be quite strong at first sight, and it will take some work to show that it follows from the hypotheses of Theorem 1.17 (it is here that the idea of enlarging the alphabet comes into play). We state our result in the next subsection and then deduce Theorem 1.17 in the following subsection. A second application of Theorem 6.3 is given in subsection 6.3, where we use it to give a short proof of a theorem of Frankl and Rödl on forbidden intersection patterns.

6.1 Statement of the general theorem

Before stating our theorem, we require the following definition, which describes a situation when for any vector $\mathbf{u}$ in some specific set (which will be given by the following definition), there are many ways of choosing a coordinate and two particular alterations of its value: one does not change the associated vector, and the other changes it by $\mathbf{u}$.

Definition 6.1. Suppose $\mathcal{V} = (\mathbf{v}^i_{j,\ell})$ is an $(n, J \times L)$-array in $\mathbb{Z}^D$. We say that $\mathbf{u}$ is an i-transfer in $\mathcal{V}$ (via (j, j') and (ℓ, ℓ')) if there are j, j' in J and ℓ, ℓ' in L with $\mathbf{v}^i_{j,\ell} - \mathbf{v}^i_{j',\ell} = \mathbf{u}$ and $\mathbf{v}^i_{j',\ell'} = \mathbf{v}^i_{j,\ell'}$.

Let $\mathcal{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_M\} \subset \mathbb{Z}^D$ and $\mathcal{P} = (P_m : m \in [M])$ for some disjoint subsets P_m of $[n]$. We say that $\mathcal{V}$ has transfers for $(\mathcal{P}, \mathcal{U})$ if $\mathbf{u}_m$ is an i-transfer in $\mathcal{V}$ for each $m \in M$ and $i \in P_m$.

We say that $\mathcal{V}$ has γ-robust transfers for $\mathcal{U}$ if it has transfers for $(\mathcal{P}, \mathcal{U})$ for some $\mathcal{P}$ such that $|P_m| \geq \gamma n$ for all $m \in [M]$.

Remark 6.2. We note that an $(n, \prod_{s\in S} J_s)$-array in $\mathbb{Z}^D$ has transfers for $(\mathcal{P}, \mathcal{U})$ if it has them as an $(n, J \times L)$-array, where $J = \prod_{s\in S'} J_s$ and $L = \prod_{s\in S\setminus S'} J_s$ for some $S' \subset S$.

We can now state our general theorem. (Recall that $\mathcal{U}$ exists by Lemma 3.4.)

Theorem 6.3. *Let $0 < n^{-1} \ll \delta \ll \zeta \ll \varepsilon, \kappa, \gamma \ll D^{-1}, M^{-1}, k^{-1}, C^{-1}$. Let S and $(J_s : s \in S)$ be sets of size at most C, and $\mathbf{R} = (R_1, \dots, R_D)$ with $\max_d R_d < n^C$. Suppose*

i. $\mu_{\mathbf{q}}$ is a κ-bounded product measure on $(\prod_{s\in S} J_s)^n$ with marginals $(\mu_{\mathbf{p}_s} : s \in S)$,
ii. $\mathcal{V} = (\mathbf{v}^i_{j_1,\dots,j_S})$ is an $\mathbf{R}$-bounded $(n, \prod_{s\in S} J_s)$-array in $\mathbb{Z}^D$,
iii. $\mathcal{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_M\} \subset \mathbb{Z}^D$ is $\mathbf{R}$-bounded and $(k, k\zeta n, \mathbf{R})$-generating,
iv. $\mathcal{V}$ has γ-robust transfers for $\mathcal{U}$,
v. $\mathbf{w} \in \mathbb{Z}^D$ with $\|\mathbf{w} - \mathcal{V}(\mu_{\mathbf{q}})\|_{\mathbf{R}} < \zeta n$.

Suppose $\mathcal{A}_s \subset J_s^n$ for $s \in S$ with $\prod_{s\in S} \mu_{\mathbf{p}_s}(\mathcal{A}_s) > (1-\delta)^n$. Then $\mu_{\mathbf{q}}((\prod_{s\in S} \mathcal{A}_s)^{\mathcal{V}}_{\mathbf{w}}) > (1-\varepsilon)^n$.

6.2 Proof of Theorem 1.17

Now we assume Theorem 6.3 and prove Theorem 1.17; in fact we prove the more general cross-intersection theorem. The strategy is to fuse together suitable co-ordinates and enlarge the alphabet.

Theorem 6.4. *Let $0 < n^{-1}, \delta \ll \zeta \ll \kappa, \gamma, \varepsilon \ll D^{-1}, C^{-1}, k^{-1}$ and $\mathbf{R} = (R_1, \dots, R_D)$ with $\max_d R_d < n^C$. Suppose*

i. $\mu_{\mathbf{q}}$ is a κ-bounded product measure on $(\{0,1\} \times \{0,1\})^n$ with marginals $(\mu_{\mathbf{p}_1}, \mu_{\mathbf{p}_2})$,
ii. $\mathcal{V} = (\mathbf{v}_i : i \in [n])$ is $\mathbf{R}$-bounded and γ-robustly $(\mathbf{R}, k)$-generating in $\mathbb{Z}^D$,

iii. $\mathbf{w} \in \mathbb{Z}^D$ *with* $\|\mathbf{w} - \mathcal{V}_\cap(\mu_\mathbf{q})\|_\mathbf{R} < \zeta n$.

Then any $\mathcal{A}_1, \mathcal{A}_2 \subset \{0,1\}^n$ *with* $\mu_{\mathbf{p}_1}(\mathcal{A}_1)\mu_{\mathbf{p}_2}(\mathcal{A}_2) > (1-\delta)^n$ *satisfy* $\mu_\mathbf{q}((\mathcal{A}_1 \times \mathcal{A}_2)^{\mathcal{V}_\cap}_\mathbf{w}) > (1-\varepsilon)^n$.

Proof. By Lemma 3.4 we can fix some $(1, \zeta n, \mathbf{R})$-generating $\mathbf{R}$-bounded $\mathcal{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_M\} \subset \mathbb{Z}^D$ with $M \le D(C+2)$. By repeatedly applying Definition 1.9, we can choose pairwise disjoint $S_{mj}, S'_{mj} \subset [n]$ for each $m \in [M]$ and $j \in [\gamma n/kM]$ with each $|S_{mj}| + |S'_{mj}| \le k$ and $\mathbf{u}_m = \sum_{i \in S_{mj}} \mathbf{v}_i - \sum_{i \in S'_{mj}} \mathbf{v}_i$. We let $N = \lfloor n/k \rfloor$ and partition $[n]$ into sets $T_1, \dots, T_N$ each of size k and a remainder set R with $0 \le |R| \le k-1$, such that each $S_{mj} \cup S'_{mj}$ is contained in some T_i. We let $\mathcal{P} = (P_m : m \in [M])$, where each P_m is the set of $i \in [N]$ such that T_i contains some $S_{mj} \cup S'_{mj}$.

We start by reducing to the case $R = \emptyset$ and $k|n$. For $R_s \subset R$ we let $\mathcal{A}_s^{R_s} = \{A \in \mathcal{A}_s : A \cap R = R_s\}$ for $s = 1, 2$. By the pigeonhole principle we can fix (R_1, R_2) so that $\mu_{\mathbf{p}_1}(\mathcal{A}_1^{R_1})\mu_{\mathbf{p}_2}(\mathcal{A}_2^{R_2}) > 2^{-2K}(1-\delta)^n > (1-2\delta)^n$. Let $\mathcal{A}'_s = \{A_s \setminus R_s : A \in \mathcal{A}_s^{R_s}\}$ and $\mathcal{V}' = (\mathbf{v}_i : i \in [n] \setminus R')$. Note that for $A_s \in \mathcal{A}'_s$ we have $A_s \cup R_s \in \mathcal{A}_s$ with $\mathcal{V}_\cap(A_1 \cup R_1, A_2 \cup R_2) = \mathcal{V}'_\cap(A_1, A_2) + \mathbf{v}'$, where $\mathbf{v}' = \sum_{i \in R_1 \cap R_2} \mathbf{v}_i$. Writing $\mathbf{w}' = \mathbf{w} - \mathbf{v}'$, we have $\mu_\mathbf{q}((\mathcal{A}_1 \times \mathcal{A}_2)^{\mathcal{V}_\cap}_\mathbf{w}) > \kappa^k \mu_\mathbf{q}((\mathcal{A}'_1 \times \mathcal{A}'_2)^{\mathcal{V}'_\cap}_{\mathbf{w}'})$, so to prove the theorem it suffices to show $\mu_\mathbf{q}((\mathcal{A}'_1 \times \mathcal{A}'_2)^{\mathcal{V}'_\cap}_{\mathbf{w}'}) > (1-\varepsilon/2)^n$.

We can naturally identify $\{0,1\}^n$ with $(\{0,1\}^k)^N$, where $A \in \{0,1\}^n$ corresponds to $(A \cap T_i : i \in [N])$ according to some fixed bijection of T_i with $[k]$. We will apply Theorem 6.3 with N in place of n, with $S = \{1, 2\}$ and $J_1 = J_2 = \{0,1\}^k$, and $\mathcal{A}'_s$ (naturally identified) in place of $\mathcal{A}_s$. We let $\mathcal{W} = (\mathbf{w}^i_{J_1,J_2})$ be the $(N, \{0,1\}^k \times \{0,1\}^k)$-array in $\mathbb{Z}^D$ defined by $\mathbf{w}^i_{J_1,J_2} = \sum_{j \in J_1 \cap J_2} \mathbf{v}_j$ for $J_1, J_2 \subset T_i$. Note that $\mathcal{V}_\cap(\mathbf{x}, \mathbf{y}) = \mathcal{W}(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y}$ in $\{0,1\}^n$ (naturally identified).

We also note that $\mathcal{W}$ has transfers for $(\mathcal{U}, \mathcal{P})$. To see this, consider $i \in P_m$ with $S_{mj} \cup S'_{mj} \subset T_i$. Let $J = S_{mj}$, $J' = S'_{mj}$, $L = S_{mj} \cup S'_{mj}$ and $L' = \emptyset$. Then $\mathbf{w}^i_{J,L'} = \mathbf{w}^i_{J',L'} = 0$ and $\mathbf{w}^i_{J,L} - \mathbf{w}^i_{J',L} = \sum_{i \in S} \mathbf{v}_i - \sum_{i \in S'} \mathbf{v}_i = \mathbf{u}_m$.

We let $\mu_{\mathbf{q}'}$ be the corresponding product measure on $(\{0,1\}^k \times \{0,1\}^k)^N$, defined by $q'^i_{\mathbf{j}^1,\mathbf{j}^2} = \prod_{i' \in T_i} q^{i'}_{j^1_{i'}, j^2_{i'}}$ for $\mathbf{j}^1$ and $\mathbf{j}^2$ in $\{0,1\}^k$, noting that $\mu_{\mathbf{q}'}$ is κ^k-bounded, and let $(\mu_{\mathbf{p}'_1}, \mu_{\mathbf{p}'_2})$ be its marginals on $(\{0,1\}^k)^N$. By construction we have $\mu_{\mathbf{p}'_s}(\mathbf{x}) = \mu_{\mathbf{p}_s}(\mathbf{x})$ and $\mu_{\mathbf{q}'}(\mathbf{x}, \mathbf{y}) = \mu_\mathbf{q}(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y}$ in $\{0,1\}^n$ (naturally identified).

To summarise, after the above reductions, we have $\mu_{\mathbf{p}'_1}(\mathcal{A}'_1)\mu_{\mathbf{p}'_2}(\mathcal{A}'_2) > (1-2\delta)^n$, and it suffices to show $\mu_{\mathbf{q}'}((\mathcal{A}'_1 \times \mathcal{A}'_2)^{\mathcal{W}}_{\mathbf{w}'}) > (1-\varepsilon/2)^n$. For $\mathbf{r} \sim \mu_{\mathbf{q}'}$ we have $\|\mathbf{w}' - \mathbb{E}\mathcal{W}(\mathbf{r}')\|_\mathbf{R} \le \|\mathbf{w} - \mathbb{E}\mathcal{V}_\cap(\mathbf{r})\|_\mathbf{R} + 2|R| < 2\zeta n$, so the theorem follows from Theorem 6.3. □

6.3 Application to a theorem of Frankl and Rödl

In this subsection we give another application of Theorem 6.3, which illustrates an additional flexibility, namely that our method allows different vectors defining the sizes of intersections from those defining the sizes of sets in the family. We will give a new proof of a theorem of Frankl and Rödl [11, Theorem 1.15] on intersection patterns in sequence spaces. (To align with notation from the rest of the paper, our notation differs from that of [11].)

Given non-negative integers $l_1, \dots, l_s$ with $\sum_i l_i = n$, let $\binom{[n]}{l_1, \dots, l_s}$ denote the set of elements $\mathbf{x} \in [s]^n$ with $|\{i \in [n] : x_i = j\}| = l_j$ for all $j \in [s]$. Given $\mathbf{x} \in \binom{[n]}{l_1, \dots, l_s}$ and $\mathbf{y} \in \binom{[n]}{k_1, \dots, k_t}$, the intersection pattern of $\mathbf{x}$ and $\mathbf{y}$ is given by an s times t matrix M, with $M_{j_1,j_2} = |\{i \in [n] : x_i = j_1, y_i = j_2\}|$ for $(j_1, j_2) \in [s] \times [t]$. For $\mathcal{A}_1 \subset \binom{[n]}{l_1, \dots, l_s}$ and $\mathcal{A}_2 \subset \binom{[n]}{k_1, \dots, k_t}$ we let $\mathcal{A}_1 \times_M \mathcal{A}_2$ denote the set of pairs $(\mathbf{x}, \mathbf{y}) \in \mathcal{A}_1 \times \mathcal{A}_2$ with intersection pattern M.

We say that M is an intersection pattern for $(l_1, \dots, l_s)$ and $(k_1, \dots, k_t)$ if each $\sum_{j_2 \in [t]} M_{j_1,j_2} = k_{j_1}$, each $\sum_{j_1 \in [s]} M_{j_1,j_2} = l_{j_2}$, and $\sum_{(j_1,j_2) \in [s] \times [t]} M_{j_1,j_2} = n$. The following result of Frankl and Rödl is the analogue of Theorem 1.1 for intersection patterns.

Theorem 6.5 (Frankl-Rödl)**.** *Given $\varepsilon, \kappa > 0$ and $s, t \in \mathbb{N}$ there is $\delta > 0$ such that the following holds. Suppose that M is an intersection pattern for $(l_1, \dots, l_s)$ and $(k_1, \dots, k_t)$ with all $M_{j_1,j_2} \ge \kappa n$. Let $\mathcal{A}_1 \subset \binom{[n]}{l_1,\dots,l_s}$ with $|\mathcal{A}_1| \ge (1-\delta)^n \binom{n}{l_1,\dots,l_s}$ and $\mathcal{A}_2 \subset \binom{[n]}{k_1,\dots,k_t}$ with $|\mathcal{A}_2| \ge (1-\delta)^n \binom{n}{k_1,\dots,k_t}$. Then $|\mathcal{A}_1 \times_M \mathcal{A}_2| \ge (1-\varepsilon)^n |\binom{n}{l_1,\dots,l_s} \times_M \binom{n}{k_1,\dots,k_t}|$.*

Proof. Fix $0 < \delta \ll \delta' \ll \varepsilon' \ll \varepsilon, \kappa$, and let $J_1 = [s]$. Let $\mathbf{e}_1, \dots, \mathbf{e}_s$ denote the standard basis for $\mathbb{Z}^s$, and let $\mathcal{V}_1 = (\mathbf{v}^i_j)$ denote the (n, J_1)-array, where each $\mathbf{v}^i_j = \mathbf{e}_j$. We can naturally identify $\binom{[n]}{l_1,\dots,l_s}$ with $(J_1^n)^{\mathcal{V}_1}_{\mathbf{z}_1}$, where $\mathbf{z}_1 = (l_1, \dots, l_s) \in \mathbb{Z}^s$. The maximum entropy measure $\mu_{\mathbf{p}_1} = \mu^{\mathcal{V}_1}_{\mathbf{z}_1}$ on J_1^n is then given by $(p_1)^i_j = l_j/n$ for all $i \in [n]$ and $j \in J_1$. Indeed, as $\mathbf{v}^i_j$ is independent of $i \in [n]$, by strict concavity of entropy (Lemma 2.5) so is $(p_1)^i_j = p_{1,j}$, and $n(p_{1,1}, \dots, p_{1,s}) = \mathbb{E}\mathcal{V}_1(\mathbf{x}) = (l_1, \dots, l_s)$. As $\mu_{\mathbf{p}_1}$ is κ-bounded we can apply Theorem 1.20 to find $\mu_{\mathbf{p}_1} \approx_{\triangle_1} \nu_1$, where ν_1 is uniform measure on $(J_1^n)^{\mathcal{V}_1}_{\mathbf{z}_1} = \binom{[n]}{l_1,\dots,l_s}$. Similarly, taking $J_2 = [t]$, we have a κ-bounded product measure $\mu_{\mathbf{p}_2}$ on J_2^n, with $\mu_{\mathbf{p}_2} \approx_{\triangle_2} \nu_2$, where ν_2 is uniform measure on $\binom{[n]}{k_1,\dots,k_t}$. Therefore $\mu_{\mathbf{p}_i}(\mathcal{A}_i) \ge (1-\delta')^n$ for $i = 1, 2$.

Similarly, we let $\mathbf{e}_{1,1}, \dots, \mathbf{e}_{s-1,t-1}$ denote the standard basis for $\mathbb{Z}^{(s-1)(t-1)}$, and let $\mathcal{V} = (\mathbf{v}^i_{j_1,j_2})$ denote the $(n, J_1 \times J_2)$-array, where $\mathbf{v}^i_{j_1,j_2} = \mathbf{e}_{j_1,j_2}$ if $(j_1, j_2) \in [s-1] \times [t-1]$ and $\mathbf{0}$ otherwise. We also let $\mathbf{w} = \sum_{(j_1,j_2) \in [s-1]\times[t-1]} M_{j_1,j_2} \mathbf{e}_{j_1,j_2}$. Note that for $\mathbf{x} \in \binom{n}{l_1,\dots,l_s}$ and $\mathbf{y} \in \binom{n}{k_1,\dots,k_t}$, we have $\mathcal{V}(\mathbf{x}, \mathbf{y}) = \mathbf{w}$ if and only if $\mathbf{x}$ and $\mathbf{y}$ have intersection pattern M. Therefore $(\mathcal{A}_1 \times \mathcal{A}_2)^{\mathcal{V}}_{\mathbf{w}} = \mathcal{A}_1 \times_M \mathcal{A}_2$.

We will apply Theorem 6.3 to estimate $(\mathcal{A}_1 \times \mathcal{A}_2)^{\mathcal{V}}_{\mathbf{w}}$ under the product measure $\mu_{\mathbf{q}}$ on $(J_1 \times J_2)^n$ defined by $q^i_{j_1,j_2} = M_{j_1,j_2}/n$. By hypothesis, $\mu_{\mathbf{q}}$ is κ-bounded, with marginals $\mu_{\mathbf{p}_1}$ and $\mu_{\mathbf{p}_2}$, and $\mathbb{E}_{(\mathbf{x},\mathbf{y}) \sim \mu_{\mathbf{q}}} \mathcal{V}(\mathbf{x}, \mathbf{y}) = \mathbf{w}$. Taking $\mathbf{R}$ to be the constant $\mathbf{1}$ vector in $\mathbb{Z}^{(s-1)(t-1)}$ we see that $\mathcal{V}$ is $\mathbf{R}$-bounded, and $\mathcal{U} = \{\mathbf{e}_{j_1,j_2}\}$ is $(st, 0, \mathbf{R})$-generating. Lastly, $\mathcal{V}$ has 1-robust transfers for $\mathcal{U}$, as for any $(j_1, j_2) \in [s-1] \times [t-1]$ we have $\mathbf{v}^i_{j_1,j_2} - \mathbf{v}^i_{s,j_2} = \mathbf{e}_{j_1,j_2}$ and $\mathbf{v}^i_{j_1,t} - \mathbf{v}^i_{s,t} = \mathbf{0}$. As $\mu_{\mathbf{p}_i}(\mathcal{A}_i) \ge (1-\delta')^n$ for $i = 1, 2$, Theorem 6.3 gives $\mu_{\mathbf{q}}((\mathcal{A}_1 \times \mathcal{A}_2)^{\mathcal{V}}_{\mathbf{w}}) = \mu_{\mathbf{q}}(\mathcal{A}_1 \times_M \mathcal{A}_2) \ge (1-\varepsilon')^n$. The theorem follows from a final application of Theorem 1.20. □

We wish to emphasize two aspects of the above proof. Firstly, it is crucial that the arrays $\mathcal{V}_1, \mathcal{V}_2$ and $\mathcal{V}$ can differ. Secondly, the arrays $\mathcal{V}_i$ are not $|J_i|^{-1}$-robustly $(\gamma, \mathbf{R})$-generic for any $\gamma > 0$ for $i = 1, 2$, so we cannot apply Lemma 4.8, but we were able to see directly that $\mu_{\mathbf{p}_1}$ and $\mu_{\mathbf{p}_2}$ are κ-bounded. Thus Theorem 6.3 has useful consequences even for arrays that are not robustly generic.

7 Correlation on product sets

In this section we will prove the following correlation inequality which will be used in the proof of Theorem 6.3; it can also be interpreted as an exponential contiguity result for product measures (see Theorem 7.2).

Theorem 7.1. *Let $0 < n^{-1}, \delta \ll \kappa, \varepsilon < 1$ and $\mu_{\mathbf{q}}$ be a κ-bounded product measure on $(\prod_{s\in S} J_s)^n$ with marginals $(\mu_{\mathbf{p}_s} : s \in S)$. Suppose $\mathcal{A}_s \subset J_s^n$ for $s \in S$ with $\prod_{s\in S} \mu_{\mathbf{p}_s}(\mathcal{A}_s) > (1-\delta)^n$. Then $\mu_{\mathbf{q}}(\prod_{s\in S} \mathcal{A}_s) > (1-\varepsilon)^n$.*

Proof of Theorem 7.1. We first consider the case $S = \{1, 2\}$. Define $f : (J_1)^n \to \mathbb{R}$ by

$$f(\mathbf{x}_1) = \log_e \left(\mu_{\mathbf{p}_1}(\mathbf{x}_1)\right) - \log_e \left(\mu_{\mathbf{q}}(\{\mathbf{x}_1\} \times \mathcal{A}_2)\right).$$

As $\mu_{\mathbf{p}_1}$ is a marginal of $\mu_{\mathbf{q}}$, we have $\mu_{\mathbf{p}_1}(\mathbf{x}_1) \ge \mu_{\mathbf{q}}(\{\mathbf{x}_1\} \times \mathcal{A}_2)$ for all $\mathbf{x}_1 \in (J_1)^n$, and so $f(\mathbf{x}_1) \ge 0$ for all $\mathbf{x}_1 \in (J_1)^n$. Note also that f is $2\log(\kappa^{-1})$-Lipschitz, as $\mu_{\mathbf{q}}$ is κ-bounded.

Let $M = \mathbb{E}_{\mu_{\mathbf{q}}}(f)$. We claim that $M \le (2\delta + \alpha)n \le 2\alpha n$. To see this, we apply a well-known concentration argument. For $I \subset \mathbb{R}$, let

$$\mathcal{B}_I = \{\mathbf{x}_1 \in (J_1)^n : f(\mathbf{x}_1) \in I\}.$$

By Lemma 2.3, letting $\alpha = 4\delta^{1/2}\log\kappa^{-1}$, we have $\mu_{\mathbf{p}_1}(\mathcal{B}_{[M-\alpha n, M+\alpha n]}) \ge 1 - 2e^{-\alpha^2 n/(8\log^2(\kappa^{-1}))} > 1 - (1-\delta)^n/2$. Now let

$$\mathcal{C} = \{\mathbf{x}_1 : \mu_{\mathbf{q}}(\{\mathbf{x}_1\} \times \mathcal{A}_2) \ge (1-\delta)^n \mu_{\mathbf{p}_1}(\mathbf{x}_1)/2\}.$$

Then $f(\mathbf{x}_1) \le 2\delta n$ for $\mathbf{x}_1 \in \mathcal{C}$, so $\mathcal{C} \subset \mathcal{B}_{[0,2\delta n]}$. However, $(1-\delta)^n \le \mu_{\mathbf{p}_2}(\mathcal{A}_2) = \mu_{\mathbf{q}}((J_1)^n \times \mathcal{A}_2) \le \mu_{\mathbf{p}_1}(\mathcal{C}^c)(1-\delta)^n/2 + \mu_{\mathbf{p}_1}(\mathcal{C})$, and so $\mu_{\mathbf{p}_1}(\mathcal{B}_{[0,2\delta n]}) \ge \mu_{\mathbf{p}_1}(\mathcal{C}) \ge (1-\delta)^n/2$. Thus $\mathcal{B}_{[0,2\delta n]} \cap \mathcal{B}_{[M-\alpha n, M+\alpha n]} \ne \emptyset$, which gives $M \le (2\delta+\alpha)n \le 2\alpha n$, as claimed.

Now set $\mathcal{B} = \mathcal{A}_1 \cap \mathcal{B}_{[0,3\alpha n]}$. As $\mu_{\mathbf{p}_1}(\mathcal{A}_1) \ge (1-\delta)^n$ and $\mu_{\mathbf{p}_1}(\mathcal{B}_{[0,3\alpha n]}) \ge \mu_{\mathbf{p}_1}(\mathcal{B}_{[M-\alpha n, M+\alpha n]}) \ge 1 - (1-\delta)^n/2$ we have $\mu_{\mathbf{p}_1}(\mathcal{B}) \ge (1-\delta)^n/2$. Therefore

$$\mu_{\mathbf{q}}(\mathcal{A}_1 \times \mathcal{A}_2) \ge \sum_{\mathbf{x}_1 \in \mathcal{B}} \mu_{\mathbf{q}}(\{\mathbf{x}_1\} \times \mathcal{A}_2) = \sum_{\mathbf{x}_1 \in \mathcal{B}} \mu_{\mathbf{p}_1}(\mathbf{x}_1) e^{-f(\mathbf{x}_1)} \ge \mu_{\mathbf{p}_1}(\mathcal{B}) e^{-3\alpha n} \ge (1-\varepsilon)^n.$$

This completes the proof in this case.

Now we deduce the general case by induction on $|S|$. Suppose the theorem is known for $|S| = k-1$ and we wish to prove it for $|S| = k$. Fix $s \in S$ and let $S' = S \setminus \{s\}$. We view $(\prod_{s \in S} J_s)^n$ as $(J_s \times J')^n$, where $J' = \prod_{s' \in S'} J_{s'}$. Let $\mu_{\mathbf{p}'}$ be the product measure on J'^n defined by $\mu_{\mathbf{p}_{S'}}(\mathbf{x}') = \mu_{\mathbf{q}}((J_1)^n \times \{\mathbf{x}'\})$. Then $\mu_{\mathbf{p}'}$ is κ-bounded and has marginals $(\mu_{\mathbf{p}_{s'}})_{s' \in S'}$, so by induction hypothesis, as $\prod_{s' \in S'} \mu_{\mathbf{p}_{s'}}(\mathcal{A}_{s'}) \ge (1-\delta)^n$ we have $\mu_{\mathbf{p}_{S'}}(\prod_{s' \in S'} \mathcal{A}_{s'}) \ge (1-\delta')^n$, where $\delta \ll \delta' \ll \varepsilon$.

Also, we can view $\mu_{\mathbf{q}}$ as a product measure on $(J_s \times J')^n$, with marginals $\mu_{\mathbf{p}_s}$ and $\mu_{\mathbf{p}'}$. Since $\mu_{\mathbf{p}_s}(\mathcal{A}_s)\mu_{\mathbf{p}'}(\prod_{s' \in S'} \mathcal{A}_{s'}) \ge (1-\delta)^n(1-\delta')^n \ge (1-2\delta')^n$, from the $|S| = 2$ case of the theorem we obtain $\mu_{\mathbf{q}}(\prod_{s \in S} \mathcal{A}_s) \ge (1-\varepsilon)^n$, as required. □

Theorem 1.16 is easily deduced from Theorem 7.1.

Proof of Theorem 1.16. Given $\zeta > 0$, taking $\mathcal{D} = \{\mathbf{r} : \|\mathcal{V}(\mathbf{r}) - \mathbb{E}\mathcal{V}\|_{\mathbf{R}} \ge \zeta n\} \subset (\{0,1\} \times \{0,1\})^n$, by Lemma 2.2 we have $\mu_{\mathbf{q}}\big(\mathcal{D}\big) \le 2De^{-\zeta^2 n/8} \le e^{-\zeta^2 n/16}$. However, provided $\delta, n^{-1} \ll \zeta, \varepsilon, \kappa$, by Theorem 7.1 any $\mathcal{A} \subset \{0,1\}^n$ with $\mu_{\mathbf{p}}(\mathcal{A}) \ge (1-\delta)^n$ satisfies $\mu_{\mathbf{q}}(\mathcal{A} \times \mathcal{A}) \ge e^{-\zeta^2 n/16} + (1-\varepsilon)^n$. Since $(\mathcal{A} \times \mathcal{A}) \cap \mathcal{D}^c = (\mathcal{A} \times \mathcal{A})_L^{\mathcal{V}_\cap}$ the result follows. □

To conclude this section we give a second application of Theorem 7.1 which will be useful in Section 12; Theorem 7.1 shows the exponential contiguity of $\mu_{\mathbf{q}}$ and $\prod_{s \in S} \mu_{\mathbf{p}_s}$, defined by $(\prod_{s \in S} \mu_{\mathbf{p}_s})(\mathbf{x}_s : s \in S) = \prod_{s \in S} \mu_{\mathbf{p}_s}(\mathbf{x}_s)$. Here the subscript Π indicates exponential contiguity relative to product sets, i.e. we apply Definition 1.19 in the case $\Omega_n = \big(\prod_{s \in S} J_s\big)^n$ and $\mathcal{F} = \Pi = (\Pi_n)_{n \in \mathbb{N}}$, where $\Pi_n = \{(\mathcal{A}_{n,s} : s \in S) :$ all $\mathcal{A}_{n,s} \in J_s^n\}$.

Theorem 7.2. *Let $0 < n^{-1} \ll \kappa \ll 1$ and $\mu_{\mathbf{q}}$ be a κ-bounded product measure on $(\prod_{s \in S} J_s)^n$ with marginals $(\mu_{\mathbf{p}_s} : s \in S)$. Then $\mu_{\mathbf{q}} \approx_\Pi \prod_{s \in S} \mu_{\mathbf{p}_s}$.*

Proof. As in the proof of Theorem 7.1, it suffices to consider the case $S = [2]$. By Theorem 7.1 we have $\mu_{\mathbf{p}_1} \times \mu_{\mathbf{p}_2} \lesssim_\Pi \mu_{\mathbf{q}}$. Conversely, consider $\mathcal{A}_s \subset J_s^n$ for $s \in [2]$. By the Cauchy-Schwarz inequality, writing $\sum$ for $\sum_{\mathbf{x}_1 \in J_1^n, \mathbf{x}_2 \in J_2^n}$, we have

$$\mu_{\mathbf{q}}(\mathcal{A}_1 \times \mathcal{A}_2)^2 = \Big(\sum \mu_{\mathbf{q}}(\mathbf{x}_1, \mathbf{x}_2) \prod_{s \in [2]} 1_{\mathbf{x}_s \in \mathcal{A}_s} \Big)^2 \leq \prod_{s \in [2]} \sum \mu_{\mathbf{q}}(\mathbf{x}_1, \mathbf{x}_2) 1_{\mathbf{x}_s \in \mathcal{A}_s} = \prod_{s \in [2]} \mu_{\mathbf{p}_s}(\mathcal{A}_s),$$

so $\mu_{\mathbf{q}} \lesssim_\Pi \mu_{\mathbf{p}_1} \times \mu_{\mathbf{p}_2}$. □

8 Proof of the general theorem

In this section we prove Theorem 6.3. We start by reducing to the case $|S| = 2$.

Lemma 8.1. *Theorem 6.3 follows from the case $|S| = 2$.*

Proof. First note that if $\mathcal{V}$ has γ-robust transfers for $\mathcal{U}$ then it has them as an $(n, L_1 \times L_2)$-array, where each $L_j = \prod_{s \in S_j} J_s$ for some partition (S_1, S_2) of S.

Now let $\mu_{\mathbf{p}_{S_1}}$ denote the product measure on L_1^n defined by $\mu_{\mathbf{p}_{S_1}}(\mathbf{x}_1) = \mu_{\mathbf{q}}(\{\mathbf{x}_1\} \times L_2^n)$; then $\mu_{\mathbf{p}_{S_1}}$ is κ-bounded. Similarly, we obtain $\mu_{\mathbf{p}_{S_2}}$ on L_2^n that is κ-bounded.

Let $\mathcal{A}_{S_j} = \prod_{s \in S_j} \mathcal{A}_s$ for $j = 1, 2$. As $\prod_{s \in S_j} \mu_{\mathbf{p}_s}(\mathcal{A}_s) \geq (1-\delta)^n$ and $\delta \ll \delta'$, by Theorem 7.1 each $\mu_{\mathbf{p}_{S_j}}(\mathcal{A}_{S_j}) \geq (1-\delta')^n$, so the case $S = \{1, 2\}$ of Theorem 6.3 applies to $(\mathcal{A}_{S_1}, \mathcal{A}_{S_2})$. □

Now we will prove a succession of special cases of Theorem 6.3, where the proof of each case builds on the previous cases, culminating in the proof of the general case. The last of these, Lemma 8.4, shows that Theorem 6.3 holds assuming $|S| = 2$, $J_1 = J_2 = \{0, 1\}$, and $q^i_{j_1, j_2} = 1/4$ for all $i \in [n]$ and $j_1, j_2 \in \{0, 1\}$; we refer to these assumptions as *the uniform binary setting.*

Lemma 8.2. *Suppose the assumptions of Theorem 6.3 hold in the uniform binary setting and $\mathcal{V}$ has transfers for $(\mathcal{P},\ \mathcal{U})$, where $\mathcal{P} = (P_m : m \in [M])$ is a partition of $[n]$. Then $(\mathcal{A} \times \mathcal{A})^{\mathcal{V}}_{\mathbf{w}} \neq \emptyset$.*

Proof. The idea of the proof is to reduce the required statement to finding two sets A and B in $\mathcal{A}$ with prescribed values of $|A \cap B \cap P_m|$ for all $m \in [M]$, where we identify $\{0, 1\}^n$ with subsets of $[n]$; this will be achieved by the Frankl-Rödl theorem and Dependent Random Choice.

We start by noting that, as $\mathbf{u}_m$ is an i-transfer for each $i \in P_m$, we can assume that $\mathbf{v}^i_{1,0} - \mathbf{v}^i_{0,0} = \mathbf{u}_m$ and $\mathbf{v}^i_{1,1} = \mathbf{v}^i_{0,1}$; indeed, this can be achieved by relabelling elements of $J_1 = J_2 = \{0, 1\}$ for each i, which preserves our hypotheses (as we are in the uniform binary setting).

Next we introduce some notation. We denote the marginals of $\mu_{\mathbf{q}}$ by $\mu_{\mathbf{p}} = \mu_{\mathbf{p}_1} = \mu_{\mathbf{p}_2}$, i.e. we have $p^i_0 = p^i_1 = \frac{1}{2}$ for all $i \in [n]$. For $K = (K_m : m \in [M])$ we let $\mathcal{B}^K$ denote the set of all $\mathbf{a} \in \{0, 1\}^n$ such that $\sum_{i \in P_m} \mathbf{a}_i = K_m$ for all $m \in [M]$.

We claim that we can fix K with $K_m = (\frac{1}{2} \pm \frac{\kappa}{4})|P_m|$ for all $m \in [M]$ such that $\mu_{\mathbf{p}}(\mathcal{A} \cap \mathcal{B}^K) > (1-\delta)^n$. Indeed, by assumption we have $\mu_{\mathbf{p}}(\mathcal{A}) > (1-\delta)^{n/2}$. Also, for $\mathbf{a} \sim \mu_{\mathbf{p}}$ and $X_m = \sum_{i \in P_m} \mathbf{a}_i$ we have $\mathbb{E} X_m = \frac{1}{2}|P_m|$, so by Chernoff's inequality $\mathbb{P}(|X_m - \mathbb{E} X_m| > \kappa |P_m|/4) \leq 2e^{-(\kappa |P_m|/4)^2 / 2|P_m|} \leq 2e^{-\kappa^2 \gamma n/32}$. There are at most n^M choices of K, so by a union bound and the pigeonhole principle there is some K with all $K_m = (\frac{1}{2} \pm \frac{\kappa}{4})|P_m|$ such that $\mu_{\mathbf{p}}(\mathcal{A} \cap \mathcal{B}^K) > n^{-M}((1-\delta)^{n/2} - 2Me^{-\kappa^2 \gamma n/32}) > (1-\delta)^n$, as claimed.

Now for $\mathbf{z} = (\mathbf{z}_m : m \in [M])$ with all $\mathbf{z}_m \in \mathbb{Z}^D$ we let $\mathcal{B}^{K, \mathbf{z}}$ denote the set of all $\mathbf{a} \in \mathcal{B}^K$ with $\sum_{i \in P_m} \mathbf{v}^i_{a_i, a_i} = \mathbf{z}_m$ for all $m \in [M]$. As all $\|\mathbf{v}^i_{j_0, j_1}\|_{\mathbf{R}} \leq 1$ and $\max_d R_d < n^C$, there are at most $(2n^{C+1})^{DM}$ possible values of $\mathbf{z}$, and so by the pigeonhole principle there is $\mathbf{z}$ so that $\mu_{\mathbf{p}}(\mathcal{A}\cap\mathcal{B}^{K,\mathbf{z}}) \ge (1-\delta)^n/(2n^{C+1})^{DM} > (1-2\delta)^n$.

We now note for any $\mathbf{a}$ and $\mathbf{a}'$ in $\mathcal{B}^{K,\mathbf{z}}$ that $\mathcal{V}(\mathbf{a},\mathbf{a}')$ is determined by the values $t_m = \sum_{i\in P_m} a_i a'_i$. Indeed, as $\mathbf{v}^i_{1,0} - \mathbf{v}^i_{0,0} = \mathbf{u}_m$ and $\mathbf{v}^i_{1,1} = \mathbf{v}^i_{0,1}$ we find that

$$\begin{aligned}\mathcal{V}(\mathbf{a},\mathbf{a}') &= \sum_{m\in[M]}\sum_{i\in P_m} \mathbf{v}^i_{a_i,a'_i} = \sum_{m\in[M]}\Big(\sum_{i\in P_m: a'_i=0}\mathbf{v}^i_{a_i,a'_i} + \sum_{i\in P_m: a'_i=1}\mathbf{v}^i_{a_i,a'_i}\Big)\\ &= \sum_{m\in[M]}\Big(\sum_{i\in P_m: a'_i=0}(\mathbf{v}^i_{a'_i,a'_i} + 1_{a_i=1}\mathbf{u}_m) + \sum_{i\in P_m: a'_i=1}\mathbf{v}^i_{a'_i,a'_i}\Big) = \sum_{m\in[M]}(\mathbf{z}_m + (K_m - t_m)\mathbf{u}_m).\end{aligned}$$

Next we claim that there are $\mathbf{b},\mathbf{b}'\in\mathcal{B}^{K,\mathbf{z}}$ such that $\mathbf{v}^* = \mathcal{V}(\mathbf{b},\mathbf{b}')$ and $\widetilde{\mathbf{v}} = \mathbb{E}\mathcal{V}(\mathbf{r})$ satisfy $\|\mathbf{v}^*-\widetilde{\mathbf{v}}\|_{\mathbf{R}} < \zeta n$, and $\mathbf{v}^* = \sum_{m\in[M]}(\mathbf{z}_m + c_m\mathbf{u}_m)$, where $c_m\in\mathbb{Z}$ with $c_m = (\frac14 \pm \frac{\kappa}{2})|P_m|$ for all $m\in[M]$. Indeed, selecting $\mathbf{r} = (\mathbf{b},\mathbf{b}')\sim\mu_{\mathbf{q}}$ and setting $X = \mathcal{V}(\mathbf{r})$ we have $\mathbb{P}(\|X - \mathbb{E}X\|_{\mathbf{R}} \ge \zeta n) \le 2De^{-\zeta^2 n/2}$ by Lemma 2.2. Also, $Y_m = \sum_{i\in P_m} b_i b'_i$ satisfies $\mathbb{E}Y_m = \frac14|P_m|$ and $\mathbb{P}(|Y_m - \mathbb{E}Y_m| > \kappa|P_m|/4) < 2e^{-\kappa^2\gamma n/128}$ by Chernoff's inequality. As $\mu_{\mathbf{p}}(\mathcal{B}^{K,\mathbf{z}}) \ge \mu_{\mathbf{p}}(\mathcal{A}\cap\mathcal{B}^{K,\mathbf{z}}) > (1-2\delta)^n$, by Theorem 7.1 we have $\mu_{\mathbf{q}}(\mathcal{B}^{K,\mathbf{z}}\times\mathcal{B}^{K,\mathbf{z}}) > (1-\delta')^n$, where $\delta\ll\delta'\ll\zeta$, so we can choose $\mathbf{b}$ and $\mathbf{b}'$ as claimed.

Now we can determine values t_m for $m\in[M]$ such that for any $\mathbf{a}$ and $\mathbf{a}'$ in $\mathcal{A}\cap\mathcal{B}^{K,\mathbf{z}}$ with $\sum_{i\in P_m} a_i a'_i = t_m$ for all $m\in[M]$ we have $\mathcal{V}(\mathbf{a},\mathbf{a}') = \mathbf{w}$. Indeed, $\|\mathbf{w}-\mathbf{v}^*\|_{\mathbf{R}} \le \|\mathbf{w}-\widetilde{\mathbf{v}}\|_{\mathbf{R}} + \|\mathbf{v}^*-\widetilde{\mathbf{v}}\|_{\mathbf{R}} < 2\zeta n$, so as $\mathcal{U}$ is $(k, k\zeta n, \mathbf{R})$-generating, we have $\mathbf{w} - \mathbf{v}^* = \sum_{m\in[M]} e_m\mathbf{u}_m$, with each $e_m\in\mathbb{Z}$ and $|e_m| \le 3k\zeta n$. Thus $\mathbf{w} = \sum_{m\in[M]}(\mathbf{z}_m + (c_m + e_m)\mathbf{u}_m)$, so we take $t_m = K_m - (c_m + e_m)$ for all $m\in[M]$. Note that our above bounds give $t_m = (\frac14 \pm \kappa)|P_m| \pm 3k\zeta n$.

It remains to show that we can find such $\mathbf{a}$ and $\mathbf{a}'$. We consider the graph $G = G_1\times\cdots\times G_M$ on $\mathcal{B}^K$, where each G_m is the graph on $\binom{P_m}{K_m}$ with $A_m A'_m \in E(G_m) \Leftrightarrow |A_m\cap A'_m| = t_m$. Recalling that $K_m = (\frac12 \pm \frac{\kappa}{4})|P_m|$ we have

$$\max(2K_m - |P_m|, 0) + \kappa|P_m| \le t_m = \big(1/4 \pm \kappa\big)|P_m| \pm 3k\zeta n \le K_m - \kappa|P_m|,$$

and so $\alpha(G_m) < (1-\delta')^n|V(G_m)|$ by Theorem 1.1. As $|V(G_m)| \ge \binom{\gamma n}{\kappa\gamma n/2} \ge |V(G_{m'})|^{\gamma\kappa}$ for all $m'\in[M]$, by Lemma 2.11 we have $\alpha(G) < (1-2\delta)^n|V(G)|$. But $|\mathcal{A}\cap\mathcal{B}^{K,\mathbf{z}}|/|\mathcal{B}^K| = \mu_{\mathbf{p}}(\mathcal{A}\cap\mathcal{B}^{K,\mathbf{z}})/\mu_{\mathbf{p}}(\mathcal{B}^K) > (1-2\delta)^n$, so $\mathcal{A}\cap\mathcal{B}^{K,\mathbf{z}}$ contains an edge of G, as required. □

Lemma 8.3. *Theorem 6.3 holds in the uniform binary setting when $\mathcal{A}_1 = \mathcal{A}_2 = \mathcal{A}$.*

Proof. Let $\mathcal{P} = (P_m : m\in[M])$ with $|P_m| = \gamma_0 n$ for all $m\in[M]$ be such that $\mathcal{V}$ has transfers for $(\mathcal{P},\ \mathcal{U})$, where $\zeta\ll\gamma_0\ll\varepsilon$. Let $B_2 = \cup_{m\in[M]}P_m$ and $B_1 = [n]\setminus B_2$. Write $\mathcal{F}^{K_1,K_2}$ for the set of all $\mathbf{a}\in\{0,1\}^n$ such that $\sum_{i\in B_j}\mathbf{a}_i = K_j$ for $j=1,2$. As in the proof of Lemma 8.2, we write $\mu_{\mathbf{p}} = \mu_{\mathbf{p}_1} = \mu_{\mathbf{p}_2}$, note that $p^i_0 = p^i_1 = 1/2$, and fix $K_j = (1/2 \pm \kappa/4)|B_j|$ for $j=1,2$ such that $\mu_{\mathbf{p}}(\mathcal{A}\cap\mathcal{F}^{K_1,K_2}) > (1-\delta)^n$.

Consider the bipartite graph G with parts $(\binom{B_1}{K_1}, \binom{B_2}{K_2})$ where $(\mathbf{b}_1,\mathbf{b}_2)\in E(G) \Leftrightarrow \mathbf{b}_1\mathbf{b}_2\in\mathcal{A}$. By Lemma 2.9 there is $\mathcal{B}\subset\binom{B_1}{K_1}$ with $|\mathcal{B}| > (1-\delta')^n\left|\binom{B_1}{K_1}\right|$, where $\delta\ll\delta'\ll\zeta$, such that for any $\mathbf{b}_1,\mathbf{b}'_1$ in $\mathcal{B}$ we have $|N_G(\mathbf{b}_1,\mathbf{b}'_1)| > (1-\delta')^n\left|\binom{B_2}{K_2}\right|$.

We will now find $\mathcal{F}\subset\mathcal{B}\times\mathcal{B}$ with $\mu_{\mathbf{q}}(\mathcal{F}) > (1-\varepsilon/2)^n$ (also writing $\mu_{\mathbf{q}}$ for its restriction to $(\{0,1\}\times\{0,1\})^{B_1}$) such that for any $(\mathbf{b}_1,\mathbf{b}'_1)\in\mathcal{F}$ there are $\mathbf{b}_2$ and $\mathbf{b}'_2$ in $N_G(\mathbf{b}_1,\mathbf{b}'_1)$, such that $\mathcal{V}(\mathbf{b}_1\mathbf{b}_2,\mathbf{b}'_1\mathbf{b}'_2) = \mathbf{w}$. This will suffice to prove the lemma, as then $\mu_{\mathbf{q}}((\mathcal{A}\times\mathcal{A})^{\mathcal{V}}_{\mathbf{w}}) \ge (1/4)^{|B_2|}\mu_{\mathbf{q}}(\mathcal{F}) > (1-\varepsilon)^n$, using $\gamma_0\ll\varepsilon$.

Let $\mathcal{V}_j = \{\mathbf{v}^i_{j_1,j_2} : i \in B_j\}$ for $j = 1,2$ and $\widetilde{\mathbf{v}}_j = \mathbb{E}\mathcal{V}_j(\mathbf{r})$, where $\mathbf{r} \sim \mu_{\mathbf{q}}$, so $\widetilde{\mathbf{v}}_1 + \widetilde{\mathbf{v}}_2 = \widetilde{\mathbf{v}} = \mathbb{E}\mathcal{V}(\mathbf{r})$. Let $\mathcal{E}$ be the set of $(\mathbf{b}_1, \mathbf{b}'_1) \in (\{0,1\} \times \{0,1\})^{B_1}$ such that $\|\mathcal{V}_1(\mathbf{b}_1, \mathbf{b}'_1) - \widetilde{\mathbf{v}}_1\|_{\mathbf{R}} > \zeta n$. Then $\mu_{\mathbf{q}}(\mathcal{E}) < 2De^{-\zeta^2 n/2}$ by Lemma 2.2. We choose $\mathcal{F} = (\mathcal{B}\times\mathcal{B})\backslash\mathcal{E}$. By Theorem 7.1 we have $\mu_{\mathbf{q}}(\mathcal{B}\times\mathcal{B}) > (1-\delta'')^{|B_1|}$ with $\delta' \ll \delta'' \ll \zeta \ll \varepsilon$, so $\mu_{\mathbf{q}}(\mathcal{F}) > (1-\varepsilon/2)^n$.

It remains to show for fixed $(\mathbf{b}_1, \mathbf{b}'_1) \in \mathcal{F}$ that there is $\mathbf{b}_2$ and $\mathbf{b}'_2$ in $N_G(\mathbf{b}_1, \mathbf{b}'_1)$ such that $\mathcal{V}_2(\mathbf{b}_2, \mathbf{b}'_2) = \mathbf{w}' := \mathbf{w} - \mathcal{V}_1(\mathbf{b}_1, \mathbf{b}'_1)$. To see this, it suffices to verify the hypotheses of Lemma 8.2, applied with $N_G(\mathbf{b}_1, \mathbf{b}'_1)$ in place of $\mathcal{A}$, restricting $\mu_{\mathbf{q}}$ to $(\{0,1\} \times \{0,1\})^{B_2}$, and with $\mathcal{V}_2$ in place of $\mathcal{V}$. We note that $\mathcal{V}_2$ has transfers for the same $(\mathcal{P}, \mathcal{U})$, and $\mathcal{P} = (P_m : m \in [M])$ is a partition of B_2. As $\|\mathbf{w}' - \widetilde{\mathbf{v}}_2\|_{\mathbf{R}} \le \|\mathbf{w} - \widetilde{\mathbf{v}}\|_{\mathbf{R}} + \|\mathcal{V}_1(\mathbf{b}_1, \mathbf{b}'_1) - \widetilde{\mathbf{v}}_1\|_{\mathbf{R}} \le 2\zeta n$, replacing ζ by 2ζ we see that all hypotheses hold, so the proof of the lemma is complete. □

Lemma 8.4. *Theorem 6.3 holds in the uniform binary setting.*

Proof. Let $\mathcal{A}'_1$ be the set of $\mathbf{a}_1 \in \mathcal{A}_1$ such that there is some $\mathbf{a}_2 = \mathbf{a}_2(\mathbf{a}_1) \in \mathcal{A}_2$ with Hamming distance $d(\mathbf{a}_1, \mathbf{a}_2) \le 2\delta' n$, where $\delta \ll \delta' \ll \zeta$. We claim that $\mu_{\mathbf{p}_1}(\mathcal{A}'_1) > (1-2\delta)^n$. This follows from the same concentration argument used in the proof of Theorem 7.1. Indeed, consider $\mathbf{r}_1 \sim \mu_{\mathbf{p}_1}$ and $X = d(\mathbf{r}_1, \mathcal{A}_2) = \min_{\mathbf{a}_2 \in \mathcal{A}_2} d(\mathbf{r}_1, \mathbf{a}_2)$. As X is 1-Lipschitz, by Lemma 2.3 we have $\mathbb{P}(|X - \mathbb{E}X| > \delta' n) < e^{-(\delta' n)^2/2n}$. This implies $\mathbb{E}X \le \delta' n$, as otherwise since $X(\mathbf{r}) = 0$ for all $\mathbf{r} \in \mathcal{A}_2$ we have $\mathbb{P}(|X - \mathbb{E}X| > \delta' n) \ge \mathbb{P}(X = 0) = \mu_{\mathbf{p}_1}(\mathcal{A}_2) = \mu_{\mathbf{p}_2}(\mathcal{A}_2) > (1-\delta)^n \gg e^{-(\delta' n)^2/2n}$, a contradiction. Therefore $\mathbb{P}(X > 2\delta' n) < e^{-(\delta' n)^2/2n}$, so the claim holds.

By the pigeonhole principle, we can fix $T \subset [n]$ with $|T| \le 2\delta' n$, a partition $T = T_{0,1} \cup T_{1,0}$ and $\mathcal{A}''_1 \subset \mathcal{A}'_1$ with $\mu_{\mathbf{p}_1}(\mathcal{A}''_1) > (1-3\delta)^n$ such that for every $\mathbf{a}_1 \in \mathcal{A}''_1$ we have $T_{1,0} = \{i \in [n] : (a_{1i}, a_2(\mathbf{a}_1)_i) = (1,0)\}$ and $T_{0,1} = \{i \in [n] : (a_{1i}, a_2(\mathbf{a}_1)_i) = (0,1)\}$.

Now let $\mathbf{w}' = \sum_{i \in T_{0,1}} (\mathbf{v}^i_{0,1} - \mathbf{v}^i_{0,0}) + \sum_{i \in T_{1,0}} (\mathbf{v}^i_{1,0} - \mathbf{v}^i_{1,1})$ and note that for any $\mathbf{a}_1$ and $\mathbf{a}'_1$ in $\mathcal{A}''_1$ with $\mathcal{V}(\mathbf{a}_1, \mathbf{a}'_1) = \mathbf{w} + \mathbf{w}'$ we have $\mathcal{V}(\mathbf{a}_1, \mathbf{a}_2(\mathbf{a}'_1)) = \mathbf{w}$. Note also that $\|\mathbf{w}'\|_{\mathbf{R}} \le |T| \le 2\delta' n$, so $\|\mathbf{w} + \mathbf{w}' - \mathbb{E}\mathcal{V}(\mathbf{r})\| \le \|\mathbf{w}'\| + \|\mathbf{w} - \mathbb{E}\mathcal{V}(\mathbf{r})\| < 2\zeta n$. Then $\mu_{\mathbf{q}}((\mathcal{A}''_1 \times \mathcal{A}''_1)^{\mathcal{V}}_{\mathbf{w}+\mathbf{w}'}) > (1-\varepsilon/2)^n$ by Lemma 8.3, so $\mu_{\mathbf{q}}((\mathcal{A}_1 \times \mathcal{A}_2)^{\mathcal{V}}_{\mathbf{w}}) \ge (1/4)^{|T|} \mu_{\mathbf{q}}((\mathcal{A}''_1 \times \mathcal{A}''_1)^{\mathcal{V}}_{\mathbf{w}+\mathbf{w}'}) > (1-\varepsilon)^n$, as required. □

Proof of Theorem 6.3. As noted earlier, we may assume $S = \{1,2\}$. By relabelling, we can also assume $\{0,1\} \subset J_1, J_2$. As $\mathcal{V}$ has γ-robust transfers for $\mathcal{U}$, there are disjoint subsets P_m of $[n]$, with $|P_m| \ge \gamma n$, so that $\mathbf{u}_m$ is an i-transfer for all $i \in P_m$. By relabelling, we can assume $\mathbf{v}^i_{1,1} - \mathbf{v}^i_{0,1} = \mathbf{u}_m$ and $\mathbf{v}^i_{1,0} = \mathbf{v}^i_{0,0}$ for all $i \in P_m$.

Next we describe an alternative method to select elements from $(\prod_{s \in S} J_s)^n$ according to $\mu_{\mathbf{q}}$. To begin, we select a random partition $[n] = S \cup T$, where each $i \in [n]$ appears in S independently with probability κ. Secondly, we randomly select $\mathbf{r}' = (\mathbf{r}'_1, \mathbf{r}'_2) \in (J_1)^T \times (J_2)^T = (J_1 \times J_2)^T$ according to a product measure $\mu_{\mathbf{q}'}$ on $(J_1 \times J_2)^T$, where $\mathbf{q}'$ will be defined below. Lastly, we select $\mathbf{s} = (\mathbf{s}_1, \mathbf{s}_2) \in \{0,1\}^S \times \{0,1\}^S = (\{0,1\} \times \{0,1\})^S$, according to the uniform measure ν on $(\{0,1\} \times \{0,1\})^S$. (We will also write ν for the uniform measure on $\{0,1\}^S$.) We obtain a random element $\mathbf{r} = \mathbf{r}' \circ \mathbf{s} \in (J_1 \times J_2)^n$, which defines a product measure $\mu_{\mathbf{q}''}$ on $(J_1 \times J_2)^n$, as $r_1, \dots, r_n$ are independent. Now we select $\mathbf{q}'$ above so that $\mathbf{q}'' = \mathbf{q}$. To determine $\mathbf{q}'$, note that if $j, j' \in \{0,1\}$ then $(q'')^i_{j,j'} = \kappa/4 + (1-\kappa)(q')^i_{j,j'}$, and otherwise $(q'')^i_{j,j'} = (1-\kappa)(q')^i_{j,j'}$. Thus we can obtain $\mathbf{q}'' = \mathbf{q}$ by setting $(q')^i_{j,j'} = (q^i_{j,j'} - \kappa/4)/(1-\kappa)$ for $i \in [n]$, $j, j' \in \{0,1\}$ and $(q')^i_{j,j'} = (q)^i_{j,j'}/(1-\kappa)$ otherwise. (Note that κ-boundedness of $\mathbf{q}$ ensures ${q'}^i_{j,j'} \in [0,1]$.)

We will analyse this alternative construction of $\mu_{\mathbf{q}}$ in two steps, where in the first step we fix a pair $\pi = (S, \mathbf{r}')$ selected above and consider the additional random choice of $\mathbf{s} = (\mathbf{s}_1, \mathbf{s}_2)$. For $j = 1,2$

we let

$$\mathcal{F}_j^\pi = \{\mathbf{s}_j \in \{0,1\}^S : \mathbf{r}' \circ \mathbf{s}_j \in \mathcal{A}_j\} \text{ and}$$
$$\mathcal{F}_\mathbf{w}^\pi = \{\mathbf{s} \in \mathcal{F}_1^\pi \times \mathcal{F}_2^\pi : \mathcal{V}(\mathbf{r}' \circ \mathbf{s}) = \mathbf{w}\}.$$

Since $\mathbf{q}'' = \mathbf{q}$, we have $\mu_{\mathbf{p}_j}(\mathcal{A}_j) = \mathbb{E}_\pi(\nu(\mathcal{F}_j^\pi))$ for $j = 1,2$ and $\mu_\mathbf{q}((\mathcal{A} \times \mathcal{A})_\mathbf{w}^\mathcal{V}) = \mathbb{E}_\pi(\nu(\mathcal{F}_\mathbf{w}^\pi))$.

In the remainder of the proof we will show that $\mathbb{P}_\pi(\nu(\mathcal{F}_\mathbf{w}^\pi) > (1-\varepsilon/2)^n) > (1-\delta')^n$, where $\delta \ll \delta' \ll \zeta$. This will imply the Theorem, as then $\mu_\mathbf{q}((\mathcal{A}_1 \times \mathcal{A}_2)_\mathbf{w}^\mathcal{V}) = \mathbb{E}_\pi(\nu(\mathcal{F}_\mathbf{w}^\pi)) > (1-\delta')^n(1-\varepsilon/2)^n > (1-\varepsilon)^n$. To achieve this, we will show that for 'good' π we can apply Lemma 8.4 to $\mathcal{F}_1^\pi$ and $\mathcal{F}_2^\pi$, with uniform product measure and the array $\mathcal{X}^\pi := (\mathbf{v}_{j,j'}^i : i \in S, j, j' \in \{0,1\})$. As $\mathcal{V}(\mathbf{r}) = \mathcal{X}^\pi(\mathbf{s}) + \mathcal{Y}^\pi(\mathbf{r}')$, we have

$$\mathcal{F}_\mathbf{w}^\pi = (\mathcal{F}_1^\pi \times \mathcal{F}_2^\pi)_{\mathbf{w}'}^{\mathcal{X}^\pi},$$

where $\mathbf{w}' := \mathbf{w} - \mathcal{Y}^\pi(\mathbf{r}')$ with $\mathcal{Y}^\pi = (\mathbf{v}_{j,j'}^i : i \in T, j \in J_1, j' \in J_2)$.

First we define some bad events for π and show that they are unlikely. Let $\mathbf{v}^\pi = \mathbb{E}[\mathcal{X}^\pi(\mathbf{s}) \mid \pi]$ and $\mathcal{B}_1$ be the event that $\|\mathbf{v}^\pi - \mathbb{E}\mathbf{v}^\pi\|_\mathbf{R} > \zeta n$. Then $\mathbb{P}(\mathcal{B}_1) \le 2De^{-\zeta^2 n/8}$ by Lemma 2.2. Similarly, the bad event $\mathcal{B}_2$ that $\|\mathcal{Y}^\pi(\mathbf{r}') - \mathbb{E}\mathcal{Y}^\pi(\mathbf{r}')\|_\mathbf{R} > \zeta n$ has $\mathbb{P}(\mathcal{B}_2) \le 2De^{-\zeta^2 n/8}$. Note that if $\mathcal{B}_1 \cup \mathcal{B}_2$ does not hold, as $\|\mathbf{w} - \mathbb{E}\mathcal{V}(\mathbf{r})\|_\mathbf{R} \le \zeta n$, we have $\|\mathbf{v}^\pi - \mathbf{w}'\|_\mathbf{R} \le 3\zeta n$.

The last bad event is that we do not have robust transfers. Let $\mathcal{P}^\pi = (P_m^\pi : m \in [M])$, where P_m^π is the set of $i \in P_m$ such that $\mathbf{u}_m$ is an i-transfer in $\mathcal{X}^\pi$. Recalling that $\mathbf{u}_m$ is an i-transfer in $\mathcal{V}$ via $(0,1)$ and $(0,1)$ for all $i \in P_m$, we have $i \in P_m^\pi$ whenever $i \in S$, so $\mathbb{E}|P_m^\pi| \ge \kappa\gamma n$. By Chernoff's inequality, the bad event $\mathcal{B}_3$ that some $|P_m^\pi| < \kappa\gamma n/2$ satisfies $\mathbb{P}(\mathcal{B}_3) < 2Me^{-\kappa^2\gamma^2 n/8}$.

Now let $\mathcal{G}$ be the good event for π that $\nu(\mathcal{F}_1^\pi)\nu(\mathcal{F}_2^\pi) > (1-\delta')^n$. By Cauchy-Schwarz and Theorem 7.1 we have

$$\mathbb{E}_\pi\nu(\mathcal{F}_1^\pi)\nu(\mathcal{F}_2^\pi) \ge (\mathbb{E}_\pi\nu(\mathcal{F}_1^\pi \times \mathcal{F}_2^\pi))^2 = \mu_\mathbf{q}(\mathcal{A}_1 \times \mathcal{A}_2)^2 > (1-\delta'/4)^{2n},$$

so $(1-\mathbb{P}(\mathcal{G}))(1-\delta')^n + \mathbb{P}(\mathcal{G}) \ge (1-\delta'/2)^n$, giving $\mathbb{P}(\mathcal{G}) > (1-\delta'/2)^n/2$. Thus with probability at least $(1-\delta')^n$ the event $\mathcal{G} \setminus \cup_{i=1}^3 \mathcal{B}_i$ holds, so we can apply Lemma 8.4 to obtain $\nu(\mathcal{F}_\mathbf{w}^\pi) = \nu((\mathcal{F}_1^\pi \times \mathcal{F}_2^\pi)_{\mathbf{w}'}^{\mathcal{X}^\pi}) > (1-\varepsilon/2)^n$, as required to prove the theorem. □

9 Proof of Theorem 1.11

In this section we will prove Theorem 1.11. Let $\mathcal{X} = (\{0,1\}^n)_\mathbf{z}^\mathcal{V}$, as in the statement of Theorem 1.11. The proof will split naturally into two pieces according to the VC-dimension of $(\mathcal{X} \times \mathcal{X})_\mathbf{w}^{\mathcal{V}_\cap}$. The next subsection shows that for high VC-dimension cases *i* or *ii* of Theorem 1.11 hold; the following subsection shows that case *iii* holds in the case of small VC-dimension.

9.1 Large VC-dimension

Here we implement the strategy discussed in subsection 1.4: we consider the maximum entropy measure $\mu_{\tilde{\mathbf{q}}}$ that represents $(\mathcal{X} \times \mathcal{X})_\mathbf{w}^{\mathcal{V}_\cap}$, and distinguish cases *i* or *ii* from Theorem 1.11 according to whether its marginals $\mu_{\tilde{\mathbf{p}}}$ are close to $\mu_\mathbf{p} := \mu_{\mathbf{p}_\mathbf{z}^\mathcal{V}}$.

We will study $\mathbf{w}$-intersections between elements in $\mathcal{X} = (\{0,1\}^n)_\mathbf{z}^\mathcal{V}$ via a new collection of vectors defined in $\mathbb{Z}^{3D}$, motivated by the observation that for any $\mathbf{a}, \mathbf{a}' \in \{0,1\}^n$ we have $\mathbf{a}, \mathbf{a}' \in \mathcal{X}$ and $\mathcal{V}_\cap(\mathbf{a}, \mathbf{a}') = \mathbf{w}$ if and only if

$$\sum_{i\in[n]} \left[a_i a_i'(\mathbf{v}_i, \mathbf{v}_i, \mathbf{v}_i) + a_i(1-a_i')(\mathbf{v}_i, \mathbf{0}, \mathbf{0}) + (1-a_i)a_i'(\mathbf{0}, \mathbf{v}_i, \mathbf{0})\right] = (\mathbf{z}, \mathbf{z}, \mathbf{w}).$$

With this in mind, we adopt the following notation throughout this subsection.

Definition 9.1. Let $J = \{0,1\} \times \{0,1\}$ and let $\widetilde{\mathcal{V}} = (\widetilde{\mathbf{v}}^i_j)$ denote the (n, J)-array in $\mathbb{Z}^{3D}$ with

$$\widetilde{\mathbf{v}}^i_{1,1} = (\mathbf{v}_i, \mathbf{v}_i, \mathbf{v}_i), \qquad \widetilde{\mathbf{v}}^i_{1,0} = (\mathbf{v}_i, \mathbf{0}, \mathbf{0}),$$
$$\widetilde{\mathbf{v}}^i_{0,1} = (\mathbf{0}, \mathbf{v}_i, \mathbf{0}), \qquad \widetilde{\mathbf{v}}^i_{0,0} = (\mathbf{0}, \mathbf{0}, \mathbf{0}) \in \mathbb{Z}^{3D},$$

where $\mathbf{0}$ denotes the zero vector in $\mathbb{Z}^D$. Let $\mathbf{z}, \mathbf{w} \in \mathbb{Z}^D$ and $\mathcal{X} = (\{0,1\}^n)^{\mathcal{V}}_{\mathbf{z}}$.

We identify $(\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}$ with $(J^n)^{\widetilde{\mathcal{V}}}_{\widetilde{\mathbf{x}}}$, where $\widetilde{\mathbf{x}} := (\mathbf{z}, \mathbf{z}, \mathbf{w})$. We define

$$\mu_{\widetilde{\mathbf{q}}} := \mu^{\widetilde{\mathcal{V}}}_{\widetilde{\mathbf{x}}} \quad \text{and} \quad \mu_{\mathbf{p}} := \mu^{\mathcal{V}}_{\mathbf{z}}.$$

We denote the marginals of $\mu_{\widetilde{\mathbf{q}}}$ by $\mu_{\widetilde{\mathbf{p}}}$ (both marginals are equal).

Next we show κ-boundedness of the above measures under our usual assumptions on $\mathcal{V}$ (and justify the final statement of the above definition).

Lemma 9.2. *Let $0 < n^{-1} \ll \kappa \ll \gamma, \gamma' \ll \lambda \ll \varepsilon, D^{-1}, C^{-1}, k^{-1}$. Let $\mathbf{R} \in \mathbb{R}^D$ with $\max_d R_d \le n^C$. and $\widetilde{\mathbf{R}} = (\mathbf{R}, \mathbf{R}, \mathbf{R}) \in \mathbb{Z}^{3D}$. Suppose $\mathcal{V} = (\mathbf{v}_i : i \in [n])$ is an $\mathbf{R}$-bounded, γ'-robustly $(\gamma, \mathbf{R})$-generic γ-robustly $(\mathbf{R}, k)$-generating array in $\mathbb{Z}^D$. Then $\widetilde{\mathcal{V}}$ is $\widetilde{\mathbf{R}}$-bounded, $(\gamma'/2)$-robustly $(\gamma^3, \widetilde{\mathbf{R}})$-generic and $(\gamma/2)$-robustly $(\widetilde{\mathbf{R}}, 3k)$-generating. Suppose also that $\dim_{VC}((\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}) \ge \lambda n$. Then $\mu_{\mathbf{p}}$ and $\mu_{\widetilde{\mathbf{q}}}$ are κ-bounded, both marginals of $\mu_{\widetilde{\mathbf{q}}}$ are $\mu_{\widetilde{\mathbf{p}}}$, and $\mu_{\widetilde{\mathbf{p}}} \in \mathcal{M}^{\mathcal{V}}_{\mathbf{z}}$.*

Proof. The proof of the first statement is 'definition chasing' so we omit it. The last statement follows from symmetry and strict concavity of $L(p)$ (see Lemma 2.5 *ii*). For κ-boundedness of $\mu_{\mathbf{p}}$ and $\mu_{\widetilde{\mathbf{q}}}$ we apply Lemma 4.8. For $\mu_{\mathbf{p}}$ this is valid as $\dim_{VC}(\mathcal{X}) \ge \lambda n$ and $\mathcal{V}$ is γ-robustly $(\gamma, \mathbf{R})$-generic. For $\mu_{\widetilde{\mathbf{q}}}$ this is valid as $\dim_{VC}((\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}) \ge \lambda n$ and $\widetilde{\mathcal{V}}$ is $(\gamma'/2)$-robustly $(\gamma^3, \widetilde{\mathbf{R}})$-generic. $\square$

Now we prove the main lemma of this subsection, which distinguishes cases *i* and *ii* according to $\|\mathbf{p} - \widetilde{\mathbf{p}}\|_1 := \sum_{i \in [n], j \in J} |p^i_j - \widetilde{p}^i_j|$.

Lemma 9.3. *Let $0 < n^{-1} \ll \delta \ll \delta_1 \ll \gamma, \gamma' \ll \lambda \ll \varepsilon, D^{-1}, C^{-1}, k^{-1}$ and let $\mathbf{R} \in \mathbb{R}^D$ with $\max_d R_d \le n^C$. Suppose $\mathcal{V} = (\mathbf{v}_i : i \in [n])$ is an $\mathbf{R}$-bounded, γ'-robustly $(\gamma, \mathbf{R})$-generic γ-robustly $(\mathbf{R}, k)$-generating array in $\mathbb{Z}^D$. Fix notation as in Definition 9.1 and suppose $\dim_{VC}((\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}) \ge \lambda n$.*

i. Suppose $\|\mathbf{p} - \widetilde{\mathbf{p}}\|_1 \le \delta_1 n$. If $\mathcal{A} \subset \mathcal{X}$ with $|\mathcal{A}| \ge (1-\delta)^n |\mathcal{X}|$ then $|(\mathcal{A} \times \mathcal{A})^{\mathcal{V}_\cap}_{\mathbf{w}}| \ge (1-\varepsilon)^n |(\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}|$.

ii. Suppose $\|\mathbf{p} - \widetilde{\mathbf{p}}\|_1 \ge \delta_1 n$. Then there is $\mathcal{B}_{full} \subset \mathcal{X}$ with $|\mathcal{B}_{full}| \le (1-\delta)^n |\mathcal{X}|$ and

$$|(\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}} \setminus (\mathcal{B}_{full} \times \mathcal{B}_{full})^{\mathcal{V}_\cap}_{\mathbf{w}}| \le (1-\delta)^n |(\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}|.$$

Furthermore, if $\mathcal{B} \subset \mathcal{B}_{full}$ with $|\mathcal{B}| \ge (1-\delta)^n |\mathcal{B}_{full}|$ then $|(\mathcal{B} \times \mathcal{B})^{\mathcal{V}_\cap}_{\mathbf{w}}| \ge (1-\varepsilon)^n |(\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}|$.

Proof. By Lemma 9.2, $\mu_{\mathbf{p}}$ and $\mu_{\widetilde{\mathbf{q}}}$ (and so $\mu_{\widetilde{\mathbf{p}}}$) are κ-bounded, where $\kappa \ll \gamma, \gamma'$. Then by Theorem 1.20, $\mu_{\mathbf{p}} \approx_{\triangle} \nu$, where ν is the uniform distribution on $\triangle_n = (\{0,1\}^n)^{\mathcal{V}}_{\mathbf{z}} = \mathcal{X}$. Also by Theorem 1.20, $\mu_{\widetilde{\mathbf{q}}} \approx_{\triangle'} \nu'$, where ν' is the uniform distribution on $\triangle'_n = (J^n)^{\widetilde{\mathcal{V}}}_{\widetilde{\mathbf{x}}} = (\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}$.

Fix constants $\delta \ll \delta_0 \ll \delta_1 \ll \delta_2 \ll \varepsilon_1 \ll \kappa$.

Case *i*: $\|\mathbf{p} - \widetilde{\mathbf{p}}\|_1 \le \delta_1 n$.

Given $\mathcal{A} \subset \mathcal{X}$ with $|\mathcal{A}| \ge (1-\delta)^n |\mathcal{X}|$, we have $\mu_{\mathbf{p}}(\mathcal{A}) \ge (1-\delta_1)^n$ by Theorem 1.20. As both $\mathbf{p}$ and $\widetilde{\mathbf{p}}$ are κ-bounded, and $\|\mathbf{p} - \widetilde{\mathbf{p}}\|_1 \le \delta_1 n$, we have $\mu_{\widetilde{\mathbf{p}}}(\mathcal{A}) \ge (1-\delta_1)^n (\kappa)^{\delta_1 n} \ge (1-\delta_2)^n$. As the hypotheses of Theorem 1.17 hold, we find $\mu_{\widetilde{\mathbf{q}}}((\mathcal{A}\times\mathcal{A})^{\mathcal{V}_\cap}_{\mathbf{w}}) \geq (1-\varepsilon_1)^n$. Theorem 1.20 applied once again for $\mu_{\widetilde{\mathbf{q}}}$ gives $|(\mathcal{A}\times\mathcal{A})^{\mathcal{V}_\cap}_{\mathbf{w}}| \geq (1-\varepsilon)^n |(J^n)^{\widetilde{\mathcal{V}}}_{\widetilde{\mathbf{x}}}| = (1-\varepsilon)^n |(\mathcal{X}\times\mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}|$.

Case *ii*: $\|\mathbf{p}-\widetilde{\mathbf{p}}\|_1 \geq \delta_1 n$.

In this case, we let

$$\mathcal{B}_{full} = \left\{\mathbf{x}\in\mathcal{X} : -\log_2 \mu_{\widetilde{\mathbf{p}}}(\mathbf{x}) = H(\mu_{\widetilde{\mathbf{p}}}) \pm \delta_1^2 n/2\right\}.$$

Note that $H(\mu_{\widetilde{\mathbf{p}}}) \leq H(\mu_{\mathbf{p}}) - \delta_1^2 n$ by Lemma 2.6, which gives $|\mathcal{B}_{full}| \leq 2^{H(\mu_{\mathbf{p}})-\delta_1^2 n/2} \leq (1-\delta)^n|\mathcal{X}|$, where we use Lemma 3.5 in the second inequality.

Next we show that almost all $\mathbf{w}$-intersections in $\mathcal{X}$ are contained in $\mathcal{B}_{full}$. We require an upper bound on the size of $\mathcal{Y} := (\mathcal{X}\times\mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}} \setminus (\mathcal{B}_{full}\times\mathcal{B}_{full})^{\mathcal{V}_\cap}_{\mathbf{w}}$. Note that $\mathcal{Y} \subset (\mathcal{X}\times(\mathcal{X}\setminus\mathcal{B}_{full}))^{\mathcal{V}_\cap}_{\mathbf{w}} \cup ((\mathcal{X}\setminus\mathcal{B}_{full})\times\mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}$. As $\mu_{\widetilde{\mathbf{p}}}$ is a marginal of $\mu_{\widetilde{\mathbf{q}}}$, this gives $\mu_{\widetilde{\mathbf{q}}}(\mathcal{Y}) \leq 2\mu_{\widetilde{\mathbf{q}}}((\mathcal{X}\times(\mathcal{X}\setminus\mathcal{B}_{full}))^{\mathcal{V}_\cap}_{\mathbf{w}}) \leq 2\mu_{\widetilde{\mathbf{p}}}(\mathcal{X}\setminus\mathcal{B}_{full})$. However, $\mu_{\widetilde{\mathbf{p}}}$ is κ-bounded, so Lemma 3.1 gives $\mu_{\widetilde{\mathbf{q}}}(\mathcal{Y}) \leq 2\mu_{\widetilde{\mathbf{p}}}(\mathcal{X}\setminus\mathcal{B}_{full}) \leq (1-\delta_0)^n$. As $\mu_{\widetilde{\mathbf{q}}} \approx_{\Delta'} \nu'$ by Theorem 1.20, this gives $|\mathcal{Y}| \leq (1-\delta)^n|(\mathcal{X}\times\mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}|$, as required.

It remains to show supersaturation relative to $\mathcal{B}_{full}$; the proof is similar to that of case i. As $\mu_{\widetilde{\mathbf{p}}}$ is κ-bounded, Lemma 3.5 gives $\log_2|\mathcal{B}_{full}| \geq H(\mu_{\widetilde{\mathbf{p}}}) - \delta_1 n$. Suppose $\mathcal{B}\subset\mathcal{B}_{full}$ with $|\mathcal{B}| \geq (1-\delta)^n|\mathcal{B}_{full}|$. Then $\mu_{\widetilde{\mathbf{p}}}(\mathcal{B}) \geq |\mathcal{B}|2^{-H(\mu_{\widetilde{\mathbf{p}}})-\delta_1^2 n/2} \geq (1-\delta_1)^n$. Theorem 6.3 gives $\mu_{\widetilde{\mathbf{q}}}((\mathcal{B}\times\mathcal{B})^{\mathcal{V}_\cap}_{\mathbf{w}}) \geq (1-\varepsilon_1)^n$. A final application of Theorem 1.20 gives $|(\mathcal{B}\times\mathcal{B})^{\mathcal{V}_\cap}_{\mathbf{w}}| \geq (1-\varepsilon)^n|(\mathcal{X}\times\mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}|$. $\square$

9.2 Small VC-dimension

To complete the proof of Theorem 1.11, it remains to show the negative result in the case that $(\mathcal{X}\times\mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}$ has small VC-dimension, i.e. that there is a large subset of $\mathcal{X}$ with no $\mathbf{w}$-intersection.

First we use universal VC-dimension (see Definition 4.3) to give a criterion for $(\mathcal{X}\times\mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}$ to have large VC-dimension (which will be used in contrapositive form). We require the following notation. Given $\mathbf{x}\in\mathcal{X}$, $j\in\{0,1\}$, $\alpha>0$ let

$$S_j(\mathbf{x}) = \{i\in[n] : x_i = j\}, \qquad \mathcal{V}^j_{\mathbf{x}} = (\mathbf{v}_i : i\in S_j(\mathbf{x})),$$
$$N^1_{\mathbf{w}}(\mathbf{x}) = (\{0,1\}^{S_1(\mathbf{x})})^{\mathcal{V}^1_{\mathbf{x}}}_{\mathbf{w}}, \qquad N^0_{\mathbf{z}-\mathbf{w}}(\mathbf{x}) = (\{0,1\}^{S_0(\mathbf{x})})^{\mathcal{V}^0_{\mathbf{x}}}_{\mathbf{z}-\mathbf{w}},$$
$$\mathcal{X}^\alpha_{\mathbf{w}} = \{\mathbf{x}\in\mathcal{X} : |N^0_{\mathbf{z}-\mathbf{w}}(\mathbf{x})| \geq (1+\alpha)^n \text{ and } |N^1_{\mathbf{w}}(\mathbf{x})| \geq (1+\alpha)^n\}.$$

An important observation is

$$\left(\mathbf{x}'\in\mathcal{X} \text{ and } \mathcal{V}_\cap(\mathbf{x},\mathbf{x}') = \mathbf{w}\right) \Leftrightarrow \left(\mathbf{x}' = \mathbf{y}_0\circ\mathbf{y}_1 \text{ with } \mathbf{y}_0\in N^0_{\mathbf{z}-\mathbf{w}}(\mathbf{x}) \text{ and } \mathbf{y}_1\in N^1_{\mathbf{w}}(\mathbf{x})\right).$$

Lemma 9.4. *Let $n^{-1} \ll \lambda \ll \gamma,\gamma' \ll \alpha \ll \varepsilon, D^{-1}, C^{-1}, k^{-1}$ and let $\mathbf{R}\in\mathbb{R}^D$ with $\max_d R_d \leq n^C$. Suppose $\mathcal{V} = (\mathbf{v}_i : i\in[n])$ where each $\mathbf{v}_i\in\mathbb{Z}^D$ is $\mathbf{R}$-bounded and $\mathcal{V}$ is γ'-robustly $(\gamma,\mathbf{R})$-generic and γ-robustly $(\mathbf{R},k)$-generating. Suppose $|\mathcal{X}| \geq (1+\varepsilon)^n$ and $|\mathcal{X}^\alpha_{\mathbf{w}}| \geq |\mathcal{X}|/2$. Then $\dim_{VC}((\mathcal{X}\times\mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}) \geq \lambda n$.*

Proof. The strategy of the proof is to find a large set S that is shattered by a subset $\mathcal{X}'$ of $\mathcal{X}$, such that if $\mathbf{x}\in\mathcal{X}'$ then $\dim_{UVC}(N^0_{\mathbf{z}-\mathbf{w}}(\mathbf{x}))$ and $\dim_{UVC}(N^1_{\mathbf{w}}(\mathbf{x}))$ are large. Then the definition of universal VC-dimension will imply that S is shattered by $(\mathcal{X}\times\mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}$, as $N^0_{\mathbf{z}-\mathbf{w}}(\mathbf{x})$ shatters $S_0(\mathbf{x})$ and $N^1_{\mathbf{w}}(\mathbf{x})$ shatters $S_1(\mathbf{x})$. First we note by Lemma 1.21 that $\mu_{\mathbf{p}}$ is κ-bounded, where $\alpha \ll \kappa \ll \varepsilon$.

Let $\mathcal{X}'$ be the set of $\mathbf{x}\in\mathcal{X}^\alpha_{\mathbf{w}}$ such that $\mathcal{V}^1_{\mathbf{x}}$ and $\mathcal{V}^0_{\mathbf{x}}$ are $(\kappa^k\gamma/2k)$-robustly $(\mathbf{R},k)$-generating. We claim that $|\mathcal{X}'| \geq |\mathcal{X}|/4$. To see this, note that for any $\mathbf{w}\in\mathbb{Z}^D$ with $\|\mathbf{w}\|_{\mathbf{R}} \leq 1$, as $\mathcal{V}$ is γ-robustly $(\mathbf{R},k)$-generating, there are $L \geq \gamma n/k$ disjoint sets $S_1,\ldots,S_L$, with $|S_\ell| \leq k$ for all $\ell\in[L]$, such that for all $\ell\in[L]$ there is a partition $S_\ell = S^1_\ell \cup S^0_\ell$ with $\sum_{i\in S^1_\ell}\mathbf{v}_i - \sum_{i\in S^0_\ell}\mathbf{v}_i = \mathbf{w}$. Given $\mathbf{x}\in\{0,1\}^n$ and

$j \in \{0,1\}$, let $L^j_{\mathbf{w}}(\mathbf{x}) = \{\ell \in [L] : S_\ell \subset S_j(\mathbf{x})\}$. As $\mu_{\mathbf{p}}$ is κ-bounded, each $\mathbb{E}_{\mathbf{x}\sim\mu_{\mathbf{p}}}(|L^j_{\mathbf{w}}(\mathbf{x})|) \geq \kappa^k L$. Let $\mathcal{B}_{\mathbf{w}}$ be the event that either $|L^0_{\mathbf{w}}(\mathbf{x})| \leq \kappa^k L/2$ or $|L^1_{\mathbf{w}}(\mathbf{x})| \leq \kappa^k L/2$. Also let $\mathcal{B}$ be the union of $\mathcal{B}_{\mathbf{w}}$ over all $\mathbf{w} \in \mathbb{Z}^D$ with $\|\mathbf{w}\|_{\mathbf{R}} \leq 1$. There are at most $(2n+1)^{CD}$ choices of $\mathbf{w}$, so by Chernoff's inequality and a union bound $\mathbb{P}_{\mathbf{x}\sim\mu_{\mathbf{p}}}(\mathcal{B}) = (1-c_\kappa)^n$ for some $c_\kappa > 0$. By Theorem 1.20, we deduce $|\mathcal{X}^\alpha_{\mathbf{w}} \setminus \mathcal{B}| \geq |\mathcal{X}|/4$. As $\mathcal{X}^\alpha_{\mathbf{w}} \setminus \mathcal{B} \subset \mathcal{X}'$ this proves the claim.

Next we claim that if $\mathbf{x} \in \mathcal{X}'$ then $\dim_{UVC}(N^0_{\mathbf{z}-\mathbf{w}}(\mathbf{x})) \geq \lambda n$ in $\{0,1\}^{S_0(\mathbf{x})}$ and $\dim_{UVC}(N^1_{\mathbf{w}}(\mathbf{x})) \geq \lambda n$ in $\{0,1\}^{S_1(\mathbf{x})}$. Indeed, as $\mathbf{x} \in \mathcal{X}^\alpha_{\mathbf{w}}$ we have $|S_1(\mathbf{x})|, |S_0(\mathbf{x})| \geq \log_2(1+\alpha)n$, so $\mathcal{V}^0_{\mathbf{x}}$ and $\mathcal{V}^1_{\mathbf{x}}$ are $(\gamma'/\log_2(1+\alpha))$-robustly $(\gamma, \mathbf{R})$-generic, and by Lemma 4.7 both $N^0_{\mathbf{z}-\mathbf{w}}(\mathbf{x})$ and $N^1_{\mathbf{w}}(\mathbf{x})$ have VC-dimension at least $\alpha' n$, where $\gamma, \gamma' \ll \alpha' \ll \alpha$. Both $\mathcal{V}^0_{\mathbf{x}}$ and $\mathcal{V}^1_{\mathbf{x}}$ are also clearly $\mathbf{R}$-bounded, and by definition of $\mathcal{X}'$ these vectors are $(\kappa^k\gamma/2k)$-robustly $(\mathbf{R}, k)$-generating, so the claim follows from Lemma 4.11.

Now we can implement the strategy outlined at the start of the proof. As $|\mathcal{X}'| \geq |\mathcal{X}|/4 \geq (1+\varepsilon)^n/4$, we have $\dim_{VC}(\mathcal{X}') \geq \lambda n$ by Lemma 4.7. Let $S \subset [n]$ with $|S| \geq \lambda n$ be shattered by $\mathcal{X}'$. We will show that S is also shattered by $(\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}} \subset (\{0,1\} \times \{0,1\})^n$. Indeed, suppose that we are given a partition $\cup_{j_1,j_2\in\{0,1\}} S_{j_1,j_2}$ of S, and wish to find sets $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ such that $\mathcal{V}_\cap(\mathbf{x}, \mathbf{x}') = \mathbf{w}$ and $\{i \in S : x_i = j_1, x'_i = j_2\} = S_{j_1,j_2}$. As $\mathcal{X}'$ shatters S, there is $\mathbf{x} \in \mathcal{X}'$ with $\{i \in S : x_i = 1\} = S_{1,0} \cup S_{1,1}$. Furthermore, using the universal VC-dimension of $N^1_{\mathbf{w}}(\mathbf{x})$ and $N^0_{\mathbf{z}-\mathbf{w}}(\mathbf{x})$, we have $\mathbf{y}_1 \in N^1_{\mathbf{w}}(\mathbf{x})$ with $\{i \in S_{1,0} \cup S_{1,1} : (\mathbf{y}_1)_i = 1\} = S_{1,1}$ and $\mathbf{y}_0 \in N^0_{\mathbf{z}-\mathbf{w}}(\mathbf{x})$ with $\{i \in S_{0,0} \cup S_{0,1} : (\mathbf{y}_2)_i = 1\} = S_{0,1}$. Now $\mathbf{x}' = \mathbf{y}_0 \circ \mathbf{y}_1 \in \mathcal{X}$ with $\mathcal{V}_\cap(\mathbf{x}, \mathbf{x}') = \mathbf{w}$ and $\{i \in S : x_i = j_1, x'_i = j_2\} = S_{j_1,j_2}$ for all $j_1, j_2 \in \{0,1\}$. Thus S is shattered by $(\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}$, and so $\dim_{VC}((\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}) \geq \lambda n$, as required. □

We conclude with the main result of this subsection, that there is a large subset of $\mathcal{X}$ with no $\mathbf{w}$-intersection.

Lemma 9.5. *Let $n^{-1} \ll \lambda \ll \gamma, \gamma' \ll \varepsilon, D^{-1}, C^{-1}, k^{-1}$ and let $\mathbf{R} \in \mathbb{R}^D$ with $\max_d R_d \leq n^C$. Suppose $\mathcal{V} = (\mathbf{v}_i : i \in [n])$ where each $\mathbf{v}_i \in \mathbb{Z}^D$ is $\mathbf{R}$-bounded and $\mathcal{V}$ is γ'-robustly $(\gamma, \mathbf{R})$-generic and γ-robustly $(\mathbf{R}, k)$-generating. Suppose $\mathbf{z} \neq \mathbf{w}$ and $\dim_{VC}((\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}) \leq \lambda n$. Then there is $\mathcal{B}_{empty} \subset \mathcal{X}$ with $|\mathcal{B}_{empty}| \geq \lfloor (1-\varepsilon)^n |\mathcal{X}| \rfloor$ and $(\mathcal{B}_{empty} \times \mathcal{B}_{empty})^{\mathcal{V}_\cap}_{\mathbf{w}} = \emptyset$.*

Proof. We may assume $|\mathcal{X}| \geq (1+\varepsilon)^n$ as otherwise we can take $\mathcal{B}_{empty} = \emptyset$. Take α and ξ such that $\gamma, \gamma' \ll \alpha \ll \xi \ll \varepsilon, D^{-1}, C^{-1}, k^{-1}$. Let $\mathcal{X}_0 = \{\mathbf{x} \in \mathcal{X} : |N^0_{\mathbf{z}-\mathbf{w}}(\mathbf{x})| \leq (1+\alpha)^n\}$ and $\mathcal{X}_1 = \{\mathbf{x} \in \mathcal{X} : |N^1_{\mathbf{w}}(\mathbf{x})| \leq (1+\alpha)^n\}$. Then $|\mathcal{X}_0 \cup \mathcal{X}_1| \geq |\mathcal{X}|/2$ by Lemma 9.4. The remainder of the proof splits into two similar cases according to which $\mathcal{X}_j$ is large; we will give full details for the case $j=1$ and then indicate the necessary modifications for $j=0$.

Suppose $|\mathcal{X}_1| \geq |\mathcal{X}|/4$. By the pigeonhole principle, we can fix $\mathcal{X}' \subset \mathcal{X}_1$ and $t \in [n]$ such that $|\mathcal{X}'| \geq |\mathcal{X}|/4n$ and $|S_1(\mathbf{x})| = t$ for all $\mathbf{x} \in \mathcal{X}'$. As $\binom{n}{t} \geq |\mathcal{X}'| \geq (1+\varepsilon)^n/4n$ we have $\xi n \leq t \leq n - \xi n$. Next we can pass to a subset $\mathcal{X}'' \subset \mathcal{X}'$ with $|\mathcal{X}''| \geq |\mathcal{X}'|/2\binom{n}{2\xi n} \geq (1-\xi^{1/2})^n|\mathcal{X}|$ that is 'well-separated', in that the Hamming distance $d(\mathbf{x}, \mathbf{x}') \geq 2\xi n$ for all distinct $\mathbf{x}, \mathbf{x}' \in \mathcal{X}''$. Indeed, we can select $\mathcal{X}''$ greedily, noting that each element of $\mathcal{X}''$ forbids at most $\sum_{i\in[0,2\xi n]} \binom{n}{i} \leq 2\binom{n}{2\xi n}$ elements from $\mathcal{X}''$. As $|S_1(\mathbf{x})| = |S_1(\mathbf{x}')| = t$, this gives $|S_1(\mathbf{x}) \setminus S_1(\mathbf{x}')| \geq \xi n$ for all distinct $\mathbf{x}, \mathbf{x}' \in \mathcal{X}''$.

Next we will define $\mathcal{B}_{empty}$. We randomly select $S \subset [n]$ with $|S| = \xi n$, and let $\mathcal{C} = \{\mathbf{x} \in \mathcal{X}'' : S \subset S_1(\mathbf{x})\}$. We say that $\mathbf{x} \in \mathcal{C}$ is *isolated* if there is no $\mathbf{x}' \in \mathcal{C}$ with $\mathcal{V}_\cap(\mathbf{x}, \mathbf{x}') = \mathbf{w}$. We let $\mathcal{B}_{empty}$ be the set of isolated $\mathbf{x} \in \mathcal{C}$. Then by definition we have $(\mathcal{B}_{empty} \times \mathcal{B}_{empty})^{\mathcal{V}_\cap}_{\mathbf{w}} = \emptyset$.

Now we will show that $\mathbb{E}|\mathcal{B}_{empty}| \geq (1-\varepsilon)^n|\mathcal{X}|$. As $\mathbb{E}(|\mathcal{C}|) = \binom{t}{\xi n}\binom{n}{\xi n}^{-1}|\mathcal{X}''|$, $|\mathcal{X}''| \geq (1-\xi^{1/2})^n|\mathcal{X}|$ and $\xi \ll \varepsilon$, it suffices to show $\mathbb{P}(\mathbf{x} \text{ is isolated} \mid \mathbf{x} \in \mathcal{C}) \geq 1/2$ for all $\mathbf{x} \in \mathcal{X}'$. To see this, we condition on $\mathbf{x} \in \mathcal{C}$ and note that S is equally likely to be any subset of $S_1(\mathbf{x})$ of size ξn. Consider any $\mathbf{x}' \in \mathcal{X}''$ with $\mathcal{V}_\cap(\mathbf{x}, \mathbf{x}') = \mathbf{w}$. Note that $\mathbf{x} \neq \mathbf{x}'$ since $\mathcal{V}(\mathbf{x}, \mathbf{x}') = \mathbf{z} \neq \mathbf{w}$, and so $S_1(\mathbf{x}) \neq S_1(\mathbf{x}')$ as both sets have size t. Furthermore $\mathbf{y} := \mathbf{x}'|_{S_1(\mathbf{x})} \in N^1_{\mathbf{w}}(\mathbf{x})$, we have $|S_1(\mathbf{x}) \setminus S_1(\mathbf{y})| \geq \xi n$ by definition of $\mathcal{X}''$, and $\mathbf{x}' \in \mathcal{C} \Leftrightarrow S \subset S_1(\mathbf{y})$. For fixed $\mathbf{y}$ we have $\mathbb{P}(S \subset S_1(\mathbf{y})) \leq \binom{t-\xi n}{\xi n}\binom{t}{\xi n}^{-1} \leq (1-\xi)^{\xi n}$. By definition of $\mathcal{X}_1$ we have a union bound over at most $(1+\alpha)^n$ choices of $\mathbf{y} \in N^1_{\mathbf{w}}(\mathbf{x})$, so as $\alpha \ll \xi$, the probability that $\mathbf{x}$ is not isolated given $\mathbf{x} \in \mathcal{C}$ is $o(1)$, so at most $1/2$, as required.

Similarly, if $|\mathcal{X}_0| \geq |\mathcal{X}|/4$, we define $\mathcal{X}'$ and $\mathcal{X}''$ in the same way for $\mathcal{X}_0$, and let $\mathcal{C} = \{\mathbf{x} \in \mathcal{X}'' : S \subset S_0(\mathbf{x})\}$. We use the same definition of $\mathcal{B}_{empty}$ as before, and bound the probability that $\mathbf{x}$ is not isolated given $\mathbf{x} \in \mathcal{C}$ by taking a union bound over at most $(1+\alpha)^n$ choices of $\mathbf{y} := \mathbf{x}'|_{S_0(\mathbf{x})} \in N^0_{\mathbf{z}-\mathbf{w}}(\mathbf{x})$. The remaining details of this case are the same, so we omit them. □

9.3 Proof of Theorem 1.11

Proof of Theorem 1.11. Take λ with $\gamma_1, \gamma_1' \ll \lambda \ll \gamma_2, \gamma_2'$. If $\dim_{VC}((\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}) \geq \lambda n$, then we can apply Lemma 9.3 with $\gamma = \gamma_1$ and $\gamma' = \gamma_1'$ to obtain case *i* or *ii* of Theorem 1.11. On the other hand, if $\dim_{VC}((\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}) \leq \lambda n$ then we apply Lemma 9.5 with $\gamma = \gamma_2$ and $\gamma' = \gamma_2'$ to obtain case *iii* of Theorem 1.11. □

10 Solution of Kalai's Conjecture

In this section we prove Theorem 1.3, which is our solution to Kalai's Conjecture 1.2. We give the proof in the first subsection, then generalise it in the following subsection to show that supersaturation of the type conjectured by Kalai is quite rare.

10.1 Proof of Theorem 1.3

As described in subsection 1.4, the supersaturation conclusion desired by Conjecture 1.2 (case *i* of Theorem 1.11) needs the maximum entropy measure $\mu_{\widetilde{\mathbf{q}}}$ that represents $(\mathcal{X} \times \mathcal{X})^{\mathcal{V}_\cap}_{\mathbf{w}}$ to have marginals $\mu_{\widetilde{\mathbf{p}}}$ close to $\mu_{\mathbf{p}} := \mu_{\mathbf{p}^{\mathcal{V}}_{\mathbf{z}}}$. Recall that in Definition 9.1 we constructed $\mu_{\widetilde{\mathbf{q}}}$ as $\mu^{\widetilde{\mathcal{V}}}_{\widetilde{\mathbf{x}}}$, where $\widetilde{\mathcal{V}}$ is a certain $(n, \{0,1\} \times \{0,1\})$-array in $\mathbb{Z}^{3D}$ and $\widetilde{\mathbf{x}} := (\mathbf{z}, \mathbf{z}, \mathbf{w})$. In this subsection we work with the Kalai vectors $\mathcal{V} = (\mathbf{v}_i)_{i \in [n]}$ with $\mathbf{v}_i = (1, i)$, so $D = 2$. In the notation of Conjecture 1.2 we have $\mathbf{z} = (k, s)$ and $\mathbf{w} = (t, w)$. Sometimes we will indicate the dependence on n as a subscript in our notation, e.g. writing $\mathbf{z}_n = (k_n, s_n) = (\lfloor \alpha_1 n \rfloor, \lfloor \alpha_2 \binom{n}{2} \rfloor)$. Our proof will use the following concrete description of the maximum entropy measures as Boltzmann distributions.

Lemma 10.1. *Let $\mathcal{V} = (\mathbf{v}^i_j)$ be an (n, J)-array in $\mathbb{Z}^D$ and $\mathbf{z} \in \mathbb{Z}^D$. Suppose $\mathbf{p} = \mathbf{p}^{\mathcal{V}}_{\mathbf{z}}$ has all $p^i_j \neq 0$. Then there is $\boldsymbol{\lambda} \in (\mathbb{R}^D)^J$ such that all $p^i_j = Z_i^{-1} e^{\boldsymbol{\lambda}_j \cdot \mathbf{v}^i_j}$, where $Z_i = \sum_{j \in J} e^{\boldsymbol{\lambda}_j \cdot \mathbf{v}^i_j}$.*

Proof. By the theory of Lagrange multipliers, $\mathbf{p}$ is a stationary point of

$$L(\mathbf{p}, \boldsymbol{\lambda}, \boldsymbol{\lambda}') = H(\mathbf{p}) \log 2 + \sum_{d \in [D]} \sum_{j \in J} \lambda_{j,d} \Big(\sum_{i \in [n]} p^i_j (v^i_j)_d - z_d \Big) + \sum_{i \in [n]} \lambda_i' \Big(1 - \sum_{j \in J} p^i_j \Big).$$

Thus $0 = -1 - \lambda_i' - \log(p^i_j) + \sum_{d \in [D]} \lambda_{j,d} (v^i_j)_d$ for each i and j, which gives the stated formula. □

When $\mathbf{p} = \mathbf{p}^{\mathcal{V}}_{\mathbf{z}}$ and $\mu_{\widetilde{\mathbf{q}}} = \mu^{\widetilde{\mathcal{V}}}_{\widetilde{\mathbf{x}}}$ are κ-bounded we can describe them explicitly using Lemma 10.1. For $\mathbf{p}$ we obtain $\boldsymbol{\lambda} = (\lambda_1, \lambda_2) \in \mathbb{R}^2$ such that $\mathbf{p} = \mathbf{p}^{\mathcal{V}}_{\mathbf{z}}$ is given by $p^i_1 = e^{\lambda_1 + \lambda_2 (i/n)} (1 + e^{\lambda_1 + \lambda_2 (i/n)})^{-1}$ (it is convenient to rescale, using λ_2/n in place of λ_2). To determine whether $\mathbf{p}$ is close to $\widetilde{\mathbf{p}}$, it will be more convenient to pass to a limit problem in which closeness is replaced by equality. With this in mind, we write

$$p_1^i = (p_{\boldsymbol{\lambda}}^{(n)})_1^i := f_{\boldsymbol{\lambda}}(i/n), \text{ where } f_{\boldsymbol{\lambda}}(x) = e^{\lambda_1+\lambda_2 x}(1+e^{\lambda_1+\lambda_2 x})^{-1}. \tag{4}$$

Similarly, Lemma 10.1 gives $\boldsymbol{\pi} = (\pi_1, \pi_1', \pi_2, \pi_2') \in \mathbb{R}^4$ (using the symmetry between $(0,1)$ and $(1,0)$) such that $\mu_{\widetilde{\mathbf{q}}} = \mu_{\widetilde{\mathbf{x}}}^{\widetilde{\mathcal{V}}}$ is given by $\widetilde{q}_{j,j'}^i = (\mathbf{q}_{\boldsymbol{\pi}}^{(n)})_{j,j'}^i := g_{j,j'}^{\boldsymbol{\pi}}(i/n)$, where

$$\begin{aligned} g_{0,0}^{\boldsymbol{\pi}}(x) &= Z_{\boldsymbol{\pi}}(x)^{-1}, & g_{0,1}^{\boldsymbol{\pi}}(x) &= g_{1,0}^{\boldsymbol{\pi}}(x) = e^{\pi_1+\pi_2 x} Z_{\boldsymbol{\pi}}(x)^{-1}, \\ g_{1,1}^{\boldsymbol{\pi}}(x) &= e^{\pi_1'+\pi_2' x} Z_{\boldsymbol{\pi}}(x)^{-1}, & &\text{with } Z_{\boldsymbol{\pi}}(x) = 1 + 2e^{\pi_1+\pi_2 x} + e^{\pi_1'+\pi_2' x}. \end{aligned} \tag{5}$$

The limit marginal problem is to characterise $\boldsymbol{\lambda}$ and $\boldsymbol{\pi}$ such that $f_{\boldsymbol{\lambda}}(x) = g_{0,1}^{\boldsymbol{\pi}}(x) + g_{1,1}^{\boldsymbol{\pi}}(x)$.

Next we formulate the constraints on $\boldsymbol{\lambda}$ and $\boldsymbol{\pi}$ defined by the parameters $\alpha_1, \alpha_2, \beta_1, \beta_2$ of Conjecture 1.2, namely $\sum_{i\in[n]} p_1^i(1,i) = (k,s) = (\alpha_1 n, \alpha_2\binom{n}{2})$ and $\sum_{i\in[n]} \widetilde{q}_{1,1}^i(1,i) = (t,w) = (\beta_1 n, \beta_2\binom{n}{2})$. The limit versions of these constraints are $h(\boldsymbol{\lambda}) = \boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ and $h^*(\boldsymbol{\pi}) = \boldsymbol{\beta} = (\beta_1, \beta_2)$, where

$$h(\boldsymbol{\lambda}) = \int_0^1 (1,2x) f_{\boldsymbol{\lambda}}(x)dx \qquad \text{and} \qquad h^*(\boldsymbol{\pi}) = \int_0^1 (1,2x) g_{1,1}^{\boldsymbol{\pi}}(x)dx. \tag{6}$$

The following lemma shows that we can think of $\boldsymbol{\lambda}$ as a reparameterisation of $\boldsymbol{\alpha}$, and that large finite instances of $[n]_{k,s}$ are well-approximated by the limit. Recall that a homeomorphism is a continuous bijection with a continuous inverse.

Lemma 10.2.

i. h is a homeomorphism between $\mathbb{R}^2$ and Λ.

ii. For $\boldsymbol{\alpha} \in \Lambda$ and large n we have $\mu_{\mathbf{z}_n}^{\mathcal{V}_n} = \mu_{\mathbf{p}}$, for some $\mathbf{p} = \mathbf{p}_{\boldsymbol{\lambda}^{(n)}}^{(n)}$ where $\boldsymbol{\lambda}^{(n)} \to \boldsymbol{\lambda} = h^{-1}(\boldsymbol{\alpha})$.

Proof. We start by noting that h is continuous. Next we claim that $h(\boldsymbol{\lambda}) \in \Lambda$ for all $\boldsymbol{\lambda} \in \mathbb{R}^2$. To see this, note that $0 \le f_{\boldsymbol{\lambda}}(x) \le 1$ for all $x \in [0,1]$. Then given $\alpha_1 = \int_0^1 f_{\boldsymbol{\lambda}}(x)dx$, we can bound $\alpha_2 = \int_0^1 2x f_{\boldsymbol{\lambda}}(x)dx$ below by $\int_0^1 2x 1_{[0,\alpha_1]}dx = \alpha_1^2$ and above by $\int_0^1 2x1_{[1-\alpha_1,1]}dx = 2\alpha_1 - \alpha_1^2$, so $(\alpha_1, \alpha_2) \in \Lambda$, as claimed.

Next we claim that the principal minors of the Jacobian of h are positive; this gives injectivity of h by the Gale-Nikaido theorem [14], and also continuity of h^{-1} by the Inverse Function Theorem. The Jacobian of h is

$$\begin{pmatrix} I(1) & 2I(x) \\ I(x) & 2I(x^2) \end{pmatrix}, \text{ where } I(g) = \int_0^1 g(x)\Big(\frac{f_{\boldsymbol{\lambda}}(x)}{1+e^{\lambda_1+\lambda_2 x}}\Big)dx.$$

All entries are positive as $f_{\boldsymbol{\lambda}}(x)$ is positive. The determinant $2(I(1)I(x^2) - I(x)^2)$ is positive by the Cauchy-Schwarz inequality. Thus the claim holds.

It remains to prove statement (ii) of the lemma. Fix $(\alpha_1, \alpha_2) \in \Lambda$. We claim that $|[n]_{k_n,s_n}| \ge (1+\gamma)^n$ for $n^{-1} \ll \gamma \ll \alpha_1, \alpha_2$. To see this, we fix $\gamma \ll \zeta \ll \theta \ll \alpha_1, \alpha_2$ and construct a θ-bounded measure $\mu_{\mathbf{p}}$ on $\{0,1\}^S$ for some $S \subset [n]$ such that $\sum_{i\in S} p_i(1,i) = (k_n, s_n) \pm \zeta(n, n^2)$; the claim then follows by Lemma 3.5. We let $S = [a - \theta k_n, a + k_n + \theta k_n]$, for some $a \in [n]$ such that $\sum_{i=a+1}^{a+k_n} i = s_n \pm n$; as $\alpha_1^2 < \alpha_2 < 2\alpha_1 - \alpha_1^2$ we have $S \subset [n]$ for small θ. Note that $\alpha_2 n^2 = 2s_n + O(n) = k_n(a + k_n/2) + O(n)$. We let $p_i = (1+2\theta)^{-1}$ for $i \in S$. Then $\sum_{i\in S} p_i = k_n + O(1)$ and $\sum_{i\in S} p_i i = s_n + O(n)$, as required to prove the claim.

Now by Lemma 1.21, $\mathbf{p}_{\mathbf{z}_n}^{\mathcal{V}_n}$ is κ-bounded, where $n^{-1} \ll \kappa \ll \gamma$, so Lemma 10.1 gives $\mathbf{p}_{\mathbf{z}_n}^{\mathcal{V}_n} = \mathbf{p}_{\boldsymbol{\lambda}^{(n)}}^{(n)}$. for some $\boldsymbol{\lambda}^{(n)}$. By κ-boundedness, $\kappa/2 \le (\mathbf{p}_{\mathbf{z}_n}^{\mathcal{V}_n})_1^i/(\mathbf{p}_{\mathbf{z}_n}^{\mathcal{V}_n})_0^i = e^{\lambda_1^{(n)}+\lambda_2^{(n)}(i/n)} \le 2\kappa^{-1}$ for all $i \in [n]$, so $\boldsymbol{\lambda}^{(n)} \in [-C,C]^2$, where $n^{-1} \ll C^{-1} \ll \kappa$. Then $(\boldsymbol{\lambda}^{(n)})$ has a convergent subsequence by compactness of $[-C,C]^2$. Furthermore, any convergent subsequence of $\boldsymbol{\lambda}^{(n)}$ has a limit $\boldsymbol{\lambda}$ that satisfies $h(\boldsymbol{\lambda}) = \alpha$, so is uniquely determined by injectivity of h. □

Next we show a limit theorem for the maximum entropy measures for (t,w)-intersections which is somewhat analogous that in Lemma 10.2 *ii* for the maximum entropy measures for $[n]_{k,s}$.

Lemma 10.3. *Suppose that* $\mathbf{g} = (\alpha_1, \alpha_2, \beta_1, \beta_2) \in [0,1]^4$ *with* $(\alpha_1, \alpha_2) \in \Lambda$. *Write* $\mu_{\widetilde{\mathbf{q}}^{(n)}} = \mu_{\widetilde{\mathbf{x}}_n}^{\widetilde{\mathcal{V}}_n}$. *Then either*

i. there is $(j,j') \in \{0,1\}^2$ *such that* $\min_{i\in[n]}(\widetilde{q}^{(n)})_{j,j'}^i \to 0$, *or*

ii. for large n *we have* $\widetilde{\mathbf{q}}^{(n)} = \mathbf{q}_{\boldsymbol{\pi}^{(n)}}^{(n)}$, *where* $\boldsymbol{\pi}^{(n)}$ *converges to some* $\boldsymbol{\pi} \in \mathbb{R}^4$.

Furthermore, the following are equivalent to case ii:

i. there is $\kappa > 0$ *such that* $\mu_{\widetilde{\mathbf{q}}^{(n)}}$ *is* κ*-bounded for large* n,

ii. there is $\lambda > 0$ *such that* $dim_{VC}([n]_{k,s} \times_{(t,w)} [n]_{k,s}) > \lambda n$ *for large* n.

Proof. We will suppose that case (i) does not apply and show that case (ii) holds. As (i) does not hold there is $\kappa' > 0$ and a subsequence $(n_m)_m$ such that each $\mu_{\widetilde{\mathbf{q}}^{(n_m)}}$ is κ'-bounded. By Lemma 10.1, we have $\boldsymbol{\pi}^{(n_m)} \in \mathbb{R}^4$ so that $\widetilde{\mathbf{q}}^{(n_m)} = \mathbf{q}_{\boldsymbol{\pi}^{(n_m)}}^{(n_m)}$ and by κ'-boundedness each $\boldsymbol{\pi}^{(n_m)} \in [-C,C]^4$ for some $C = C(\kappa') > 0$. We can pass to a convergent subsequence by compactness and so we can assume $\boldsymbol{\pi}^{(n_m)} \to \boldsymbol{\pi} \in \mathbb{R}^4$, relabelling if necessary.

Next we make some observations on the functions $g_{j,j'}^{\boldsymbol{\pi}}(x)$ defined in (5). We note for each $j,j' \in \{0,1\}$ that $(\boldsymbol{\pi}', x) \mapsto g_{j,j'}^{\boldsymbol{\pi}'}(x)$ defines a function on $[-C,C]^4 \times [0,1]$ that is continuous, and so uniformly continuous. As $\boldsymbol{\pi}^{(n_m)} \to \boldsymbol{\pi}$ we deduce that $g_{j,j'}^{\boldsymbol{\pi}^{(n_m)}}$ converges uniformly to $g_{j,j'}^{\boldsymbol{\pi}}$ on $[0,1]$, and so $(\mathbf{q}^{(n)})_{j,j'}^i = (\mathbf{q}_{\boldsymbol{\pi}}^{(n)})_{j,j'}^i = g_{j,j'}^{\boldsymbol{\pi}}(i/n)$ satisfies

$$\max_{i\in[n_m]} |(\widetilde{q}^{(n_m)})_{j,j'}^i - (q^{(n_m)})_{j,j'}^i| \to 0 \text{ as } m \to \infty. \tag{7}$$

Also, as $\widetilde{\mathbf{q}}_{\boldsymbol{\pi}^{(n_m)}}^{(n_m)}$ is κ'-bounded,

$$\kappa' \le \left(\widetilde{q}_{\boldsymbol{\pi}^{(n_m)}}^{(n_m)}\right)_{j,j'}^{\lfloor x n_m \rfloor} = g_{j,j'}^{\boldsymbol{\pi}^{(n_m)}}(x) + o(1) \to g_{j,j'}^{\boldsymbol{\pi}}(x) \text{ as } m \to \infty,$$

using continuity and $\boldsymbol{\pi}^{(n_m)} \to \boldsymbol{\pi}$, so $g_{j,j'}^{\boldsymbol{\pi}}(x) \ge \kappa'$ for all $x \in [0,1]$. We also claim that

$$\int_0^1 (1,2x) g_{1,1}^{\boldsymbol{\pi}} dx = (\beta_1, \beta_2) \quad \text{and} \quad \int_0^1 (1,2x) g_{1,0}^{\boldsymbol{\pi}} dx = (\alpha_1 - \beta_1, \alpha_2 - \beta_2) = \int_0^1 (1,2x) g_{0,1}^{\boldsymbol{\pi}} dx. \tag{8}$$

To see this, we use $\widetilde{\mathcal{V}}_n(\mu_{\widetilde{\mathbf{q}}^{(n)}}) = \widetilde{\mathbf{x}}_n = (\alpha_1 n, \alpha_2\binom{n}{2}, \alpha_1 n, \alpha_2\binom{n}{2}, \beta_1 n, \beta_2\binom{n}{2}) \pm \mathbf{1}$. For example, for the first coordinate in (8), as $\sum_i (\widetilde{q}^{(n_m)})_{1,1}^i = \beta_1 n \pm 1$, by (7) we have $\beta_1 + o(1) = \frac{1}{n}\sum_i g_{1,1}^{\boldsymbol{\pi}}(i/n) = \int_0^1 g_{1,1}^{\boldsymbol{\pi}}(x)dx + o(1)$. The other coordinates are similar, so (8) holds.

Now we claim that $\mu_{\widetilde{\mathbf{q}}^{(n)}}$ is κ-bounded with $n^{-1} \ll \kappa \ll \kappa'$. To see this, take $n^{-1} \ll \lambda \ll \kappa$ so that $\|\widetilde{\mathcal{V}}_n(\mu_{\mathbf{q}^{(n)}}) - \widetilde{\mathbf{x}}_n\|_{\widetilde{\mathbf{R}}_n} \le \lambda n$; this is possible by the calculations in the proof of (8). By Lemma 9.2 and properties (1) and (2) of the Kalai vectors, $\widetilde{\mathcal{V}}_n$ is (0.05)-robustly $(\widetilde{\mathbf{R}}_n, 21)$-generating and $\gamma/2$-robustly

$(\gamma^3/8, \widetilde{\mathbf{R}}_n)$-generic for any $\gamma > 0$. To prove the claim, by Theorem 4.6 (implication $v \Longrightarrow i$) it suffices to show that $(\widetilde{\mathcal{V}}_n, \widetilde{\mathbf{R}}_n, \widetilde{\mathbf{x}}_n)$ is λ-feasible. To see this, consider any $\mathbf{x}'_n \in \mathbb{Z}^D$ with $\|\mathbf{x}'_n - \widetilde{\mathbf{x}}_n\|_{\widetilde{\mathbf{R}}_n} \le \lambda n$ and $\mathcal{V}'_{n'}$ obtained from $\widetilde{\mathcal{V}}_n$ by deleting at most λn co-ordinates. Then $(J^{n'})^{\mathcal{V}'_{n'}}_{\mathbf{x}'_n} \ne 0$, by Lemma 3.5 applied to the $\widetilde{\mathbf{R}}_n$-bounded 0.04-robustly $(\widetilde{\mathbf{R}}_n, 21)$-generating array $\mathcal{V}'_{n'}$ and the (κ'-bounded) restriction μ' of $\mu_{\mathbf{q}^{(n)}}$, as $\|\mathcal{V}'_{n'}(\mu') - \mathbf{x}'_n\|_{\widetilde{\mathbf{R}}} \le 2\lambda n$. This proves the claim.

Our next claim is that $n^{-1}H(\mu_{\widetilde{\mathbf{q}}^{(n)}})$ converges to $H^*(\boldsymbol{\pi})$. To see this, first note that as $g^{\boldsymbol{\pi}^{(n_m)}}_{j,j'}$ converges uniformly to $g^{\boldsymbol{\pi}}_{j,j'}$ on $[0,1]$, as $m \to \infty$ we have

$$\begin{aligned} n_m^{-1}H(\mu_{\widetilde{\mathbf{q}}^{(n_m)}}) &= \mathbb{E}_{i\in[n_m]} \sum_{j,j'\in\{0,1\}} -g^{\boldsymbol{\pi}^{(n_m)}}_{j,j'}(i/n_m)\log_2 g^{\boldsymbol{\pi}^{(n_m)}}_{j,j'}(i/n_m) \\ &\to H^*(\boldsymbol{\pi}) := \sum_{j,j'\in\{0,1\}} -\int_0^1 g^{\boldsymbol{\pi}}_{j,j'}(x)\log_2 g^{\boldsymbol{\pi}}_{j,j'}(x)dx. \end{aligned}$$

By the same calculation with $\boldsymbol{\pi}$ in place of $\boldsymbol{\pi}^{(n_m)}$ we deduce $\lim_n n^{-1}H(\mu_{\mathbf{q}^{(n)}}) = \lim_m n_m^{-1}H(\mu_{\mathbf{q}^{(n_m)}}) = \lim_m n_m^{-1}H(\mu_{\widetilde{\mathbf{q}}^{(n_m)}}) = H^*(\boldsymbol{\pi})$. On the other hand, Lemma 3.5 gives $H(\mu_{\widetilde{\mathbf{q}}^{(n)}}) \pm o(n) = |(J^n)^{\widetilde{\mathcal{V}}_n}_{\widetilde{x}_n}| \ge H(\mu_{\mathbf{q}^{(n)}}) - o(n)$, so $\liminf_n n^{-1}H\big(\mu_{\widetilde{\mathbf{q}}^{(n)}}\big) \ge H^*(\boldsymbol{\pi})$. This estimate applies to any accumulation point $\boldsymbol{\pi}$ of $\{\boldsymbol{\pi}^{(n)}\}_n$. Considering any subsequence (n_m) with $n_m^{-1}H(\mu_{\widetilde{\mathbf{q}}^{(n_m)}}) \to \limsup_n n^{-1}H\big(\mu_{\widetilde{\mathbf{q}}^{(n)}}\big)$ we obtain $\liminf_n n^{-1}H\big(\mu_{\widetilde{\mathbf{q}}^{(n)}}\big) \ge \limsup_n n^{-1}H\big(\mu_{\widetilde{\mathbf{q}}^{(n)}}\big)$, so the claim holds.

Now we will prove $\boldsymbol{\pi}^{(n)} \to \boldsymbol{\pi}$, using stability of maximum entropy measures. Suppose for contradiction that there are subsequences $\boldsymbol{\pi}^{(n_m)} \to \boldsymbol{\pi}$ and $\boldsymbol{\pi}^{(s_m)} \to \boldsymbol{\pi}'$ with $\boldsymbol{\pi} \ne \boldsymbol{\pi}'$. By the previous claim, $H^*(\boldsymbol{\pi}) = H^*(\boldsymbol{\pi}')$. Take $m^{-1} \ll \lambda_3 \ll \lambda_2 \ll \lambda_1 \ll \|\boldsymbol{\pi} - \boldsymbol{\pi}'\|_1$ so that the non-zero continuous functions $g^{\boldsymbol{\pi}}_{j,j'} - g^{\boldsymbol{\pi}'}_{j,j'}$ satisfy $\|g^{\boldsymbol{\pi}}_{j,j'} - g^{\boldsymbol{\pi}'}_{j,j'}\|_1 \ge \lambda_1$ for all $j, j' \in \{0,1\}$. As m is large, this gives $\|\mathbf{q}^{(s_m)} - \widetilde{\mathbf{q}}^{(s_m)}\|_1 \ge \lambda_1 s_m/2$. Since $n^{-1}\|\widetilde{\mathcal{V}}_n(\mu_{\mathbf{q}^{(n)}}) - \widetilde{\mathbf{x}}_n\|_{\widetilde{\mathbf{R}}_n} \to 0$ as $n \to \infty$, we also have $\widetilde{\mathcal{V}}_{s_m}(\mu_{\mathbf{q}^{(s_m)}}) = \widetilde{\mathbf{x}}_{s_m} \pm \lambda_3 \widetilde{\mathbf{R}}_{s_m}$. As in the proof of Lemma 4.10, we can modify $\mathbf{q}^{(s_m)}$ to obtain $\mathbf{q}^*$ with $\widetilde{\mathcal{V}}_{s_m}(\mu_{\mathbf{q}^*}) = \widetilde{\mathbf{x}}_{s_m}$, $\|\mathbf{q}_{s_m} - \mathbf{q}^*\|_1 < \lambda_2 s_m$, and $(s_m)^{-1}H(\mu_{\mathbf{q}^*}) > H^*(\boldsymbol{\pi}) - \lambda_2/2 = H^*(\boldsymbol{\pi}') - \lambda_2/2 \ge (s_m)^{-1}H(\mu_{\widetilde{\mathbf{q}}^{(s_m)}}) - \lambda_2$. But then $\widetilde{\mathbf{q}}^{(s_m)}, \mathbf{q}^* \in \mathcal{M}^{\widetilde{\mathcal{V}}_{s_m}}_{\widetilde{\mathbf{x}}_{s_m}}$, which gives the required contradiction by Lemma 2.6, since $\mu_{\widetilde{\mathbf{q}}^{(n)}} = \mu^{\widetilde{\mathcal{V}}_n}_{\widetilde{\mathbf{x}}_n}$ for all n, $\|\widetilde{\mathbf{q}}^{(s_m)} - \mathbf{q}^*\|_1 \ge \lambda_1 s_m/4$ and $H(\mathbf{q}^*) \ge H(\widetilde{\mathbf{q}}^{(s_m)}) - \lambda_2 s_m$.

Lastly the first equivalence is immediate from our proof, and the second from Theorem 4.6. □

Our next lemma explains the characterisation of the set Γ that appears in Theorem 1.3: it is the set of $\mathbf{g} = (\alpha_1, \alpha_2, \beta_1, \beta_2)$ with $(\alpha_1, \alpha_2) \in \Lambda$ such that the limit marginal problem has a solution. First we complete the definition of Γ by defining the functions $\beta_1, \beta_2 : \Lambda \to \mathbb{R}$ that appear in the definition of Γ_1, using the definitions of h in (6) and $f_{\boldsymbol{\lambda}}$ in (4). Suppose $\boldsymbol{\alpha} \in \Lambda$ with $\alpha_1 \ne \alpha_2$ and let $\boldsymbol{\lambda} = h^{-1}(\boldsymbol{\alpha})$. We define

$$(\beta_1(\alpha_1,\alpha_2), \beta_2(\alpha_1,\alpha_2)) = \int_0^1 (1, 2x) f_{\boldsymbol{\lambda}}(x)^2 dx. \tag{9}$$

Lemma 10.4. *Suppose* $\mathbf{g} = (\alpha_1, \alpha_2, \beta_1, \beta_2) \in [0,1]^4$ *with* $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in \Lambda$. *Let* $\boldsymbol{\lambda} = h^{-1}(\boldsymbol{\alpha})$. *Then* $\mathbf{g} \in \Gamma$ *if and only if there is* $\boldsymbol{\pi} = (\pi_1, \pi'_1, \pi_2, \pi'_2) \in \mathbb{R}^4$ *with* $h^*(\boldsymbol{\pi}) = (\beta_1, \beta_2)$ *and* $f_{\boldsymbol{\lambda}}(x) = g^{\boldsymbol{\pi}}_{0,1}(x) + g^{\boldsymbol{\pi}}_{1,1}(x)$. *Furthermore, if* $\mathbf{g} \in \Gamma$ *there is a unique such* $\boldsymbol{\pi}$, *which we denote* $\boldsymbol{\pi}^{\boldsymbol{\lambda}}$, *and* $\log_2\big|[n]_{k,s} \times_{(t,w)} [n]_{k,s}\big| = H(\mu_{\mathbf{q}'}) + o(n)$, *where* $\mathbf{q}' = \mathbf{q}^{(n)}_{\boldsymbol{\pi}^{\boldsymbol{\lambda}}}$.

Proof. First suppose that there is $\boldsymbol{\pi}$ with $h^*(\boldsymbol{\pi}) = (\beta_1, \beta_2)$ and $f_{\boldsymbol{\lambda}}(x) = g^{\boldsymbol{\pi}}_{0,1}(x) + g^{\boldsymbol{\pi}}_{1,1}(x)$. Then we also have $1 - f_{\boldsymbol{\lambda}}(x) = g^{\boldsymbol{\pi}}_{0,0}(x) + g^{\boldsymbol{\pi}}_{1,0}(x)$. Setting $y = e^x$ and rearranging, we find the polynomial equality

$$e^{\pi_1} y^{\pi_2} + e^{\pi'_1} y^{\pi'_2} = e^{\lambda_1} y^{\lambda_2} + e^{\lambda_1+\pi_1} y^{\lambda_2+\pi_2} \tag{10}$$

for all $y \in [1, e]$. Thus one of the following two conditions holds:

(a) $\pi_2 = \lambda_2$ and $\pi'_2 = \lambda_2 + \pi_2$, (b) $\lambda_2 = \pi'_2 = 0$.

Suppose that $\lambda_2 \neq 0$ and so case (a) holds, giving $\pi_2 = \lambda_2$ and $\pi'_2 = 2\lambda_2$. Equating coefficients in (10) gives $\lambda_1 = \pi_1$ and $\pi'_1 = 2\lambda_1$, so $\boldsymbol{\pi} = \boldsymbol{\pi}^{\boldsymbol{\lambda}} = (\lambda_1, 2\lambda_1, \lambda_2, 2\lambda_2)$. Then $g^{\boldsymbol{\pi}}_{1,1}(x) = (e^{\pi_1+\pi_2 x})^2(1 + e^{\pi_1+\pi_2 x})^{-2} = f_{\boldsymbol{\lambda}}(x)^2$, so ${\mathbf{q}'}^i_{1,1} = g^{\boldsymbol{\pi}}_{1,1}(i/n) = f_{\boldsymbol{\lambda}}(i/n)^2 = ((\mathbf{p}^{(n)}_{\boldsymbol{\lambda}})^i_1)^2$, i.e. $\mu_{\mathbf{q}'}$ is the product of its two marginals $\mu_{\mathbf{p}^{(n)}_{\boldsymbol{\lambda}}}$. Note that $\alpha_1 \neq \alpha_2$, as h is injective and $h(\lambda_1, 0) = (\alpha_1, \alpha_1)$. Thus $\mathbf{g} \in \Gamma_1$.

Now suppose that $\lambda_2 = \pi_2 = 0$. Then the right hand side of (10) is constant, so $\pi'_2 = 0$, and so $e^{\pi_1} + e^{\pi'_1} = e^{\lambda_1} + e^{\lambda_1+\pi_1}$. Thus $\alpha_1 = \alpha_2 = f_{\boldsymbol{\lambda}}(x) = e^{\lambda_1}(1 + e^{\lambda_1})^{-1}$, and $\beta_1 = \beta_2 = g^{\boldsymbol{\pi}}_{1,1}(x) = e^{\pi'_1}(1 + 2e^{\pi_1} + e^{\pi'_1})^{-1}$. Furthermore, $\beta_1 = g^{\boldsymbol{\pi}}_{1,1}(x) < g^{\boldsymbol{\pi}}_{1,1}(x) + g^{\boldsymbol{\pi}}_{0,1}(x) = \alpha_1$ and $2\alpha_1 - 1 = 2(e^{\pi_1} + e^{\pi'_1})(1 + 2e^{\pi_1} + e^{\pi'_1})^{-1} - 1 < e^{\pi'_1}(1 + 2e^{\pi_1} + e^{\pi'_1})^{-1} = \beta_1$. Thus $\mathbf{g} \in \Gamma_2$.

It remains to consider $\lambda_2 = 0$ and $\pi_2 \neq 0$. Then case (b) must hold, so $\pi'_2 = 0$. Equating coefficients in (10) gives $\lambda_1 = \pi'_1$ and $\lambda_1 + \pi_1 = \pi_1$, so $\lambda_1 = \pi'_1 = 0$. Then $\alpha_1 = \alpha_2 = f_{\boldsymbol{\lambda}}(x) = 1/2$. We also have $g^{\boldsymbol{\pi}}_{1,1}(x) = e^{\pi_1+\pi_2 x}(2+2e^{\pi_1+\pi_2 x})^{-1} = f_{\widetilde{\boldsymbol{\pi}}}(x)/2$, where $\widetilde{\boldsymbol{\pi}} = (\pi_1, \pi_2)$. We deduce $(\beta_1, \beta_2) = h^*(\boldsymbol{\pi}) = \int_0^1 (1, 2x) g^{\boldsymbol{\pi}}_{1,1}(x) dx = \frac{1}{2} \int_0^1 (1, 2x) f_{\widetilde{\boldsymbol{\pi}}}(x) dx = h(\widetilde{\boldsymbol{\pi}})/2$, so $2(\beta_1, \beta_2) \in \Lambda$. Thus $\mathbf{g} \in \Gamma_3$.

We conclude that if $h^*(\boldsymbol{\pi}) = (\beta_1, \beta_2)$ and $f_{\boldsymbol{\lambda}}(x) = g^{\boldsymbol{\pi}}_{0,1}(x) + g^{\boldsymbol{\pi}}_{1,1}(x)$ then $\mathbf{g} \in \Gamma$. Conversely, if $\mathbf{g} \in \Gamma$ then the analysis of each case above exhibits the unique $\boldsymbol{\pi} = \boldsymbol{\pi}^{\boldsymbol{\lambda}}$ satisfying these conditions. Indeed, if $\mathbf{g} \in \Gamma_1$ we have $\boldsymbol{\pi} = (\lambda_1, 2\lambda_1, \lambda_2, 2\lambda_2)$, if $\mathbf{g} \in \Gamma_2$ we have $\boldsymbol{\pi} = (\pi_1, \pi'_1, 0, 0)$ where $g^{\boldsymbol{\pi}}_{1,1} = \beta_1$ and $g^{\boldsymbol{\pi}}_{0,1}(x) = \alpha_1 - \beta_1$ (when $\alpha_1 \neq \beta_1$ this gives two linear equations for e^{π_1} and $e^{\pi'_1}$ that have a unique solution), and if $\mathbf{g} \in \Gamma_3$ we have $\boldsymbol{\pi} = (\pi_1, 0, \pi_2, 0)$, where $(\pi_1, \pi_2) = h^{-1}(2\beta_1, 2\beta_2)$

Finally, let $\mathbf{q}' = \mathbf{q}^{(n)}_{\boldsymbol{\pi}^{\boldsymbol{\lambda}}}$, and note that $\mu_{\mathbf{q}'} \in \mathcal{M}^{\widetilde{\mathcal{V}}}_{\widetilde{\mathbf{x}}}$, so $H(\mu_{\mathbf{q}'}) \leq H(\mu^{\widetilde{\mathcal{V}}}_{\widetilde{\mathbf{x}}}) \leq \log_2 \left|[n]_{k,s} \times_{(t,w)} [n]_{k,s}\right| + o(n)$ by Lemma 3.5. For the inequality in the other direction we consider each Γ_i separately.

If $\mathbf{g} \in \Gamma_1$ we let $\mathbf{p}' = \mathbf{p}^{(n)}_{\boldsymbol{\lambda}}$, note that $H(\mu_{\mathbf{q}'}) = 2H(\mu_{\mathbf{p}'})$ and $\log_2 |[n]_{k,s}| = H(\mu_{\mathbf{p}'}) + o(n)$ by Lemma 3.5. Then $\log_2 \left|[n]_{k,s} \times_{(t,w)} [n]_{k,s}\right| \leq 2 \log_2 |[n]_{k,s}| = H(\mu_{\mathbf{q}'}) + o(n)$.

If $\mathbf{g} \in \Gamma_2$ we have $\left|[n]_{k,s} \times_{(t,w)} [n]_{k,s}\right| \leq \left|\binom{[n]}{k} \times_t \binom{[n]}{k}\right| = \binom{n}{t, k-t, k-t, n-2k+t} = 2^{H(\mu_{\mathbf{q}'})+o(n)}$.

If $\mathbf{g} \in \Gamma_3$ we note that if $(A, B) \in [n]_{k,s} \times_{(t,w)} [n]_{k,s}$, where $k = \left\lfloor \frac{1}{2} n \right\rfloor$, $s = \left\lfloor \frac{1}{2} \binom{n}{2} \right\rfloor$, $t = \lfloor \beta_1 n \rfloor$ and $w = \left\lfloor \beta_2 \binom{n}{2} \right\rfloor$, then $C := (A \cap B) \cup (\overline{A} \cap \overline{B}) \in [n]_{k',s'}$, where $k' = 2\beta_1 n + O(1)$ and $s' = 2\beta_2 \binom{n}{2} + O(n)$. By Lemma 3.5, $\log_2 |[n]_{k',s'}| = H(\mu_{\mathbf{p}''}) + o(n)$, where $\mathbf{p}'' = \mathbf{p}^{(n)}_{\widetilde{\boldsymbol{\lambda}}}$ with $\widetilde{\boldsymbol{\lambda}} = h^{-1}(2\beta_1, 2\beta_2)$. By the form of $\boldsymbol{\pi}$ calculated above we have $(q')^i_{1,1} = (q')^i_{0,0} = \frac{1}{2}(p'')^i_1$ and $(q')^i_{1,0} = (q')^i_{0,1} = \frac{1}{2}(p'')^i_0$, so

$$H(\mu_{\mathbf{q}'}) = 2 \sum_{j=0,1} -\tfrac{1}{2}(p'')^i_j \log_2 \tfrac{1}{2}(p'')^i_j = \sum_{j=0,1} ((p'')^i_j - (p'')^i_j \log_2 (p'')^i_j) = n + H(\mu_{\mathbf{p}''}).$$

Given C, there are at most 2^n choices for (A, B), so $\log_2 \left|[n]_{k,s} \times_{(t,w)} [n]_{k,s}\right| \leq n + \log_2 |[n]_{k',s'}| = H(\mu_{\mathbf{q}'}) + o(n)$. Thus in all cases we have the required bound. □

We conclude this subsection with the solution to Kalai's conjecture.

Proof of Theorem 1.3. We first fix parameters $n^{-1} \ll \lambda \ll \delta \ll \delta_1 \ll \delta_2 \ll \delta_3 \ll \varepsilon \ll \theta \ll C^{-1} \ll \varepsilon', \kappa \ll \mathbf{g}$. Suppose $\mathbf{g} = (\alpha_1, \alpha_2, \beta_1, \beta_2) \in [0, 1]^4$ with $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in \Lambda$ and let $\boldsymbol{\lambda} = h^{-1}(\boldsymbol{\alpha})$. By

Lemma 10.2 *ii*, the maximum entropy measure $\mu_{\mathbf{p}}$ for $[n]_{k,s}$ is given by $\mathbf{p} = \mathbf{p}^{(n)}_{\boldsymbol{\lambda}^{(n)}}$, where $\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^{(n)}\|_1 \leq \delta$. In particular, $\mathbf{p}$ is κ-bounded, since $\kappa \ll (\alpha_1, \alpha_2)$.

First we consider case *i*, i.e. we suppose $d(\mathbf{g}, \Gamma) \leq \delta$ and prove that $\mathbf{g}$ is (n, δ, ε)-Kalai. Consider $\mathcal{A} \subset [n]_{k,s}$ with $|\mathcal{A}| \geq (1-\delta)^n |[n]_{k,s}|$. Then $\mu_{\mathbf{p}}(\mathcal{A}) \geq (1-\delta_1)^n$ by Theorem 1.20, so setting $\mathbf{p}' = \mathbf{p}^{(n)}_{\boldsymbol{\lambda}}$ we have $\mu_{\mathbf{p}'}(\mathcal{A}) \geq (1-2\delta_1)^n$ as $\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^{(n)}\|_1 \leq \delta$. Next we fix $\mathbf{g}' \in \Gamma$ with $\|\mathbf{g}' - \mathbf{g}\|_1 < d(\mathbf{g}, \Gamma) + \delta < 2\delta$, apply Lemma 10.4 to obtain $\boldsymbol{\pi}^{\boldsymbol{\lambda}} \in \mathbb{R}^4$, and let $\mathbf{q}' = \mathbf{q}^{(n)}_{\boldsymbol{\pi}^{\boldsymbol{\lambda}}}$. Note that $\mu_{\mathbf{q}'}$ is κ-bounded as $\kappa \ll \mathbf{g}$, that $\mu_{\mathbf{q}'}$ has marginals $\mu_{\mathbf{p}'}$ and $\mathbb{E}_{(A,B)\sim\mu_{\mathbf{q}'}}(|A \cap B|, \sum(A \cap B)) = (t, w) \pm 4\delta(n, n^2)$, as $\|\mathbf{g}' - \mathbf{g}\|_1 < 2\delta$. As $\mu_{\mathbf{p}'}(\mathcal{A}) \geq (1-2\delta_1)^n$, Theorem 1.17 gives $\mu_{\mathbf{q}'}(\mathcal{A} \times_{(t,w)} \mathcal{A}) \geq (1-\delta_2)^n$. Lemma 3.1 then gives $|\mathcal{A} \times_{(t,w)} \mathcal{A}| \geq (1-\delta_3)^n 2^{H(\mathbf{q}')} \geq (1-\varepsilon)^n \big|[n]_{k,s} \times_{(t,w)} [n]_{k,s}\big|$ by the final part of Lemma 10.4. Thus $\mathbf{g}$ is (n, δ, ε)-Kalai, completing the proof of *i*.

We now consider case *ii*, i.e. we suppose $d(\mathbf{g}, \Gamma) \geq \varepsilon'$ and $|[n]_{k,s} \times_{(t,w)} [n]_{k,s}| \geq (1-\varepsilon)^{-n}$, and prove that $\mathbf{g}$ is not (n, δ, ε)-Kalai. We adopt the running notation $\mu_{\widetilde{\mathbf{q}}} = \mu_{\widetilde{\mathbf{q}}^{(n)}} = \mu^{\widetilde{\mathcal{V}}_n}_{\widetilde{\mathbf{x}}_n}$ introduced at the beginning of the section. By the equivalences in Lemma 10.3, for large n we either have (a) $\widetilde{\mathbf{q}}^{(n)}$ is κ-bounded, or (b) $\dim_{VC}([n]_{k,s} \times_{(t,w)} [n]_{k,s}) \leq \lambda n$. We can assume that (b) does not hold, as here Lemma 9.5 gives $\mathcal{A} \subset [n]_{k,s}$ with $|\mathcal{A}| \geq (1-\delta)^n |[n]_{k,s}|$ and $\mathcal{A} \times_{(t,w)} \mathcal{A} = \emptyset$, and so $|\mathcal{A} \times_{(t,w)} \mathcal{A}| < 1 \leq \lfloor (1-\varepsilon)^n [n]_{k,s} \times_{(t,w)} [n]_{k,s} \rfloor$. Thus $\mathbf{g}$ is not (n, δ, ε)-Kalai if (b) holds, so we can assume that (a) holds, i.e. $\mu_{\widetilde{\mathbf{q}}}$ is κ-bounded for large n. By Lemma 10.1 there is $\boldsymbol{\pi} \in \mathbb{R}^4$ such that $\widetilde{\mathbf{q}} = \mathbf{q}^{(n)}_{\boldsymbol{\pi}}$, where as $\mu_{\widetilde{\mathbf{q}}}$ is κ-bounded we have $\boldsymbol{\pi} \in [-C, C]^4$.

We will now prove that $\mu_{\mathbf{p}}$ is not close to the marginal distribution of $\mu_{\widetilde{\mathbf{q}}}$ and conclude by applying Lemma 9.5. Consider the continuous function $\phi : [-C, C]^4 \to \mathbb{R}$, where

$$\phi(\boldsymbol{\sigma}) := \|(\beta_1, \beta_2) - h^*(\boldsymbol{\sigma})\|_1 + \int_0^1 |f_{\boldsymbol{\lambda}}(x) - g^{\boldsymbol{\sigma}}_{1,0}(x) - g^{\boldsymbol{\sigma}}_{1,1}(x)| dx.$$

As $\|\mathbf{g} - \mathbf{g}'\|_1 \geq \varepsilon'$ for all $\mathbf{g}' \in \Gamma$, by Lemma 10.4, the continuity of ϕ and the compactness of $[-C, C]^4$ we have $\phi(\boldsymbol{\sigma}) \geq \theta > 0$ for all $\boldsymbol{\sigma} \in [-C, C]^4$; in particular, this holds for $\boldsymbol{\sigma} = \boldsymbol{\pi}$. As $\mathbb{E}_{(A,B)\sim\mu_{\widetilde{\mathbf{q}}}}\big(|A \cap B|, \sum(A \cap B)\big) = (\beta_1, \beta_2\binom{n}{2}) \pm (1, 1)$, we have $\|(\beta_1, \beta_2) - h^*(\boldsymbol{\pi})\|_1 \ll \theta/2$ for large n, so $\int_0^1 |f_{\boldsymbol{\lambda}}(x) - g^{\boldsymbol{\pi}}_{1,0}(x) - g^{\boldsymbol{\pi}}_{1,1}(x)| dx \geq \theta/2$.

We now translate from the limit setting back to the finite setting. For large n, the previous inequality implies

$$\theta n/4 \leq \sum_{i=1}^n |f_{\boldsymbol{\lambda}}(i/n) - g^{\boldsymbol{\pi}}_{1,0}(i/n) - g^{\boldsymbol{\pi}}_{1,1}(i/n)| = \sum_{i=1}^n |(p')^i_1 - \widetilde{p}^i_1|,$$

where $\mathbf{p}' = \mathbf{p}^{(n)}_{\boldsymbol{\lambda}}$ and $\widetilde{\mathbf{q}} = \mathbf{q}^{(n)}_{\boldsymbol{\pi}}$ has marginals $\mu_{\widetilde{\mathbf{p}}}$. As $\mathbf{p} = \mathbf{p}^{(n)}_{\boldsymbol{\lambda}^{(n)}}$ and $\|\boldsymbol{\lambda} - \boldsymbol{\lambda}^{(n)}\|_1 \leq \delta$, we deduce $\|\mathbf{p} - \widetilde{\mathbf{p}}\|_1 \geq \theta n/8$. Finally, we apply Lemma 9.3 *ii* (with $\delta_1 = \theta/2$ and $\delta = \varepsilon$) to find $\mathcal{B}_{full} \subset [n]_{k,s}$ so that $\mathcal{A} = [n]_{k,s} \setminus \mathcal{B}_{full}$ satisfies $|\mathcal{A}| \geq (1 - o(1))|[n]_{k,s}| > (1-\delta)^n |[n]_{k,s}|$ and $|\mathcal{A} \times_{(t,w)} \mathcal{A}| < (1-\varepsilon)^n |[n]_{k,s} \times_{(t,w)} [n]_{k,s}|$. Thus $\mathbf{g}$ is not (n, δ, ε)-Kalai, which completes the proof. □

10.2 Uniqueness in higher dimensions

In this subsection we illustrate how the method used to prove Theorem 1.3 can be applied in a broader context. Throughout this subsection we work with the following setting.

- Fix $\boldsymbol{\alpha} = (\alpha_d)_{d\in[D]}$ and $\boldsymbol{\beta} = (\beta_d)_{d\in[D]}$ in $(0,1)^D$.
  For all $n \in \mathbb{N}$ let $\mathbf{z}_n = (\lfloor \alpha_d n^2 \rfloor)_{d\in[D]}$ and $\mathbf{w}_n = (\lfloor \alpha_d n^2 \rfloor)_{d\in[D]}$.

- Suppose $\mathcal{V}_n = (\mathbf{v}_i^{(n)})_{i\in[n]}$ are $(n,\{0,1\})$-arrays in $[n]^D$ such that $(\mathcal{V}_n, \mathbf{z}_n)$ is robustly generating and robustly generic.
- Write $\mathcal{X}_n = (\{0,1\}^n)^{\mathcal{V}_n}_{\mathbf{z}_n})$ and suppose that $|\mathcal{X}_n| > (1+\eta)^n$, where $\eta = \eta(\boldsymbol{\alpha}) > 0$ is fixed.
- The arrays $\mathcal{V}_n$ have a 'scaling limit': there is a positive measurable function $p : [0,1]^D \to \mathbb{R}$ with $\int_{[0,1]^D} p(\mathbf{x}) d\mathbf{x} = 1$ such that for any measurable set $B \subset [0,1]^D$ we have

$$\lim_{n\to\infty} n^{-1}\big|\{i \in [n] : n^{-1}\mathbf{v}_i^{(n)} \in B\}\big| = \int_B p(\mathbf{x}) d\mathbf{x}.$$

The assumption that $(\mathcal{V}_n, \mathbf{z}_n)$ is robustly generic is in fact redundant, as it can be shown to follow from the scaling limit assumption, but for the sake of brevity we omit this deduction.

We say that $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is (n, δ, ε)-*good* if the corresponding $\mathcal{V}_n$-intersection problem exhibits 'full supersaturation' analogous to that in Conjecture 1.2, i.e. any $\mathcal{A} \subset \mathcal{X}_n$ with $|\mathcal{A}| \ge (1-\delta)^n|\mathcal{X}_n|$ satisfies $\big|(\mathcal{A}\times\mathcal{A})^{(\mathcal{V}_n)\cap}_{\mathbf{w}_n}\big| \ge (1-\varepsilon)^n\big|(\mathcal{X}_n\times\mathcal{X}_n)^{(\mathcal{V}_n)\cap}_{\mathbf{w}_n}\big|$. We will outline the proof of the following analogue of Theorem 1.3, which shows that if we exclude the case of 'uniformly random sets' (i.e. $\boldsymbol{\alpha} \ne \big(1/2\big)_{d\in[D]}$) then 'full supersaturation' only occurs for one specific value of $\boldsymbol{\beta}$.

Theorem 10.5. *In the above setting, if $\boldsymbol{\alpha} \ne \big(1/2\big)_{d\in[D]}$ then there is $\boldsymbol{\beta}^* = \boldsymbol{\beta}^*(\boldsymbol{\alpha}) \in (0,1)^D$ such that for $n^{-1} \ll \delta \ll \varepsilon \ll \varepsilon' \ll \boldsymbol{\alpha}$,*
i. if $\|\boldsymbol{\beta} - \boldsymbol{\beta}^*\|_1 \le \delta$ then $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is (n,δ,ε)-good, and
ii. if $\|\boldsymbol{\beta} - \boldsymbol{\beta}^*\|_1 \ge \varepsilon'$ and $\big|(\mathcal{X}_n\times\mathcal{X}_n)^{(\mathcal{V}_n)\cap}_{\mathbf{w}_n}\big| \ge (1-\varepsilon)^{-n}$ then $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is not (n,δ,ε)-good.

Similarly to the previous subsection, we wish to determine when $\mu_{\widetilde{\mathbf{q}}} = \mu^{\widetilde{\mathcal{V}}}_{\widetilde{\mathbf{x}}}$ (with $\widetilde{\mathcal{V}}$ and $\widetilde{\mathbf{x}}$ as in Definition 9.1) has marginals close to $\mu_{\mathbf{p}} = \mu^{\mathcal{V}}_{\mathbf{z}}$ (here we are omitting the subscript n from our notation). If these measures are κ-bounded, Lemma 10.1 gives $\boldsymbol{\lambda} \in \mathbb{R}^D$ such that

$$p_1^i = (p^{(n)}_{\boldsymbol{\lambda}})_1^i := f_{\boldsymbol{\lambda}}(\mathbf{v}_i^{(n)}/n), \text{ where } f_{\boldsymbol{\lambda}}(\mathbf{x}) = e^{\boldsymbol{\lambda}\cdot\mathbf{x}}(1+e^{\boldsymbol{\lambda}\cdot\mathbf{x}})^{-1},$$

and $\boldsymbol{\pi}_1, \boldsymbol{\pi}_2 \in \mathbb{R}^D$ such that $\widetilde{q}^i_{j,j'} = (\mathbf{q}^{(n)}_{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2})^i_{j,j'} := g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{j,j'}(\mathbf{v}_i^{(n)}/n)$, where

$$g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{0,0}(\mathbf{x}) = Z_{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}(\mathbf{x})^{-1}, \qquad g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{0,1}(\mathbf{x}) = g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{1,0}(\mathbf{x}) = e^{\boldsymbol{\pi}_1\cdot\mathbf{x}} Z_{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}(\mathbf{x})^{-1},$$
$$g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{1,1}(\mathbf{x}) = e^{\boldsymbol{\pi}_2\cdot\mathbf{x}} Z_{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}(\mathbf{x})^{-1}, \qquad \text{with } Z_{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}(\mathbf{x}) = 1 + 2e^{\boldsymbol{\pi}_1\cdot\mathbf{x}} + e^{\boldsymbol{\pi}_2\cdot\mathbf{x}}.$$

Again we study the marginal problem for $\mu_{\widetilde{\mathbf{q}}}$ and $\mu_{\mathbf{p}}$ via the limit marginal problem of characterising $\boldsymbol{\lambda}$ and $\boldsymbol{\pi}$ such that $f_{\boldsymbol{\lambda}}(\mathbf{x}) = g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{0,1}(\mathbf{x}) + g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{1,1}(\mathbf{x})$. The constraints are $\mathbf{z}_n = \sum_{i\in[n]} p_1^i \mathbf{v}_\mathbf{i}^{(\mathbf{n})}$ and $\mathbf{w}_n = \sum_{i\in[n]} \widetilde{q}^i_{1,1}\mathbf{v}_\mathbf{i}^{(\mathbf{n})}$. The limit versions are $h(\boldsymbol{\lambda}) = \boldsymbol{\alpha}$ and $h^*(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2) = \boldsymbol{\beta}$, where

$$h(\boldsymbol{\lambda}) = \int_{[0,1]^D} \mathbf{x} f_{\boldsymbol{\lambda}}(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \qquad \text{and} \qquad h^*(\boldsymbol{\pi}_1,\boldsymbol{\pi}_2) = \int_{[0,1]^D} \mathbf{x} g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{1,1}(\mathbf{x}) p(\mathbf{x}) d\mathbf{x}.$$

Our next lemma is analogous to Lemma 10.2.

Lemma 10.6.
i. h is a homeomorphism between $\mathbb{R}^D$ and $h(\mathbb{R}^D)$.
ii. For large n we have $\mu^{\mathcal{V}_n}_{\mathbf{z}_n} = \mu_{\mathbf{p}}$, for some $\mathbf{p} = \mathbf{p}^{(n)}_{\boldsymbol{\lambda}^{(n)}}$ where $\boldsymbol{\lambda}^{(n)} \to \boldsymbol{\lambda} = h^{-1}(\boldsymbol{\alpha})$.

We omit the proof of Lemma 10.6, as it is the same as that of Lemma 10.2, except in one detail which we will now check, namely that the principal minors of the Jacobian of h are positive. To see this, note that the Jacobian J has entries

$$J_{i,j} = \int_{\mathbf{x}\in[0,1]^D} x_i x_j \Big(\frac{f_{\boldsymbol{\lambda}}(\mathbf{x})}{1+e^{\boldsymbol{\lambda}.\mathbf{x}}}\Big) p(\mathbf{x}) d\mathbf{x}.$$

For any $\mathbf{y} \in \mathbb{R}^D$ we have $\mathbf{y}^T J \mathbf{y} = \int_{\mathbf{x}\in[0,1]^D} |\langle \mathbf{x}, \mathbf{y}\rangle|^2 f_{\boldsymbol{\lambda}}(\mathbf{x})(1+e^{\boldsymbol{\lambda}.\mathbf{x}})^{-1} p(\mathbf{x}) d\mathbf{x}$. As $f_{\boldsymbol{\lambda}}$ and p are positive, we have $\mathbf{y}^T J \mathbf{y} > 0$ whenever $\mathbf{y} \neq 0$, as required. We also have the following analogue of Lemma 10.3; again, we omit the similar proof.

Lemma 10.7. *Write* $\mu_{\widetilde{\mathbf{q}}^{(n)}} = \mu^{\widetilde{\mathcal{V}}_n}_{\widetilde{\mathbf{x}}_n}$. *Then either*

i. there is $(j, j') \in \{0,1\}^2$ *such that* $\min_{i\in[n]} (\widetilde{q}^{(n)})^i_{j,j'} \to 0$, *or*

ii. for large n *we have* $\widetilde{\mathbf{q}}^{(n)} = \mathbf{q}^{(n)}_{\boldsymbol{\pi}_1^{(n)}, \boldsymbol{\pi}_2^{(n)}}$, *where* $(\boldsymbol{\pi}_1^{(n)}, \boldsymbol{\pi}_2^{(n)})$ *converges to some* $(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2) \in \mathbb{R}^{2D}$.

Furthermore, the following are equivalent to case ii:

i. there is $\kappa > 0$ *such that* $\mu_{\widetilde{\mathbf{q}}^{(n)}}$ *is* κ*-bounded for large* n,

ii. there is $\lambda > 0$ *such that* $dim_{VC}((\mathcal{X}_n \times \mathcal{X}_n)^{(\mathcal{V}_n)\cap}_{\mathbf{w}_n}) > \lambda n$ *for large* n.

The uniqueness in Theorem 10.5 is explained by the following lemma which solves the limit marginal problem.

Lemma 10.8. *Suppose* $h(\boldsymbol{\lambda}) = \boldsymbol{\alpha} \neq (1/2)_{d\in[D]}$. *Then there is unique* $\boldsymbol{\beta}^* \in [0,1]^D$ *such that there is* $(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2) \in \mathbb{R}^D \times \mathbb{R}^D$ *with* $h^*(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2) = \boldsymbol{\beta}^*$ *and* $f_{\boldsymbol{\lambda}}(\mathbf{x}) = g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{0,1}(\mathbf{x}) + g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{1,1}(\mathbf{x})$. *Furthermore,* $(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2)$ *is unique, and* $\log_2 \left|(\mathcal{X}_n \times \mathcal{X}_n)^{(\mathcal{V}_n)\cap}_{\mathbf{w}_n}\right| = H(\mu_{\mathbf{q}'}) + o(n)$, *where* $\mathbf{q}' = \mathbf{q}^{(n)}_{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}$.

Proof. As h is injective and $h(\mathbf{0}) = (1/2)_{d\in[D]} \neq \boldsymbol{\alpha}$, we have $\boldsymbol{\lambda} \neq \mathbf{0}$. Rearranging $1 - f_{\boldsymbol{\lambda}}(\mathbf{x}) = g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{0,0}(\mathbf{x}) + g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{0,1}(\mathbf{x})$ gives $e^{\boldsymbol{\pi}_1\cdot\mathbf{x}} + e^{\boldsymbol{\pi}_2\cdot\mathbf{x}} = e^{\boldsymbol{\lambda}\cdot\mathbf{x}} + e^{(\boldsymbol{\pi}_1+\boldsymbol{\lambda})\cdot\mathbf{x}}$, so (a) $\boldsymbol{\pi}_1 = \boldsymbol{\lambda}$ and $\boldsymbol{\pi}_2 = \boldsymbol{\pi}_1 + \boldsymbol{\lambda}$, or (b) $\boldsymbol{\pi}_1 = \boldsymbol{\pi}_1 + \boldsymbol{\lambda}$ and $\boldsymbol{\lambda} = \boldsymbol{\pi}_2$. However, (b) cannot hold, as $\boldsymbol{\lambda} \neq \mathbf{0}$. Thus $\boldsymbol{\pi}_1 = \boldsymbol{\lambda}$ and $\boldsymbol{\pi}_2 = 2\boldsymbol{\pi}_1$, so $g^{\boldsymbol{\pi}_1,\boldsymbol{\pi}_2}_{1,1}(\mathbf{x}) = f_{\boldsymbol{\lambda}}(\mathbf{x})^2$ and $\boldsymbol{\beta}^* = \int_{\mathbf{x}\in[0,1]^D} \mathbf{x} f_{\boldsymbol{\lambda}}(\mathbf{x})^2 p(\mathbf{x}) d\mathbf{x}$. Uniqueness of $(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2)$ is clear, and the final estimate follows in the same way as the case $\mathbf{g} \in \Gamma_1$ of Lemma 10.4. □

Given the above lemmas, the proof of Theorem 10.5 is very similar to that of Theorem 1.3, so we omit the details.

11 Optimal supersaturation

In this section we characterise the optimal level of supersaturation for $\mathcal{V}$-intersections in terms of a certain optimisation problem; as outlined in subsection 1.4, this corresponds to the optimal choice of measure satisfying the hypotheses of Theorem 1.17, i.e. determining H_{max} in the following setting, which we adopt throughout this section.

- Let $0 < n^{-1}, \delta \ll \kappa \ll \gamma_1, \gamma_1' \ll \gamma_2, \gamma_2' \ll \varepsilon \ll \alpha, D^{-1}, C^{-1}, k^{-1}$.
- Suppose $\mathcal{V} = (\mathbf{v}_i : i \in [n])$ and each $\mathbf{v}_i \in \mathbb{Z}^D$ is $\mathbf{R}$-bounded, where $\mathbf{R} \in \mathbb{R}^D$ with $\max_d R_d < n^C$.
- Suppose $\mathcal{V}$ is γ_i'-robustly $(\gamma_i, \mathbf{R})$-generic and γ_i-robustly $(\mathbf{R}, k)$-generating for $i = 1, 2$.
- Let $\mathbf{z} \in \mathbb{Z}^D$ and $\mathcal{X} = (\{0,1\}^n)^{\mathcal{V}}_{\mathbf{z}}$ with $|\mathcal{X}| \geq (1+\alpha)^n$. Write $\mu_{\mathbf{p}} := \mu^{\mathcal{V}}_{\mathbf{z}}$.

- Let $\mathbf{w} \in \mathbb{Z}^D$ and let $\mathbf{Q}$ denote the set of $\mathbf{q}$ such that $\mu_\mathbf{q}$ is a κ-bounded product measure on $(\{0,1\} \times \{0,1\})^n$ with both marginals $\mu_\mathbf{p}$ and $\mathcal{V}_\cap(\mu_\mathbf{q}) = \mathbf{w}$.

- Let $H_{max} = \max_{\mathbf{q} \in \mathbf{Q}} H(\mu_\mathbf{q})$ if $\mathbf{Q} \neq \emptyset$ or $H_{max} = 0$ if $\mathbf{Q} = \emptyset$.

The main result of this section is as follows.

Theorem 11.1. *In the above setting,*
i. if $\mathcal{A} \subset \mathcal{X}$ with $|\mathcal{A}| > (1-\delta)^n|\mathcal{X}|$ then $|(\mathcal{A} \times \mathcal{A})^{\mathcal{V}_\cap}_\mathbf{w}| \geq \lfloor (1-\varepsilon)^n 2^{H_{max}} \rfloor$.
ii. there is $\mathcal{A} \subset \mathcal{X}$ with $|\mathcal{A}| > (1-\varepsilon)^n|\mathcal{X}|$ and $|(\mathcal{A} \times \mathcal{A})^{\mathcal{V}_\cap}_\mathbf{w}| \leq (1+\varepsilon)^n 2^{H_{max}}$.

We start by giving the short deduction of statement i from Theorem 1.17. The hypotheses of the latter hold by Lemma 1.21, and Theorem 1.20 gives $\mu_\mathbf{p}(\mathcal{A}) > (1-\delta')^n$, where $\delta \ll \delta' \ll \kappa$. We can assume $\mathbf{Q} \neq \emptyset$, and Theorem 1.17 applied to $\mathbf{q} \in \mathbf{Q}$ gives $\mu_\mathbf{q}((\mathcal{A} \times \mathcal{A})^{\mathcal{V}_\cap}_\mathbf{w}) > (1-\varepsilon/2)^n$. Also, by Lemma 3.1, $\mathcal{B} := \{\mathbf{x} \in (\{0,1\} \times \{0,1\})^n : \log_2 \mu_\mathbf{q}(\mathbf{x}) \notin -H_{max} \pm \delta n\}$ has $\mu_\mathbf{q}(\mathcal{B}) \leq (1-\delta^3)^n$, so $|(\mathcal{A} \times \mathcal{A})^{\mathcal{V}_\cap}_\mathbf{w}| \geq 2^{H_{max} - \delta n} \mu_\mathbf{q}((\mathcal{A} \times \mathcal{A})^{\mathcal{V}_\cap}_\mathbf{w} \setminus \mathcal{B}) \geq (1-\varepsilon)^n 2^{H_{max}}$, as required.

The remainder of the section will be occupied with the proof of statement ii. A key idea is the use of 'empirical measures', which we will now introduce. First we note by Lemma 1.21 that $\mathbf{p}$ is κ'-bounded, where

$$\gamma_1, \gamma_1' \ll \kappa' \ll \gamma_2, \gamma_2'.$$

We fix a partition of $[n]$ into sets $S_1, \dots, S_M$ so that

$$|p_i - p_j| \leq \kappa \text{ and } \|\mathbf{v}_i - \mathbf{v}_j\|_\mathbf{R} \leq \kappa \text{ if } i, j \in S_m \text{ for some } m \in [M].$$

This can be achieved with $M \leq (2\kappa^{-1} + 2)^{D+1}$. We define the *type* of $A \subset [n]$ as $\mathbf{k}(A) = (|A \cap S_1|, \dots, |A \cap S_M|)$. We let $\mathbf{k} = (k_1, \dots, k_M)$ be the most common type of sets in $\mathcal{X}$, write

$$\mathcal{B} = \{A \in \mathcal{X} : \mathbf{k}(A) = \mathbf{k}\},$$

and define the empirical measure $\mu_{\mathbf{p}'}$ by

$$(p')^i_1 = k_m/|S_m| \text{ for all } m \in [M], i \in S_m.$$

Note that $|\mathcal{B}| \geq |\mathcal{X}|/n^M$. The following lemma shows that $\mu_{\mathbf{p}'}$ is a good approximation to the maximum entropy measure $\mu_\mathbf{p}$.

Lemma 11.2. $\|\mathbf{p} - \mathbf{p}'\|_1 \leq \kappa n$.

Proof. Let $\mathcal{E}$ be the set of $A \subset [n]$ such that some $||A \cap S_m| - \sum_{i \in S_m} p_i| > \frac{\kappa n}{2M}$. Let $\kappa \ll \kappa_1 \ll \kappa_2 \ll \kappa'$. Then $\mu_\mathbf{p}(\mathcal{E}) \leq (1-\kappa_1)^n$ by Chernoff's inequality, so $|\mathcal{E} \cap \mathcal{X}| < (1-\kappa_2)^n |\mathcal{X}| < |\mathcal{B}|$ by Theorem 1.20. We deduce $|k_m - \sum_{i \in S_m} p_i| \leq \kappa n/2M$ for all $m \in [M]$, so $\sum_{m \in [M]} |k_m - \sum_{i \in S_m} p_i| \leq \kappa n/2$; the lemma follows. □

We use a similar construction of an empirical measure that represents $\mathbf{w}$-intersections. Let G be the graph with $V(G) = \mathcal{B}$ where $AB \in E(G)$ if $|A \cap B|_\mathcal{V} = \mathbf{w}$. We define the *type* of $AB \in E(G)$ as $\mathbf{t}_{AB} = (t_1, \dots, t_M)$, where $t_m = |A \cap B \cap S_m|$ for $m \in [M]$. A type $\mathbf{t}$ gives rise to a measure $\mu_{\mathbf{q}(\mathbf{t})}$, where for $i \in S_m$ we define

$$q(\mathbf{t})^i_{1,1} = t_m/|S_m|, \; q(\mathbf{t})^i_{1,0} = q(\mathbf{t})^i_{0,1} = (k_m - t_m)/|S_m| \text{ and } q(\mathbf{t})^i_{0,0} = (|S_m| - 2k_m + t_m)/|S_m|.$$

Note that each $\mu_{\mathbf{q}(\mathbf{t})}$ has both marginals $\mu_{\mathbf{p}'}$ and $\|\mathcal{V}_\cap(\mu_{\mathbf{q}(\mathbf{t})}) - \mathbf{w}\|_\mathbf{R} \leq \|\mathbf{p} - \mathbf{p}'\|_1 \leq \kappa n$.

We can assume

$$H(\mu_{\mathbf{q}}) < \log_2 |E(G)| - \varepsilon n/2 \text{ for all } \mathbf{q} \in \mathbf{Q}, \tag{11}$$

otherwise the proof is complete. We fix a type $\widetilde{\mathbf{t}}$ occurring at least $e(G)/n^{2M}$ times and set $\widetilde{\mathbf{q}} = \mathbf{q}(\widetilde{\mathbf{t}})$. Then $H(\mu_{\widetilde{\mathbf{q}}}) \geq \log_2(e(G)/n^{2M})$, so $\widetilde{\mathbf{q}} \notin \mathbf{Q}$ by (11). The following lemma will show that all empirical measures associated to edges of G are close to $\widetilde{\mathbf{q}}$; we will then use this and $\widetilde{\mathbf{q}} \notin \mathbf{Q}$ in Lemma 11.5 to find a large independent set in G, which will complete the proof of Theorem 11.1.

We fix λ with $\gamma_1, \gamma_1' \ll \lambda \ll \kappa'$.

Lemma 11.3. *Suppose $CD \in E(G)$ has type $\mathbf{t}'$. Then $\|\widetilde{\mathbf{q}} - \mathbf{q}(\mathbf{t}')\|_1 < \lambda n$.*

For the proof we require the following bound analogous to (11) for a wider class of measures.

Lemma 11.4. *Let $\mu_{\mathbf{q}'}$ be a λ-dense product measure on $(\{0,1\} \times \{0,1\})^n$ with marginals $\mu_{\mathbf{p}'}$ and $\mathcal{V}_\cap(\mu_{\mathbf{q}'}) = \mathbf{w}'$ with $\|\mathbf{w} - \mathbf{w}'\|_{\mathbf{R}} \leq \kappa n$. Then $H(\mu_{\mathbf{q}'}) < \log_2 |E(G)| - \varepsilon n/3$.*

Proof. We will obtain the required bound from (11) a measure in $\mathbf{Q}$ close to $\mu_{\mathbf{q}'}$. Recall that $\mathbf{p}'$ is κ'-bounded and $\|\mathbf{p} - \mathbf{p}'\|_1 \leq \kappa n$ Consider $\mathbf{q}''$ that minimises $\|\mathbf{q}'' - \mathbf{q}'\|_1$ subject to $\mu_{\mathbf{q}}$ being κ-bounded and having marginals $\mu_{\mathbf{p}}$. For each i we can construct $\{(q'')^i_{j,j'}\}$ from $\{(q')^i_{j,j'}\}$ by moving probability mass $|p_i - p_i'|$ to create the correct marginals, and moving a further mass of at most 2κ while maintaining the same marginals to ensure κ-boundedness. Therefore $\|\mathbf{q}'' - \mathbf{q}'\|_1 \leq 6\kappa n$.

Now we will perturb $\mathbf{q}''$ to obtain $\mathbf{q} \in \mathbf{Q}$, i.e. we maintain κ-boundedness and the same marginals $\mu_{\mathbf{p}}$, and obtain $\mathcal{V}_\cap(\mu_{\mathbf{q}}) = \mathbf{w}$.

As $\mu_{\mathbf{q}'}$ is λ-dense and $\kappa \ll \lambda$ there is $S \subset [n]$ with $|S| \geq \lambda n/2$ such that $(q'')^i_{j,j'} \geq \lambda/2$ for all $i \in S$ and $j, j' \in \{0,1\}$. As $\mathcal{V}$ is γ_1'-robustly $(\gamma_1, \mathbf{R})$-generic, and $\lambda \gg \gamma_1'$ we can find $M \geq |S|/2D \geq \lambda n/4D$ disjoint sets $I_1, \dots, I_M \subset S$, with $|I_m| = D$ and $|\det(\mathcal{V}_{I_m})| \geq \gamma_1 R_1 \cdots R_D$ for all $m \in [M]$.

Write $\mathcal{V}_\cap(\mu_{\mathbf{q}''}) = \mathbf{w}''$, and note that $\|\mathbf{w}'' - \mathbf{w}\|_{\mathbf{R}} \leq \|\mathbf{q}'' - \mathbf{q}'\|_1 \leq 6\kappa n$. Then $\mathbf{u} = (\mathbf{w} - \mathbf{w}'')/M$ has $\|\mathbf{u}\|_{\mathbf{R}} \leq 24D\kappa\lambda^{-1} < \sqrt{\kappa}$. Applying Cramer's rule as in Lemma 4.8, for each $m \in [M]$ we find coefficients b_i with $\sum_{i \in I_m} b_i \mathbf{v}_i = \mathbf{u}$ and $|b_i| \leq \sqrt{\kappa} D! \gamma_1^{-1}$.

Now we obtain $\mathbf{q}$ from $\mathbf{q}''$ where for each $i \in \cup_{m \in [M]} I_m$ we let $q^i_{1,1} = (q'')^i_{1,1} + b_i$, $q^i_{0,1} = q^i_{1,0} = (q'')^i_{1,0} - b_i$ and $q^i_{1,1} = (q'')^i_{1,1} + b_i$, and $q^i_{j,j'} = (q'')^i_{j,j'}$ otherwise. By construction $\mathbf{q} \in \mathbf{Q}$ and $\|\mathbf{q}' - \mathbf{q}\|_1 \leq \|\mathbf{q}' - \mathbf{q}''\|_1 + \|\mathbf{q}'' - \mathbf{q}\|_1 < \kappa^{1/3} n$. The lemma now follows from (11). □

Proof of Lemma 11.3. Suppose for a contradiction that $\|\widetilde{\mathbf{q}} - \mathbf{q}(\mathbf{t}')\|_1 \geq \lambda n$. Consider the interpolation $\mathbf{q}' = \lambda \mathbf{q}(\mathbf{t}') + (1-\lambda)\widetilde{\mathbf{q}}$. Recall that any $\mu_{\mathbf{q}(\mathbf{t})}$ has marginals $\mu_{\mathbf{p}'}$ and satisfies $\|\mathcal{V}_\cap(\mu_{\mathbf{q}(\mathbf{t})}) - \mathbf{w}\|_{\mathbf{R}} \leq \kappa n$, so $\mu_{\mathbf{q}'}$ has the same properties. Also, as $H(\mu_{\widetilde{\mathbf{q}}}) \geq \log_2(e(G)/n^{2M})$ we have $H(\mu_{\mathbf{q}'}) > \log_2 |E(G)| - \varepsilon n/3$.

As $\|\widetilde{\mathbf{q}} - \mathbf{q}(\mathbf{t}')\|_1 \geq \lambda n$ we can find $S \subset [n]$ with $|S| \geq \lambda n/2$ such that $\sum_{j,j' \in \{0,1\}} |\widetilde{q}^i_{j,j'} - q(\mathbf{t}')^i_{j,j'}| \geq \lambda/2$ for all $i \in S$. As $\widetilde{\mathbf{q}}$ and $\mathbf{q}(\mathbf{t}')$ have the same marginals $\mu_{\mathbf{p}'}$ we have $|\widetilde{q}^i_{j,j'} - q(\mathbf{t}')^i_{j,j'}| \geq \lambda/8$ for all $i \in S$ and $j, j' \in \{0,1\}$. For each such i, j, j' we deduce $(q')^i_{j,j'} \geq \lambda^2/8$. However, this contradicts Lemma 11.4 (with $\lambda^2/8$ in place of λ). □

The following lemma completes the proof of Theorem 11.1.

Lemma 11.5. *There is $\mathcal{A} \subset \mathcal{X}$ with $|\mathcal{A}| \geq (1-\varepsilon)^n |\mathcal{X}|$ and $(\mathcal{A} \times \mathcal{A})^{\mathcal{V}_\cap}_{\mathbf{w}} = \emptyset$.*

Proof. We can assume $e(G) \geq (1+\varepsilon/2)^n |\mathcal{B}|$, as otherwise by Turán's theorem ([30], see also [5, IV.2]) G contains an independent set $\mathcal{A}$ of order $(1+\varepsilon/2)^{-n}|\mathcal{B}|/2 \geq (1-\varepsilon)^n |\mathcal{X}|$. As $\log_2 |\mathcal{X}| \geq H(\mu_{\mathbf{p}}) - \kappa n$ by Lemma 3.5 and $\|\mathbf{p} - \mathbf{p}'\|_1 \leq \kappa n$ by Lemma 11.2 we deduce $H(\mu_{\widetilde{\mathbf{q}}}) \geq \log_2(e(G)/n^{2M}) \geq H(\mu_{\mathbf{p}'}) + \varepsilon n/4$.

We have $H(\mu_{\widetilde{\mathbf{q}}}) = \sum_{i\in[n]} H(\widetilde{\mathbf{q}}^i)$ and $H(\mu_{\mathbf{p}'}) = \sum_{i\in[n]} H(\mathbf{p}'^i)$, where each $H(\widetilde{\mathbf{q}}^i) \le \log_2 4 = 2$, so there is $T \subset [n]$ with $|T| \ge \varepsilon n/16$ such that $H(\widetilde{\mathbf{q}}^i) \ge H(\mathbf{p}'^i) + \varepsilon/16$. As $\mu_{\widetilde{\mathbf{q}}}$ has marginals $\mu_{\mathbf{p}'}$ we deduce $\widetilde{\mathbf{q}}^i_{0,1} = \widetilde{\mathbf{q}}^i_{1,0} > \varepsilon^2$ for all $i \in T$. Let $T_1 = \{i \in T : \widetilde{q}^i_{1,1} < \lambda\}$ and $T_0 = \{i \in T : \widetilde{q}^i_{0,0} < \lambda\}$. By Lemma 11.4 we have $|T_1| \ge |T|/4$ or $|T_0| \ge |T|/4$.

Case 1: $|T_1| \ge |T|/4$.

Let $\mathcal{B}^* = \{B \in \mathcal{B} : |B \cap T_1| \ge \kappa'|T_1|/2\}$. As $\mathbf{p}$ is κ'-bounded and $|T_1| \ge \varepsilon n/64$ we have $\mu_{\mathbf{p}}(\mathcal{B} \setminus \mathcal{B}^*) \le (1 - c_{\kappa'})^n$, which by Theorem 1.20 gives $|\mathcal{B}^*| \ge |\mathcal{B}|/2$. Let $G^* = G[B^*]$ denote the induced subgraph of G with vertex set $\mathcal{B}^*$.

We claim that for all $AB \in E(G^*)$ we have $|A \cap B \cap T_1| < 4\lambda^{1/2}n$. Indeed, suppose for a contradiction that $|A \cap B \cap T_1| \ge 4\lambda^{1/2}n$. Let J be the set of $m \in [M]$ with $|T_1 \cap S_m| \ge 2\lambda^{1/2}|S_m|$. Then $\sum_{m\in J} |S_m| \ge 2\lambda^{1/2}n$. For all $i \in \bigcup_{m\in J} S_m$ we have $q(\mathbf{t}_{AB})^i_{1,1} - \widetilde{q}^i_{1,1} \ge 2\lambda^{1/2} - \lambda > \lambda^{1/2}$, by definition of J and T_1. But then $\|\widetilde{\mathbf{q}} - \mathbf{q}(\mathbf{t}_{AB})\|_1 \ge \lambda^{1/2} \sum_{m\in J} |S_m| > \lambda n$. This contradicts Lemma 11.3, so the claim holds.

Therefore, for any $U \subset T_1$ of size $u = \lceil 4\lambda^{1/2}n \rceil$, the family $\mathcal{A}_U := \{B \in \mathcal{B}^* : U \subset B\}$ forms an independent set in G^*. Consider a uniformly random choice of such U. For any $B \in \mathcal{B}^*$, as $|B \cap T_1| \ge \kappa'|T_1|/2$ we have $\mathbb{P}(B \in \mathcal{A}_U) \ge (\kappa'/4)^u \ge (1 - \lambda^{1/3})^n$, as $\lambda \ll \kappa'$. Therefore $\mathbb{E}_U|\mathcal{A}_U| = \sum_{B\in\mathcal{B}^*} \mathbb{P}(B \in \mathcal{A}_U) \ge (1-\varepsilon)^n|\mathcal{X}|$. Thus for some U we obtain an independent set $\mathcal{A}_U$ of at least this size, which completes the proof of Case 1.

Case 2: $|T_0| \ge |T|/4$.

The proof of this case is similar to that of Case 1, so we just outline the differences. Now we let $G^* = G[B^*]$, where $\mathcal{B}^* = \{B \in \mathcal{B} : |T_0 \setminus B| \ge \kappa'|T_0|/2\}$. Similarly to Case 1, we have $|\mathcal{B}^*| \ge |\mathcal{B}|/2$, and there is no edge $AB \in E(G^*)$ with $|T_0 \setminus (A \cup B)| \ge 4\lambda^{1/2}n$. Thus for any $U \subset T_0$ with $|U| = u$, the family $\mathcal{A}_U := \{B \in \mathcal{B}^* : U \cap B = \emptyset\}$ is an independent set in G^*. Consider a uniformly random choice of such U. For any $B \in \mathcal{B}^*$, as $|T_0 \setminus B| \ge \kappa'|T_0|/2$ we have $\mathbb{P}(B \in \mathcal{A}_U) \ge (\kappa'/4)^u \ge (1-\lambda^{1/3})^n$, as $\lambda \ll \kappa'$. Therefore for some U we obtain an independent set $\mathcal{A}_U$ with size at least the expectation, which is at least $(1-\varepsilon)^n|\mathcal{X}|$. □

12 Exponential continuity

In this section we recast our results using the following notion of continuity that arises naturally when comparing distributions according to exponential contiguity.

Definition 12.1. Let $\Omega = (\Omega_n)_{n\in\mathbb{N}}$ and $\mu = (\mu_n)_{n\in\mathbb{N}}$, where each μ_n is a probability measure on Ω_n. Let $\mathcal{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ where each $\mathcal{F}_n$ is a set of measurable subsets of Ω_n. We say that $\mathcal{B} = (\Omega, \mathcal{F})$ is an exponential probability space and write $\mathcal{M}(\mathcal{B}) = \mathcal{M}(\Omega)$ for the set of such μ. We write $\nu \approx \mu$ when $\nu \approx_{\mathcal{F}} \mu$. Given exponential probability spaces $\mathcal{B} = (\Omega, \mathcal{F})$, $\mathcal{B}' = (\Omega', \mathcal{F}')$ we say that $f : \mathcal{M}(\Omega) \to \mathcal{M}(\Omega')$ is exponentially continuous at $\mu \in \mathcal{M}(\Omega)$ if $\mu' \approx \mu \Rightarrow f(\mu') \approx f(\mu)$.

Theorem 12.2. *Let $0 < n^{-1} \ll \zeta \ll \kappa, \gamma \ll D^{-1}, M^{-1}, C^{-1}, k^{-1}$. Suppose*

i. $\mathcal{B}_s = (\Omega_s, \mathcal{F}_s)$ are exponential probability spaces with $\Omega_{s,n} = J_s^n$ for $s \in S$,

ii. $\mu_{\mathbf{q}}$ is κ-bounded product measure on Ω_n,

iii. $\mathcal{V} = (\mathbf{v}^i_{j_1,\dots,j_S})$ is an $(n, \prod_{s\in S} J_s)$-array in $\mathbb{Z}^D$,

iv. all $\|\mathbf{v}^i_{j_1,\dots,j_S}\|_{\mathbf{R}} \le 1$, where $\mathbf{R} = (R_1, \dots, R_D)$ with $\max_d R_d < n^C$,

v. $\mathcal{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_M\} \subset \mathbb{Z}^D$ is $\mathbf{R}$-bounded and $(k, k\zeta n, \mathbf{R})$-generating,

vi. $\mathcal{V}$ has γ-robust transfers for $\mathcal{U}$,

vii. $\mathbf{w} \in \mathbb{Z}^D$ *with* $\|\mathbf{w} - \mathcal{V}(\mu_\mathbf{q})\|_\mathbf{R} < \zeta n$.
Let $\mathcal{B} = (\Omega, \mathcal{F}) = \prod_{s\in S} \mathcal{B}_s$ *and* $\mathcal{B}' = (\Omega', \mathcal{F}')$, *where* $\Omega'_n = (\Omega_n)^{\mathcal{V}}_{\mathbf{w}}$ *and* $\mathcal{F}'_n = \{\mathcal{A} \cap \Omega'_n : \mathcal{A} \in \mathcal{F}\}$. *Let* f *be restriction of measure from* $\mathcal{M}(\Omega)$ *to* $\mathcal{M}(\Omega')$. *Then* f *is exponentially continuous at* $\mu_\mathbf{q}$.

Proof. Let $\mu_\mathbf{q}$ have marginals $(\mu_{\mathbf{p}_s} : s \in S)$ and suppose $\mu_{\mathbf{q}'} \approx \mu_\mathbf{q}$ with marginals $(\mu_{\mathbf{p}'_s} : s \in S)$. Suppose $n^{-1} \ll \delta \ll \zeta \ll \varepsilon \ll \kappa, \gamma$. We want to show for $\mathcal{A} = \prod_{s\in S} \mathcal{A}_s \in \mathcal{F}'_n$ that $f(\mu_\mathbf{q})(\mathcal{A}) > (1-\delta)^n \Rightarrow f(\mu_{\mathbf{q}'})(\mathcal{A}) > (1-\varepsilon)^n$ and $f(\mu_{\mathbf{q}'})(\mathcal{A}) > (1-\delta)^n \Rightarrow f(\mu_\mathbf{q})(\mathcal{A}) > (1-\varepsilon)^n$. As $f(\mu_\mathbf{q})(\mathcal{A}) = \mu_\mathbf{q}(\mathcal{A})/\mu_\mathbf{q}(\Omega'_n)$ and $f(\mu_{\mathbf{q}'})(\mathcal{A}) = \mu_{\mathbf{q}'}(\mathcal{A})/\mu_{\mathbf{q}'}(\Omega'_n)$, it suffices to show that $\mu_\mathbf{q}(\Omega'_n), \mu_{\mathbf{q}'}(\Omega'_n) > (1-\varepsilon')^n$ with $\varepsilon' \ll \varepsilon$. This holds for $\mu_\mathbf{q}(\Omega'_n)$ by Theorem 6.3, and so for $\mu_{\mathbf{q}'}$ by exponential contiguity. □

Remark 12.3. In the setting of the above theorem, if $\mu_\mathbf{q}$ has marginals $(\mu_{\mathbf{p}_s} : s \in S)$ $\mu_{\mathbf{q}'}$ has marginals $(\mu_{\mathbf{p}'_s} : s \in S)$, and each $\mathcal{F}_{s,n}$ is the set of subsets of some $\Delta_{s,n} \subset \Omega_{s,n}$, then we have $\mu_\mathbf{q} \approx \mu_{\mathbf{q}'}$ precisely when each $\mu_{\mathbf{p}_s} \approx \mu_{\mathbf{p}'_s}$: this holds by Theorem 7.2 and the following lemma.

Lemma 12.4. *Suppose* $\mu = (\mu_n)_{n\in\mathbb{N}}$ *and* $\nu = (\nu_n)_{n\in\mathbb{N}}$ *where each* μ_n *and* ν_n *is a probability measure on* Ω_n. *Suppose also* $\mu' = (\mu'_n)_{n\in\mathbb{N}}$ *and* $\nu' = (\nu'_n)_{n\in\mathbb{N}}$ *where each* μ'_n *and* ν'_n *is a probability measure on* Ω'_n. *Let* $\Delta = (\Delta_n)_{n\in\mathbb{N}}$ *with each* $\Delta_n \subset \Omega_n$ *and* $\Delta' = (\Delta'_n)_{n\in\mathbb{N}}$ *with each* $\Delta'_n \subset \Omega'_n$. *Then* $\mu \times \mu' \approx_{\Delta\times\Delta'} \nu \times \nu'$ *if and only if* $\mu \approx_\Delta \nu$ *and* $\mu' \approx_{\Delta'} \nu'$.

Proof. Let $n^{-1} \ll \delta \ll \varepsilon$. Suppose first that $\mu \times \mu' \approx_{\Delta\times\Delta'} \nu \times \nu'$. Consider $A^1_n \subset \Delta_n$ with $\mu_n(A^1_n) > (1-\delta)^n$. Let $A_n = A^1_n \times \Delta'_n$. Then $(\mu_n \times \mu'_n)(A_n) = \mu_n(A^1_n) > (1-\delta)^n$, so $\nu_n(A^1_n) = (\nu_n \times \nu'_n)(A_n) > (1-\varepsilon)^n$ by assumption, i.e. $\mu \lesssim_\Delta \nu$. Similarly $\nu \lesssim_\Delta \mu$, so $\mu \approx_\Delta \nu$, and similarly $\mu' \approx_{\Delta'} \nu'$. Now suppose $\mu \approx_\Delta \nu$ and $\mu' \approx_{\Delta'} \nu'$. Let $B_n = \{(\mathbf{x},\mathbf{y}) \in \Delta_n \times \Delta'_n : (\nu_n \times \nu'_n)(\mathbf{x},\mathbf{y}) < (1-\varepsilon)^n (\mu_n \times \mu'_n)(\mathbf{x},\mathbf{y})\}$. We have $B_n \subset (B^1_n \times \Delta'_n) \cup (\Delta_n \times B^2_n)$, where $B^1_n = \{\mathbf{x} \in \Delta_n : \nu_n(\mathbf{x}) < (1-\varepsilon)^{n/2}\mu_n(\mathbf{x})\}$ and $B^2_n = \{\mathbf{y} \in \Delta'_n : \nu'_n(\mathbf{y}) < (1-\varepsilon)^{n/2}\mu'_n(\mathbf{y})\}$. By assumption, $\mu_n(B^1_n) \le (1-2\delta)^n$ and $\mu'_n(B^2_n) \le (1-2\delta)^n$. Therefore $(\mu_n \times \mu'_n)(B_n) \le 2(1-2\delta)^n < (1-\delta)^n$, i.e. $\mu \times \mu' \lesssim_{\Delta\times\Delta'} \nu \times \nu'$. Similarly, $\nu \times \nu' \lesssim_{\Delta\times\Delta'} \mu \times \mu'$, so $\mu \times \mu' \approx_{\Delta\times\Delta'} \nu \times \nu'$. □

13 Concluding remarks

There are several natural directions in which to explore potential generalisations of our results: instead of associating vectors in $\mathbb{Z}^D$ to each coordinate we may consider values in another (abelian) group G, and we may consider more general functions of the coordinate values, e.g. a (low degree) polynomial (e.g. a quadratic for application to the Borsuk conjecture) rather than a linear function (is there a 'local' version of Kim-Vu [24] polynomial concentration?). Even for linear functions in one dimension, our setting seems somewhat related to some open problems in Additive Combinatorics, such as the independence number of Paley graphs, but here our assumptions seem too restrictive (one cannot use transfers). We may also ask when better bounds hold, e.g. for $G = \mathbb{Z}/6\mathbb{Z}$ we recall an open problem of Grolmusz [15]: is there a subexponential bound for set systems where the size of each set is divisible by 6 but each pairwise intersection is not divisible by 6?

Our results may interpreted as giving robust statistics in the theory of social choice. Suppose that we represent a voter by an opinion vector $\mathbf{x} \in J^n$, where each x_i represents an opinion on the ith issue, for example, when $|J| = 2$ each issue could be a question with a yes/no answer. Then we can represent a population of voters by a probability measure μ on J^n, where $\mu(\mathbf{x})$ is the proportion of a voters with opinion $\mathbf{x}$. Now suppose that we want to compare two (or more) voters. One natural measure of comparison is to assign a score to each opinion and calculate the total score on opinions where they agree. If this is too simplistic, then we could assign score vectors in some $\mathbb{R}^D$, where D is small enough to give a genuine compression of the data, but large enough to capture the varied nature of the issues: we compare $\mathbf{x}$ and $\mathbf{x}'$ according to $\mathcal{V}_\cap(\mathbf{x}, \mathbf{x}')$. Taking the perspective of robust statistics (see [16]), it is natural to ask whether this statistic is sensitive to our uncertainty in the probability measure that represents the population as a whole: Theorem 12.2 (with the remark following it) gives one possible answer.

Acknowledgements: The authors would like to thank the referees for a careful reading of the paper and many helpful comments.

References

[1] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley, 2008.

[2] L. Babai and P. Frankl, *Linear algebra methods in combinatorics*, Department of Computer Science, University of Chicago, preliminary version, September 1992.

[3] A. Barvinok and J. Hartigan, Maximum entropy Gaussian approximations for the number of integer points and volumes of polytopes, *Adv. Appl. Math.* **45** (2010), 252-289.

[4] B. Bollobás, *Combinatorics*, Cambridge University Press, 1986.

[5] B. Bollobás, *Modern Graph Theory*, Springer, 1998.

[6] H. Buhrman, R. Cleve and A. Wigderson, Quantum vs. classical communication and computation, *Proceedings of 30th STOC* (1998), 63–68.

[7] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications and Signal Processing, 2006.

[8] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Springer, 2009.

[9] P. Erdős, C. Ko and R. Rado, Intersection theorems for systems of finite sets, *Quart. J. Math. Oxford Ser. (2)* **12** (1961), 313–320.

[10] J. Fox and B. Sudakov, Dependent random choice, *Random Structures Algorithms* **38** (2011), 68–99.

[11] P. Frankl and V. Rödl, Forbidden intersections, *Trans. Amer. Math. Soc.* **300** (1987), 259–286.

[12] P. Frankl and V. Rödl, A partition property of simplices in Euclidean space, *J. Amer. Math. Soc.* **3** (1990), 1–7.

[13] P. Frankl and R.M. Wilson, Intersection theorems with geometric consequences, *Combinatorica* **1** (1981), 357–368.

[14] D. Gale and H. Nikaido, The Jacobian matrix and global univalence of mappings, *Math. Ann.* **159** (1965), 81-93.

[15] V. Grolmusz, Superpolynomial size set-systems with restricted intersections mod 6 and explicit Ramsey graphs, *Combinatorica* **20** (2000), 71–85.

[16] P.J. Huber and E.M. Ronchetti, *Robust statistics*, Wiley, 2009.

[17] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley-Interscience, 2000.

[18] E.T. Jaynes, Information Theory and Statistical Mechanics, *Physical Review Series II* **106** (1957), 620–630.

[19] S. Jukna, *Extremal combinatorics*, Springer, 2011.

[20] J. Kahn and G. Kalai, A counterexample to Borsuk's conjecture, *Bull. Amer. Math. Soc.* **29** (1993), 60–62.

[21] G. Kalai, Some old and new problems in combinatorial geometry I: Around Borsuk's problem, in *Surveys in Combinatorics* 2005, 147–174, London Math. Soc. Lecture Note Ser., 424, Cambridge Univ. Press, 2015.

[22] G.O.H. Katona, Intersection theorems for systems of finite sets, *Acta Math. Acad. Sci. Hung.* **15** (1964), 329–337.

[23] P. Keevash and E. Long, Frankl–Rödl type theorems for codes and permutations, *Trans. Amer. Math. Soc.* **369** (2017), 1147–1162.

[24] J.H. Kim and V. Vu, Concentration of multivariate polynomials and applications, *Combinatorica* **20** (2000), 417–434.

[25] C. McDiarmid, Concentration, in: Probabilistic Methods for Algorithmic Discrete Mathematics, *Alg. Combin.* 16:195–248 (1998).

[26] D.H.J. Polymath, A new proof of the density Hales-Jewett theorem, *Ann. of Math.* **175** (2012), 1283–1327.

[27] N. Sauer, On the density of families of sets, *J. Combin. Theory Ser. A* **13** (1972), 145–147.

[28] J. Sgall, Bounds on pairs of families with restricted intersections, *Combinatorica* **19**(4), (1999), 555-566.

[29] S. Shelah, A combinatorial problem; stability and order for models and theories in infinitary languages, *Pacific J. Math.* **41** (1972), 247–261.

[30] P. Turán, On an extremal problem in graph theory (in Hungarian), *Mat. Fiz. Lapok* **48** (1941), 436–452.

[31] V.N. Vapnik and A.Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theory Probab. Appl.* **16** (1971), 264–280.