Global hypercontractivity and its applications

Keevash, Peter; Lifshitz, Noam; Long, Eoin; Minzer, Dor

Global hypercontractivity and its applications

Peter Keevash* Noam Lifshitz† Eoin Long‡ Dor Minzer§

Abstract

The classical hypercontractive inequality for the noise operator on the discrete cube plays a crucial role in many of the fundamental results in the Analysis of Boolean functions, such as the KKL (Kahn-Kalai-Linial) theorem, Friedgut's junta theorem and the invariance principle of Mossel, O'Donnell and Oleszkiewicz. In these results the cube is equipped with the uniform (1/2-biased) measure, but it is desirable, particularly for applications to the theory of sharp thresholds, to also obtain such results for general p-biased measures. However, simple examples show that when p is small there is no hypercontractive inequality that is strong enough for such applications.

In this paper, we establish an effective hypercontractive inequality for general p that applies to 'global functions', i.e. functions that are not significantly affected by a restriction of a small set of coordinates. This class of functions appears naturally, e.g. in Bourgain's sharp threshold theorem, which states that such functions exhibit a sharp threshold. We demonstrate the power of our tool by strengthening Bourgain's theorem, thereby making progress on a conjecture of Kahn and Kalai and by establishing a p-biased analog of the seminal invariance principle of Mossel, O'Donnell, and Oleszkiewicz.

Our sharp threshold results also have significant applications in Extremal Combinatorics. Here we obtain new results on the Turán number of any bounded degree uniform hypergraph obtained as the expansion of a hypergraph of bounded uniformity. These are asymptotically sharp over an essentially optimal regime for both the uniformity and the number of edges and solve a number of open problems in the area. In particular, we give general conditions under which the crosscut parameter asymptotically determines the Turán number, answering a question of Mubayi and Verstraëte. We also apply the Junta Method to refine our asymptotic results and obtain several exact results, including proofs of the Huang–Loh–Sudakov conjecture on cross matchings and the Füredi–Jiang–Seiver conjecture on path expansions.

*Mathematical Institute, University of Oxford, Oxford, UK. Email: keevash@maths.ox.ac.uk.
Research supported in part by ERC Consolidator Grant 647678.
†Einstein Institute for Mathematics, Hebrew University of Jerusalem. Email: noamlifshitz@gmail.com.
Research supported in part by ERC advanced grant 834735.
‡University of Birmingham, Birmingham, UK. Email: e.long@bham.ac.uk.
§Department of Mathematics, Massachusetts Institute of Technology. Email: minzer.dor@gmail.com.
Some of the work was done while the author was a postdoc in the Institute for Advanced Study, Princeton, supported NSF grant CCF-1412958 and Rothschild Fellowship.

Introduction

The field of Analysis of Boolean functions is centered around the study of functions on the discrete cube $\{0,1\}^n$, via their Fourier–Walsh expansion, often using the classical hypercontractive inequality for the noise operator, obtained independently by Bonami [11], Gross [44] and Beckner [4]. In particular, the fundamental 'KKL' theorem of Kahn, Kalai and Linial [50] applies hypercontractivity to obtain structural information on Boolean valued functions with small 'total influence' / 'edge boundary' (see Section I.1.2); such functions cannot be 'global': they must have a co-ordinate with large influence.

The theory of sharp thresholds is closely connected (see Section I.2) to the structure of Boolean functions of small total influence, not only in the KKL setting of uniform measure on the cube, but also in the general p-biased setting. However, we will see below that the hypercontractivity theorem is ineffective for small p. This led Friedgut [36], Bourgain [36, appendix], and Hatami [45] to develop new ideas for proving p-biased analogs of the KKL theorem. The theme of these works can be roughly summarised by the statement: an effective analog of the KKL theorem holds for a certain class of 'global' functions. However, these theorems were incomplete in two important respects:

- *Sharpness:* Unlike the KKL theorem, they are not sharp up to constant factors.
- *Applicability:* They are only effective in the 'dense setting' when $\mu_p(f)$ is bounded away from 0 and 1, whereas the 'sparse setting' $\mu_p(f) = o(1)$ is needed for many important open problems.

A sparse analogue of the KKL theorem was a key missing ingredient in a strategy suggested by Kahn and Kalai [49] for their well-known conjecture relating critical probabilities to expectation thresholds.

Main contribution

The most fundamental new result of this paper is a hypercontractive theorem for functions that are 'global' (in a sense made precise below). This has many applications, of which the most significant are as follows.

- We strengthen Bourgain's Theorem by obtaining an analogue of the KKL theorem that is both quantitively tight and applicable in the sparse regime.
- We obtain a sharp threshold result for global monotone functions in the spirit of the Kahn-Kalai conjecture, bounding the ratio between the critical probability (where $\mu_p(f) = \frac{1}{2}$) and the smallest p for which $\mu_p(f)$ is non-negligible.
- We obtain a p-biased generalisation of the seminal invariance principle of Mossel, O'Donnell and Oleszkiewicz [75] (itself a generalisation of the Berry-Esseen theorem from linear functions to polynomials of bounded degree), thus opening the door to p-biased versions of its many striking applications in Hardness of Approximation and Social Choice Theory (see O'Donnell [77, Section 11.5]) and Extremal Combinatorics (see Dinur–Friedgut–Regev [16]).
- We obtain strong new estimates on a wide class of hypergraph Turán numbers, which are central and challenging parameters in Extremal Combinatorics. Our results apply to bounded degree uniform hypergraphs which are obtained as an expansion of a hypergraph of bounded uniformity, and allow us to solve a number of open problems in the area.

Structure of the paper

To facilitate navigation between various topics we have divided this paper into five parts. The first part is an extended synopsis in which we motivate and state our main results. The second part introduces global hypercontractivity, our fundamental new contribution that underpins the applications in the subsequent three parts of the paper. After presenting the basic form of our hypercontractivity inequality needed for the later applications, we continue to develop the general theory, as this has independent interest, and also has further applications via our generalised invariance principle. The third part contains our analytic applications to results on sharp thresholds and isoperimetric stability. We then move to combinatorial applications in the final two parts, where for the benefit of any reader whose primary interest lies in these applications, we would highlight that these draw upon the earlier parts in a 'black-box' manner, which are therefore not pre-requisite reading for the final parts. The fourth part concerns pseudorandomness notions for set systems and their application to approximation by juntas, which is the basis of the Junta Method in Extremal Combinatorics. We then apply these results in the fifth part to obtain several exact results on hypergraph Turán numbers.

Subsequent work

Since the first appearance of this paper, our global hypercontractivity inequality has found several further applications. Noise sensitivity of sparse sets is related to small-set expansion on graphs, which has found many applications in Computer Science. Here the interpretation of Theorem I.3.2 is that although not all small sets in the p-biased cube expand, global small sets do expand. Results of a similar nature were proved for the Grassman graph (see [60]) and the Johnson graph (see [59]). The former result was essential in the proof of the 2-to-2 Games Conjecture, a prominent problem in the field of hardness of approximation. Both these works involve long calculations, and have sub-optimal parameters. In subsequent works [22, 27, 28, 54] hypercontractive results for global functions are proven for various domains by reducing to the p-biased cube and using Theorem I.1.3. The results of [22, 27] imply the corresponding results about small expanding sets in the Grassman/Johnson graph with optimal parameters. A similar result was also established for a certain noise operator on the symmetric group [28].

Part I
Results

This part is an extended synopsis of our paper, in which we motivate and state our results. Section I.1 concerns our results on global hypercontractivity. We consider its applications to Analysis in the two subsequent sections. In Section I.2 we discuss sharp thresholds and the Kahn–Kalai Conjecture. Section I.3 concerns noise sensitivity, which gives an alternative approach to sharp thresholds and is of interest in its own right. We conclude the part by discussing our applications to Extremal Combinatorics in Section I.4.

I.1 Hypercontractivity of global functions

Before formally stating our main theorem, we start by recalling (the p-biased version of) the classical hypercontractive inequality. Let $p \in \left(0, \frac{1}{2}\right]$ (the case $p > \frac{1}{2}$ is similar). For $r \geq 1$ we write $\|\cdot\|_r$ (suppressing p from our notation) for the norm on $L^r(\{0,1\}^n, \mu_p)$.

Definition I.1.1 (Noise operator)**.** For $x \in \{0,1\}^n$ we define the ρ-correlated distribution $N_\rho(x)$ on $\{0,1\}^n$: a sample $\mathbf{y} \sim N_\rho(x)$ is obtained by, independently for each i setting $\mathbf{y}_i = x_i$ with probability ρ, or otherwise (with probability $1-\rho$) we resample $\mathbf{y}_i$ with $\mathbb{P}(\mathbf{y}_i = 1) = p$. We define the noise operator T_ρ on $L^2(\{0,1\}^n, \mu_p)$ by

$$\mathrm{T}_\rho\left(f\right)\left(x\right) = \mathbb{E}_{\boldsymbol{y}\sim N_\rho(x)}\left[f\left(\boldsymbol{y}\right)\right].$$

Hölder's inequality gives $\|f\|_r \leq \|f\|_s$ whenever $r \leq s$. The hypercontractivity theorem gives an inequality in the other direction after applying noise to f; for example, for $p = 1/2$, $r = 2$ and $s = 4$ we have

$$\|\mathrm{T}_\rho f\|_4 \leq \|f\|_2$$

for any $\rho \leq \frac{1}{\sqrt{3}}$. A similar inequality also holds when $p = o(1)$, but the correlation ρ has to be so small that it is not useful in applications; e.g. if $f(x) = x_1$ (the 'dictator' or 'half cube'), then $\|f\|_2 = \sqrt{\mu_p(f)} = \sqrt{p}$ and $\mathrm{T}_\rho f(x) = \mathbb{E}_{\boldsymbol{y}\sim N_\rho(x)}\mathbf{y}_1 = \rho x_1 + (1-\rho)p$, so $\|\mathrm{T}_\rho f\|_4 > (\mathbb{E}[\rho^4 x_1^4])^{1/4} = \rho p^{1/4}$. Thus we need $\rho = O(p^{1/4})$ to obtain any hypercontractive inequality for general f.

I.1.1 Local and global functions

To resolve this issue, we note that the tight examples for the hypercontractive inequality are *local*, in the sense that a small number of coordinates can significantly influence the output of the function. On the other hand, many functions of interest are *global*, in the sense that a small number of coordinates can change the output of the function only with a negligible probability; such global functions appear naturally in Random Graph Theory [2], Theoretical Computer Science [36] and Number Theory [37]. Our hypercontractive inequality will show that constant noise suffices for functions that are global in a sense captured by *generalised influences*, which we will now define.

Let $f\colon \{0,1\}^n \to \mathbb{R}$. For $S \subset [n]$ and $x \in \{0,1\}^S$, we write $f_{S\to x}$ for the function obtained from f by restricting the coordinates of S according to x (if $S = \{i\}$ is a singleton we simplify notation to $f_{i\to x}$). We write $|x|$ for the number of ones in x. For $i \in [n]$, the ith influence is $\mathrm{I}_i(f) = \|f_{i\to 1} - f_{i\to 0}\|_2^2$, where the norm is with respect to the implicit measure μ_p. In general, we define the influence with respect to any $S \subset [n]$ by sequentially applying the operators $f \mapsto f_{i\to 1} - f_{i\to 0}$ for all $i \in S$, as follows.

Definition I.1.2. For $f\colon \{0,1\}^n \to \mathbb{R}$ and $S \subset [n]$ we let (suppressing p in the notation)

$$\mathrm{I}_S\left(f\right) = \mathbb{E}_{\mu_p}\Big[\Big(\sum_{x\in\{0,1\}^S} (-1)^{|S|-|x|}\, f_{S\to x}\Big)^2\Big].$$

We say f has *β-small generalised influences* if $\mathrm{I}_S(f) \le \beta\, \mathbb{E}[f^2]$ for all $S \subseteq [n]$.

The reader familiar with the KKL theorem and the invariance principle may wonder why it is necessary to introduce generalised influences rather than only considering influences (of singletons). The reason is that under the uniform measure the properties of having small influences or small generalised influences are qualitatively equivalent, but this is no longer true in the p-biased setting for small p (consider $f(x) = \frac{x_1x_2+\cdots+x_{n-1}x_n}{\|x_1x_2+\cdots+x_{n-1}x_n\|}$).

We are now ready to state our main theorem, which shows that global[1] functions are hypercontractive for a noise operator with a constant rate. Moreover, our result applies to general L^r norms and product spaces (see Section II.1), but for simplicity here we just highlight the case of $(4,2)$-hypercontractivity in the cube.

Theorem I.1.3. *Let $p \in \left(0, \frac{1}{2}\right]$. Suppose $f \in L^2\left(\{0,1\}^n, \mu_p\right)$ has β-small generalised influences (for p). Then $\|\mathrm{T}_{1/5} f\|_4 \le \beta^{1/4}\|f\|_2$.*

We now move on to demonstrate the power of global hypercontractivity in the contexts of isoperimetry, noise sensitivity, sharp thresholds, and invariance. We emphasise that Theorem I.1.3 is the only new ingredient required for these applications, so we expect that it will have many further applications to generalising results proved via usual hypercontractivity on the cube with uniform measure.

I.1.2 Isoperimetry and influence

Stability of isoperimetric problems is a prominent open problem at the interface of Geometry, Analysis and Combinatorics. This meta-problem is to characterise sets whose boundary is close to the minimum possible given their volume; there are many specific problems obtained by giving this a precise meaning. Such results in Geometry were obtained for the classical setting of Euclidean Space by Fusco, Maggi and Pratelli [42] and for Gaussian Space by Mossel and Neeman [74].

The relevant setting for our paper is that of the cube $\{0,1\}^n$, endowed with the p-biased measure μ_p. We refer to this problem as the (p-biased) edge-isoperimetric stability problem. We identify any subset of $\{0,1\}^n$ with its characteristic Boolean function $f : \{0,1\}^n \to \{0,1\}$, and define its 'boundary' as the (total) influence[2]

$$\mathrm{I}[f] = \sum_{i=1}^{n} \mathrm{I}_i[f], \text{ where each } \mathrm{I}_i[f] = \Pr_{\boldsymbol{x}\sim\mu_p}\left[f(\boldsymbol{x} \oplus e_i) \neq f(\boldsymbol{x})\right],$$

i.e. the *ith influence* $\mathrm{I}_i[f]$ of f is the probability that f depends on bit i at a random input according to μ_p. (The notion of influence for real-valued functions, given in Section I.1, coincides with this notion for Boolean-valued functions). When $p = 1/2$ the total influence corresponds to the classical combinatorial notion of edge-boundary[3].

The KKL theorem of Kahn, Kalai and Linial [50] concerns the structure of functions $f : \{0,1\}^n \to \{0,1\}$, considering the cube under the uniform measure, with variance bounded away from 0 and 1 and with total influence is upper bounded by some number K. It states that f has a coordinate with influence at least $e^{-O(K)}$. The tribes example of Ben-Or and Linial [5] shows that this is sharp.

I.1.3 p-biased versions

The p-biased edge-isoperimetric stability problem is somewhat understood in the *dense regime* (where $\mu_p(f)$ is bounded away from 0 and 1) especially for Boolean functions f that are *monotone* (satisfy

[1]Strictly speaking, our assumption is stronger than the most natural notion of global functions: we require all generalised influences to be small, whereas a function should be considered global if it has small generalised influences $I_S(f)$ for small sets S. However, in practice, the hypercontractivity Theorem is typically applied to low-degree truncations of Boolean functions (see Section II.1), when there is no difference between these notions, as $I_S(f) = 0$ for large S.

[2]Everything depends on p, which we fix and suppress in our notation.

[3]For the vertex boundary, stability results showing that approximately isoperimetric sets are close to Hamming balls were obtained independently by Keevash and Long [56] and by Przykucki and Roberts [78].

$f(x) \leq f(y)$ whenever all $x_i \leq y_i$). Roughly speaking, most edge-isoperimetric stability results in the dense regime say that Boolean functions of small influence have some 'local' behaviour (see the seminal works of Friedgut–Kalai [38], Friedgut [35, 36], Bourgain [36, Appendix], and Hatami [45]). In particular, Bourgain (see also [77, Chapter 10]) showed that for any monotone Boolean function f with $\mu_p(f)$ bounded away from 0 and 1 and $pI[f] \leq K$ there is a set J of $O(K)$ coordinates such that $\mu_p(f_{J\to 1}) \geq \mu_p(f) + e^{-O(K^2)}$. This result is often interpreted as 'almost isoperimetric (dense) subsets of the p-biased cube must be local' or on the contrapositive as 'global functions have large total influence'. Indeed, if a restriction of a small set of coordinates can significantly boost the p-biased measure of a function, then this intuitively means that it is of a local nature.

For monotone functions, the conclusion in Bourgain's theorem is equivalent (see Section III.1) to having some set J of size $O(K)$ with $\mathrm{I}_J(f) \geq e^{-O(K^2)}$. Thus Bourgain's theorem can be viewed as a p-biased analog of the KKL theorem, where influences are replaced by generalised influences. However, unlike the KKL Theorem, Bourgain's result is not sharp, and the anti-tribes example of Ben-Or and Linial only shows that the K^2 term in the exponent cannot drop below K.

As a first application of our hypercontractivity theorem we replace the term $e^{-O(K^2)}$ by the term $e^{-O(K)}$, which is sharp by Ben-Or and Linial's example, see Section III.1.

Theorem I.1.4. *Let $p \in \left(0, \frac{1}{2}\right]$, and let $f\colon \{0,1\}^n \to \{0,1\}$ be a monotone Boolean function with $\mu_p(f)$ bounded away from 0 and 1 and $\mathrm{I}[f] \leq \frac{K}{p}$. Then there is a set J of $O(K)$ coordinates such that $\mu_p(f_{J\to 1}) \geq \mu_p(f) + e^{-O(K)}$.*

For general functions we prove a similar result, where the conclusion $\mu_p(f_{J\to 1}) \geq \mu_p(f) + e^{-O(K)}$ is replaced with $\mathrm{I}_J(f) \geq e^{-O(K)}$.

I.1.4 The sparse regime

On the other hand, the *sparse regime* (where we allow any value of $\mu_p(f)$) seemed out of reach of previous methods in the literature. Here Russo [79], and independently Kahn and Kalai [49], gave a proof of the p-biased isoperimetric inequality: $p\mathrm{I}[f] \geq \mu_p(f)\log_p(\mu_p(f))$ for every f. They also showed that equality holds only for the monotone sub-cubes. Kahn and Kalai posed the problem of determining the structure of monotone Boolean functions f that they called *d-optimal*, meaning that $p\mathrm{I}[f] \leq d\mu_p(f)\log_p(\mu_p(f))$, i.e. functions with total influence within a certain multiplicative factor of the minimal value guaranteed by the isoperimetric inequality. They conjectured in [49, Conjecture 4.1(a)] that for any constant $C > 0$ there are constants $K, \delta > 0$ such that if f is $C\log(1/p)$-optimal then there is a set J of $\leq K \log \frac{1}{\mu_p(f)}$ coordinates such that $\mu_p(f_{J\to 1}) \geq (1+\delta)\mu_p(f)$.

The corresponding result with a similar conclusion was open even for C-optimal functions! Our second theorem is a variant of the Kahn–Kalai conjecture which applies to $C\log(1/p)$-optimal functions when C is sufficiently small (whereas the conjecture requires an arbitrary constant C). We compensate for our stronger hypothesis in the following result by obtaining a much stronger conclusion than that asked for by Kahn and Kalai; for example, if f is $\frac{\log(1/p)}{100C}$-optimal then $\mu_p(f_{J\to 1}) \geq \mu_p(f)^{0.01}$. We will also show that our result is sharp up to the constant factor C.

Theorem I.1.5. *Let $p \in \left(0, \frac{1}{2}\right]$, $K \geq 1$ and let f be a Boolean function with $p\mathrm{I}[f] < K\mu_p(f)$. Then there is a set J of $\leq CK$ coordinates, where C is an absolute constant, such that $\mu_p(f_{J\to 1}) \geq e^{-CK}$.*

I.2 Sharp thresholds

The results of Friedgut and Bourgain mentioned above also had the striking consequence that any 'global' Boolean function has a sharp threshold, which was a breakthrough in the understanding of this phenomenon, as it superceded many results for specific functions.

The sharp threshold phenomenon concerns the behaviour of $\mu_p(f_n)$ for p around the critical probability, defined as follows. Consider any sequence $f_n\colon \{0,1\}^n \to \{0,1\}$ of monotone Boolean functions.

For $t \in [0,1]$ let $p_n(t) = \inf\{p : \mu_p(f_n) \geq t\}$. In particular, $p_n^c := p_n(1/2)$ is commonly known as the 'critical probability' (which we think of as small in this paper). A classical theorem of Bollobás and Thomason [10] shows that for any $\varepsilon > 0$ there is $C > 0$ such that $p_n(1-\varepsilon) \leq Cp_n(\varepsilon)$. This motivates the following definition: we say that the sequence (f_n) has a *coarse threshold* if for each $\varepsilon > 0$ the length of the interval $[p_n(\varepsilon), p_n(1-\varepsilon)]$ is $\Theta(p_n^c)$, otherwise we say that it has a *sharp threshold*.

The classical approach for understanding sharp thresholds is based on the Margulis–Russo formula $\frac{d\mu_p(f)}{dp} = \mathrm{I}_{\mu_p}(f)$, see [70] and [79]. Here we note that if f has a coarse threshold, then by the Mean Value Theorem there is a constant $\epsilon > 0$, some p with $\mu_p(f) \in (\epsilon, 1-\epsilon)$ and $p\mathrm{I}_{\mu_p}(f) = \Theta(1)$, so one can apply various results mentioned in Section I.1.2. Thus Bourgain's Theorem implies that there is a set J of $O(K)$ coordinates such that $\mu_{p'}(f_{J\to 1}) \geq \mu_{p'}(f) + e^{-O(K^2)}$. While this approach is useful for studying the behaviour of f around the critical probability, it rarely gives any information regarding the location of the critical probability. Indeed, many significant papers are devoted to locating the critical probability of specific interesting functions, see e.g. the breakthroughs of Johansson, Kahn and Vu [48] and Montgomery [72].

A general result was conjectured by Kahn and Kalai for the class of Boolean functions of the form $f_n \colon \{0,1\}^{\binom{[n]}{2}} \to \{0,1\}$, whose input is a graph G and whose output is 1 if G contains a certain fixed graph H. For such functions there is a natural 'expectation heuristic' p_n^E for the critical probability, namely the least value of p such that the expected number of copies of any subgraph of H in $G(n,p)$ is at least $1/2$. Markov's inequality implies $p_n^c \geq p_n^E$, and the hope of the Kahn–Kalai Conjecture is that there is a corresponding upper bound up to some multiplicative factor. They conjectured in [49, Conjecture 2.1] that $p_n^c = O\left(p_n^E \log n\right)$, but this is widely open, even if $\log n$ is replaced by $n^{o(1)}$.

The proposed strategy of Kahn and Kalai to this conjecture via isoperimetric stability is as follows.

- Prove a lower bound on $\mu_{p_n^E}(f_n)$.
- Show (e.g. via Russo's lemma) that if $|[p_E, p_c]|$ is too large, then the p-biased total influence at some point in the interval $[p_E, p_c]$ must be relatively small.
- Prove an edge-isoperimetric stability result that rules out the latter possibility.

Theorem I.1.5 makes progress on the third ingredient. Combining it with Russo's Lemma, we obtain the following result that can be used to bound the critical probability. Let f be a monotone Boolean function. We say that f is *M-global* in an interval I if for each set J of size $\leq M$ and each $p \in I$ we have $\mu_p(f_{J\to 1}) \leq \mu_p(f)^{0.01}$.

Theorem I.2.1. *There exists an absolute constant C such that the following holds for any monotone Boolean function f with critical probability p_c and $p \leq p_c$. Suppose for some $M > 0$ that f is M-global in the interval $[p, p_c]$ and that $\mu_p(f) \geq e^{-M/C}$. Then $p_c \leq M^C p$.*

To see the utility of Theorem I.2.1, imagine that one wants to bound the critical probability as $p_n^c \leq p$, but instead of showing $\mu_p(f_n) \geq \frac{1}{2}$ one can only obtain a weaker lower bound $\mu_p(f) \geq e^{-M/C}$, where f is M-global; then one can still bound the critical probability as $p_n^c \leq M^{O(1)}p$.

I.3 Noise sensitivity

Studying the effect of 'noise' on a Boolean function is a fundamental paradigm in various contexts, including hypercontractivity (as in Section I.1) and Gaussian isoperimetry (via the invariance principle, see Section II.3). Roughly speaking, a function f is 'noise sensitive' if $f(x)$ and $f(y)$ are approximately independent for a random input x and random small perturbation y of x; an equivalent formulation (which we adopt below) is that the 'noise stability' of f is small (compared to $\mu_p(f)$). Formally, we use the following definition.

Definition I.3.1. The noise stability $\mathrm{Stab}_\rho(f)$ of $f \in L^2(\{0,1\}^n, \mu_p)$ is defined by

$$\mathrm{Stab}_\rho(f) = \langle f, \mathrm{T}_\rho f\rangle = \mathbb{E}_{\boldsymbol{x}\sim\mu_p}\left[f(\boldsymbol{x})\,\mathrm{T}_\rho f(\boldsymbol{x})\right].$$

A sequence f_n of Boolean functions is said to be noise sensitive if for each fixed ρ we have $\mathrm{Stab}_\rho(f_n) = \mu_p(f_n)^2 + o(\mu_p(f_n))$.

Note that everything depends on p, but this will be clear from the context, so we suppress p from the notation Stab_ρ. Kahn, Kalai, and Linial [50] (see also [77, Section 9]) showed that sparse subsets of the uniform cube are noise sensitive, where we recall that the sequence (f_n) is *sparse* if $\mu_p(f_n) = o(1)$ and *dense* if $\mu_p(f_n) = \Theta(1)$.

The relationship between noise and influence in the cube under the uniform measure was further studied by Benjamini, Kalai, and Schramm [8] (with applications to percolation), who gave a complete characterisation: a sequence (f_n) of monotone dense Boolean functions is noise sensitive if and only if the sum of the squares of the influences of f_n is $o(1)$. Schramm and Steif [80] proved that any dense Boolean function on n variables that can be computed by an algorithm that reads $o(n)$ of the input bits is noise sensitive. Their result had the striking application that the set of exceptional times in dynamical critical site percolation on the triangular lattice, in which an infinite cluster exists, is of Hausdorff dimension in the interval $\left[\frac{1}{6}, \frac{31}{36}\right]$. Ever since, noise sensitivity was considered in many other contexts (see e.g. the recent results and open problems of Lubetzky–Steif [68] and Benjamini [7]).

In contrast to the uniform setting, in the p-biased setting for small p it is no longer true that sparse sets are noise sensitive (e.g. consider dictators). Our main contribution to the theory of noise sensitivity is showing that 'global' sparse sets are noise sensitive. Formally, we say that a sequence f_n of sparse Boolean functions is *weakly global* if for any $\varepsilon, C > 0$ there is $n_0 > 0$ so that $\mu_p((f_n)_{J\to 1}) < \varepsilon$ for all $n > n_0$ and J of size at most C.

Theorem I.3.2. *Any weakly global sequence of Boolean functions is noise sensitive.*

We will deduce the following sharp threshold result, which will underpin our combinatorial applications discussed in the next section.

Theorem I.3.3. *For any $\alpha > 0$ there is $C > 0$ such that for any $\varepsilon, p, q \in (0, 1/2)$ with $q \geq (1+\alpha)p$, writing $r = C\log\varepsilon^{-1}$ and $\delta = 10^{-3r-1}\varepsilon^3$, any monotone (r,δ)-global Boolean function f with $\mu_p(f) \leq \delta$ satisfies $\mu_q(f) \geq \mu_p(f)/\varepsilon$.*

I.4 Hypergraph Turán numbers

A longstanding and challenging direction of research in Extremal Combinatorics, initiated by Turán in the 1940's, is that of determining the maximum size of a k-graph (k-uniform hypergraph) $\mathcal{H} \subset \binom{[n]}{k}$ on n vertices not containing some fixed k-graph F; this is the Turán number, denoted $\mathrm{ex}(n, F)$. Turán numbers of graphs (the case $k = 2$) are quite well-understood (if F is not bipartite), but there are very few results even for specific hypergraphs, let alone general results for families of hypergraphs (see the survey [52]). Here we prove a number of general results on Turán numbers for the family of bounded degree expanded hypergraphs (to be defined below), thus solving several open problems. Our proofs build upon our new sharp threshold theorems and the Junta Method of Keller and Lifshitz [57] (which greatly extended the applications of an approach initiated by Dinur and Friedgut [15]). A striking feature of our results is their applicability across an essentially optimal range of uniformities and sizes, which previously seemed entirely out of reach.

I.4.1 Cross matchings

Before introducing the general setting of expanded hypergraphs, we first consider an important case, which is in itself a source of many significant problems, namely the problem of finding matchings. In both theory and application, a wide range of significant questions can be recast as existence questions for matchings (see e.g. the books [67, 81] and the survey [53]).

Perhaps the most well-known open question concerning matchings, due to Erdős [25], asks how large a family $\mathcal{F} \subset \binom{[n]}{k}$ can be if it does not contain an s-matching, i.e. sets $\{A_1, \dots, A_s\}$ with $A_i \cap A_j = \emptyset$ for all distinct $i, j \in [s]$. Two natural families of such $\mathcal{F}$ are stars $\mathcal{S}_{n,k,s-1} := \left\{A \in \binom{[n]}{k} : A \cap [s-1] \neq \emptyset\right\}$ and cliques $\mathcal{C}_{k,s-1} := \binom{[ks-1]}{k}$. Erdős conjectured that one of these families is always extremal.

Conjecture I.4.1 (Erdős Matching Conjecture)**.** *Let $n \geq ks$ and suppose that $\mathcal{F} \subset \binom{[n]}{k}$ does not contain an s-matching. Then $|\mathcal{F}| \leq \max\left\{|\mathcal{S}_{n,k,s-1}|, |\mathcal{C}_{k,s-1}|\right\}$.*

This conjecture remains open, despite an extensive literature, of which we will mention a few highlights. The case $s = 2$ is the classical Erdős–Ko–Rado theorem [26]. Erdős and Gallai [24] confirmed the conjecture for $k = 2$. The case $k = 3$ was proven by Luczak and Mieczkowska [69] for large s and by Frankl [33] for all s. Bollobás, Daykin and Erdős [9] proved the conjecture provided $n = \Omega(k^3 s)$, which was reduced to $n = \Omega(k^2 s)$ by Huang, Loh and Sudakov [46] and finally to $n = \Omega(ks)$ by Frankl [29] (in fact to $n \gtrsim 2ks$, recently improved by Frankl and Kupavskii [31] to $n \geq 5ks/3$ for large s), which is the optimal order of magnitude for the extremal family to be a star rather than a clique – or even to just contain s disjoint k-sets.

Our first result in this context is a cross version of that of Frankl, which proves (a strengthened form of) a conjecture of Huang, Loh and Sudakov [46]. Here we say that families $\mathcal{F}_1, \dots, \mathcal{F}_s$ cross contains a hypergraph $\{A_1, \dots, A_s\}$ (e.g. an s-matching) if $A_i \in \mathcal{F}_i$ for each $i \in [s]$.

Theorem I.4.2. *There is a constant $C > 0$ so that if $n, s, k_1, \dots, k_s \in \mathbb{N}$ with $k_i \leq n/s$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $|\mathcal{F}_i| \geq |\mathcal{S}_{n,k_i,s-1}|$ for all $i \in [s]$, either $\mathcal{F}_1, \dots, \mathcal{F}_s$ cross contain an s-matching, or there is $J \subset [n]$ with $|J| = s - 1$ such that each $\mathcal{F}_i = \mathcal{S}_{n,k_i,J} := \{A \in \binom{[n]}{k_i} : A \cap J \neq \emptyset\}$.*

Remark I.4.3. Theorem I.4.2 in the case that all $k_i = k$ was proved by Huang, Loh and Sudakov [46] for $n = \Omega(k^2 s)$ and recently by Frankl and Kupavskii [32] for $n = \Omega(ks \log s)$; our result applies to $n = \Omega(ks)$, which is the optimal order of magnitude. Moreover, we obtain a strong stability result (see Theorem V.1.1 below) which gives structural information even if we only assume that the size of each family is within a constant factor of that of a star: either there is a cross matching or some family correlates strongly with a star. Besides having independent interest, this stability result will play a key role in the proof of our general Turán results.

I.4.2 Expanded hypergraphs

As mentioned above, there are very few general results on Turán numbers for a family of hypergraphs. One family for which there has been substantial progress is that of expanded graphs (see the survey [76]). Given an r-graph G and $k \geq r$, the k-expansion $G^+ = G^+(k)$ is the k-uniform hypergraph obtained from G by adding $k - r$ new vertices to each edge, i.e. G^+ has edge set $\{e \cup S_e : e \in E(G)\}$ where $|S_e| = k - r$, $S_e \cap V(G) = \emptyset$ and $S_e \cap S_{e'} = \emptyset$ for all distinct $e, e' \in E(G)$. In particular, a k-graph s-matching is the k-expansion of a graph s-matching.

When G is a graph (the case $r = 2$), in the non-degenerate case when k is less than the chromatic number $\chi(G)$ the Turán numbers $\mathrm{ex}(n, G^+(k))$ are well-understood (see [76, Section 2]), so the main focus for ongoing research is the degenerate case $k \geq \chi(G)$. Here Frankl and Füredi [30] introduced the following important parameter and corresponding construction that seems to often determine the asymptotics of the Turán number. For any r-graph G, we call $S \subset V(G^+)$ a *crosscut* if $|E \cap S| = 1$ for all $E \in G^+$. The crosscut $\sigma(G)$ of G is the size of the minimal such set, i.e.

$$\sigma(G) := \min\left\{|S| : S \subset V(G^+) \text{ with } |E \cap S| = 1 \text{ for all } E \in G^+\right\}.$$

It is easy to see that $\sigma(G)$ exists for $k \geq r + 1$ and is independent of k. Clearly,

$$\mathcal{S}^{(1)}_{n,k,\sigma(G)-1} := \left\{A \in \binom{[n]}{k} : |A \cap [\sigma(G) - 1]| = 1\right\}$$

is G^+-free. Moreover, this simple construction determines the asymptotics of $\mathrm{ex}(n, G^+(k))$ for $n > n_0(k, G)$ for several graphs G, including paths [41, 61], cycles [40, 61] and trees [39, 62]. Given this phenomenon, according to Mubayi and Verstraëte [76], one of the major open problems on expansions is to decide when the Turán number is asymptotically determined by the crosscut construction. Our next result resolves this problem for all bounded degree r-graphs (so in particular for graphs) in a range of parameters that is optimal up to constant factors. Moreover, we also obtain a strong structural approximation for any family that is close to extremal (see Theorem I.4.8 below).

Theorem I.4.4. *For any $r, \Delta \geq 2$ and $\varepsilon > 0$ there is $C > 0$ so that the following holds for any r-graph G with s edges, maximum degree $\Delta(G) \leq \Delta$ and $\sigma(G) \geq 2$. For any $k, n \in \mathbb{N}$ with $C \leq k \leq n/Cs$ we have $ex(n, G^+(k)) = (1 \pm \varepsilon)|\mathcal{S}^{(1)}_{n,k,\sigma(G)-1}|$.*

Remark I.4.5. Some lower bound on k is necessary to obtain the conclusion in Theorem I.4.4. Indeed, we have already mentioned that the non-degenerate case $k \leq \chi(G)$ when G is a graph exhibits different behaviour (a complete partite k-graph shows that $\mathrm{ex}(n, G^+) = \Omega(n/k)^k$), and moreover, examples in [76] show that some lower bound on k may be necessary even if G is bipartite (e.g. if $G = K_{9,9}$ then consider the 3-graph of triangles in a suitably dense random graph made G-free by edge deletions). The upper bound on k in our result is also necessary up to the constant factor by space considerations, as even the complete k-graph $\binom{[n]}{k}$ can only contain $G^+(k)$ if $n \geq |V(G^+)| = |V(G)| + (k-2)s$. With the exception of Frankl's matching theorem [29], Theorem I.4.4 appears to be the only known Turán result in which both the uniformity k and the size s can vary over such a wide range.

Next we consider conditions under which we can refine the asymptotic result of Theorem I.4.4 and determine the Turán number $\mathrm{ex}(n, G^+)$ exactly. One complication here is that crosscuts may be beaten by stars $\mathcal{S}_{n,k,\tau(G)-1}$, where

$$\tau(G) := \min\left\{|S| : |S \cap e| \geq 1 \text{ for all } e \in E(G)\right\}$$

is the transversal number of G. Clearly $\tau(G) \leq \sigma(G)$. For fixed s, crosscuts cannot be beaten by smaller stars, but this may not hold when s grows with n, as then edges with more than one vertex in the base of the star are significant. Another complication is that lower order correction terms are necessary for certain G, e.g. for k-expanded paths $P_\ell^+(k)$ of length ℓ for $n > n_0(k, \ell)$ we have $\mathrm{ex}(n, P_3^+(k)) = \binom{n-1}{k-1} = |\mathcal{S}_{n,k,1}|$, as predicted by the crosscut/star construction, but $\mathrm{ex}(n, P_4^+(k)) = \binom{n-1}{k-1} + \binom{n-3}{k-2}$, as we can add all sets containing some fixed pair of vertices. This is analogous to the familiar situation in extremal graph theory where we only expect exact results for graphs that are critical with respect to the key parameter of the extremal construction. Accordingly, we introduce the following analogous concept of criticality for expanded hypergraphs with respect to crosscuts and stars: we say that G is *critical* if it has an edge e such that

$$\sigma(G \setminus e) = \tau(G \setminus e) < \tau(G) = \sigma(G).$$

We obtain the following general exact result for Turán numbers.

Theorem I.4.6. *For any $r, \Delta \geq 2$ there is $C > 0$ such that for any critical r-graph G with s edges, maximum degree $\Delta(G) \leq \Delta$ and $C \leq k \leq n/Cs$ we have $ex(n, G^+(k)) = |\mathcal{S}_{n,k,\sigma(G)-1}|$.*

This result applies to many graphs considered in the previous literature, such as paths of odd length. Paths of even length are not critical, but satisfy a generalised criticality property: deleting one edge does not reduce the transversal number, but deleting two edges (whether disjoint or intersecting) does reduce the crosscut number. Thus we have the following natural construction for excluding any expanded path P_ℓ^+ of length ℓ. Let $\mathcal{F}^*_{n,k,\ell} = \mathcal{S}_{n,k,J}$ with $|J| = \sigma(P_\ell) - 1$ if ℓ is odd, or if ℓ is even obtain $\mathcal{F}^*_{n,k,\ell}$ from $\mathcal{S}_{n,k,J}$ by adding $\{A \in \binom{[n]}{k} : T \subset A\}$ for some $T \in \binom{[n] \setminus J}{2}$. Clearly $\mathcal{F}^*_{n,k,\ell}$ is P_ℓ^+-free. Füredi, Jiang and Seiver [41] showed that $\mathrm{ex}(n, P_\ell^+) = |\mathcal{F}^*_{n,k,\ell}|$ provided $n \gg n_0(k, s)$, and conjectured that this holds provided $n \geq Cks$. We prove this conjecture.

Corollary I.4.7. *There is $C > 0$ so that if $n, k, \ell \in \mathbb{N}$ and $C \leq k \leq n/C\ell$ then $ex(n, P_\ell^+) = |\mathcal{F}^*_{n,k,\ell}|$.*

I.4.3 The Junta Method

In recent years, the Analysis of Boolean functions has found significant application in Extremal Combinatorics, via the connection provided by the Margulis-Russo formula between the sharp threshold phenomenon and influences of Boolean functions. This approach was initiated by Dinur and Friedgut [15], who applied a theorem of Friedgut [35] on Boolean functions of small influence to prove that large uniform intersecting families can be approximated by juntas, i.e. families that depend only on a few coordinates. This connection has since played a key role in intersection theorems for a variety of settings, including graphs [19], permutations [20] and sets [21, 23].

The approach of Dinur and Friedgut was substantially generalised by Keller and Lifshitz [57] to apply to a variety of Turán problems on expanded hypergraphs. At a very high level, their Junta Method is a version of the Stability Method in Extremal Combinatorics, in that it consists of two steps: an approximate step that determines the rough structure of families that are close to optimal, and an exact step that refines the structure and determines the optimal construction. Their approximate step consisted of showing that any G^+-free family is approximately contained in a G^+-free junta. This is also true in our approach, but the crucial difference is that they required the number s of edges in G to be fixed, whereas we allow it to grow as a function of n. Friedgut's theorem can no longer be applied in this setting, as we require a threshold result for Boolean functions $f : \{0,1\}^n \to \{0,1\}$ according to the p-biased measure μ_p in the sparse regime where both p and $\mu_p(f)$ may be functions of n that approach zero. Our new sharp threshold theorems, as in subsection I.2, provide the needed improvement on the analytic side which, when combined with a number of additional combinatorial ideas, allow us to obtain the following junta approximation theorem.

Theorem I.4.8. *For any $r, \Delta \geq 2$ and $\varepsilon > 0$ there are $c, C > 0$ so that, given an r-graph G with s edges and maximum degree $\Delta(G) \leq \Delta$, for any G^+-free $\mathcal{F} \subset \binom{[n]}{k}$ with $C \leq k \leq \frac{n}{Cs}$, there is $J \subset V(G)$ with $|J| \leq \sigma(G) - 1$ and $|\mathcal{F} \setminus \mathcal{S}_{n,k,J}| \leq \varepsilon |\mathcal{S}_{n,k,\sigma(G)-1}|$.*

We note that Theorem I.4.4 is immediate from Theorem I.4.8, as for $k \geq C \gg \varepsilon^{-1}$ we have

$$\mathrm{ex}(n, G^+) \geq |\mathcal{S}^{(1)}_{n,k,\sigma(G)-1}| \geq (1-\varepsilon)|\mathcal{S}_{n,k,\sigma(G)-1}|.$$

Part II

Hypercontractivity of global functions

This part concerns our theory of global hypercontractivity, which underpins all the results of this paper. We start in Section II.1 by proving Theorem I.1.3, which is the form of our result that suffices for our subsequent applications. In the remainder of the part (which could be omitted by a reader primarily interested in these applications) we investigate the theory more deeply, as this has independent interest and further applications. Section II.2 generalises our hypercontractivity result in two directions: we consider general norms and general product spaces. We conclude this part in Section II.3 by proving our p-biased version of the Invariance Principle and remarking on some of its applications (we omit the details of these for the sake of brevity).

Notation

Here we summarise some notation and basic properties of Fourier analysis on the cube. We fix $p \in (0,1)$ and suppress it in much of our notation, i.e. we consider $\{0,1\}^n$ to be equipped with the p-biased measure μ_p, unless otherwise stated. We let $\sigma = \sqrt{p(1-p)}$ (the standard deviation of a p-biased bit). For each $i \in [n]$ we define $\chi_i \colon \{0,1\}^n \to \mathbb{R}$ by $\chi_i(x) = \frac{x_i - p}{\sigma}$ (so χ_i has mean 0 and variance 1). We use the orthonormal Fourier basis $\{\chi_S\}_{S\subset[n]}$ of $L^2(\{0,1\}^n, \mu_p)$, where each $\chi_S := \prod_{i\in S}\chi_i$. Any $f : \{0,1\}^n \to \mathbb{R}$ has a unique expression $f = \sum_{S\subset[n]} \hat{f}(S)\chi_S$ where $\{\hat{f}(S)\}_{S\subset[n]}$ are the *p-biased Fourier coefficients of f*. Orthonormality gives the Plancherel identity $\langle f, g\rangle = \sum_{S\subset[n]} \hat{f}(S)\hat{g}(S)$. In particular, we have the Parseval identity $\mathbb{E}[f^2] = \|f\|_2^2 = \langle f, f\rangle = \sum_{S\subset[n]} \hat{f}(S)^2$. For $\mathcal{F} \subset \{0,1\}^n$ we define the $\mathcal{F}$-truncation $f^{\mathcal{F}} = \sum_{S\in\mathcal{F}} \hat{f}(S)\chi_S$. Our truncations will always be according to some degree threshold r, for which we write $f^{\le r} = \sum_{|S|\le r} \hat{f}(S)\chi_S$.

For $i \in [n]$, the *i-derivative* f_i and *i-influence* $\mathrm{I}_i(f)$ of f are

$$f_i = \mathrm{D}_i[f] = \sigma\big(f_{i\to 1} - f_{i\to 0}\big) = \sum_{S: i\in S} \hat{f}(S)\chi_{S\setminus\{i\}}, \text{ and}$$
$$\mathrm{I}_i(f) = \|f_{i\to 1} - f_{i\to 0}\|_2^2 = \sigma^{-2}\mathbb{E}[f_i^2] = \tfrac{1}{p(1-p)} \sum_{S: i\in S} \hat{f}(S)^2.$$

The *influence* of f is

$$\mathrm{I}(f) = \sum_i \mathrm{I}_i(f) = (p(1-p))^{-1} \sum_S |S|\hat{f}(S)^2. \tag{1}$$

In general, for $S \subset [n]$, the *S-derivative* of f is obtained from f by sequentially applying D_i for each $i \in S$, i.e.

$$\mathrm{D}_S(f) = \sigma^{|S|} \sum_{x\in\{0,1\}^S} (-1)^{|S|-|x|} f_{S\to x} = \sum_{T: S\subset T} \hat{f}(T)\chi_{T\setminus S}.$$

The *S-influence* of f (as in Definition I.1.2) is

$$\mathrm{I}_S(f) = \sigma^{-2|S|}\|\mathrm{D}_S(f)\|_2^2 = \sigma^{-2|S|} \sum_{E: S\subset E} \hat{f}(E)^2. \tag{2}$$

Recalling that a function f has α-small generalised influences if $\mathrm{I}_S(f) \le \alpha\mathbb{E}[f^2]$ for all $S \subset [n]$, we see that this is equivalent to $\mathbb{E}[\mathrm{D}_S(f)^2] \le \alpha\sigma^{2|S|}\mathbb{E}[f^2]$ for all $S \subset [n]$.

II.1 Hypercontractivity and generalised influences

In this section we prove our hypercontractive inequality (Theorem I.1.3), which is the fundamental result that underpins all of the results in this paper.

The idea of the proof is to reduce hypercontractivity in μ_p to hypercontractivity in $\mu_{1/2}$ via the 'replacement method' (the idea of Lindeberg's proof of the Central Limit Theorem, and of the proof of Mossel, O'Donnell and Oleszkiewicz [75] of the invariance principle). Throughout this section we fix $f : \{0,1\}^n \to \mathbb{R}$ and express f in the p-biased Fourier basis as $\sum_S \hat{f}(S)\chi_S^p$, where $\chi_S^p = \prod_{i\in S}\chi_i^p$ and $\chi_i^p(x) = \frac{x_i - p}{\sigma}$ (the same notation as above, except that we introduce the superscript p to distinguish the p-biased and uniform settings).

For $0 \le t \le n$ we define $f_t = \sum_S \hat{f}(S)\chi_S^t$, where

$$\chi_S^t = \prod_{i\in S\cap[t]}\chi_i^{1/2}(x)\prod_{i\in S\setminus[t]}\chi_i^p(x) \in L^2(\{0,1\}^{[t]},\mu_{1/2}) \times L^2(\{0,1\}^{[n]\setminus[t]},\mu_p).$$

Thus f_t interpolates from $f_0 = f \in L^2(\{0,1\}^n,\mu_p)$ to $f_n = \sum_S \hat{f}(S)\chi_S^{1/2} \in L^2(\{0,1\}^n,\mu_{1/2})$. As $\{\chi_S^t : S \subset [n]\}$ is an orthonormal basis we have $\|f_t\|_2 = \|f\|_2$ for all t.

We also define noise operators $\mathrm{T}_{\rho',\rho}^t$ on $L^2(\{0,1\}^{[t]},\mu_{1/2}) \times L^2(\{0,1\}^{[n]\setminus[t]},\mu_p)$ by $\mathrm{T}_{\rho',\rho}^t(g)(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{y}\sim N_{\rho',\rho}(\boldsymbol{x})}[f(\boldsymbol{y})]$, where to sample $\boldsymbol{y}$ from $N_{\rho',\rho}(\boldsymbol{x})$, for $i \le t$ we let $y_i = x_i$ with probability ρ' or otherwise we resample y_i from $\mu_{1/2}$, and for $i > t$ we let $y_i = x_i$ with probability ρ or otherwise we resample y_i from μ_p. Thus $\mathrm{T}_{\rho',\rho}^t$ interpolates from $\mathrm{T}_{\rho',\rho}^0 = \mathrm{T}_\rho$ (for μ_p) to $\mathrm{T}_{\rho',\rho}^n = \mathrm{T}_{\rho'}$ (for $\mu_{1/2}$).

We record the following estimate for 4-norms of p-biased characters:

$$\lambda := \mathbb{E}[(\chi_i^p)^4] = \sigma^{-4}(p(1-p)^4 + (1-p)p^4) = \sigma^{-2}((1-p)^3 + p^3) \le \sigma^{-2}.$$

The core of our argument by replacement is the following lemma which controls the evolution of $\mathbb{E}[(\mathrm{T}_{2\rho,\rho}^t f_t)^4] = \|\mathrm{T}_{2\rho,\rho}^t f_t\|_4^4$ for $0 \le t \le n$.

Lemma II.1.1. $\mathbb{E}[(\mathrm{T}_{2\rho,\rho}^{t-1}f_{t-1})^4] \le \mathbb{E}[(\mathrm{T}_{2\rho,\rho}^t f_t)^4] + 3\lambda\rho^4\mathbb{E}[(\mathrm{T}_{2\rho,\rho}^t((\mathrm{D}_t f)_t))^4]$.

Proof. We write

$$\begin{aligned} f_t &= \chi_t^{1/2}g + h \quad \text{and} \quad f_{t-1} = \chi_t^p g + h, \quad \text{where} \\ g &= (\mathrm{D}_t f)_t = \sum_{S:t\in S}\hat{f}(S)\chi_{S\setminus\{t\}}^t = \sum_{S:t\in S}\hat{f}(S)\chi_{S\setminus\{t\}}^{t-1} = (\mathrm{D}_t f)_{t-1}, \quad \text{and} \\ h &= \mathbb{E}_{x_t\sim\mu_{1/2}}f_t = \sum_{S:t\notin S}\hat{f}(S)\chi_S^t = \sum_{S:t\notin S}\hat{f}(S)\chi_S^{t-1} = \mathbb{E}_{x_t\sim\mu_p}f_{t-1}. \end{aligned}$$

We also write

$$\mathrm{T}_{2\rho,\rho}^t f_t = 2\rho\chi_t^{1/2}d + e \quad \text{and} \quad \mathrm{T}_{2\rho,\rho}^{t-1}f_{t-1} = \rho\chi_t^p d + e, \quad \text{where}$$
$$d = \mathrm{T}_{2\rho,\rho}^t g = \mathrm{T}_{2\rho,\rho}^{t-1}g \quad \text{and} \quad e = \mathrm{T}_{2\rho,\rho}^t h = \mathrm{T}_{2\rho,\rho}^{t-1}h.$$

We can calculate the expectations in the statement of the lemma by conditioning on all coordinates other than x_t, i.e. $\mathbb{E}_{\boldsymbol{x}}[\cdot] = \mathbb{E}_{\boldsymbol{x}'}[\mathbb{E}_{x_t}[\cdot \mid \boldsymbol{x}']]$ where $\boldsymbol{x}'$ is obtained from $\boldsymbol{x} = (x_1,\dots,x_n)$ by removing x_t. It therefore suffices to establish the required inequality for each fixed $\boldsymbol{x}'$ with expectations over the choice of x_t; thus we can treat d and e as constants, and it suffices to show

$$\mathbb{E}_{x_t}[(\rho d\chi_t^p + e)^4] \le \mathbb{E}_{x_t}[(2\rho d\chi_t^{1/2} + e)^4] + 3\lambda\rho^4 d^4. \tag{3}$$

As χ_t^p has mean 0, we can expand the left hand side of (3) as

$$(\rho d)^4\mathbb{E}[(\chi_t^p)^4] + 4e(\rho d)^3\mathbb{E}[(\chi_t^p)^3] + 6e^2(\rho d)^2\mathbb{E}[(\chi_t^p)^2] + e^4 \le 3\lambda(d\rho)^4 + 8(de\rho)^2 + e^4,$$

where we bound the second term using Cauchy-Schwarz then AM-GM by

$$4\cdot\mathbb{E}[(d\rho\chi_t^p)^4]^{1/2}\cdot\mathbb{E}[(de\rho\chi_t^p)^2]^{1/2}\le 2\left(\mathbb{E}[(d\rho\chi_t^p)^4]+\mathbb{E}[(de\rho\chi_t^p)^2]\right)=2(\lambda(d\rho)^4+(de\rho)^2).$$

Similarly, as $\mathbb{E}[\chi_t^{1/2}]=\mathbb{E}[(\chi_t^{1/2})^3]=0$, we can expand the first term on the right hand side of (3) as

$$(2\rho d)^4\mathbb{E}[(\chi_t^{1/2})^4]+6e^2(2\rho d)^2\mathbb{E}[(\chi_t^{1/2})^2]+e^4=(2\rho d)^4+6(2\rho de)^2+e^3\ge 8(de\rho)^2+e^4.$$

The lemma follows. □

Now we apply the previous lemma inductively to prove the following estimate.

Lemma II.1.2. $\|\mathrm{T}^i_{2\rho,\rho}f_i\|_4^4\le\sum_{S\subset[n]\setminus[i]}(3\lambda\rho^4)^{|S|}\|\mathrm{T}^n_{2\rho,\rho}((\mathrm{D}_Sf)_n)\|_4^4$ *for all* $0\le i\le n$.

Proof. We prove the inequality by induction on $n-i$ simultaneously for all functions f. If $n=i$ then equality holds trivially. Now suppose that $i<n$. By Lemma II.1.1 with $t=i+1$, and the induction hypothesis applied to f and D_tf with i replaced by t, we have

$$\begin{aligned}\|\mathrm{T}^i_{2\rho,\rho}f_i\|_4^4&\le\|\mathrm{T}^t_{2\rho,\rho}f_t\|_4^4+3\lambda\rho^4\|\mathrm{T}^t_{2\rho,\rho}((\mathrm{D}_tf)_t)\|_4^4\\&\le\sum_{S\subset[n]\setminus[t]}(3\lambda\rho^4)^{|S|}\|\mathrm{T}^n_{2\rho,\rho}((\mathrm{D}_Sf)_n)\|_4^4+3\lambda\rho^4\sum_{S\subset[n]\setminus[t]}(3\lambda\rho^4)^{|S|}\|\mathrm{T}^n_{2\rho,\rho}((\mathrm{D}_S\mathrm{D}_tf)_n)\|_4^4\\&=\sum_{S\subset[n]\setminus[i]}(3\lambda\rho^4)^{|S|}\|\mathrm{T}^n_{2\rho,\rho}((\mathrm{D}_Sf)_n)\|_4^4.\end{aligned}$$ □

In particular, recalling that $\mathrm{T}^0_{2\rho,\rho}=\mathrm{T}_\rho$ on μ_p and $\mathrm{T}^n_{2\rho,\rho}=\mathrm{T}_{2\rho}$ on $\mu_{1/2}$, the case $i=0$ of Lemma II.1.2 is as follows.

Proposition II.1.3. $\|\mathrm{T}_\rho f\|_4^4\le\sum_{S\subset[n]}(3\lambda\rho^4)^{|S|}\|\mathrm{T}_{2\rho}((\mathrm{D}_Sf)_n)\|_4^4$.

The 4-norms on the right hand side of Proposition II.1.3 are with respect to the uniform measure $\mu_{1/2}$, where we can apply standard hypercontractivity (the 'Beckner-Bonami Lemma') for $\rho\le 1/2\sqrt{3}$ to obtain $\|\mathrm{T}_{2\rho}((\mathrm{D}_Sf)_n)\|_4^4\le\|(\mathrm{D}_Sf)_n\|_2^4=\|\mathrm{D}_Sf\|_2^4=\sigma^{4|S|}\mathrm{I}_S(f)^2$. Recalling that $\lambda\le\sigma^{-2}$, we deduce the following bound for $\|\mathrm{T}_\rho f\|_4^4$ in terms of the generalised influences of f.

Theorem II.1.4. *If* $\rho\le 1/\sqrt{12}$ *then* $\|\mathrm{T}_\rho f\|_4^4\le\sum_{S\subset[n]}(3\lambda\rho^4)^{|S|}\|\mathrm{D}_Sf\|_2^4\le\sum_{S\subset[n]}(3\sigma^2\rho^4)^{|S|}\mathrm{I}_S(f)^2$.

Now we deduce our hypercontractivity inequality. It is convenient to prove a slightly stronger statement, which implies Theorem I.1.3 using $\|\mathrm{D}_Sf\|_2^2=\sigma^{2|S|}\mathrm{I}_S(f)\le\lambda^{-|S|}\mathrm{I}_S(f)$ and $\|\mathrm{T}_{1/5}f\|_4\le\|\mathrm{T}_{1/\sqrt{24}}f\|_4$ (any T_ρ is a contraction in L^p for any $p\ge 1$).

Theorem II.1.5. *Let* $f\in L^2(\{0,1\}^n,\mu_p)$ *with all* $\|\mathrm{D}_Sf\|_2^2\le\beta\lambda^{-|S|}\mathbb{E}[f^2]$. *Then* $\|\mathrm{T}_{1/\sqrt{24}}f\|_4\le\beta^{1/4}\|f\|_2$.

Proof. By Theorem II.1.4 applied to $\mathrm{T}_{1/\sqrt{2}}f$ with $\rho=1/\sqrt{12}$ we have

$$\|\mathrm{T}_{1/\sqrt{24}}f\|_4^4\le\sum_{S\subset[n]}(3\lambda\rho^4)^{|S|}\|\mathrm{D}_S\mathrm{T}_{1/\sqrt{2}}f\|_2^4.$$

As $\|\mathrm{D}_S\mathrm{T}_{1/\sqrt{2}}f\|_2^2=\sum_{E:S\subset E}2^{-|E|}\hat{f}(E)^2\le\sum_{E:S\subset E}\hat{f}(E)^2=\|\mathrm{D}_Sf\|_2^2\le\beta\lambda^{-|S|}\mathbb{E}[f^2]$ we deduce

$$\|\mathrm{T}_{1/\sqrt{24}}f\|_4^4\le\sum_{S\subset[n]}\sum_{E:S\subset E}\beta\mathbb{E}[f^2]2^{-|E|}\hat{f}(E)^2=\beta\mathbb{E}[f^2]\sum_E\hat{f}(E)^2=\beta\|f\|_2^4.$$ □

Hypercontractivity in practice

We will mostly use the following application of the hypercontractivity theorem.

Lemma II.1.6. *Let f be a function of degree r. Suppose that $\mathrm{I}_S(f) \le \delta$ for all $|S| \le r$. Then*

$$\|f\|_4 \le 5^{\frac{3r}{4}} \delta^{\frac{1}{4}} \|f\|_2^{0.5}.$$

The proof uses the following lemma, which is immediate from the Fourier expression in (2).

Lemma II.1.7. $\mathrm{I}_S(f^{\le r}) \le \mathrm{I}_S(f)$ *for all $S \subset [n]$ and $\mathrm{I}_S(f^{\le r}) = 0$ if $|S| > r$.*

Proof of Lemma II.1.6. Write $f = \mathrm{T}_{1/5}(h)$, where $h = \sum_{|T| \le r} 5^{|T|} \hat{f}(T) \chi_T$. We will bound the 4-norm of f by applying Theorem I.1.3 to h, so we need to bound the generalised influences of h.

By Lemma II.1.7, for $S \subset [n]$ we have $\mathrm{I}_S(h) = 0$ if $|S| > r$. For $|S| \le r$, we have

$$\mathrm{I}_S(h) = \sigma^{-2|S|} \sum_{T: S \subset T, |T| \le r} 5^{2|T|} \hat{f}(T)^2 \le 5^{2r} \mathrm{I}_S(f) \le 5^{2r}\delta = \alpha \|h\|_2^2,$$

where $\alpha = 5^{2r}\delta / \|h\|_2^2$. By Theorem I.1.3, we have

$$\|f\|_4 = \|\mathrm{T}_{1/5} h\|_4 \le \alpha^{\frac{1}{4}} \|h\|_2 = 5^{r/2} \delta^{\frac{1}{4}} \sqrt{\|h\|_2} \le 5^{\frac{3r}{4}} \delta^{\frac{1}{4}} \sqrt{\|f\|_2}.$$

In the final inequality we used $\|h\|_2 \le 5^r \|f\|_2$, which follows from Parseval. □

II.2 General hypercontractivity

In this section we generalise Theorem I.1.3 in two different directions. One direction is showing hypercontractivity from general q-norms to the 2-norm (rather than merely treating the case $q = 4$); the other is replacing the cube by general product spaces.

II.2.1 Hypercontractivity with general norms

We start by describing a more convenient general setting in which we replace characters on the cube by arbitrary random variables. To motivate this setting, we remark that one can extend the proof of Theorem II.1.4 to any random variable of the form

$$f = \sum_{S \subset [n]} a_S \prod_{i \in S} \mathbf{Z}_i, \tag{4}$$

where $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$ are independent real-valued random variables having expectation 0, variance 1 and 4th moment at most σ^{-2}. To motivate the analogous setting for general $q > 2$, we note that the characters χ_i^p satisfy

$$\mathbb{E}[|\chi_i^p|^q] \le \|\chi_i^p\|_\infty^{q-2} \|\chi_i^p\|_2^2 = \sigma^{2-q}.$$

This suggests replacing the 4th moment condition by $\|\mathbf{Z}_i\|_q^q \le \sigma^{2-q}$. Given f as in (4), we define the (generalised) derivatives by substituting the random variables Z_i for the characters χ_i^p in our earlier Fourier formulas, i.e.

$$\mathrm{D}_i[f] = \sum_{S:\, i \in S} a_S \prod_{j \in S \setminus \{i\}} \mathbf{Z}_i \quad \text{and} \quad \mathrm{D}_T(f) = \sum_{S:\, T \subset S} a_S \prod_{j \in S \setminus T} \mathbf{Z}_i,$$

Similarly, we adopt analogous definitions of the generalised influences and noise operator, i.e.

$$\mathrm{I}_S[f] = \|\frac{1}{\sigma} \mathrm{D}_S[f]\|_2^2 \quad \text{and} \quad \mathrm{T}_\rho[f] = \sum_S \rho^{|S|} a_S \prod_{i \in S} \mathbf{Z}_i.$$

We prove the following hypercontractive inequality.

Theorem II.2.1. *Let $q \geq 2$ and $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$ be independent real-valued random variables satisfying*

$$\mathbb{E}[\mathbf{Z}_i] = 0, \quad \mathbb{E}[\mathbf{Z}_i^2] = 1, \quad \textit{and} \quad \mathbb{E}[|\mathbf{Z}_i|^q] \leq \sigma^{2-q}.$$

Let $f = \sum_{S \subset [n]} a_S \prod_{i \in S} \mathbf{Z}_i$ and $\rho < \frac{1}{2q^{1.5}}$. Then

$$\|\mathrm{T}_\rho f\|_q^q \leq \sum_{S \subset [n]} \sigma^{(2-q)|S|} \|\mathrm{D}_S(f)\|_2^q.$$

Theorem II.2.1 is a qualitative generalisation of Theorem II.1.4 (with smaller ρ, which we do not attempt to optimise). The following generalised variant of Theorem I.1.3 follows by repeating the proof in Section II.1.

Theorem II.2.2. *Let $q > 2$, let $f = \sum_{S \subset [n]} a_S \prod_{i \in S} \mathbf{Z_i}$ let $\delta > 0$, and let $\rho \leq (2q)^{-1.5}$. Suppose that $\mathrm{I}_S[f] \leq \beta \|f\|_2^2$ for all $S \subset [n]$. Then*

$$\|\mathrm{T}_\rho[f]\|_q \leq \beta^{\frac{q-2}{2q}} \|f\|_2.$$

We now begin with the ingredients of the proof of Theorem II.2.1, following that of Theorem II.1.4. For $0 \leq t \leq n$ let

$$f_t = \sum_S a_S \chi_S^t, \quad \text{where } \chi_S^t = \prod_{i \in S \cap [t]} \chi_i^{1/2} \prod_{i \in S \setminus [t]} \mathbf{Z}_i.$$

Here, just as in Section II.1, the function f_t interpolates from the original function $f_0 = f$ to $f_n = \sum_S a_S \chi_S^{1/2} \in L^2(\{0,1\}^n, \mu_{1/2})$. As $\{\chi_S^t : S \subset [n]\}$ are orthonormal we have $\|f_t\|_2 = \|f\|_2$ for all t.

As before, we define the noise operators $\mathrm{T}^t_{\rho',\rho}$ on a function $f = \sum_S a_S \chi_S^t$ by

$$\mathrm{T}^t[f] = \sum_S \rho'^{|S \cap [t]|} \rho^{|S \setminus [t]|} a_S \chi_S^t.$$

Thus $\mathrm{T}^t_{\rho',\rho}$ interpolates from $\mathrm{T}^0_{\rho',\rho} = \mathrm{T}_\rho$ (for the original function) to $\mathrm{T}^n_{\rho',\rho} = \mathrm{T}_{\rho'}$ (for $\mu_{1/2}$).

Our goal will now be to adjust Lemma II.1.1 to the general setting, which is similar in spirit to the 4-norm case, although somewhat trickier. It turns out that the case $n = 1$ poses the main new difficulties, so we start with this in the next lemma.

Lemma II.2.3. *Let $q > 2$ and $\mathbf{Z}$ be a random variable satisfying $\mathbb{E}[\mathbf{Z}] = 0, \mathbb{E}[\mathbf{Z}^2] = 1, \mathbb{E}[|\mathbf{Z}|^q] \leq \sigma^{2-q}$. Let $e, d \in \mathbb{R}$ and $\rho \in (0, \frac{1}{2q})$. Then $\|e + \rho d \mathbf{Z}\|_q^q \leq \|e + d\chi^{\frac{1}{2}}\|_q^q + \sigma^{2-q} d^q$.*

Proof. If $e = 0$ then the lemma is trivial. Therefore we may rescale and assume that $e = 1$. It will be convenient to consider both sides of the inequality as functions of d: we write

$$f(d) = \|1 + \rho d \mathbf{Z}\|_q^q \quad \text{and} \quad g(d) = \|1 + d\chi^{\frac{1}{2}}\|_q^q + \sigma^{2-q} d.$$

As $f(0) = g(0)$, it suffices to show that $f'(0) = g'(0)$ and $f'' \leq g''$ everywhere.

Let us compute the derivatives. We note that the function $x \mapsto |x^q|$ has derivative $q|x|^{q-1}\mathrm{sign}(x)$, which is in turn continuously differentiable for $q > 2$. Thus

$$f' = \mathbb{E}[q\,|1 + \rho d \mathbf{Z}|^{q-1}\,\mathrm{sign}(1 + \rho d \mathbf{Z}) \rho \mathbf{Z}] = \rho q \mathbb{E}[|1 + \rho d \mathbf{Z}|^{q-1} \mathrm{sign}(1 + \rho d \mathbf{Z}) \mathbf{Z}] \quad \text{and}$$
$$f'' = (q-1) q \rho^2 \mathbb{E}[|1 + \rho d \mathbf{Z}|^{q-2} \mathbf{Z}^2].$$

Differentiating g we obtain

$$g' = q\mathbb{E}\left[\left|1 + d\chi^{\frac{1}{2}}\right|^{q-1} \mathrm{sign}(1 + d\chi^{\frac{1}{2}}) \chi^{\frac{1}{2}}\right] + q\sigma^{2-q} d^{q-1} \quad \text{and}$$
$$g'' = q(q-1)\mathbb{E}\left[\left|1 + d\chi^{\frac{1}{2}}\right|^{q-2} \left(\chi^{\frac{1}{2}}\right)^2\right] + q(q-1) d^{q-2} \sigma^{2-q} \geq q(q-1)/2 + q(q-1) d^{q-2} \sigma^{2-q}.$$

Thus $g'(0) = f'(0) = 0$ and it remains to show $f'' \le g''$ everywhere. Our strategy for bounding f'' is to decompose the expectation over two complementary events E_1 and E_2, where E_1 is the event that $|1 + \rho d\mathbf{Z}| \le |d\mathbf{Z}|$ (and E_2 is its complementary event). We write $f'' = f_1'' + f_2''$, where each

$$f_i'' = (q-1)q\rho^2 \mathbb{E}[|1 + \rho d\mathbf{Z}|^{q-2}\mathbf{Z}^2 \mathbf{1}_{E_i}].$$

First we note the bound

$$f_1'' \le q(q-1)\rho^2 d^{q-2}\mathbb{E}[|\mathbf{Z}|^q] \le q(q-1)d^{q-2}\sigma^{2-q}.$$

Given the above lower bound on g'', it remains to show $f_2'' \le q(q-1)/2$. On the event E_2 we have

$$|d\mathbf{Z}| \le |1 + \rho d\mathbf{Z}| \le 1 + |\rho d\mathbf{Z}|.$$

Rearranging, we obtain $|\rho d\mathbf{Z}|(\rho^{-1} - 1) \le 1$. Since $\rho^{-1} \ge 2q$, we get

$$1 + |\rho d\mathbf{Z}| \le 1 + \frac{1}{2q-1}.$$

Using $\mathbb{E}[\mathbf{Z}^2] = 1$ this yields

$$f_2'' \le q(q-1)\rho^2\Big(1 + \frac{1}{2q-1}\Big)^{q-2} \le e\rho^2 q(q-1) \le q(q-1)/2.$$

Hence $f'' = f_1'' + f_2'' \le g''$ for any value of d. This completes the proof of the lemma. □

We are now ready to show the replacement step.

Lemma II.2.4. $\mathbb{E}[(\mathrm{T}^{t-1}_{2q\rho,\rho} f_{t-1})^q] \le \mathbb{E}[(\mathrm{T}^t_{2q\rho,\rho} f_t)^q] + \sigma^{2-q}\mathbb{E}[(\mathrm{T}^t_{2q\rho,\rho}((\mathrm{D}_t f)_t))^q]$.

Proof. We write

$$\begin{aligned} f_t &= \chi_t^{1/2} g + h \quad \text{and} \quad f_{t-1} = \chi_t^p g + h, \quad \text{where} \\ g &= (\mathrm{D}_t f)_t = \sum_{S: t\in S} \hat{f}(S)\chi^t_{S\setminus\{t\}} = \sum_{S:t\in S} \hat{f}(S)\chi^{t-1}_{S\setminus\{t\}} = (\mathrm{D}_t f)_{t-1}, \quad \text{and} \\ h &= \mathbb{E}_{x_t\sim\mu_{1/2}} f_t = \sum_{S:t\notin S}\hat{f}(S)\chi^t_S = \sum_{S:t\notin S}\hat{f}(S)\chi^{t-1}_S = \mathbb{E}_{\mathbf{Z}_t} f_{t-1}. \end{aligned}$$

We also write

$$\begin{aligned} \mathrm{T}^t_{2q\rho,\rho} f_t = 2q\rho\chi_t^{1/2} d + e \quad &\text{and} \quad \mathrm{T}^{t-1}_{2q\rho,\rho} f_{t-1} = \rho\mathbf{Z}_t d + e, \quad \text{where} \\ d = \mathrm{T}^t_{2q\rho,\rho} g = \mathrm{T}^{t-1}_{2q\rho,\rho} g \quad &\text{and} \quad e = \mathrm{T}^t_{2q\rho,\rho} h = \mathrm{T}^{t-1}_{2q\rho,\rho} h. \end{aligned}$$

As before, we can calculate the expectations in the statement of the lemma by conditioning on all coordinates other than $\mathbf{Z}_t$ and $\chi_t^{\frac{1}{2}}$, so the lemma follows from Lemma II.2.3, with $2qd$ in place of d. □

From now on, everything is similar to Section II.1. We may apply the previous lemma inductively to obtain.

Lemma II.2.5. $\|\mathrm{T}^i_{2q\rho,\rho} f_i\|_q^q \le \sum_{S\subset[n]\setminus[i]} \sigma^{(2-q)|S|} \|\mathrm{T}^n_{2q\rho,\rho}((\mathrm{D}_S f)_n)\|_q^q$ *for all* $0 \le i \le n$.

In particular, recalling that $\mathrm{T}^0_{2q\rho,\rho} = \mathrm{T}_\rho$ on the original function and $\mathrm{T}^n_{2q\rho,\rho} = \mathrm{T}_{2q\rho}$ on $\mu_{1/2}$, the case $i = 0$ of Lemma II.2.5 is as follows.

Proposition II.2.6. $\|\mathrm{T}_\rho f\|_q^q \le \sum_{S\subset[n]} \sigma^{(2-q)|S|}\|\mathrm{T}_{2q\rho}((\mathrm{D}_S f)_n)\|_q^q$.

The q-norms on the right hand side of Proposition II.2.6 are with respect to the uniform measure $\mu_{1/2}$, where we can apply standard hypercontractivity with noise rate $\leq 1/\sqrt{q-1}$ to obtain

$$\|\mathrm{T}_{2q\rho}((\mathrm{D}_S f)_n)\|_q^q \leq \|(\mathrm{D}_S f)_n\|_2^q = \|\mathrm{D}_S f\|_2^q.$$

This completes the proof of Theorem II.2.1.

In the case where the $\mathbf{Z}_i$ have different qth moments, the proof can be adjusted to give a better upper bound. We write

$$\mathbb{E}[\mathbf{Z}_i^q] = \sigma_i^{2-q}, \quad \sigma_S = \prod_{i\in S}\sigma_i \quad \text{and} \quad \mathrm{I}_S[f] = \|\frac{1}{\sigma_S}\mathrm{D}_S[f]\|_2^2. \tag{5}$$

The proof of Theorem II.2.1 yields the following variant of Theorem II.1.4.

Theorem II.2.7. *Let $q \geq 2$, let $\rho \leq (2q)^{-1.5}$, and let $f = \sum a_S \prod_{i\in S}\mathbf{Z}_i$ with Z_i as in (5). Then*

$$\|\mathrm{T}_\rho f\|_q^q \leq \sum_{S\subset[n]} \sigma_S^{2-q}\|\mathrm{D}_S[f]\|_2^q.$$

The following variant of Theorem I.1.3 follows from Theorem II.2.7. The proof is similar to the one given in Section II.1, where Theorem I.1.3 is deduced from Theorem II.1.4.

Theorem II.2.8. *Let $q > 2$, $\beta > 0$ and $\rho \leq (2q)^{-1.5}$. Suppose $f = \sum_{S\subset[n]} a_S \prod_{i\in S}\mathbf{Z}_i$ with Z_i as in (5) has $\mathrm{I}_S[f] \leq \beta\|f\|_2^2$ for all $S \subset [n]$. Then*

$$\|\mathrm{T}_\rho f\|_q \leq \beta^{\frac{q-2}{2q}}\|f\|_2.$$

Finally, we state the following variant of Lemma II.1.6, which is easy to deduce from Theorem II.2.8.

Lemma II.2.9. *Let $q > 2$ and $\delta > 0$. Suppose $f = \sum_{S\subset[n]} a_S \prod_{i\in S}\mathbf{Z}_i$ with Z_i as in (5) has $\mathrm{I}_S[f] \leq \delta$ for all $|S| \leq r$. Then*

$$\|f\|_q \leq (2q)^{1.5r}\delta^{\frac{q-2}{2q}}\|f\|_2^{\frac{2}{q}}.$$

II.2.2 A hypercontractive inequality for product spaces

Now we consider the setting of a general discrete product space $(\Omega,\nu) = \prod_{t=1}^n(\Omega_t,\nu_t)$. We assume $p_t = \min_{\omega_t\in\Omega_t}\nu_t(\omega_t) \in (0,1/2)$ for each $t \in [n]$, and we write $p = \min_t p_t$. We recall the projections E_J on $L^2(\Omega,\nu)$ defined by $(\mathrm{E}_J f)(\omega) = \mathbb{E}_{\omega_J}[f(\omega) \mid \omega_{\overline{J}}]$, the generalised Laplacians L_S defined by composing L_t for all $t \in S$, where $\mathrm{L}_t f = f - \mathrm{E}_t f$, and the generalised influences $\mathrm{I}_S(f) = \mathbb{E}[\mathrm{L}_S(f)^2]\prod_{i\in S}\sigma_i^{-2}$, where $\sigma_i^2 = p_i(1-p_i)$.

We will require the theory of orthogonal decompositions in product spaces, which we summarise following the exposition in [77, Section 8.3]. For $f \in L^2(\Omega,\nu)$ and $J, S \subset [n]$ we write $f^{\subset J} = \mathrm{E}_{\overline{J}} f$ and define $f^{=S} = \sum_{J\subset S}(-1)^{|S\setminus J|}f^{\subset J}$ (inclusion-exclusion for $f^{\subset J} = \sum_{S\subset J} f^{=S}$). This decomposition is known as the Efron–Stein decomposition [18]. The key properties of $f^{=S}$ are that it only depends on coordinates in S and it is orthogonal to any function that depends only on some set of coordinates not containing S; in particular, $f^{=S}$ and $f^{=S'}$ are orthogonal for $S \neq S'$. We note that $f = f^{\subset[n]} = \sum_S f^{=S}$. We have similar Plancherel / Parseval relations as for Fourier decompositions, namely $\langle f, g\rangle = \sum_S f^{=S}g^{=S}$, so $\mathbb{E}[f^2] = \sum_S (f^{=S})^2$.

Our goal in this section is to prove an hypercontractive inequality for the Efron–Stein decomposition in the spirit of Theorem II.1.4. The noise operator is defined by $\mathrm{T}_\rho[f] = \sum_{S\subset[n]}\rho^{|S|}f^{=S}$. It also has a combinatorial interpretation, which is similar to the usual one on the p-biased setting. Given $x \in \Omega$, a sample $\mathbf{y} \sim N_\rho(x)$ is chosen by independently setting y_i to x_i with probability ρ and resampling it from (Ω_i,ν_i) with probability $1-\rho$. In the general product space setting there are no good analogs to $\mathrm{D}_i[f]$ and $\mathrm{D}_S(f)$, and we instead work with the Laplacians, which have similar Fourier formulas: $\mathrm{L}_i[f] = \sum_{S\colon i\in S} f^{=S}$, and $\mathrm{L}_T[f] = \sum_{S\colon T\subset S} f^{=S}$. In the special case where $\Omega_i = \{0,1\}$ we have $\|\mathrm{L}_S[f]\|_2 = \|\mathrm{D}_S[f]\|_2$. It will be convenient to write $\sigma_S = \prod_{i\in S}\sigma_i$.

The main result of this section is the following theorem.

Theorem II.2.10. *Let $f \in L^2(\Omega,\nu)$, let $q > 2$ be an even integer, and let $\rho \le \frac{1}{8q^{1.5}}$. Then*

$$\|\mathrm{T}_\rho f\|_q^q \le \sum_{S\subset[n]} \sigma_S^{2-q}\|\mathrm{L}_S[f]\|_2^q.$$

The idea of the proof is as follows. We encode our function $f \in L^2(\Omega,\nu)$ as a function $\tilde{f} := \sum_S \|f^{=S}\|_2\chi_S$ for appropriate $\chi_S = \prod_{i\in S}\chi_i$ (in fact, these will be biased characters on the cube). We then bound $\|\mathrm{T}_\rho f\|_q$ by $\|\mathrm{T}_\rho \tilde{f}\|_q$ and use Theorem II.2.8 to bound the latter norm.

The main technical component of the theorem is the following proposition.

Proposition II.2.11. *Let $g \in L^2(\Omega,\nu)$ let $\chi_S = \prod_{i\in S}\chi_i$, where χ_i are independent random variables having expectation 0, variance 1, and satisfying $\mathbb{E}[\chi_S^j] \ge \sigma_S^{2-j}$ for each integer $j \in (2,q]$. Let $\tilde{g} = \sum_{S\subset[n]}\|g^{=S}\|_2\chi_S$. Then*

$$\|g\|_q \le \|\tilde{g}\|_q.$$

Below, we fix χ_S as in the proposition, and let $\tilde{\circ}$ denote the operator mapping a function $g \in L^2(\Omega,\nu)$ to the function $\sum_{S\subset[n]} g^{=S}\chi_S$.

To prove the proposition, we will expand out $\|g\|_q^q$ and $\|\tilde{g}\|_q^q$ according to their definitions and compare similar terms: namely, we show that a term of the form $\mathbb{E}[\prod_{i=1}^q g^{=S_i}]$ is bounded by the corresponding term in $\|\tilde{g}\|_q^q$, i.e. $\prod_{i=1}^q \|g^{=S_i}\|_2\mathbb{E}[\prod_{i=1}^q \chi_{S_i}]$. We now establish such a bound.

We begin with identifying cases in which both terms are equal to 0, and for that we use the orthogonality of the decomposition $\{g^{=S}\}_{S\subset[n]}$. Afterwards, we only rely on the fact that $g^{=S}$ depends only on the coordinates in S.

Lemma II.2.12. *Let q be some integer, let $g \in L^2(\Omega,\nu)$, and let $S_1,\dots,S_q \subset [n]$ be some sets. Suppose that some $j \in [n]$ belongs to exactly one of the sets $S_1,\dots,S_q$. Then*

$$\mathbb{E}\left[\prod_{i=1}^q g^{=S_i}\right] = 0 \quad \textit{and} \quad \mathbb{E}\left[\prod_{i=1}^q \chi_{S_i}\right] = 0.$$

Proof. Assume without loss of generality that $j \in S_1$. The second equality $\mathbb{E}\left[\prod_{i=1}^q \chi_{S_i}\right] = 0$ follows by taking expectation over χ_j, using the independence between the random variables χ_i. For the first equality, observe that the function $\prod_{i=2}^q g^{=S_i}$ depends only on coordinates in $S_2 \cup \cdots, S_q \subset [n]\setminus\{j\}$. Hence the properties of the Efron–Stein decomposition imply

$$0 = \left\langle g^{=S_1}, \prod_{i=2}^q g^{=S_i}\right\rangle = \mathbb{E}\left[\prod_{i=1}^q g^{=S_i}\right]. \qquad \square$$

Thus we only need to consider terms corresponding to $S_1,\dots,S_q$ in which each coordinate appears in at least two sets. To facilitate our inductive proof we work with general functions f_i that depend only on coordinates of S_i (rather than only with the functions of the form $g^{=S_i}$).

Lemma II.2.13. *Let $f_1,\dots,f_q \in L^2(\Omega,\nu)$ be functions that depend on sets $S_1,\dots,S_q$ respectively. Let T_i for $i = 3,\dots,q$ be the set of coordinates covered by the sets $S_1,\dots,S_q$ exactly i times. Then*

$$\left|\mathbb{E}\left[\prod_{i=1}^q f_i\right]\right| \le \prod_{i=1}^q \|f_i\|_2 \cdot \prod_{j=3}^q \sigma_{T_j}^{2-j}.$$

Proof. The proof is by induction on n, simultaneously for all functions. We start with the case $n = 1$, which we prove by reducing to the case that all f_i are equal.

The case $n = 1$.

Here each f_i either depends on a single input or is constant and depends only on the empty set. We may assume that none of the f_i's is constant, as otherwise we may eliminate it from the inequality by dividing by $|f_i|$. By the generalised Hölder inequality we have

$$\left| \mathbb{E}\left[\prod_{i=1}^{q} f_i\right] \right| \le \prod_{i=1}^{q} \|f_i\|_q.$$

Hence the case $n = 1$ of the lemma will follow once we prove it assuming all the f_i are equal.

The $n = 1$ case with equal f_i's

We show that if (Ω, ν) is a discrete probability space in which any atom has probability at least p, then $\|f\|_q^q \le \|f\|_2^q \sigma^{2-q}$, where $\sigma = \sqrt{p(1-p)}$.

While the inequality $\|f\|_2 \le \|f\|_q$ holds in any probability space, the reverse inequality holds in any measure space where each atom has measure at least 1. Accordingly, we consider the measure $\tilde{\nu}$ on Ω defined by $\tilde{\nu}(x) = \nu(x)p^{-1}$. Then

$$\|f\|_{q,\nu}^q = p\|f\|_{q,\tilde{\nu}}^q \le p\|f\|_{2,\tilde{\nu}}^q = p^{1-\frac{q}{2}}\|f\|_{2,\nu}^q \le \sigma^{2-q}\|f\|_{2,\nu}^q.$$

This completes the proof of the $n = 1$ case.

The inductive step

Let $f_1, \ldots, f_q \in L^2(\Omega, \nu)$ be functions. Let $\mathbf{x} \sim \prod_{i=1}^{n-1}(\Omega_i, \nu_i)$. By the $n = 1$ case we have:

$$\left| \mathbb{E}\left[\prod_{i=1}^{q} f_i\right] \right| = \left| \mathbb{E}_{\mathbf{x}}\left[\mathbb{E}\left[\prod_{i=1}^{q} (f_i)_{[n-1]\to\mathbf{x}}\right]\right] \right| \le \mathbb{E}_{\mathbf{x}}\left[\prod_{i=1}^{q} \|(f_i)_{[n-1]\to\mathbf{x}}\|_2 \sigma_n^j\right],$$

where j is $2 - i$ if $n \in T_i$ for $i \ge 3$, and otherwise 0. The lemma now follows by applying the inductive hypothesis on the functions $\mathbf{x} \to \|(f_i)_{[n-1]\to\mathbf{x}}\|$ and using $\left\| \left\| (f_i)_{[n-1]\to\mathbf{x}} \right\|_2 \right\|_{2,\mathbf{x}} = \|f_i\|_2$. □

Proof of Proposition II.2.11. We wish to upper bound

$$\mathbb{E}[g^q] = \sum_{S_1,\ldots,S_q} \mathbb{E}\left[\prod_{i=1}^{q} g^{=S_i}\right]$$

by

$$\sum_{S_1,\ldots,S_q} \mathbb{E}\left[\prod_{i=1}^{q} \chi_{S_i}\right] \prod_{i=1}^{q} \|g^{=S_i}\|_2.$$

We upper bound each term participating in the expansion of g^q by the corresponding term in $\tilde{g}^q$. In the case the sets S_i cover some element exactly once, Lemma II.2.12 implies that both terms are 0. Otherwise, the sets S_i cover each element either 0 times or at least 2 times; let T_i be the set of elements of $S_1, \ldots, S_q$ appearing in exactly i of the sets (as in Lemma II.2.13). By the assumption of the proposition, we have $\mathbb{E}\left[\prod_{i=1}^{q} \chi_{S_i}\right] \ge \prod_{i=3}^{q} \sigma_{T_i}^{2-|T_i|}$. The proof is concluded by combining this with the upper bound on $\mathbb{E}\left[\prod_{i=1}^{q} g^{=S_i}\right]$ following from Lemma II.2.13 with $f_i = g^{=S_i}$. □

Proof of Theorem II.2.10. Let $\sigma_i' = \sqrt{p_i/4(1 - p_i/4)}$. We choose χ_i to be the $\frac{p_i}{4}$-biased character, $\chi_i = \frac{x_i - p_i/4}{\sigma_i'}$. Clearly χ_i has mean 0 and variance 1, and a direct computation shows that $\mathbb{E}\left[\chi_i^j\right] \ge (\sigma_i)^{2-j}$ for all integer $j > 2$, hence all of the conditions of Proposition II.2.11 hold.

Denote $\sigma'_S = \prod_{i \in S} \sigma'_i$ and set $h = T_{\frac{1}{4}} f$, $g = \mathrm{T}_{\frac{1}{2q^{1.5}}} h$. By Proposition II.2.11 and Theorem II.2.7 we have

$$\|\mathrm{T}_{\frac{1}{8q^{1.5}}} f\|_q^q = \|g\|_q^q \le \|\tilde{g}\|_q^q \le \sum_S (\sigma'_S)^{2-q} \|\mathrm{D}_S[\tilde{h}]\|_2.$$

We note that by Parseval, the 2-norm of $\tilde{h}$ and its derivatives are equal to the 2-norm of h and its Laplacians, and thus the last sum is equal to

$$\sum_S (\sigma'_S)^{2-q} \|\mathrm{L}_S[h]\|_2^q \le \sum_S (\sigma_S)^{2-q} \|\mathrm{L}_S[f]\|_2^q.$$

In the last inequality we used $\sigma'_S \ge 2^{-|S|}\sigma_S$ and $\|\mathrm{L}_S[h]\|^q \le 2^{-q|S|}\|\mathrm{L}_S[f]\|_2^q$ (which follows from Parseval). This completes the proof of the theorem. □

II.3 An invariance principle for global functions

Invariance (also known as Universality) is a fundamental paradigm in Probability, describing the phenomenon that many random processes converge to a specific distribution that is the same for many different instances of the process. The prototypical example is the Berry-Esseen Theorem, giving a quantitative version of the Central Limit Theorem (see e.g. [77, Section 11.5]). More sophisticated instances of the phenomenon that have been particularly influential on recent research in several areas of Mathematics include the universality of Wigner's semicircle law for random matrices (see [71]) and of Schramm–Loewner evolution (SLE) e.g. in critical percolation (see [82]).

In the context of the cube, the Invariance Principle is a powerful tool developed by Mossel, O'Donnell and Oleszkiewicz [75] while proving their 'Majority is Stablest' Theorem, which can be viewed as an isoperimetric theorem for the noise operator. Roughly speaking, the result (in a more general form due to Mossel [73]) is that 'majority functions' (characteristic functions of Hamming balls) minimise noise sensitivity among functions that are 'far from being dictators'. The Invariance Principle converts many problems on the cube to equivalent problems in Gaussian Space; in particular, 'Majority is Stablest' is converted into an isoperimetric problem in Gaussian Space which was solved by a classical theorem of Borell [12] (half-spaces are isoperimetric).

In this section we will establish an invariance principle for global functions that has several applications analogous to those of the classical invariance principle, such as the following variant of 'majority is stablest'. We define the *p-biased α-Hamming ball* on $\{0,1\}^n$ as the function H_α whose value is 1 on an input x if and only if x has at least t coordinates equal to 1, and t is chosen so that $\mu_p(H_\alpha)$ is as close to α as possible.

Corollary II.3.1. *For each $\epsilon > 0$, there exists $\delta > 0$, such that the following holds. Let $\rho \in (\epsilon, 1-\epsilon)$, let $n > \delta^{-1}$, and let $f, g \in L^2(\{0,1\}^n, \mu_p)$. Suppose that $\mathrm{I}_S[f] \le \delta$ and that $\mathrm{I}_S[g] \le \delta$ for each set S of at most δ^{-1} coordinates. Then*

$$\langle \mathrm{T}_\rho f, g \rangle \le \left\langle \mathrm{T}_\rho H_{\mu_p(f)}, H_{\mu_p(g)} \right\rangle + \epsilon.$$

We omit the proof of this result as it is very similar to that in [73]. For the sake of brevity we also omit discussion of other applications of our invariance principle, including a sharp threshold for almost monotone Boolean functions, which is analogous to results of Lifshitz [65].

In the basic form (see [77, Section 11.6]) of the Invariance Principle, we consider a multilinear real-valued polynomial f of degree $\le k$ and wish to compare $f(\boldsymbol{x})$ to $f(\boldsymbol{y})$, where $\boldsymbol{x}$ and $\boldsymbol{y}$ are random vectors each having independent coordinates, according a smooth (to third order) test function ϕ. (Comparison of the cumulative distributions requires ϕ to be a step function, but this can be handled by smooth approximation.) The version of [77, Remark 11.66] shows that if the coordinates x_i have mean 0, variance 1 and are suitably hypercontractive (satisfy $\|a+\rho b x_i\|_3 \le \|a+bx_i\|_2$ for any $a, b \in \mathbb{R}$), and similarly for y_i, then

$$\left|\mathbb{E}[\phi(f(\boldsymbol{x}))] - \mathbb{E}[\phi(f(\boldsymbol{y}))]\right| \le \tfrac{1}{3}\|\phi'''\|_\infty \rho^{-3k} \sum_{i\in[n]} \mathrm{I}_i(f)^{3/2}. \tag{6}$$

The hypercontractivity assumption applies e.g. if the coordinates are standard Gaussians or p-biased bits (renormalised to have mean 0 and variance 1) with p bounded away from 0 or 1, but if $p = o(1)$ then we need $\rho = o(1)$, in which case their theorem becomes ineffective. We will apply our hypercontractivity inequality to obtain an invariance principle that is effective for small probabilities and functions with small generalised influences. We adopt the following setup.

Setup II.3.2. Let $\sigma_1, \dots, \sigma_n > 0$, let $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$ and $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_n)$ be random vectors with independent coordinates, where each X_i and Y_i are real-valued random variable with mean 0, variance 1, and satisfy $\|X_i\|_3^3 \le \sigma_i^{-1}$ and $\|Y_i\|_3^3 \le \sigma_i^{-1}$. Let $f \in \mathbb{R}[v]$ be a multilinear polynomial of degree d in n variables $v = (v_1, \dots, v_n)$. Let $\phi : \mathbb{R} \to \mathbb{R}$ be smooth.

For $S \subset [n]$ we write $\hat{f}(S)$ for the coefficient in f of $v_S = \prod_{i\in S} v_i$. We write $W_S(f) = \sum_{J:S\subset J} \hat{f}(J)^2$ and similarly to Section II.2.1 we define the generalised influences by $\mathrm{I}_S(f) = W_S(f)\prod_{i\in S}\sigma_i^{-2}$.

We write $\mathrm{T}_\rho[f] = \sum_{S\subset[n]} \rho^{|S|}\hat{f}(S)v_S$.

Now we state our invariance principle, which compares $f(\mathbf{X})$ to $f(\mathbf{Y})$.

Theorem II.3.3. *Under Setup II.3.2, if $\mathrm{I}_S[f] \le \epsilon$ for each nonempty set S, then*

$$|\mathbb{E}[\phi(f(\mathbf{X}))] - \mathbb{E}[\phi(f(\mathbf{Y}))]| \le 2^{5d}\|\phi'''\|_\infty W_\emptyset(f)\sqrt{\epsilon}.$$

The term $W_\emptyset(f)$ can be replaced by either $\mathbb{E}[f(\mathbf{X})^2]$ or $\mathbb{E}[f(\mathbf{Y})^2]$ as they are all equal.

Theorem II.3.3 can be informally interpreted as saying that if a multilinear, low degree polynomial f is global, then the distribution of $f(\mathbf{X})$ does not really depend on the distribution of $\mathbf{X}$ except for the mean and variance of each coordinate.

In particular, it implies that plugging in the p-biased characters into f results in a fairly similar distribution to the one resulting from plugging in the uniform characters into f. A posteriori, this may be seen as an intuitive explanation for Theorem I.1.3, as the standard hypercontractivity theorem holds in the uniform cube.

Next, we set up some notations and preliminary observations for the proof of Theorem II.3.3. Throughout we fix $\mathbf{X}$, $\mathbf{Y}$, f, and ϕ as in Setup II.3.2. We write $\mathbf{X}_S = \prod_{i\in S}\mathbf{X}_i$, and similarly for $\mathbf{Y}$. Recall that $f = \sum_S \hat{f}(S)v_S$ is a (formal) multilinear polynomial in $\mathbb{R}[v]$ of degree d. Note that $f(\mathbf{X}) = \sum_S \hat{f}(S)\mathbf{X}_S$ has $\mathbb{E}[f(\mathbf{X})^2] = \sum_S \hat{f}(S)^2$, as $\mathbb{E}\mathbf{X}_S^2 = 1$ and $\mathbb{E}[\mathbf{X}_S\mathbf{X}_T] = 0$ for $S \ne T$. The random variable $f(\mathbf{X})$ has the orthogonal decomposition $f = \sum_S f^{=S}$ with each $f^{=S} = \hat{f}(S)\mathbf{X}_S$. Further note that $\mathrm{L}_S f(\mathbf{X}) = \sum_{J:S\subset J}\hat{f}(J)\mathbf{X}_J$ so we have the identities

$$\mathrm{I}_S(f)\prod_{i\in S}\sigma_i^2 = \mathbb{E}[(\mathrm{L}_S f(\mathbf{X}))^2] = \mathbb{E}[(\mathrm{L}_S f(\mathbf{Y}))^2] = \sum_{J:S\subset J}\hat{f}(J)^2 = W^{S^\uparrow}(f).$$

We apply the replacement method as in Section II.1 (and as in the proof of the original invariance principle by Mossel, O'Donnell and Oleszkiewicz [75]). For $0 \le t \le n$, define $\mathbf{Z}^{:t} = (\mathbf{Z}_1^{:t}, \dots, \mathbf{Z}_n^{:t}) = (\mathbf{Y}_1, ..., \mathbf{Y}_t, \mathbf{X}_{t+1}, ..., \mathbf{X}_n)$, and note that $f(\mathbf{Z}^{:t})$ has the orthogonal decomposition $f(\mathbf{Z}^{:t}) = \sum_S f(\mathbf{Z}^{:t})^{=S}$ with

$$f(\mathbf{Z}^{:t})^{=S} = \hat{f}(S)\mathbf{Z}_S = \hat{f}(S)\mathbf{Y}_{S\cap[t]}\mathbf{X}_{S\setminus[t]}.$$

Proof of Theorem II.3.3. We adapt the exposition in [77, Section 11.6]. As $\mathbf{Z}^{:0} = \mathbf{X}$ and $\mathbf{Z}^{:n} = \mathbf{Y}$ we have by telescoping and the triangle inequality

$$|\mathbb{E}[\phi(f(\mathbf{X}))] - \mathbb{E}[\phi(f(\mathbf{Y}))]| \leq \sum_{t=1}^{n} |\mathbb{E}[\phi(f(\mathbf{Z}^{:t-1}))] - \mathbb{E}[\phi(f(\mathbf{Z}^{:t}))]|.$$

Consider any $t \in [n]$ and write

$$f(\mathbf{Z}^{:t-1}) = U_t + \Delta_t \mathbf{Y}_t \quad \text{and} \quad f(\mathbf{Z}^{:t}) = U_t + \Delta_t \mathbf{X}_t, \quad \text{where}$$

$$U_t = \mathrm{E}_t f(\mathbf{Z}^{:t-1}) = \mathrm{E}_t f(\mathbf{Z}^{:t}) \quad \text{and} \quad \Delta_t = \mathrm{D}_t f(\mathbf{Z}^{:t-1}) = \mathrm{D}_t f(\mathbf{Z}^{:t}).$$

Both of the functions U_t and Δ_t are independent of the random variables X_t and Y_t.

By Taylor's Theorem,

$$\begin{aligned} \phi(f(\mathbf{Z}^{:t-1})) &= \phi(U_t) + \phi'(U_t)\Delta_t \mathbf{Y}_t + \tfrac{1}{2}\phi''(U_t)(\Delta_t \mathbf{Y}_t)^2 + \tfrac{1}{6}\phi'''(A)(\Delta_t \mathbf{Y}_t)^3, \quad \text{and} \\ \phi(f(\mathbf{Z}^{:t})) &= \phi(U_t) + \phi'(U_t)\Delta_t \mathbf{X}_t + \tfrac{1}{2}\phi''(U_t)(\Delta_t \mathbf{X}_t)^2 + \tfrac{1}{6}\phi'''(A')(\Delta_t \mathbf{X}_t)^3, \end{aligned}$$

for some random variables A and A'. As $\mathbf{X}_t$ and $\mathbf{Y}_t$ have mean 0 and variance 1 we have $0 = \mathbb{E}[\phi'(U_t)\Delta_t \mathbf{Y}_t] = \mathbb{E}[\phi'(U_t)\Delta_t \mathbf{X}_t]$ and $\mathbb{E}[\phi''(U_t)(\Delta_t)^2] = \mathbb{E}[\phi''(U_t)(\Delta_t \mathbf{Y}_t)^2] = \mathbb{E}[\phi''(U_t)(\Delta_t \mathbf{X}_t)^2]$, so

$$|\mathbb{E}[\phi(f(\mathbf{Z}^{:t-1}))] - \mathbb{E}[\phi(f(\mathbf{Z}^{:t}))]| \leq \tfrac{1}{6}\|\phi'''\|_\infty (\mathbb{E}[|\Delta_t \mathbf{X}_t|^3] + \mathbb{E}[|\Delta_t \mathbf{Z}_t|^3]) \leq \tfrac{1}{3}\|\phi'''\|_\infty \sigma_t^{-1} \|\Delta_t\|_3^3.$$

The function Δ_t is the function $\mathrm{D}_t[f]$ applied on random variables satisfying the hypothesis of Lemma II.2.9. Moreover, $\mathrm{I}_S[\mathrm{D}_t[f]]$ is either 0 when $t \in S$, or $\sigma_t^2 \mathrm{I}_{S\cup\{t\}}[f]$ when $t \notin S$, in which case $\mathrm{I}_S[f] \leq \sigma_t^2 \epsilon$. Hence, by Lemma II.2.9 (with $q = 3$), we obtain

$$\|\Delta_t\|_3^3 \leq 6^{4.5d} \sigma_t \sqrt{\epsilon} \|\Delta_t\|_2^2 = 6^{4.5d} \sigma_t \sqrt{\epsilon} \cdot \sum_{S \ni t} \hat{f}(S)^2.$$

Hence,

$$\sum_{t=0}^{n} \tfrac{1}{3}\|\phi'''\|_\infty \sigma_t^{-1} \|\Delta_t\|_3^3 \leq 6^{4.5d} \sqrt{\epsilon} \tfrac{1}{3} \|\phi'''\|_\infty \sum_{S} |S| \hat{f}(S)^2 \leq 6^{4.5d} \sqrt{\epsilon} \tfrac{d}{3} \|\phi'''\|_\infty W_\emptyset(f).$$

This completes the proof of the theorem since $6^{4.5d}\frac{d}{3} \leq 2^{12d}$. $\square$

Part III

Sharp thresholds

In this part we apply the hypercontractivity results of the previous part to obtain new results in the theory of sharp thresholds. To prepare for the analysis, we start in Section III.1 by establishing the equivalence between the two notions of globalness introduced earlier, namely control of generalised influences and insensitivity of the measure under restriction to a small set of coordinates. Section III.2 concerns the total influence of global functions, and includes the proofs of our stability results for the isoperimetric inequality (Theorems I.1.4 and I.1.5) and our first sharp threshold result (Theorem I.2.1). In Section III.3 we prove our result on noise sensitivity and apply this to deduce an alternative sharp threshold result.

III.1 Characterising global functions

Above we have introduced two notions of what it means for a Boolean function f to be global. The first globalness condition, which appears e.g. in Theorem I.1.4, is that the measure of f is not sensitive to restrictions to small sets of coordinates. The second condition is a bound on generalised influences $I_S(f)$ for small sets S. In this section we show that we can move freely between these notions for two classes of Boolean functions: namely sparse ones and monotone ones.

Throughout we assume $p \leq 1/2$, which does not involve any loss in generality in our main results; indeed, if $p > 1/2$ we can consider the dual $f^*(x) = 1 - f(1-x)$ of any Boolean function f, for which $\mu_{1-p}(f^*) = 1 - \mu_p(f)$ and $\mathrm{I}_{\mu_{1-p}}(f^*) = \mathrm{I}_{\mu_p}(f)$.

We start by formalising our first notion of globalness.

Definition III.1.1. We say that a Boolean function f is (r,δ)-*global* if $\mu_p\left(f_{J\to 1}\right) \leq \mu_p\left(f\right) + \delta$ for each set J of size at most r.

We remark that Definition III.1.1 is a rather weak notion of globalness, so it is quite surprising that it suffices for Theorems I.1.5 and I.3.2, where one might have expected to need the stricter notion that $\mu_p(f_{J\to 1})$ is close to $\mu_p(f)$.

The following lemma shows that if a sparse Boolean function is global in the sense of Definition III.1.1 then it has small generalised influences.

Lemma III.1.2. *Suppose that $f : \{0,1\}^n \to \{0,1\}$ is an (r,δ)-global Boolean function with $\mu_p(f) \leq \delta$. Then $\mathrm{I}_S\left(f^{\leq r}\right) \leq \mathrm{I}_S\left(f\right) \leq 8^r\delta$ for all $S \subset [n]$ with $|S| \leq r$.*

Proof. The first inequality is from Lemma II.1.7. Next, we estimate

$$\sqrt{\mathrm{I}_S\left(f\right)} = \left\| \sum_{x\in\{0,1\}^S} (-1)^{|S|-|x|}\, f_{S\to x} \right\|_2 \leq \sum_{x\in\{0,1\}^S} \|f_{S\to x}\|_2 = \sum_{x\in\{0,1\}^S} \sqrt{\mu_p(f_{S\to x})}. \tag{7}$$

Next we fix $x \in \{0,1\}^S$ and claim that $\mu_p(f_{S\to x}) \leq 2^r\delta$. By substituting this bound in (7) we see that this suffices to complete the proof. Let T be the set of all $i \in S$ such that $x_i = 1$. Since f is nonnegative, we have $\mu_p(f_{T\to 1}) \geq (1-p)^{|S\setminus T|}\,\mu_p(f_{S\to x})$. As f is (r,δ)-global and $\mu_p(f) \leq \delta$, we have $\mu_p\left(f_{T\to 1}\right) \leq 2\delta$, so $\mu_p(f_{S\to x}) \leq (1-p)^{|T|-r} 2\delta \leq 2^r\delta$, where for the last inequality we can assume $T \neq \emptyset$, as $\mu_p\left(f_{T\to 1}\right) = \mu_p(f) \leq \delta \leq 2^r\delta$. This completes the proof. □

Next we show an analogue of the previous lemma replacing the assumption that f is sparse by the assumption that f is monotone.

Lemma III.1.3. *Let $f\colon \{0,1\}^n \to \{0,1\}$ be a monotone Boolean (r,δ)-global function. Then $\mathrm{I}_S\left(f\right) \leq 8^r\delta$ for every nonempty S of size at most r.*

The proof is based on the following lemma showing that globalness is inherited (with weaker parameters) under restriction of a coordinate.

Lemma III.1.4. *Suppose that f is a monotone (r,δ)-global function. Then for each i:*

1. *$f_{i\to 1}$ is $(r-1,\delta)$-global,*
2. *$\mu_p\left(f_{i\to 0}\right) \geq \mu_p\left(f\right) - \frac{p\delta}{1-p}$,*
3. *$f_{i\to 0}$ is $\left(r-1,\frac{\delta}{1-p}\right)$-global.*

Proof. To see (1), note that for any J with $|J|\leq r-1$ we have $\mu_p((f_{i\to 1})_{J\to 1}) = \mu_p(f_{J\cup\{i\}\to 1}) \leq \mu_p(f)+\delta \leq \mu_p(f_{i\to 1})+\delta$, where the last inequality holds as f is monotone. Statement (2) follows from the upper bound $\mu_p\left(f_{i\to 1}\right)\leq \mu_p\left(f\right)+\delta$ and $\mu_p\left(f_{i\to 0}\right) = \frac{\mu_p(f)-p\mu_p(f_{i\to 1})}{(1-p)}$.

For (3), we note that by monotonicity $\mu_p\left((f_{i\to 0})_{S\to 1}\right)\leq \mu_p\left(f_{\{i\}\cup S\to 1}\right)$. As f is (r,δ)-global,

$$\mu_p\left(f_{S\cup\{i\}\to 1}\right) \leq \mu_p\left(f\right)+\delta \leq \mu_p\left(f_{i\to 0}\right)+\delta+\frac{p\delta}{1-p} = \mu_p\left(f_{i\to 0}\right)+\frac{\delta}{1-p},$$

using (2). Hence, $f_{i\to 0}$ is $\left(r,\frac{\delta}{1-p}\right)$-global. ∎

Proof of Lemma III.1.3. We argue by induction on r. In the case where $r=1$, Lemma III.1.4 and monotonicity of f imply (using $p\leq 1/2$)

$$\mathrm{I}_i\left(f\right) = \mu_p\left(f_{i\to 1}\right)-\mu_p\left(f_{i\to 0}\right) \leq \delta+\frac{p\delta}{1-p}\leq 2\delta.$$

Now we bound $\mathrm{I}_{S\cup\{i\}}\left(f\right)$ for $r>1$ and S of size $r-1$ with $i\notin S$.

Note that $\mathrm{D}_{S\cup\{i\}}\left(f\right) = \mathrm{D}_S\left[\mathrm{D}_i(f)\right]$. By the triangle inequality, we have

$$\sqrt{\mathrm{I}_{S\cup\{i\}}\left(f\right)} = \sigma^{-r}\|\mathrm{D}_{S\cup\{i\}}(f)\|_2 = \sigma^{1-r}\|\mathrm{D}_S(f_{i\to 1})-\mathrm{D}_S(f_{i\to 0})\|_2 \leq \sqrt{\mathrm{I}_S\left(f_{i\to 1}\right)}+\sqrt{\mathrm{I}_S(f_{i\to 0})}.$$

By the induction hypothesis and Lemma III.1.4 the right hand side is at most

$$\sqrt{8^{r-1}\delta}+\sqrt{8^{r-1}2\delta}\leq \sqrt{8^r\delta}.$$

Taking squares, we obtain $\mathrm{I}_{S\cup\{i\}}\left(f\right)\leq 8^r\delta$. ∎

We conclude this section by showing the converse direction of the equivalence between our two notions of globalness, i.e. that if the generalised influences of a function f are small then f is global in the sense of its measure being insensitive to restrictions to small sets. (We will not use the lemma in the sequel but include the proof for completeness.)

Lemma III.1.5. *Let $f\colon\{0,1\}^n\to\{0,1\}$ be a Boolean function and let $r>0$. Suppose that $\mathrm{I}_S[f]\leq\delta$ for each nonempty set S of at most r coordinates. Then f is $(r,4^r\delta)$-global.*

Proof. To facilitate a proof by induction on r we prove the slightly stronger statement that f is $(r,\sum_{i=1}^r 4^{i-1}\delta)$-global. Suppose first that $r=1$. Our goal is to show that if $\mathrm{I}_i[f]<\delta$, then $\mu_p(f_{i\to 1})-\mu_p(f_{i\to 0})<\delta$, and indeed,

$$\mu_p(f_{i\to 1})-\mu_p(f_{i\to 0})\leq \Pr[f_{i\to 1}\neq f_{i\to 0}] = \|f_{i\to 1}-f_{i\to 0}\|_2^2 = \|\mathrm{D}_i[f]\|_2^2 = \mathrm{I}_i[f]<\delta.$$

Now suppose that $r>1$ and that the lemma holds with $r-1$ in place of r. The lemma will follow once we show that for all i and all nonempty sets S of size at most $r-1$, we have $\mathrm{I}_S[f_{i\to 1}]\leq 4\delta$. Indeed, the induction hypothesis and the $n = 1$ case will imply that for each set S of size at most r and each $i \in S$ we have $\mu_p(f_{S\to 1}) \le \mu_p(f_{i\to 1}) + \sum_{j=1}^{r-1} 4^{j-1} \cdot 4\delta \le \mu_p(f) + \sum_{j=1}^{r} 4^{j-1}\delta$.

We now turn to showing the desired upper bound on the generalised influences of $f_{i\to 1}$. Let S be a set of size at most $r-1$. Recall that $\mathrm{I}_S[f_{i\to 1}] = \|\mathrm{D}_S[f_{i\to 1}]\|_2^2$. We may assume that $i \notin S$ for otherwise the generalised influence $\mathrm{I}_S[f_{i\to 1}]$ is 0. We make two observations. Firstly, we have

$$\mathrm{D}_{S\cup\{i\}}[f] = \mathrm{D}_S[f_{i\to 1}] - \mathrm{D}_S[f_{i\to 0}].$$

Secondly, conditioning on the ouput of the coordinate i we have

$$\|\mathrm{D}_S[f]\|_2^2 = p\|\mathrm{D}_S[f_{i\to 1}]\|_2^2 + (1-p)\|\mathrm{D}_S[f_{i\to 0}]\|_2^2,$$

which implies $\|\mathrm{D}_S[f_{i\to 0}]\|_2 \le \sqrt{2}\|\mathrm{D}_S[f]\|_2$. We may now apply the triangle inequality on the first observation and use the second observation to obtain

$$\sqrt{\mathrm{I}_S[f]} = \|\mathrm{D}_S[f_{i\to 1}]\|_2 \le \|\mathrm{D}_{S\cup\{i\}}[f]\|_2 + \|\mathrm{D}_S[f_{i\to 0}]\|_2 \le \sqrt{\delta} + \sqrt{2}\|\mathrm{D}_S[f]\|_2 \le 2\sqrt{\delta}.$$

Taking squares, we obtain the desired upper bound on the generalised influences of $f_{i\to 1}$. □

III.2 Total influence of global functions

In this section we show that our hypercontractive inequality (Theorem I.1.3) implies our stability results for the isoperimetric inequality, namely Theorems I.1.4 and I.1.5. We also deduce our first sharp threshold result, Theorem I.2.1.

III.2.1 The spectrum of sparse global sets

The key step in the proofs of Theorems I.1.5 and I.3.2 is to show that the Fourier spectrum of global sparse subsets of the p-biased cube is concentrated on the high degrees. We recall first a proof that in the uniform cube (i.e. cube with uniform measure), *all* sparse sets have this behaviour (not just the global ones). Our proof is based on ideas from Talagrand [83] and Bourgain and Kalai [13].

Theorem III.2.1. *Let f be a Boolean function on the uniform cube, and let $r > 0$. Then*

$$\left\|f^{\le r}\right\|_2^2 \le 3^r \mu_{1/2}(f)^{1.5}.$$

The idea of the proof is to bound $\left\|f^{\le r}\right\|_2^2 = \left\langle f^{\le r}, f\right\rangle$ via Hölder by $\left\|f^{\le r}\right\|_4 \|f\|_{4/3}$, bound the 4-norm via hypercontractivity and express the 4/3-norm in terms of the measure of f using the assumption that f is Boolean. For future reference, we decompose the argument into two lemmas, the first of which applies also to the p-biased settting and the second of which requires hypercontractivity, and so is specific to the uniform setting. Theorem III.2.1 follows immediately from Lemmas III.2.2 and III.2.3 below.

In the following lemma we consider $\{-1, 0, 1\}$-valued functions so that it can be applied to either a Boolean function or its discrete derivative.

Lemma III.2.2. *Let $f\colon \{0,1\}^n \to \{0, 1, -1\}$, let $\mathcal{F}$ be a family of subsets of $[n]$, and let $g(x) = f^{\mathcal{F}} = \sum_{S\in\mathcal{F}} \hat{f}(S)\chi_S(x)$. Then $\|g\|_2^2 \le \|g\|_4 \|f\|_2^{1.5}$, where the norms can be taken with respect to an arbitrary p-biased measure.*

Proof. By Plancherel and Hölder's inequality, $\mathbb{E}[g^2] = \langle f, g\rangle \le \|f\|_{4/3}\|g\|_4$, where $\|f\|_{4/3} = \mathbb{E}[f^2]^{3/4} = \|f\|_2^{1.5}$ as f is $\{-1, 0, 1\}$-valued. □

Applying Lemma III.2.2 with $g = f^{\le r}$, we obtain a lower bound on the 4-norm of g. We now upper bound it by appealing to the Hypercontractivity Theorem.

Lemma III.2.3. *Let g be a function of degree r on the uniform cube. Then $\|g\|_4 \leq \sqrt{3}^r \|g\|_2$.*

Proof. Let h be the function, such that $\mathrm{T}_{1/\sqrt{3}}h = g$, i.e. $h = \sum_{|S|\leq r} \sqrt{3}^{|S|} \hat{g}(S)\chi_S$. Then the Hypercontractivity Theorem implies that $\|g\|_4 \leq \|h\|_2$, and by Parseval $\|h\|_2 \leq \sqrt{3}^r \|g\|_2$. □

We shall now adapt the proof of Theorem III.2.1 to global functions on the p-biased cube. The only part in the above proof that needs an adjustment is Lemma III.2.3, and in fact we have already provided the required adjustment in Section II.1 in the form of Lemma II.1.6.

Theorem III.2.4. *Let $r \geq 1$, and let $f\colon \{0,1\}^n \to \{0,1,-1\}$. Suppose that $\mathrm{I}_S[f^{\leq r}] \leq \delta$ for each set S of size at most r. Then $\mathbb{E}[(f^{\leq r})^2] \leq 5^r \delta^{\frac{1}{3}} \mathbb{E}\left[f^2\right]$.*

Proof. Applying Lemma II.1.6 with $g = f^{\leq r}$, we obtain the upper bound $\|g\|_4 \leq 5^{\frac{3r}{4}} \delta^{\frac{1}{4}} \|g\|_2^{0.5}$. Since the function f takes values only in the set $\{0,1,-1\}$, we may apply Lemma III.2.2. Combining it with the upper bound on the 4-norm of g, we obtain

$$\|g\|_2^2 \leq \|g\|_4 \|f\|_2^{1.5} \leq 5^{\frac{3r}{4}} \delta^{\frac{1}{4}} \|g\|_2^{0.5} \|f\|_2^{1.5}.$$

Rearranging, and raising everything to the power $\frac{4}{3}$, we obtain $\|g\|_2^2 \leq 5^r \delta^{\frac{1}{3}} \|f\|_2^2$. □

Let us say that f is *ϵ-concentrated* above degree r if $\|f^{\leq r}\|_2^2 \leq \epsilon \|f\|_2^2$. The significance of Theorem III.2.4 stems from the fact that it implies the following result showing that for each $r, \epsilon > 0$ there exists a $\delta > 0$ such that any sparse (r,δ)-global function is ϵ-concentrated above degree r.

Corollary III.2.5. *Let $r \geq 1$. Suppose that f is an (r,δ)-global Boolean function with $\mu_p(f) < \delta$. Then $\mathbb{E}[(f^{\leq r})^2] \leq 10^r \delta^{\frac{1}{3}} \mu_p(f)$.*

Proof. By Lemma III.1.2, for each S of size r we have $\mathrm{I}_S\left(f^{\leq r}\right) \leq \mathrm{I}_S(f) < 8^r \delta$. Then Theorem III.2.4 implies $\|f^{\leq r}\|_2^2 \leq 10^r \delta^{\frac{1}{3}} \|f\|_2^2$, where since f is Boolean we have $\|f\|_2^2 = \mu_p(f)$. □

III.2.2 Isoperimetric stability

We are now ready to prove our variant of the Kahn-Kalai Conjecture and sharp form of Bourgain's Theorem, both of which can be thought of as isoperimetric stability results. Both proofs closely follow existing proofs and substitute our new hypercontractivity inequality for the standard hypercontractivity theorem: for the first we follow a proof of the isoperimetric inequality, and for the second the proof of KKL given by Bourgain and Kalai [13] (their main idea is to apply the argument we gave in Theorem III.2.1 for each of the derivatives of f).

Proof of Theorem I.1.5. We prove the contrapositive statement that for a sufficiently large absolute constant C, if f is a Boolean function such that $\mu_p(f_{J\to 1}) \leq e^{-CK}$ for all J of size at most CK, then $p\mathrm{I}[f] > K\mu_p(f)$. Let f be such a function, and set $\delta = e^{-CK}$. Provided that $C > 2$, f is $(2K,\delta)$-global, and has p-biased measure at most δ. By Corollary III.2.5, we have

$$\|f^{\leq 2K}\|_2^2 \leq 10^{2K} \delta^{\frac{1}{3}} \mu_p(f) \leq \mu_p(f)/2,$$

provided that C is sufficiently large. Hence,

$$\|f^{>2K}\|_2^2 = \|f\|_2^2 - \|f^{\leq 2K}\|_2^2 \geq \mu_p(f)/2.$$

By (1) we obtain $p(1-p)\mathrm{I}[f] \geq 2K\|f^{>2K}\|_2^2$, so $p\mathrm{I}[f] > K\mu_p(f)$. □

Next we require the following lemma which bounds the norm of a low degree truncation in terms of the total influence.

Lemma III.2.6. *Let $r \geq 0$. Suppose that for each nonempty set S of size at most r, $\mathrm{I}_S\left(f^{\leq r}\right) \leq \delta$. Then*

$$\|f^{\leq r}\|_2^2 \leq \mu_p(f)^2 + 5^{r-1}\delta^{\frac{1}{3}}\sigma^2\mathrm{I}[f].$$

Proof. Let $g_i := f_{i\to 1} - f_{i\to 0}$. Then for each S of size at most $r-1$, with $i \notin S$ we have

$$\mathrm{I}_S(g_i^{\leq r}) = \mathrm{I}_{S\cup\{i\}}(f^{\leq r}) \leq \delta,$$

and for each S containing i we have $\mathrm{I}_S((g_i)^{\leq r}) = 0$. By Lemma III.2.4, $\mathbb{E}[((g_i)^{\leq r})^2] \leq 5^{r-1}\delta^{\frac{1}{3}}\mathbb{E}[g_i^2]$. The lemma now follows by summing over all i, using $\sum_i \mathbb{E}[g_i^2] = \mathrm{I}(f)$:

$$\begin{aligned}\|f^{\leq r}\|_2^2 &= \sum_{|S|\leq r}\hat{f}(S)^2 \leq \hat{f}(\emptyset)^2 + \sum_{|S|\leq r}|S|\hat{f}(S)^2 \\ &= \mu_p(f)^2 + \sigma^2\sum_i \mathbb{E}[((g_i)^{\leq r})^2] \leq \mu_p(f)^2 + 5^{r-1}\delta^{1/3}\sigma^2\mathrm{I}(f). \qquad \square\end{aligned}$$

We now establish a variant of Bourgain's Theorem for general Boolean functions, in which we replace the conclusion on the measure of a restriction by finding a large generalised influence.

Theorem III.2.7. *Let $f\colon \{0,1\}^n \to \{0,1\}$. Suppose that $p\mathrm{I}[f] \leq K\mu_p(f)(1-\mu_p(f))$. Then there exists an S of size $2K$, such that $\mathrm{I}_S(f) \geq 5^{-8K}$.*

Proof. Let $r = 2K$ and let $\delta = 5^{-8K}$. Suppose for contradiction that $\mathrm{I}_S(f) \leq \delta$ for each set S of size at most r. By Lemma III.2.6,

$$\|f^{\leq r}\|_2^2 - \mu_p(f)^2 \leq 5^{r-1}\delta^{1/3}\sigma^2 I(f) < p\mathrm{I}[f]/2K \leq \mu_p(f)(1-\mu_p(f))/2.$$

On the other hand, by Parseval

$$\|f - f^{\leq r}\|_2^2 = \sum_{|S|\geq r}\hat{f}(S)^2 \leq r^{-1}\sum_{|S|\geq r}|S|\hat{f}(S)^2 \leq r^{-1}p(1-p)\mathrm{I}(f) \leq \mu_p(f)(1-\mu_p(f))/2.$$

However, these bounds contradict the fact that

$$\mu_p(f)(1-\mu_p(f)) = \|f\|_2^2 - \mu_p(f)^2 = \|f^{\leq r}\|_2^2 - \mu_p(f)^2 + \|f - f^{\leq r}\|_2^2. \qquad \square$$

Proof of Theorem I.1.4. The theorem follows immediately from Theorem III.2.7 and Lemma III.1.3. $\square$

III.2.3 Sharpness examples

We now give two examples showing sharpness of the theorems in this section, both based on the tribes function of Ben-Or and Linial [5].

Example III.2.8. We consider the anti-tribes function $f = f_{s,w} : \{0,1\}^n \to \{0,1\}$ defined by s disjoint sets $T_1, \ldots, T_s \subset [n]$ each of size w, where $f(x) = \prod_{j=1}^s \max_{i\in T_j} x_i$, i.e. $f(x) = 1$ if for every j we have $x_i = 1$ for some $i \in T_j$, otherwise $f(x) = 0$. We have $\mu_p(f) = (1-(1-p)^w)^s$ and $\mathrm{I}[f] = \mu_p(f)' = sw(1-p)^{w-1}(1-(1-p)^w)^{s-1}$. We choose s, w with $s(1-p)^w = 1$ (ignoring the rounding to integers) so that $\mu_p(f) = (1-s^{-1})^s$ is bounded away from 0 and 1, and $K = (1-p)p\mathrm{I}[f] = pw(1-s^{-1})^{-1}\mu_p(f) = \Theta(pw)$. Thus $\log s = w\log(1-p)^{-1} = \Theta(K)$. However, for any $J \subset [n]$ with $|J| = t \leq s$ we have $\mu_p(f_{J\to 1}) \leq (1-s^{-1})^{s-t} \leq 2^{t/s}\mu_p(f)$, so to obtain a density bump of $e^{-o(K)}$ we need $t = e^{-o(K)}s = e^{\Omega(K)} \gg K$. Thus Theorem I.1.4 is sharp.

Example III.2.9. Let $f(x) = f_{s,w}(x)\prod_{i\in T} x_i$ with $f_{s,w}$ as in Example III.2.8 and $T \subset [n]$ a set of size t disjoint from $\cup_j T_j$. We have $\mu_p(f) = p^t(1-(1-p)^w)^s$ and $I[f] = \mu_p(f)' = tp^{t-1}(1-(1-p)^w)^s + p^t sw(1-p)^{w-1}(1-(1-p)^w)^{s-1}$. We fix $K > 1$ and choose s, w with $s(1-p)^w = K$, so that $\mu_p(f) = p^t(1-K/s)^s = p^t e^{-\Theta(K)}$ for $s > 2K$ and $p(1-p)I[f] = \mu_p(f)((1-p)t + pwK(1-K/s)^{-1}) = \mu_p(f)\Theta(K)$ if $pw = \Theta(1)$ and $t = O(K)$. For any $J \subset [n]$ with $|J| = t + u \leq t + s$ we have $\mu_p(f_{J\to 1}) \leq (1-K/s)^{s-u} \leq e^{-K(1-u/s)} \leq e^{-K/2}$ unless $u > s/2 = \Theta(K)$. Thus Theorem I.1.5 is sharp.

III.2.4 Sharp thresholds: the traditional approach

In this section we deduce Theorem I.2.1 from our edge-isoperimetric stability results and the Margulis–Russo Lemma. Recall that a monotone Boolean function is M-global in an interval if $\mu_p\left(f_{J\to 1}\right) \le \mu_p\left(f\right)^{0.01}$ for each p in the interval and set J of size M. We prove the following slightly stronger version of Theorem I.2.1.

Theorem III.2.10. *There exists an absolute constant C such that the following holds for any monotone Boolean function f that is M-global in some interval $[p,q]$: if $q \le p_c$ and $\mu_p\left(f\right) \ge e^{-M/C}$ then*

$$\mu_q\left(f\right) \ge \mu_p(f)^{\left(\frac{p}{q}\right)^{1/C}}. \tag{8}$$

In particular, $q \le M^C p$.

Proof. By Theorem I.1.5, since f is M-global throughout the interval, there exists a constant C such that $\mathrm{I}_x\left[f\right] \ge \frac{\mu_x(f)\log(\frac{1}{\mu_x(f)})}{Cx}$ for all x in the interval $[p,q]$. By the Margulis-Russo lemma,

$$\frac{d}{dx}\log\left(-\log(\mu_x\left(f\right))\right) = \frac{\mu_x(f)'}{\mu_x(f)\log(\mu_x\left(f\right))} = \frac{I_x[f]}{\mu_x(f)\log(\mu_x\left(f\right))} \le \frac{-1}{Cx}$$

in all of the interval $[p,q]$. Hence,

$$\log\left(-\log(\mu_q(f))\right) \le \log(-\log(\mu_p(f))) - \frac{\log(\frac{q}{p})}{C}.$$

The first part of the theorem follows by taking exponentials, multiplying by -1 then taking exponentials again. To see the final statement, note that $q \le p_c$ implies $\mu_q\left(f\right) \le \frac{1}{2}$. We cannot have $q \ge M^c p$, as then the right hand side in (8) would be larger than $e^{-\frac{1}{C}} > 1/2$ for large C. To obtain Theorem I.2.1 we substitute $q = p_c$. □

III.3 Noise sensitivity and sharp thresholds

We start this section by showing that sparse global functions are noise sensitive; Theorem I.3.2 follows immediately from Theorem III.3.1.

Theorem III.3.1. *Let $\rho \in (0,1)$, and let $\epsilon > 0$. Let $r = \frac{\log(2/\epsilon)}{\log(1/\rho)}$, and let $\delta = 10^{-3r-1}\epsilon^3$. Suppose that f is an (r,δ)-global Boolean function with $\mu_p\left(f\right) < \delta$. Then*

$$\mathrm{Stab}_\rho\left(f\right) \le \epsilon\mu_p\left(f\right).$$

Proof. We have

$$\langle \mathrm{T}_\rho f, f\rangle \le \sum_{|S|\le r}\hat{f}\left(S\right)^2 + \rho^r \sum_{|S|>r}\hat{f}\left(S\right)^2 \le \mathbb{E}\left[\left(f^{\le r}\right)^2\right] + \frac{\varepsilon}{2}\mu_p(f).$$

The statement now follows from Corollary III.2.5, which gives $\mathbb{E}[(f^{\le r})^2] \le 10^r\delta^{1/3}\mathbb{E}[f^2] < \varepsilon\mu_p(f)/2$. □

In the remainder of this section, following [65], we deduce sharp thresholds from noise sensitivity via the following *directed noise operator*, which is implicit in the work of Ahlberg, Broman, Griffiths and Morris [3] and later studied in its own right by Abdullah and Venkatasubramanian [1].

Definition III.3.2. Let $D\left(p,q\right)$ denote the unique distribution on pairs $(\boldsymbol{x},\boldsymbol{y}) \in \{0,1\}^n \times \{0,1\}^n$ such that $\boldsymbol{x} \sim \mu_p$, $\boldsymbol{y} \sim \mu_q$, all $\boldsymbol{x}_i \le \boldsymbol{y}_i$ and $\{(\boldsymbol{x}_i,\boldsymbol{y}_i) : i \in [n]\}$ are independent. We define a linear operator $\mathrm{T}^{p\to q} : L^2(\{0,1\}^n,\mu_p) \to L^2(\{0,1\}^n,\mu_q)$ by

$$\mathrm{T}^{p\to q}\left(f\right)\left(y\right) = \mathbb{E}_{(\boldsymbol{x},\boldsymbol{y})\sim D(p,q)}\left[f\left(\boldsymbol{x}\right) \mid \boldsymbol{y} = y\right].$$

The directed noise operator $\mathrm{T}^{p\to q}$ is a version of the noise operator where bits can be flipped only from 0 to 1. The associated notion of directed noise stability, i.e. $\langle f, \mathsf{T}^{p\to q} f\rangle_{\mu_q}$, is intuitively a measure of how close a Boolean function f is to being monotone. Indeed, for any $(\mathbf{x}, \mathbf{y})$ with all $x_i \le y_i$ we have $f(\boldsymbol{x}) f(\boldsymbol{y}) \le f(\boldsymbol{x})$, with equality if f is monotone, so

$$\langle f, \mathsf{T}^{p\to q} f\rangle = \mathbb{E}_{(\mathbf{x},\mathbf{y})\sim D(p,q)}\left[f(\boldsymbol{x}) f(\boldsymbol{y})\right] \le \mathbb{E}_{(\mathbf{x},\mathbf{y})\sim D(p,q)}\left[f(\boldsymbol{x})\right] = \mu_p(f),$$

with equality if f is monotone[4]. We note that the adjoint operator $(\mathrm{T}^{p\to q})^\star : L^2(\{0,1\}^n, \mu_q) \to L^2(\{0,1\}^n, \mu_p)$ defined by $\langle \mathrm{T}^{p\to q} f, g\rangle = \left\langle f, (\mathrm{T}^{p\to q})^\star g\right\rangle$ satisfies $(\mathrm{T}^{p\to q})^\star = \mathrm{T}^{q\to p}$, where

$$\mathrm{T}^{q\to p}(g)(x) = \mathbb{E}_{(\boldsymbol{x},\boldsymbol{y})\sim D(p,q)}\left[g(\boldsymbol{y}) \mid \boldsymbol{x} = x\right].$$

The following simple calculation relates these operators to the noise operator.

Lemma III.3.3. *Let $0 < p < q < 1$ and $\rho = \frac{p(1-q)}{q(1-p)}$. Then $(\mathrm{T}^{p\to q})^\star \mathrm{T}^{p\to q} = \mathrm{T}_\rho$ on $L^2(\{0,1\}^n, \mu_p)$.*

Proof. We need to show that the following distributions on pairs of p-biased bits $(\mathbf{x}, \mathbf{x}')$ are identical: (a) let $\mathbf{x}$ be a p-biased bit, with probability ρ let $\mathbf{x}' = \mathbf{x}$, otherwise let $\mathbf{x}'$ be an independent p-biased bit, (b) let $(\mathbf{x}, \mathbf{y}) \sim D(p,q)$ and then $(\mathbf{x}', \mathbf{y}) \sim D(p,q) \mid y$. It suffices to show $\mathbb{P}(x \ne x')$ is the same in both distributions. We condition on x. Consider $x = 1$. In distribution (a) we have $\mathbb{P}(\mathbf{x}' = 0) = (1-\rho)(1-p)$. In distribution (b) we have $\mathbb{P}(\mathbf{y} = 1) = 1$ and then $\mathbb{P}(\mathbf{x}' = 0) = 1 - p/q = (1-\rho)(1-p)$, as required. Now consider $\mathbf{x} = 0$. In distribution (a) we have $\mathbb{P}(\mathbf{x}' = 1) = (1-\rho)p$. In distribution (b) we have $\mathbb{P}(\mathbf{y} = 1) = \frac{q-p}{1-p}$ and then $\mathbb{P}(\mathbf{x}' = 1 \mid \mathbf{y} = 1) = p/q$, so $\mathbb{P}(\mathbf{x}' = 1) = \frac{p(q-p)}{q(1-p)} = (1-\rho)p$, as required. □

We now give an alternative way to deduce sharp threshold results, using noise sensitivity, rather than the traditional approach via total influence (as in the proof of Theorem III.2.10). Our alternative approach has the following additional nice features, both of which have been found useful in Extremal Combinatorics (see [65]).

1. To deduce a sharp threshold result in an interval $[p, q]$ it is enough to show that f is global only according to the p-biased distribution. This is a milder condition than the one in the traditional approach, that requires globalness throughout the entire interval.
2. The monotonicity requirement may be relaxed to "almost monotonicity".

Proposition III.3.4. *Let $f\colon \{0,1\}^n \to \{0,1\}$ be a monotone Boolean function. Let $0 < p < q < 1$ and $\rho = \frac{p(1-q)}{q(1-p)}$. Then $\mu_q(f) \ge \mu_p(f)^2/\mathrm{Stab}_\rho(f)$.*

Proof. By Cauchy–Schwarz and Lemma III.3.3,

$$\mu_p(f)^2 = \langle \mathsf{T}^{p\to q} f, f\rangle_{\mu_q}^2 \le \langle \mathsf{T}^{p\to q} f, \mathsf{T}^{p\to q} f\rangle_{\mu_q} \langle f, f\rangle_{\mu_q} = \langle \mathrm{T}_\rho f, f\rangle_{\mu_p} \mu_q(f).$$

□

The above proof works not only for monotone functions, but also for functions where the first equality above is replaced by approximate equality (which is a natural notion for a function to be "almost monotone"). We conclude this part by recalling the following sharp threshold theorem for global functions, and noting that its proof is immediate from Theorem III.3.1 and Proposition III.3.4.

Theorem I.3.3. *For any $\alpha > 0$ there is $C > 0$ such that for any $\varepsilon, p, q \in (0, 1/2)$ with $q \ge (1+\alpha)p$, writing $r = C\log\varepsilon^{-1}$ and $\delta = 10^{-3r-1}\varepsilon^3$, any monotone (r,δ)-global Boolean function f with $\mu_p(f) \le \delta$ satisfies $\mu_q(f) \ge \mu_p(f)/\varepsilon$.*

[4] The starting point for [65] is the observation that this inequality is close to an equality if f is almost monotone.

Part IV
Pseudorandomness and junta approximation

The first main result proved in this part will be our junta approximation theorem, Theorem I.4.8, which we will now restate, using the notation $\mathcal{G}(r, s, \Delta)$ for the family of all r-graphs G with s edges and maximum degree $\Delta(G) \le \Delta$. We recall that $S \subset V(G^+)$ is a *crosscut* if $|E \cap S| = 1$ for all $E \in G^+$, and $\sigma(G)$ denotes the minimum size of a crosscut.

Theorem IV.0.1. *Let $G \in \mathcal{G}(r, s, \Delta)$ and $C \gg r\Delta\varepsilon^{-1}$. Then for any G^+-free $\mathcal{F} \subset \binom{[n]}{k}$ with $C \le k \le n/Cs$, there is $J \subset V(G)$ with $|J| \le \sigma(G) - 1$ and $|\mathcal{F} \setminus \mathcal{S}_{n,k,J}| \le \varepsilon|\mathcal{S}_{n,k,\sigma(G)-1}|$.*

The set J in Theorem IV.0.1 will consist of all vertices of suitably large degree. Thus $\mathcal{F}^{\emptyset}_J := \mathcal{F} \setminus \mathcal{S}_{n,k,J}$ does not have any vertices of large degree, which we think of a pseudorandomness property, called 'globalness', due to its interpretation as globalness of the corresponding characteristic Boolean function.

An important theme of this part, treated in its first section, will be the interplay between two pseudorandomness notions: globalness and another, called uncapturability. We will see that globalness implies uncapturability, and that uncapturability can be 'upgraded' to globalness by taking appropriate restrictions.

In the proof of Theorem IV.0.1 we will consider separately the two steps of showing $|J| \le \sigma(G) - 1$ and $|\mathcal{F} \setminus \mathcal{S}_{n,k,J}| \le \varepsilon|S_{n,k,\sigma(G)-1}|$. For both steps we consider a two step embedding strategy for G^+, where in the first step we embed[5] G in the 'fat shadow' of $\mathcal{F}$ (meaning that the image of every edge has many extensions to an edge of $\mathcal{F}$) and in the second step we 'lift' edges from the fat shadow to the original family.

This proof strategy is implemented at the end of the first section, assuming results that will be proved in later sections. The lifting step requires results on cross matchings presented in Section IV.2, which will also be used for the proof of the Huang–Loh–Sudakov Conjecture in Section V.1. The analysis of fat shadows and the embedding steps will be carried out in Section IV.3.

After proving Theorem I.4.8, in Section IV.4 we prove the following refined junta approximation result, in which we improve the bound on $|\mathcal{F}^{\emptyset}|$; besides being of interest in its own right, this bound is needed for the proofs of our exact Turán results in the next part.

Theorem IV.0.2. *Let $G \in \mathcal{G}(r, s, \Delta)$, $0 < C^{-1} \ll \delta \ll \varepsilon \ll (r\Delta)^{-1}$ and $C \le k \le n/Cs$. Then for any G^+-free $\mathcal{F} \subset \binom{[n]}{k}$ with $|\mathcal{F}| > |\mathcal{S}_{n,k,\sigma(G)-1}| - \delta\binom{n-1}{k-1}$ there is $J \in \binom{[n]}{\sigma(G)-1}$ with $|\mathcal{F} \setminus \mathcal{S}_{n,k,J}| \le \varepsilon\binom{n-1}{k-1}$.*

Throughout the remainder of the paper it will often be convenient to assume that G belongs to the subset $\mathcal{G}'(r, s, \Delta)$ of $\mathcal{G}(r, s, \Delta)$ consisting of its r-partite r-graphs. There is no loss of generality in this assumption, as $G^+(r\Delta)$ is $r\Delta$-partite for any $G \in \mathcal{G}(r, s, \Delta)$. To see this, consider a greedy algorithm in which we assign vertices of G sequentially to $r\Delta$ parts, ensuring for every edge that all of its vertices are in distinct parts. Clearly this algorithm can be completed. Then the expansion vertices can be assigned so that each edge of G^+ has one vertex in each part.

IV.1 Globalness and uncapturability

This section introduces the key concepts that will underpin this part of the paper. After introducing some basic definitions that run throughout the paper in the first subsection, we will define and analyse our two pseudorandomness notions in the second subsection. We conclude in the third section by proving our junta approximation theorem, assuming two embedding lemmas that will be proved in Section IV.3.

[5] For simplicity in this overview we are only describing the embedding strategy used to bound $|\mathcal{F}^{\emptyset}_J|$; the strategy for bounding $|J|$ is similar, but adapted so that J can play the role of a crosscut in G.

IV.1.1 Definitions

Given $m, n \in \mathbb{N}$ with $m \leq n$ we let $[n] = \{1, 2, \ldots, n\}$ and $[m, n] = \{m, m+1, \ldots, n\}$. We write $\{0,1\}^X$ for the power set (set of subsets) of a set X (identifying sets with their characteristic $0/1$ vectors) and $\binom{X}{k} = X^{(k)} = \{A \subset X : |A| = k\}$. We call $\mathcal{F} \subset \{0,1\}^X$ a family or a hypergraph on the vertex set X, and the elements of $\mathcal{F}$ are called edges. We say $\mathcal{F}$ is k-uniform if $\mathcal{F} \subset \binom{X}{k}$; we also call $\mathcal{F}$ a k-graph on X.

Given a family $\mathcal{F} \subset \{0,1\}^X$ and $B \subset J \subset X$ we write $\mathcal{F}_J^B$ for the family

$$\mathcal{F}_J^B := \{A \in \{0,1\}^{X \setminus J} : A \cup B \in \mathcal{F}\} \subset \{0,1\}^{X \setminus J}.$$

Clearly $\mathcal{F}_J^B$ is $(k - |B|)$-uniform if $\mathcal{F}$ is k-uniform. If either B or J has a single element $\{j\}$ then we will often suppress the bracket, e.g. $\mathcal{F}_v^v = \mathcal{F}_{\{v\}}^{\{v\}}$.

We refer to $\mathcal{F}_v^v$ as the *exclusive link* of v in $\mathcal{F}$. The *inclusive link* of v in $\mathcal{F}$ is $\mathcal{F} * v := \{E \in \mathcal{F} : v \in E\}$. The *degree* of a vertex v in $\mathcal{F}$ is $d_{\mathcal{F}}(v) = |\mathcal{F}_v^v| = |\mathcal{F} * v|$. The minimum and maximum degrees of $\mathcal{F}$ are $\delta(\mathcal{F}) = \min_{v \in V(\mathcal{F})} d_{\mathcal{F}}(v)$ and $\Delta(\mathcal{F}) = \max_{v \in V(\mathcal{F})} d_{\mathcal{F}}(v)$.

Let $\mathcal{H}_1, \ldots, \mathcal{H}_s \subset \{0,1\}^V$. We say that $\mathcal{F}_1, \ldots, \mathcal{F}_s \subset \{0,1\}^X$ *cross contain* $\mathcal{H}_1, \ldots, \mathcal{H}_s$ if there is an injection $\phi : V \to X$ such that $\phi(\mathcal{H}_i) \subset \mathcal{F}_i$ for all $i \in [s]$. Here we write $\phi(\mathcal{H}_i) = \{\phi(e) : e \in \mathcal{H}_i\}$ with each $\phi(e) = \{\phi(x) : x \in e\}$.

We simply say that $\mathcal{F}_1, \ldots, \mathcal{F}_s$ cross contain $\mathcal{H}$ if they cross contain any ordering of the edges of $\mathcal{H}$. Thus a single hypergraph $\mathcal{F}$ contains $\mathcal{H}$ if $\mathcal{F}_1, \ldots, \mathcal{F}_s$ cross contain $\mathcal{H}$, where $\mathcal{F}_i = \mathcal{F}$ for all $i \in [s]$.

Given an r-graph G and $k \geq r$, we recall that the k-expansion $G^+ = G^+(k)$ is the k-uniform hypergraph obtained from G by adding $k - r$ new vertices to each edge, i.e. G^+ has edge set $\{e \cup S_e : e \in E(G)\}$ where $|S_e| = k - r$, $S_e \cap V(G) = \emptyset$ and $S_e \cap S_{e'} = \emptyset$ for all distinct $e, e' \in E(G)$.

When embedding expanded hypergraphs in uniform families, we may allow the uniformity of our families to vary, defining cross containment of G^+ in the obvious way: the edge of G^+ embedded in the family $\mathcal{F}_i \subset \binom{[n]}{k_i}$ is obtained from an edge of G by adding $k_i - r$ new vertices.

A family $\mathcal{F} \subset \{0,1\}^X$ is said to be *monotone* if given $F \in \mathcal{F}$ and $F \subset F' \subset X$ we also have $F' \in \mathcal{F}$. Given $\mathcal{F} \subset \{0,1\}^X$ the *up closure* of $\mathcal{F}$ is the monotone family $\mathcal{F}^{\uparrow} = \{B \subset X : A \subset B \text{ for some } A \in \mathcal{F}\} \subset \{0,1\}^X$. The ℓ-shadow of $\mathcal{F}$ is $\partial^{\ell}(\mathcal{F}) := \{F \in \binom{X}{\ell} : F \subset G \text{ for some } G \in \mathcal{F}\}$.

Given $\mathcal{F} \subset \binom{X}{k}$ we will write $\mu(\mathcal{F}) = |\mathcal{F}|/\binom{|X|}{k}$. Some of our results are more naturally stated with $|\mathcal{F}|$ and others with $\mu(\mathcal{F})$, so we will freely move between these settings. Given $p \in [0,1]$ we will use μ_p to denote the p-biased measure on $\{0,1\}^n$, where a set $\boldsymbol{A} \sim \mu_p$ is selected by including each $i \in [n]$ independently with probability p. We extend this notation to families $\mathcal{F} \subset \{0,1\}^n$ by $\mu_p(\mathcal{F}) := \Pr_{\boldsymbol{A} \sim \mu_p}[\boldsymbol{A} \in \mathcal{F}]$. We often identify a family $\mathcal{F}$ with its characteristic Boolean function $f : \{0,1\}^n \to \{0,1\}$ and apply the above terminology freely in either setting, e.g. we call f monotone if $\mathcal{F}$ is monotonem and write $\mu_p(f)$ for the expectation of f under μ_p.

To pass between these measures we note the following simple properties that will be henceforth used without further comment. For any $\mathcal{F} \subset \{0,1\}^n$ and $J \subset [n]$, we have the union bound estimate

$$\mu_p(\mathcal{F}) \leq \mu_p(\mathcal{F}_J^{\emptyset}) + p \sum_{j \in J} \mu_p(\mathcal{F}_j^j) \leq \mu_p(\mathcal{F}_J^{\emptyset}) + |J|p.$$

The same estimate holds replacing μ_p by uniform measures μ for $\mathcal{F} \subset \binom{[n]}{k}$ with $k = pn$, remembering to use the correct normalisations: we have $\mu(\mathcal{F}) = |\mathcal{F}|\binom{n}{k}^{-1}$ and $\mu(\mathcal{F}_j^j) = |\mathcal{F}_j^j|\binom{n-1}{k-1}^{-1}$.

In the other direction, we have the bounds

$$\begin{aligned} &\mu_p(\mathcal{F}) \geq (1-p)^{|J|}\mu_p(\mathcal{F}_J^{\emptyset}) \text{ for } \mathcal{F} \subset \{0,1\}^n, \text{ and} \\ &\mu(\mathcal{F}) \geq \tbinom{n}{k}^{-1}\tbinom{n-|J|}{k}\mu(\mathcal{F}_J^{\emptyset}) \geq \left(1 - \tfrac{|J|}{n-k}\right)^k \mu(\mathcal{F}_J^{\emptyset}) \text{ for } \mathcal{F} \subset \tbinom{[n]}{k}. \end{aligned}$$

Throughout $a \ll b$ or $a^{-1} \gg b^{-1}$ will mean that the following statement holds provided a is sufficiently small as a function of b.

IV.1.2 Pseudorandomness

Here we define our two key notions of pseudorandomness for set systems, namely uncapturability and globalness, and explore some of their basic properties.

Definition IV.1.1. Let $\mathcal{F} \subset \{0,1\}^n$ and μ be a measure on $\{0,1\}^n$.

We say $\mathcal{F}$ is (μ, a, ε)-uncapturable if $\mu(\mathcal{F}_J^{\emptyset}) \geq \varepsilon$ whenever $J \subset [n]$ with $|J| \leq a$.

We say $\mathcal{F}$ is (μ, a, ε)-global if $\mu(\mathcal{F}_J^J) \leq \varepsilon$ whenever $J \subset [n]$ with $|J| \leq a$.

We say $\mathcal{F}$ is (μ, a, ε)-capturable if it is not (μ, a, ε)-uncapturable, or (μ, a, ε)-local if it is not (μ, a, ε)-global. We omit μ from the notation if it is clear from the context, i.e. if $\mathcal{F} \subset \binom{[n]}{k}$ with uniform measure or $\mathcal{F} \subset \{0,1\}^n$ with p-biased measure μ_p, where p is clear from the context.

We now establish some basic properties of these definitions. For each property we state two lemmas that apply when μ is uniform or $\mu = \mu_p$. We only give proofs in the uniform setting, as those in the p-biased setting are essentially the same. The following pair of lemmas shows that globalness is preserved by restrictions.

Lemma IV.1.2. *If $\mathcal{F} \subset \binom{[n]}{k}$ is (a, ε)-global and $I \subset J \subset [n]$ with $|I| < a$ and $|J| < n/2k$ then $\mathcal{F}_J^I$ is $(a - |I|, 2\varepsilon)$-global.*

Lemma IV.1.3. *If $\mathcal{F} \subset \{0,1\}^n$ under μ_p is (a, ε)-global and $I \subset J \subset [n]$ with $|I| < a$ and $|J| < 1/2p$ then $\mathcal{F}_J^I$ is $(a - |I|, 2\varepsilon)$-global.*

Proof of Lemma IV.1.2. For any $K \subset [n] \setminus J$ with $|K| \leq a - |I|$, we have $\mu(\mathcal{F}_{I \cup K}^{I \cup K}) \leq \varepsilon$, so $\mu((\mathcal{F}_J^I)_K^K) \leq \left(1 - \frac{|J \setminus I|}{n-k}\right)^{-k} \varepsilon < 2\varepsilon$. $\square$

The next pair shows that globalness implies uncapturability.

Lemma IV.1.4. *If $\mathcal{F} \subset \binom{[n]}{k}$ is $(1, \varepsilon)$-global with $\varepsilon = \mu(\mathcal{F})n/2ak$ then $\mathcal{F}$ is $(a, \mu(\mathcal{F})/2)$-uncapturable.*

Lemma IV.1.5. *If $\mathcal{F} \subset \{0,1\}^n$ under μ_p is $(1, \varepsilon)$-global with $\varepsilon = \mu_p(\mathcal{F})/2ap$ then $\mathcal{F}$ is $(a, \mu_p(\mathcal{F})/2)$-uncapturable.*

Proof of Lemma IV.1.4. If $|J| \leq a$ then $\mu(\mathcal{F}_J^{\emptyset}) \geq \mu(\mathcal{F}) - \varepsilon ak/n \geq \mu(\mathcal{F})/2$. $\square$

Uncapturability does not imply globalness, but we do have a partial converse: by taking restrictions we can upgrade uncapturable families to families that are global or large.

Lemma IV.1.6. *Suppose $\beta \in (0, .1)$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $2r < k_i < \beta n/2rm$ are (rm, δ_i)-uncapturable for $i \in [m]$. Then there are pairwise disjoint $S_1, \ldots, S_m$ with each $|S_i| \leq r$ such that, setting $\mathcal{G}_i = (\mathcal{F}_i)_S^{S_i}$ where $S = \bigcup_i S_i$, whenever $\mu(\mathcal{G}_i) < \beta$ we have $S_i = \emptyset$ and $\mathcal{G}_i$ is $(r, 2\beta)$-global with $\mu(\mathcal{G}_i) > \delta_i$.*

Lemma IV.1.7. *Suppose $\beta \in (0, .1)$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $k_i < \beta n/2rm$ are (rm, δ_i)-uncapturable for $i \in [m]$. Then there are pairwise disjoint $S_1, \ldots, S_m$ with each $|S_i| \leq r$ such that, setting $\mathcal{G}_i = (\mathcal{F}_i^{\uparrow})_S^{S_i}$ where $S = \bigcup_i S_i$ and $p_i = k_i/(n - |S|)$, whenever $\mu_{p_i}(\mathcal{G}_i) < \beta$ we have $S_i = \emptyset$ and $\mathcal{G}_i$ is $(r, 2\beta)$-global with $\mu_{p_i}(\mathcal{G}_i) > \delta_i/4$.*

Proof of Lemma IV.1.6. Let $(S_i : i \in I)$ be a maximal collection of pairwise disjoint sets with $|S_i| \leq r$ and $\mu((\mathcal{F}_i)_{S_i}^{S_i}) > 1.5\beta$. Let $S = \bigcup_{i \in I} S_i$ and $\mathcal{G}_i = (\mathcal{F}_i)_S^{S_i}$ for each $i \in [m]$, where $S_i = \emptyset$ for $i \in [m] \setminus I$. For any $i \in I$ we have $\mu(\mathcal{G}_i) > \mu((\mathcal{F}_i)_{S_i}^{S_i}) - |S \setminus S_i| k_i/n > \beta$. Now consider i with $\mu(\mathcal{G}_i) < \beta$. Then $i \notin I$, so $S_i = \emptyset$ and $\mu(\mathcal{G}_i) > \delta_i$ by uncapturability. Furthermore, for any $R \subset [n] \setminus S$ with $|R| \leq r$ we have $\mu((\mathcal{F}_i)_R^R) \leq 1.5\beta$, so $(\mathcal{G}_i)_R^R = ((\mathcal{F}_i)_R^R)_S^{\emptyset}$ has $\mu((\mathcal{G}_i)_R^R) \leq \left(1 - \frac{|S|}{n-k_i}\right)^{-k_i} \mu((\mathcal{F}_i)_R^R) < 2\beta$. $\square$

We conclude this subsection with a lemma on decomposing any family according to its vertex degrees, where to make an analogy with the regularity method we think of high degree vertex links as 'structured' and the low degree remainder as 'pseudorandom'.

Lemma IV.1.8. *Let $\mathcal{F} \subset \binom{[n]}{k}$ and $J = \{i : \mu(\mathcal{F}_i^i) > \varepsilon\}$. If $|J| < n/2k$ then $\mathcal{G} = \mathcal{F}_J^\emptyset$ is $(1, 2\varepsilon)$-global, and so $(a, \mu(\mathcal{G})/2)$-uncapturable with $a = \mu(\mathcal{G})n/4k\varepsilon$,*

Proof. If $j \in [n] \setminus J$ then $\mu(\mathcal{F}_j^j) \le \varepsilon$ by definition of J, so $\mu(\mathcal{G}_j^j) \le \left(1 - \frac{|J|}{n-k}\right)^{-k} \mu(\mathcal{F}_j^j) < 2\varepsilon$. The lemma follows by Definition IV.1.1 and Lemma IV.1.4. $\square$

IV.1.3 Embeddings

Here we will prove Theorem I.4.8 assuming two fundamental embedding results, which will be proved in Section IV.3. The first of these shows that sufficiently large families contain a cross copy of any expanded hypergraph G^+. Our bound on $\mu(\mathcal{F}_i)$ is sharper for larger k_i: when $k_i = O(1)$ it is a constant, which is relatively weak (but still useful), whereas when $k_i \gg \log n$ it is $O(sk_i/n) = O(\sigma(G)k_i/n)$, which is tight up to the constant factor.

Lemma IV.1.9. *Given $G \in \mathcal{G}(r, s, \Delta)$, $C \gg r\Delta$ and $C \le k_i \le n/Cs$ for all $i \in [s]$, any $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with all $\mu(\mathcal{F}_i) \ge e^{-k_i/C} + Csk_i/n$ cross contain G^+.*

When the uniformities k_i are small we cannot improve this cross containment result, as below density $e^{-\Omega(k_i)}$ the families $\mathcal{F}_i$ may have disjoint supports. However, when finding G^+ in a single family $\mathcal{F}$ we can get a much better bound on the density, and moreover it suffices to assume that $\mathcal{F}$ is sufficiently uncapturable, as follows.

Lemma IV.1.10. *Given $G \in \mathcal{G}(r, s, \Delta)$, $C \gg C_1 \gg C_2 \gg r\Delta$ and $C \le k \le n/Cs$, any $(C_1 s, sk/C_2 n)$-uncapturable $\mathcal{F} \subset \binom{[n]}{k}$ contains G^+.*

We conclude this section by deducing our junta approximation theorem from the above lemmas.

Proof of Theorem I.4.8. Let $G \in \mathcal{G}(r, s, \Delta)$ and $C \gg C_1 \gg C_2 \gg r\Delta\varepsilon^{-1}$. Consider any G^+-free $\mathcal{F} \subset \binom{[n]}{k}$ with $C \le k \le \frac{n}{Cs}$. Let $J = \{i \in [n] : \mu(\mathcal{F}_i^i) \ge \beta\}$, where $\beta := e^{-k/C_1} + C_1 sk/n$. We need to show $|J| \le \sigma(G) - 1$ and $|\mathcal{F}_J^\emptyset| \le \varepsilon |\mathcal{S}_{n,k,\sigma(G)-1}|$.

The bound on $|J|$ follows from Lemma IV.1.9. Indeed, supposing for a contradiction $|J| \ge \sigma(G)$, we may fix a minimal crosscut S of G^+ and distinct $j_s \in J$ for each $s \in S$. Let $I = \{i_s : s \in S\}$ and $\mathcal{F}_s = \mathcal{F}_I^{i_s}$ for $s \in S$. By definition of J, for each $s \in S$ we have $\mu(\mathcal{F}_s) > \beta - |I|k/n > \beta/2$, so by Lemma IV.1.9 the families $(\mathcal{F}_s : s \in S)$ cross contain the exclusive links $((G^+)_s^s : s \in S)$. However, this contradicts $\mathcal{F}$ being G^+-free.

As $|J| < s \le n/Ck$ we can apply Lemma IV.1.8 to see that $\mathcal{G} = \mathcal{F}_J^\emptyset$ is $(a, \mu(\mathcal{G})/2)$-uncapturable with $a = \mu(\mathcal{G})n/4k\beta$. However, by Lemma IV.1.10 $\mathcal{G}$ is $(C_1 s, sk/C_2 n)$-capturable, so we must have $\mu(\mathcal{G})/2 < sk/C_2 n$, or $a < C_1 s$, so again $\mu(\mathcal{G}) < 4\beta C_1 sk/n < sk/C_2 n$. As $\mu(\mathcal{S}_{n,k,\sigma(G)-1}) > .9(\sigma(G) - 1)k/n$ and $s \le \Delta\sigma(G)$ we deduce $|\mathcal{F}_J^\emptyset| = |\mathcal{G}| < \varepsilon |\mathcal{S}_{n,k,\sigma(G)-1}|$. $\square$

IV.2 Matchings

The main result of this section is the following lemma on cross containment of matchings in uncapturable families, which will be used for 'lifting' (as described in the previous section) and also in the proof of the Huang–Loh–Sudakov Conjecture.

Lemma IV.2.1. *Let $C \gg C_1 \gg C_2 \gg 1$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $k_i \le n/Cs$ for $i \in [s]$. Suppose $\mathcal{F}_i$ is $(C_1 m, mk_i/C_2 n)$-uncapturable for $i \in [m]$ and $\mu(\mathcal{F}_i) > C_1 sk_i/n$ for $i > m$. Then $\mathcal{F}_1, \ldots, \mathcal{F}_s$ cross contain a matching.*

We start in the first subsection by recalling some basic probabilistic tools, and also our new sharp threshold result from Part III. Next we present some extremal results on cross matchings in the second subsection. We conclude by proving the uncapturability result in the third subsection.

IV.2.1 Probabilistic tools and sharp thresholds

We start with the following lemma that will be used to pass between the uniform and p-biased measures.

Lemma IV.2.2. *Let $n, k \in \mathbb{N}$ with $k = pn \le n$. Then $\mathbb{P}\big(\operatorname{Bin}(n,p) \ge k\big) \ge 1/4$. Thus if $\mathcal{A} \subset \binom{[n]}{k}$ we have $\mu_p(\mathcal{A}^\uparrow) \ge \mu(\mathcal{A})/4$.*

Proof. The first statement appears in [43]. The second holds as $\big|\mathcal{A}^\uparrow \cap \binom{[n]}{j}\big| \ge \alpha\binom{n}{j}$ for $j \ge k$ by the LYM inequality, and so $\mu_p(\mathcal{A}^\uparrow) \ge \sum_{j=k}^n \mathbb{P}\big(\operatorname{Bin}(n,p) = j\big)\mu\big(\mathcal{A}^\uparrow \cap \binom{[n]}{j}\big) \ge \mathbb{P}\big(\operatorname{Bin}(n,p) \ge k\big)\alpha \ge \alpha/4$. $\square$

We will also need the following well-known Chernoff bound (see [47, Theorem 2.8]), as applied to sums of Bernoulli random variables, i.e. random variables which take values in $\{0,1\}$; if these are identically distributed then we obtain a binomial variable. The inequality can also be applied to a hypergeometric random variable (see [47, Remark 2.11]), i.e. $|S \cap T|$ with $S \in \binom{X}{s}$ and uniformly random $T \in \binom{X}{t}$ for some X, s and t.

Lemma IV.2.3. *Let X be a sum of independent Bernoulli random variables and $0 < a < 3/2$. Then $\mathbb{P}\big[|X - \mathbb{E}X| \ge a\mathbb{E}X\big] \le 2e^{-\frac{a^2}{3}\mathbb{E}X}$.*

Next we recall our sharp threshold result for global functions that we proved in Part III which will play a crucial role in this section, and so for all subsequent applications of Lemma IV.2.1.

Theorem I.3.3. *For any $\alpha > 0$ there is $C > 0$ such that for any $\varepsilon, p, q \in (0, 1/2)$ with $q \ge (1+\alpha)p$, writing $r = C\log\varepsilon^{-1}$ and $\delta = 10^{-3r-1}\varepsilon^3$, any monotone (r,δ)-global Boolean function f with $\mu_p(f) \le \delta$ satisfies $\mu_q(f) \ge \mu_p(f)/\varepsilon$.*

We will apply the following two consequences of this result.

Theorem IV.2.4. *Suppose $\mathcal{F} \subset \{0,1\}^n$ is monotone with $\mu_p(\mathcal{F}) = \mu$.*

1. *If $\mu \ll r^{-1} \ll \varepsilon$ then there is $R \subset [n]$ with $|R| \le r$ and $\mu_{2p}(\mathcal{F}_R^R) \ge \mu/\varepsilon$.*
2. *If $p \ll K^{-1} \ll \eta \ll 1$ then there is $R \subset [n]$ with $|R| \le K\log\mu^{-1}$ and $\mu_{Kp}(\mathcal{F}_R^R) \ge \mu^\eta$.*

Proof. Let f be the monotone Boolean characteristic function of $\mathcal{F}$.

For (1) we apply Theorem I.3.3 with $\alpha = 1$ and the same ε, If f is not (r,δ)-global then for some R with $|R| \ge r$ we have $\mu_{2p}(\mathcal{F}_R^R) \ge \mu_p(\mathcal{F}_R^R) = \mu_p(f_{R\to 1}) \ge \delta \ge \mu/\varepsilon$. On the other hand, if f is (r,δ)-global then we can take $R = \emptyset$, as Theorem I.3.3 gives $\mu_{2p}(\mathcal{F}) \ge \mu/\varepsilon$.

For (2), we repeatedly apply Theorem I.3.3 with $\alpha = 1$ and $\varepsilon = \mu^{\eta^2}$, so $r = C\log\varepsilon^{-1} = C\eta^2\log\mu^{-1}$ and $\delta = 10^{-3r-1}\varepsilon^3 \ge \mu^\eta$, as we may assume $\eta \ll C^{-1}$. We can assume that f is (r,δ)-global, otherwise we immediately obtain R as required, so $\mu_{2p}(\mathcal{F}) \ge \mu/\varepsilon = \mu^{1-\eta^2}$. Repeating the argument, if we do not find R then after $t \le \eta^{-2}$ iterations we reach $\mu_{2^t p}(\mathcal{F}) \ge \delta \ge \mu^\eta$, so we can take $R = \emptyset$. $\square$

IV.2.2 Extremal results

In this subsection we adapt the method of [46, Lemma 3.1] to prove a variant form of the following result of Huang, Loh and Sudakov [46].

Lemma IV.2.5. *Let $k_1, \dots, k_s, n \in \mathbb{N}$ with $\sum_{i\in[s]} k_i \le n$. Suppose $\mathcal{F}_i \subset \binom{[n]}{k_i}$ for $i \in [s]$ do not cross contain a matching. Then $\mu(\mathcal{F}_i) \le k_i(s-1)/n$ for some $i \in [s]$.*

We will prove the following variant that allows a few families to be significantly smaller.

Lemma IV.2.6. *Let $1 \le m \le s$, $k_1, \dots, k_s \ge 0$ and $n \ge \sum_{i\in[s]} k_i$. Suppose $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $\mu(\mathcal{F}_i) > 2k_i m/n$ for $i \in [m]$ and $\mu(\mathcal{F}_i) > 2k_i s/n$ for $i \in [m+1, s]$. Then $\{\mathcal{F}_i\}_{i\in[s]}$ cross contain a matching.*

We also require the following version for the p-biased measure, which we will deduce from Lemma IV.2.6 by a limit argument similar to those in [17, 34].

Lemma IV.2.7. *Let $m \le s$ and $p_1, \dots, p_s > 0$ with $\sum_{i\in[s]} p_i \le 1/2$. Suppose that $\mathcal{F}_1, \dots \mathcal{F}_s \subset \{0,1\}^n$ are monotone families with $\mu_{p_i}\big(\mathcal{F}_i\big) \ge 3mp_i$ for $i \in [m]$ and $\mu_{p_i}\big(\mathcal{F}_i\big) \ge 3sp_i$ for $i \in [m+1, s]$. Then $\{\mathcal{F}_i\}_{i\in[s]}$ cross contain a matching.*

We introduce the following terminology. Given $\mathbf{a} = (a_1, \dots, a_s) \in \mathbb{R}^s$ and $n, k_1, \dots, k_s \ge 0$ we say $\mathbf{a}$ is *forcing* for $(n, k_1, \dots, k_s)$ if any families $\mathcal{F}_1, \dots, \mathcal{F}_s$ with $\mathcal{F}_i \subset \binom{[n]}{k_i}$ and $\mu(\mathcal{F}_i) > \frac{a_i k_i}{n}$ for all $i \in [s]$ cross contain an s-matching. We say $\mathbf{a} = (a_1, \dots, a_s) \in \mathbb{R}^s$ is *forcing* if it is forcing for $(n, k_1, \dots, k_s)$ whenever $n \ge \sum_{i\in[s]} k_i$ and *exactly forcing* if it is forcing for $(n, k_1, \dots, k_s)$ whenever $n = \sum_{i\in[s]} k_i$. Any forcing sequence is clearly exactly forcing; we establish the converse.

Lemma IV.2.8. *A sequence $\mathbf{a} \in \mathbb{R}^s$ is forcing if and only if it is exactly forcing.*

We require the following compression operators. Given distinct $i, j \in [n]$ and $F \subset [n]$, we let

$$C_{i,j}(F) := \begin{cases} (F \setminus \{j\}) \cup \{i\} & \text{if } j \in F, i \notin F; \\ F & \text{otherwise.} \end{cases}$$

Given $\mathcal{F} \subset \{0,1\}^n$, we let $C_{i,j}(\mathcal{F}) = \{C_{i,j}(F) : F \in \mathcal{F}\} \cup \{F \in \mathcal{F} : C_{i,j}(F) \in \mathcal{F}\}$. We say $\mathcal{F}$ is $C_{i,j}$-compressed if $C_{i,j}(\mathcal{F}) = \mathcal{F}$.

Proof of Lemma IV.2.8. A forcing sequence is clearly exactly forcing, so it remains to prove the converse. We argue by induction on s; the base case $s = 1$ is clear. Suppose that $\mathbf{a} \in \mathbb{R}^s$ is exactly forcing. We fix $k_1, \dots, k_s \ge 0$ and show by induction on $n \ge \sum_{i\in[s]} k_i$ that $\mathbf{a}$ is forcing for $(n, k_1, \dots, k_s)$, i.e. any families $\mathcal{F}_1, \dots, \mathcal{F}_s$ with $\mathcal{F}_i \subset \binom{[n]}{k_i}$ and $\mu(\mathcal{F}_i) > \frac{a_i k_i}{n}$ for all $i \in [s]$ cross contain an s-matching. The base case $n = \sum_{i\in[s]} k_i$ holds as $\mathbf{a}$ is exactly forcing.

First suppose $k_i = 0$ for some $i \in [s]$; without loss of generality $i = s$. Then $\mathbf{a}' = (a_1, \dots, a_{s-1})$ is exactly forcing, and so forcing by induction on s. Thus $\mathcal{F}_1, \dots, \mathcal{F}_{s-1}$ cross contain an $(s-1)$-matching. Combined with $\emptyset \in \mathcal{F}_s$ we find a cross s-matching in $\mathcal{F}_1, \dots, \mathcal{F}_s$, as required.

We may now assume $k_i \ge 1$ for all $i \in [s]$. We suppose for contradiction that $\mathcal{F}_1, \dots, \mathcal{F}_s$ do not cross contain an s-matching. Let $\mathcal{G}_1, \dots, \mathcal{G}_s$ be obtained from $\mathcal{F}_1, \dots, \mathcal{F}_s$ by successively applying the compression operators $C_{1,n}, C_{2,n}, \dots, C_{n-1,n}$. As is well-known (e.g. see [46, Lemma 2.1 (iii)]), $\mathcal{G}_1, \dots, \mathcal{G}_s$ do not cross contain an s-matching and are $C_{j,n}$-compressed for all $j \in [n-1]$. For each $i \in [s]$ let

$$\mathcal{G}_i(n) := \big\{A \subset [n-1] : A \cup \{n\} \in \mathcal{G}_i\big\} \subset \tbinom{[n-1]}{k_i - 1};$$
$$\mathcal{G}_i(\overline{n}) := \big\{A \subset [n-1] : A \in \mathcal{G}_i\big\} \subset \tbinom{[n-1]}{k_i}.$$

We now claim that if $I \subset [s]$ then $\{\mathcal{H}_i\}_{i\in[s]}$ are cross free of an s-matching, where $\mathcal{H}_i = \mathcal{G}_i(n)$ for $i \in I$ and $\mathcal{H}_i = \mathcal{G}_i(\overline{n})$ for $i \notin I$. For contradiction, suppose $\{A_i\}_{i\in[s]}$ is such a cross matching in $\{\mathcal{H}_i\}_{i\in[s]}$. Then $A_i \cup \{n\} \in \mathcal{G}_i$ for all $i \in I$ and $A_i \in \mathcal{G}_i$ for $i \notin I$. However, as $\mathcal{G}_i$ is $C_{j,n}$-compressed for all $j \in [n-1]$ and $n > \sum_{i\in[s]} k_i$, there are distinct $j_i \in [n] \setminus \big(\cup_{i\in[s]} A_i\big)$ for all $i \in I$ such that $A_i \cup \{j_i\} \in \mathcal{G}_i$. Then $\{A_i \cup \{j_i\}\}_{i\in I} \cup \{A_i\}_{i\in[s]\setminus I}$ is a cross s-matching in $\{\mathcal{G}_i\}_{i\in[s]}$, a contradiction. Thus the claim holds.

By induction on n, it now suffices to show that for each $i \in [s]$ either $\mu(\mathcal{G}_i(n)) > a_i(k_i - 1)/(n-1)$ or $\mu(\mathcal{G}_i(\overline{n})) > a_i k_i/(n-1)$; indeed, we then obtain the required contradiction by setting $I = \{i \in [s] : \mu(\mathcal{G}_i(n)) > a_i(k_i - 1)/(n-1)\}$ in the above claim. But this is clear, as otherwise

$$\frac{a_i k_i}{n} < \mu(\mathcal{G}_i) = \Big(\frac{n-k_i}{n}\Big)\mu(\mathcal{G}_i(\overline{n})) + \Big(\frac{k_i}{n}\Big)\mu(\mathcal{G}_i(n)) \le \Big(\frac{n-k_i}{n}\Big)\Big(\frac{a_i k_i}{n-1}\Big) + \Big(\frac{k_i}{n}\Big)\Big(\frac{a_i(k_i-1)}{n-1}\Big) = \frac{a_i k_i}{n},$$

a contradiction. This completes the proof. $\square$

We conclude this subsection by deducing Lemmas IV.2.6 and IV.2.7.

Proof of Lemma IV.2.6. By Lemma IV.2.8 it suffices to prove the statement under the assumption $n = \sum_{i\in[s]} k_i$. Note first that if $n = 0$ then $\mathcal{F}_i = \{\emptyset\}$ for all $i \in [s]$ which clearly cross contain an s-matching. Thus we may assume $n > 0$. For any $i \in [m]$ we have $2k_i m/n < \mu(\mathcal{F}_i) \leq 1$, so $k_i < n/2m$, and similarly $k_i < n/2s$ for $i \in [m+1, s]$. But now $n = \sum_{i\in[s]} k_i < m \cdot n/2m + (s-m)\cdot n/2s < n$ is a contradiction. □

Proof of Lemma IV.2.7. Let $N^{-1} \ll \varepsilon \ll \min_{i\in[s]} p_i$ and $\mathcal{G}_i = \mathcal{F}_i \times \{0,1\}^{[N]\setminus[n]} \subset \{0,1\}^N$ for each $i \in [s]$. Then each $\mu_{p_i}(\mathcal{G}_i) = \mu_{p_i}(\mathcal{F}_i)$. Writing $I_i = \big[(1-\varepsilon)Np_i, (1+\varepsilon)Np_i\big]$, by Lemma IV.2.3 each $\mu_{p_i}\big(\cup_{k\notin I_i}\binom{[N]}{k}\big) < \varepsilon$, so there are $k_i \in I_i$ such that each $\mu\big(\mathcal{G}_i \cap \binom{[N]}{k_i}\big) > \mu_{p_i}(\mathcal{F}_i) - \varepsilon$, which is at least $2mk_i/N$ for $i \in [m]$ and $2sk_i/N$ for $i \in [m+1, s]$. The result now follows from Lemma IV.2.6. □

IV.2.3 Capturability

In this subsection we conclude this section by proving its main lemma on cross matchings in uncapturable families. The idea of the proof is to take suitable restrictions that boost the measure of the families so that we can apply the extremal result from the previous subsection. However, uncapturability is not preserved by restrictions, so we first upgrade to globalness, which is preserved by restrictions. We also pass from the setting of uniform families to that of biased measures, which allows us to apply our sharp threshold result, and also has the technical advantage that we do not need to assume any lower bound on the uniformity of our families.

Proof of Lemma IV.2.1. Let $C \gg C_1 \gg C_2 \gg 1$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $k_i \leq n/Cs$ for $i \in [s]$. Suppose $\mathcal{F}_i$ is $(C_1 m, mk_i/C_2 n)$-uncapturable for $i \in [m]$ and $\mu(\mathcal{F}_i) > C_1 sk_i/n$ for $i > m$. We need to show that $\mathcal{F}_1, \ldots, \mathcal{F}_s$ cross contain a matching.

We start by upgrading uncapturability to globalness and moving to biased measures. By Lemma IV.1.7 with $r = C_1$ and $\beta = C_1^{-2}$ there are pairwise disjoint $S_1, \ldots, S_m$ with each $|S_i| \leq r$ such that, setting $\mathcal{G}_i = (\mathcal{F}_i^{\uparrow})_S^{S_i}$ where $S = \bigcup_i S_i$ and $p_i = k_i/(n - |S|)$, whenever $\mu_{p_i}(\mathcal{G}_i) < C_1^{-2}$ we have $S_i = \emptyset$ and $\mathcal{G}_i$ is $(C_1, 2C_1^{-2})$-global with $\mu_{p_i}(\mathcal{G}_i) > mk_i/4C_2 n > mp_i/5C_2$. We note by Lemma IV.1.5 that $\mathcal{G}_i$ is $(a, mp_i/10C_2)$-uncapturable, where $a = (mp_i/5C_2)/(4p_i C_1^{-2}) > C_1 m$.

Next we will choose pairwise disjoint $R_1, \ldots, R_m \subset [n] \setminus S$ with each $|R_i| < C_1/8$, write $R_{<j} = \bigcup_{i<j} R_i$, and define families $\mathcal{G}_i^j$ by $\mathcal{G}_i^j = (\mathcal{G}_i)_{R_{<j}}^{\emptyset}$ for $i \geq j$ or $\mathcal{G}_i^j = (\mathcal{G}_i)_{R_{<j}}^{R_i}$ for $i < j$.

We claim that we can choose each R_i to ensure $\mu_{2p_i}(\mathcal{G}_i^i) \geq 7mp_i$. To see this, first note that $\mathcal{G}_i^{i-1} = (\mathcal{G}_i)_{R_{<i}}^{\emptyset}$ has $\mu_{p_i}(\mathcal{G}_i^{i-1}) \geq mp_i/10C_2$ by uncapturability. If $\mu_{p_i}(\mathcal{G}_i^{i-1}) \geq 7mp_i$ we let $R_i = \emptyset$ to obtain $\mu_{2p_i}(\mathcal{G}_i^i) = \mu_{2p_i}(\mathcal{G}_i^{i-1}) \geq \mu_{p_i}(\mathcal{G}_i^{i-1}) \geq 7mp_i$. Otherwise, as $mp_i < 2C^{-1} \ll C_1^{-1} \ll C_2^{-1}$ we can apply Theorem IV.2.4.1 with $\varepsilon^{-1} = 70C_2$ and $r = C_1/8$ to choose R_i with $|R_i| \leq r$ so that $\mathcal{G}_i^i = (\mathcal{G}_i^{i-1})_{R_i}^{R_i}$ has $\mu_{2p_i}(\mathcal{G}_i^i) > \mu_{p_i}(\mathcal{G}_i^{i-1})/\varepsilon \geq 7mp_i$. Either way the claim holds.

By Lemma IV.1.3 each $\mathcal{G}_i^i$ with $i \in [m]$ is $(C_1/2, 4C_1^{-2})$-global, so $\mathcal{G}_i^m = (\mathcal{G}_i^i)_{\bigcup_{j>i} R_j}^{\emptyset}$ has $\mu_{2p_i}(\mathcal{G}_i^m) \geq \mu_{2p_i}(\mathcal{G}_i^i) - m(C_1/8)\cdot 4C_1^{-2} \cdot 2p_i \geq 3m(2p_i)$. For $i > m$ we have $\mu(\mathcal{F}_i) > C_1 sk_i/n$, so $\mu_{p_i}(\mathcal{G}_i^i) > \mu_{p_i}(\mathcal{F}_i)/4 - m(C_1/8)p_i > 3sp_i$. By Lemma IV.2.7, $\mathcal{G}_1^m, \ldots, \mathcal{G}_s^m$ cross contain a matching; hence so do $\mathcal{F}_1, \ldots, \mathcal{F}_s$. □

IV.3 Shadows and embeddings

In this section we will complete the proof of our junta approximation theorem by implementing the strategy described above of finding embeddings in fat shadows. We start in the first subsection by defining and analysing fat shadows. In the second subsection we find shadow embeddings. We then conclude in the final subsection with lifted embeddings (using the lifting result from the previous section) that prove Lemmas IV.1.9 and IV.1.10, thus proving Theorem IV.0.1.

IV.3.1 Fat shadows

In this subsection we present various lower bounds on the density of fat shadows, defined as follows.

Definition IV.3.1. The c-fat r-shadow of $\mathcal{F} \subset \binom{[n]}{k}$ is $\partial^r_c \mathcal{F} := \{A \in \binom{[n]}{r} : \mu(\mathcal{F}^A_A) \ge c\}$.
The c-fat shadow of $\mathcal{F}$ is $\partial_c \mathcal{F} := \bigcup_{r \le k} \partial^r_c \mathcal{F}$.

The following simple 'Markov' bound is useful when $\mathcal{F}$ is nearly complete.

Lemma IV.3.2. *If $\mu(\mathcal{F}) \ge 1 - cc'$ then $\mu(\partial^r_{1-c}\mathcal{F}) \ge 1 - c'$.*

Proof. Consider uniformly random $A \subset B \subset [n]$ with $|A| = r$ and $|B| = k$. For any $A \notin \partial^r_{1-c}\mathcal{F}$ we have $\mathbb{P}(B \notin \mathcal{F} \mid A) \ge c$, so $cc' \ge \mathbb{P}(B \notin \mathcal{F}) \ge c \cdot \mathbb{P}(A \notin \partial^r_{1-c}\mathcal{F}) = c(1 - \mu(\partial^r_{1-c}\mathcal{F}))$. □

Another bound is given the following Fairness Proposition of Keller and Lifshitz [57].

Proposition IV.3.3 (Fairness Proposition). *Let $C \gg r/\varepsilon$ and $\mathcal{F} \subset \binom{[n]}{k}$ with $k \ge r$ and $\mu(\mathcal{F}) \ge e^{-k/C}$. For $c = (1-\varepsilon)\mu(\mathcal{F})$ we have $\mu(\partial^r_c \mathcal{F}) \ge 1 - \varepsilon$.*

When the above bounds are not applicable we rely on the following lemma, whose proof will occupy the remainder of this subsection.

Lemma IV.3.4. *Let $\mathcal{F} \subset \binom{[n]}{k}$, $r < \ell \le k$ and $\mathcal{H} = \{B \in \binom{[n]}{\ell} : \partial^r B \subset \partial^r_c \mathcal{F}\}$, where $c = \mu(\mathcal{F})/2\binom{\ell}{r}$. Then $\mu(\mathcal{H}) \ge \mu(\mathcal{F})/2$. Thus $\mu(\partial^r_c\mathcal{F}) \ge (\mu(\mathcal{F})/2)^{r/\ell}$. Furthermore, if $G \in \mathcal{G}'(r,s,\Delta)$, $C \gg r\Delta$ and $\partial^r_c\mathcal{F}$ is G-free then $\mu(\partial^r_c\mathcal{F}) \ge \big((\mu(\mathcal{F})/2 - (s/n)^{\ell/C})n/s\ell^2\big)^{r/(\ell-1)}$.*

We require several further lemmas for the proof of Lemma IV.3.4. We start by stating a consequence of the Lovász form [66] of the Kruskal–Katona theorem [51, 63].

Lemma IV.3.5. *If $1 \le \ell \le k \le n$ and $\mathcal{A} \subset \binom{[n]}{k}$ then $\mu(\partial^\ell(\mathcal{A})) \ge \mu(\mathcal{A})^{\ell/k}$.*

Proof. We define $\beta \in [0,1]$ by $|\mathcal{A}| = \binom{\beta n}{k}$, so that $\mu(\mathcal{A}) = \prod_{i=0}^{k-1}(\beta - i/n)$. By the Lovász form of Kruskal–Katona, we have $|\partial^\ell \mathcal{A}| \ge \binom{\beta n}{\ell}$, so $\mu(\partial^\ell(\mathcal{A}))^k \ge \prod_{i=0}^{\ell-1}(\beta - i/n)^k \ge \mu(\mathcal{A})^\ell$. □

Next we require an estimate on the Turán numbers of r-partite r-graphs, which follows from [14, Theorem 2] due to Conlon, Fox and Sudakov. (Recall that $\mathcal{G}'(r,s,\Delta)$ is the family of r-partite r-graphs with s edges and maximum degree Δ.)

Theorem IV.3.6. *Let $F \in \mathcal{G}'(r,s,\Delta)$ and $C \gg r\Delta$. Then any F-free $\mathcal{H} \subset \binom{[n]}{r}$ with $n > Cs$ has $\mu(\mathcal{H}) < (s/n)^{1/C}$.*

We note that the following lemma is immediate from Theorem IV.3.6 and Lemma IV.3.5.

Lemma IV.3.7. *Let $G \in \mathcal{G}'(r,s,\Delta)$, $C \gg r\Delta$, $C \le k \le n/Cs$ and $\mathcal{F} \subset \binom{[n]}{k}$. If $\partial^r\mathcal{F}$ is G-free then $\mu(\mathcal{F}) \le (s/n)^{k/C}$.*

Our next lemma is an adaptation of one due to Kostochka, Mubayi and Verstraëte [61].

Lemma IV.3.8. *Suppose $G \in \mathcal{G}'(r,s,\Delta)$, $C \gg r\Delta$ and $\mathcal{F}$ is a G^+-free k-graph on $[n]$. Then $\mu(\partial\mathcal{F}) \ge (\mu(\mathcal{F}) - (s/n)^{k/C})n/sk^2$.*

Proof. We define $\mathcal{G} \subset \mathcal{F}$ by starting with $\mathcal{G} = \mathcal{F}$ and then repeating the following procedure: if there is any $A \in \partial\mathcal{G}$ with $|\mathcal{G}^A_A| \le ks$ then remove from $\mathcal{G}$ all edges containing A. This terminates with some $\mathcal{G}$ such that $|\mathcal{G}^A_A| > rs$ for all $A \in \partial\mathcal{G}$ and $|\mathcal{G}| \ge |\mathcal{F}| - ks|\partial\mathcal{F}|$, so $\mu(\partial\mathcal{F}) \ge (\mu(\mathcal{F}) - \mu(\mathcal{G}))n/sk^2$.

We will now show that $\partial_r\mathcal{G}$ is G-free, which will complete the proof due to Lemma IV.3.7. To see this, we suppose that $\phi(G)$ is a copy of G in $\partial_r\mathcal{G}$ and will obtain a contradiction by finding a copy of G^+ in $\mathcal{G}$. To do so, we start by fixing for each edge A of G an edge e_A of $\mathcal{G}$ containing $\phi(A)$. Then we repeat the following procedure: while some e_A contains some $\phi(x)$ with $x \notin A$, replace e_A by some edge $(e_A \setminus \{\phi(x)\}) \cup \{v\}$ with $v \notin \operatorname{Im}\phi$. As $|\mathcal{G}^A_A| > ks$ for all $A \in \partial\mathcal{G}$ we can always choose v as required. The procedure terminates with a copy of G^+, so the proof is complete. □

We conclude this subsection with the proof of its main lemma.

Proof of Lemma IV.3.4. Consider uniformly random (A,B,C) with $C \subset B \subset A \subset [n]$ and $|C| = r$, $|B| = \ell$, $|A| = k$. Write $p = \mathbb{P}(A \in \mathcal{F}, C \notin \partial^r_c \mathcal{F})$ and $q = \mathbb{P}(A \in \mathcal{F}, B \notin \mathcal{H})$.

For any $C \notin \partial^r_c \mathcal{F}$ we have $\mathbb{P}(A \in \mathcal{F} \mid C) = \mu(\mathcal{F}^C_C) \le c$, so $p \le c$. On the other hand, $p \ge q\binom{\ell}{r}^{-1}$, as for any $A \in \mathcal{F}$ and $B \notin \mathcal{H}$ we have $\mathbb{P}(C \notin \partial^r_c \mathcal{F} \mid A, B) \ge \binom{\ell}{r}^{-1}$. We deduce $q \le \binom{\ell}{r} c = \mu(\mathcal{F})/2$.

Thus $\mu(\mathcal{H}) = \mathbb{P}(B \in \mathcal{H}) \ge \mathbb{P}(A \in \mathcal{F}) - q \ge \mu(\mathcal{F})/2$.

As $\partial^r \mathcal{H} \subset \partial^r_c \mathcal{F}$, Lemma IV.3.5 gives $\mu(\partial^r_c \mathcal{F}) \ge (\mu(\mathcal{F})/2)^{r/\ell}$.

Now suppose $G \in \mathcal{G}'(r,s,\Delta)$ and $\partial^r_c \mathcal{F}$ is G-free. Then $\mathcal{H}$ is G^+-free, so Lemma IV.3.8 gives $\mu(\partial \mathcal{H}) \ge (\mu(\mathcal{H}) - (s/n)^{\ell/C}) n / s\ell^2$. As $\partial^r \partial \mathcal{H} \subset \partial^r_c \mathcal{F}$, Lemma IV.3.5 gives the required bound. □

IV.3.2 Shadow embeddings

The following lemma implements a simple greedy algorithm for cross embedding any bounded degree r-graph in a collection of nearly complete r-graphs (more generally, we also allow smaller edges).

Lemma IV.3.9. *Let $0 < \eta \ll (r\Delta)^{-1}$ and $G = \{e_1, \dots, e_s\}$ be a hypergraph of maximum degree Δ with each $|e_i| = r_i \le r$. Suppose for each $i \in [s]$ that $\mathcal{G}_i$ is an r_i-graph on $[n]$, where $n \ge 2rs$ and $\mu(\mathcal{G}_i) > 1 - \eta$. Then $\mathcal{G}_1, \dots, \mathcal{G}_s$ cross contain G.*

Proof. Write $V(G) = \{v_1, \dots, v_m\}$. We may assume that G has no isolated vertices, so $m \le \sum_i d_G(v_i) \le rs \le n/2$. We will construct an injection $\phi : V(G) \to [n]$ such that each $\phi(e_j) \in \mathcal{G}_j$. To do so, we define ϕ sequentially so that, for each $0 \le t \le m$ the definition of ϕ on $V_t := \{v_i : i \le t\}$ is *t-good*, meaning that for each edge e_j we have

$$\phi(e_j \cap V_t) \in \partial_{c_{jt}} \mathcal{G}_j, \text{ where } c_{jt} = 1 - \eta(2\Delta)^{|e_j \cap V_t|}. \tag{9}$$

Note that (9) holds whenever $e_j \cap V_t = \emptyset$, as $\mu(\mathcal{G}_j) > 1 - \eta$; in particular, (9) holds when $t = 0$.

It remains to show for any $0 \le t < m$ that we can extend any t-good embedding ϕ to a $(t+1)$-good embedding. To see this, first note that we only need to check (9) when e_j is one of at most Δ edges containing v_{t+1}. Fix any such edge e_j, let $f = \phi(e_j \cap V_t)$, and let B_j be the set of $x \in [n]$ such that choosing $\phi(v_{t+1}) = x$ would give $\phi(e_j \cap V_{t+1}) = f \cup \{x\} \notin \partial_{c_{j(t+1)}} \mathcal{G}_j$. Then

$$|B_j| \eta (2\Delta)^{|f|+1} \le \sum_{x \in B} \left(1 - \mu\big((\mathcal{G}_j)^{f \cup \{x\}}_{f \cup \{x\}}\big)\right) \le n(1 - \mu((\mathcal{G}_j)^f_f)) < n\eta(2\Delta)^{|f|},$$

so $|B_j| < n/2\Delta$. Summing over at most Δ choices of j forbids fewer than $n/2$ choices of x. The requirement that ϕ be injective also forbids fewer than $n/2$ vertices, so we can extend ϕ as required. □

IV.3.3 Lifted embeddings

We conclude this section by proving the two embedding lemmas assumed above, thus completing the proof of Theorem IV.0.1.

Proof of Lemma IV.1.9. Suppose $n, s, k_1, \dots, k_s \in \mathbb{N}$ with $C \le k_i \le \frac{n}{Cs}$ for all $i \in [s]$, and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with each $\mu(\mathcal{F}_i) \ge e^{-k_i/C} + Csk_i/n$. Let η be as in Lemma IV.3.9. We can assume C is large enough so that Proposition IV.3.3 gives $\mu(\mathcal{G}_i) \ge 1 - \eta$ for each $i \in [s]$, where $\mathcal{G}_i$ is the r-graph on $[n]$ consisting of all $e \in \binom{[n]}{r}$ with $\mu((\mathcal{F}_i)^e_e) \ge Csk_i/2n$. By Lemma IV.3.9 we can find $R_1, \dots, R_s$ forming a copy of G with $R_i \in \mathcal{G}_i$ for all $i \in [s]$. Let $R = R_1 \cup \dots \cup R_s$. By the union bound, each $\mu\big((\mathcal{F}_i)^{R_i}_R\big) \ge \mu\big((\mathcal{F}_i)^{R_i}_{R_i}\big) - |R|k_i/n \ge Csk_i/4n$ for $C \ge 8$, so Lemma IV.2.5 gives a cross matching $E_1, \dots, E_s$ in $(\mathcal{F}_1)^{R_1}_R, \dots, (\mathcal{F}_s)^{R_s}_R$. Now $\mathcal{F}_1, \dots, \mathcal{F}_s$ cross contain a copy of G^+ with edges $R_1 \cup E_1, \dots, R_s \cup E_s$. □

Proof of Lemma IV.1.10. Let $G \in \mathcal{G}(r,s,\Delta)$ and $C \gg C_1 \gg C_2 \gg r\Delta$. Suppose for a contradiction that $\mathcal{F} \subset \binom{[n]}{k}$ with $C \le k \le n/Cs$ is $(C_1 s, sk/C_2 n)$-uncapturable but G^+-free.

Let $\mathcal{B}$ be a maximal collection of pairwise disjoint sets where each $B \in \mathcal{B}$ has $|B| \le r+1$ and $\mu(\mathcal{F}_B^B) > \beta := e^{-k/C_1} + C_1 sk/n$. We claim that $|\mathcal{B}| < s$. To see this, suppose for a contradiction that we have distinct $B_1, \dots, B_s$ in $\mathcal{B}$. Let $B = \bigcup_{i=1}^s B_i$ and $\mathcal{F}_i = \mathcal{F}_B^{B_i}$ for $i \in [s]$. Then each $\mu(\mathcal{F}_i) > \beta - |B|k/n > e^{-k/C_1} + C_1 sk/2n$. Now Lemma IV.1.9 gives a cross copy of G^+ in $\mathcal{F}_1, \dots, \mathcal{F}_s$, contradicting $\mathcal{F}$ being G^+-free, so $|\mathcal{B}| < s$, as claimed.

Now let $\mathcal{G} = \mathcal{F}_B^\emptyset$ with $B = \bigcup \mathcal{B}$. Then $\mathcal{G}$ is $(r+1, 2\beta)$-global by definition of $\mathcal{B}$ and $\mu(\mathcal{G}) > sk/C_2 n$ by uncapturability of $\mathcal{F}$. Let $\mathcal{H} = \{B \in \binom{[n]}{C_2} : \partial^r B \subset \partial_c^r \mathcal{G}\}$, where $c = \mu(\mathcal{G})/2\binom{C_2}{r} > sk/nC_2^{2r}$. We have $\mu(\mathcal{H}) \ge \mu(\mathcal{G})/2$ by Lemma IV.3.4. We will show that $\partial^r \mathcal{H}$ is G-free. Then Lemma IV.3.7 with $C_2/2 \gg r\Delta$ in place of C will give the contradiction $sk/C_2 n < \mu(\mathcal{G}) \le 2\mu(\mathcal{H}) \le (s/n)^2$.

It remains to show that $\partial^r \mathcal{H}$ is G-free. Suppose for a contradiction that $A_1, \dots, A_s$ is a copy of G in $\partial^r \mathcal{H}$. Let $A = \bigcup_{i=1}^s A_i$ and $\mathcal{G}_i = \mathcal{G}_A^{A_i}$ for $i \in [s]$. Then each $\mathcal{G}_i$ is $(1, 4\beta)$-global by Lemma IV.1.2 with $\mu(\mathcal{G}_i) > c - |A| \cdot 2\beta k/n > c/2$. Now each $\mathcal{G}_i$ is $(C_1 s, c/4)$-uncapturable by Lemma IV.1.4, so $\mathcal{G}_1, \dots, \mathcal{G}_s$ cross contain a matching by Lemma IV.2.1 with $m = s$. However, this contradicts $\mathcal{F}$ being G^+-free. □

IV.4 Refined junta approximation

In this final section of the part we will prove Theorem IV.0.2, our refined junta approximation result, which will play a key role in the proofs of our results in the next part. We start in the first subsection by setting out the strategy of the proof and implementing it assuming an embedding lemma, whose proof will then occupy the remainder of the section.

IV.4.1 Strategy

Our embedding strategy considers a setup below that blends the two embedding strategies used in the proof of Theorem I.4.8: it has elements of Lemma IV.1.9 (mapping a crosscut to a junta) and of Lemma IV.1.10 (embedding in the fat shadow and lifting via uncapturability).

Setup IV.4.1. Let $G \in \mathcal{G}'(r,s,\Delta)$. Let S be a crosscut in $G^+(r+1)$ with $|S| = \sigma := \sigma(G)$. Suppose $S_1 \subset S$ with $|S_1| = \sigma_1 \le \sigma$ and $\{G_x^x : x \in S_1\}$ vertex disjoint. Let $H_1, \dots, H_{\sigma_1}$ be the inclusive links $G * x = \{e \in G : x \in e\}$ for $x \in S_1$ and $H_{\sigma_1+1}, \dots, H_\sigma$ be the exclusive links G_x^x for $x \in S \setminus S_1$. Let $V_1 = \bigcup_{i=1}^{\sigma_1} V(H_i)$ and suppose $\{j : V(H_j) \cap V_1 \ne \emptyset\} = [\sigma_2]$. Let $H_i' = H_i$ for $i \in [\sigma_1]$ and $H_i' = \{e \cap V_1 : e \in H_i\}$ for $i \in [\sigma_1+1, \sigma_2]$.

We note that $\sigma \le s \le \Delta\sigma$. To use Setup IV.4.1 for embedding G^+ in $\mathcal{F} \subset \binom{[n]}{k}$ it suffices to find $J = \{j_{\sigma_1+1}, \dots, j_\sigma\} \subset [n]$ and a cross copy of $H_1^+, \dots, H_\sigma^+$ in $\mathcal{F}_1, \dots, \mathcal{F}_\sigma$, where $\mathcal{F}_i = \mathcal{F}_J^\emptyset$ for $i \in [\sigma_1]$ and $\mathcal{F}_i = \mathcal{F}_J^{j_i}$ for $i \in [\sigma_1+1, \sigma]$. This will be achieved by the following lemma.

Lemma IV.4.2. *Let $C \gg C_1 \gg \theta^{-1} \gg \varepsilon^{-1} \gg r\Delta$ and $C < k < n/Cs$. Let $G, H_1, \dots, H_\sigma$ be as in Setup IV.4.1 with $\sigma_1 \le \theta\sigma$. Let $\mathcal{F}_i \subset \binom{[n]}{k}$ for $i \in [\sigma_1]$ and $\mathcal{F}_i \subset \binom{[n]}{k-1}$ for $i \in [\sigma_1+1, \sigma]$. Suppose $\mathcal{F}_i$ is $(C_1\sigma_1, \varepsilon\sigma_1 k/n)$-uncapturable for $i \in [\sigma_1]$, that $\mu(\mathcal{F}_i) \ge 1-\theta$ for $i \in [\sigma_1+1, \sigma_2]$, and $\mu(\mathcal{F}_i) \ge \beta := e^{-k/C_1} + C_1 sk/n$ for $i \in [\sigma_2+1, \sigma]$. Then $\mathcal{F}_1, \dots, \mathcal{F}_\sigma$ cross contain $H_1^+, \dots, H_\sigma^+$.*

Next we deduce Theorem IV.0.2 from Lemma IV.4.2.

Proof of Theorem IV.0.2. Let $G \in \mathcal{G}(r,s,\Delta)$ with $\sigma(G) = \sigma$ and $C \gg C_1 \gg \theta^{-1} \gg \delta^{-1} \gg \varepsilon^{-1} \gg r\Delta$. Suppose $\mathcal{F} \subset \binom{[n]}{k}$ with $C \le k \le n/Cs$ is G^+-free with $|\mathcal{F}| > |\mathcal{S}_{n,k,\sigma-1}| - \delta\binom{n-1}{k-1}$. We need to find $J \in \binom{[n]}{\sigma-1}$ with $|\mathcal{F}_J^\emptyset| \le \varepsilon\binom{n-1}{k-1}$.

As in the proof of Theorem I.4.8 we let $J = \{i \in [n] : \mu(\mathcal{F}_i^i) \ge \beta\}$, where $\beta := e^{-k/C_1} + C_1 sk/n$. We recall that $|J| \le \sigma - 1$ and $\mathcal{F}_J^\emptyset$ is $(a, \mu(\mathcal{F}_J^\emptyset)/2)$-uncapturable with $a = \mu(\mathcal{F}_J^\emptyset)n/4k\beta$. Replacing '$\varepsilon$' in that proof by $.1\theta^2$ we obtain $|\mathcal{F}_J^\emptyset| \le .1\theta^2|\mathcal{S}_{n,k,\sigma-1}| \le .2\theta^2(\sigma-1)\binom{n-1}{k-1}$. We may assume $\sigma \ge 2\theta^{-1}$, otherwise $|\mathcal{F}_J^\emptyset| \le \theta\binom{n-1}{k-1}$. As $|\mathcal{F}_J^\emptyset| \ge |\mathcal{F}| - |\mathcal{S}_{n,k,J}| \ge (.9(\sigma-1-|J|)-\delta)\binom{n-1}{k-1}$ we deduce $|J| > (1-.3\theta^2)(\sigma-1)$, so $1 \le \sigma_1 := \sigma - |J| \le 1 + .3\theta^2\sigma \le \theta\sigma$.

Now we let $S, S_1, H_1, \dots, H_\sigma$ be as in Setup IV.4.1, where we can greedily choose $S_1 \subset S$ with $|S_1| = \sigma_1$ such that $\{G_x^x : x \in S_1\}$ are vertex disjoint, as any partial choice of S_1 forbids at most $\sigma_1(\Delta r)^2 < \sigma$ vertices of S. We write $J = \{j_{\sigma_1+1}, \dots, j_\sigma\}$, let $\mathcal{F}_i = \mathcal{F}_J^\emptyset$ for $i \in [\sigma_1]$ and $\mathcal{F}_i = \mathcal{F}_J^{j_i}$ for $i \in [\sigma_1+1, \sigma]$, where we can assume $|\mathcal{F}_{\sigma_1+1}| \ge \dots \ge |\mathcal{F}_\sigma|$. We note that $\mu(\mathcal{F}_{\sigma_2}) > 1-\theta$, as otherwise we would have the contradiction $|\mathcal{F}| < |\mathcal{F}_J^\emptyset| + \big(\sigma_2 - \sigma_1 + (\sigma-\sigma_2)(1-\theta)\big)\binom{n-1}{k-1} < \big((1+.2\theta^2)\sigma - \sigma_1 - \theta(\sigma-\sigma_2)\big)\binom{n-1}{k-1} < |\mathcal{S}_{n,k,\sigma-1}| - \delta\binom{n-1}{k-1}$.

Now we must have $\mu(\mathcal{F}_J^\emptyset) \le \varepsilon\sigma_1 k/n$; otherwise $\mathcal{F}_J^\emptyset$ is $(C_1\sigma_1, \varepsilon\sigma_1 k/2n)$-uncapturable, so $\mathcal{F}_1, \dots, \mathcal{F}_\sigma$ cross contain $H_1^+, \dots, H_\sigma^+$ by Lemma IV.4.2, contradicting $\mathcal{F}$ being G^+-free. As $|\mathcal{F}_J^\emptyset| \ge |\mathcal{F}| - |\mathcal{S}_{n,k,J}| \ge (.9(\sigma_1-1)-\delta)\binom{n-1}{k-1}$ we deduce $.9(\sigma_1-1)-\delta \le \varepsilon\sigma_1$, so $\sigma_1 = 1$ and $\mu(\mathcal{F}_J^\emptyset) \le \varepsilon k/n$. □

The remainder of the section will be devoted to the proof of Lemma IV.4.2. Similarly to the proofs of our previous embedding results (Lemmas IV.1.9 and IV.1.10), the strategy will be to find shadow embeddings and then lifting embeddings. However, there are further technical challenges to overcome in the current setting, particularly when the uniformity k of our families is small, when we need to 'pause' the shadow embedding after embedding $H_i' = H_i$ for $i \in [\sigma_1]$, then lift this part of the embedding, then complete the shadow embedding, and finally lift the remainder of the embedding. The shadow embedding lemma will be presented in the next subsection. The third subsection contains further results on upgrading uncapturability to globalness, which we call 'enhanced upgrading', as they obtain globalness parameters that are significantly stronger than one might expect, and this will be a crucial technical ingredient of the proof. In the fourth subsection we establish an improved lifting result that allows for a much weaker uncapturability assumption than that in Lemma IV.2.1. We conclude with the proof of Lemma IV.4.2 in the final subsection.

IV.4.2 Shadow embeddings

Here we extend the argument used in Lemma IV.3.9 to prove the following lemma that will be applied to show that the fat shadows of $\mathcal{F}_1, \dots, \mathcal{F}_\sigma$ as in Lemma IV.4.2 cross contain $H_1, \dots, H_\sigma$. Whereas before we were embedding into nearly complete hypergraphs, now many of our hypergraphs will be quite sparse, which makes the embedding more challenging: the idea is to replace the naive greedy arguments by Theorem IV.3.6, here making key use of our observation that we can assume G is r-partite.

Lemma IV.4.3. *Let $C \gg \eta^{-1} \gg K \gg r\Delta$ and $0 < \theta < \eta$. Let $G, H_1, \dots, H_\sigma$ be as in Setup IV.4.1 and $\mathcal{G}_1, \dots, \mathcal{G}_\sigma \subset \binom{[n]}{r}$ with $n > C\sigma$. Suppose $\mu(\mathcal{G}_i) \ge 1-\eta$ for $i \in [\sigma_2+1, \sigma]$, $\mu(\mathcal{G}_i) \ge 1-\theta$ for $i \in [\sigma_1+1, \sigma_2]$ and $\mu(\mathcal{G}_i) \ge \theta^{1/2r} + n^{-1/K} + r\Delta\sigma_1/n$ for $i \in [\sigma_1]$. Let $c = 1-\theta^{1/r}$. Then $\partial_c\mathcal{G}_1, \dots, \partial_c\mathcal{G}_{\sigma_2}$ cross contain $H_1', \dots, H_{\sigma_2}'$ and $\mathcal{G}_1, \dots, \mathcal{G}_\sigma$ cross contain $H_1, \dots, H_\sigma$.*

Proof. For each $i \in [\sigma_1+1, \sigma_2]$ we define $\mathcal{G}_i^r, \dots, \mathcal{G}_i^0$ recursively by $\mathcal{G}_i^r = \mathcal{G}_i$ and $\mathcal{G}_i^{j-1} = \partial^{j-1}_{1-\theta^{1/r}}\mathcal{G}_i^j$ for $j \in [r]$. Clearly each $\mathcal{G}_i^j \subset \partial_{c_j}\mathcal{G}_i$ where $c_j = 1-(r-j)\theta^{1/r}$.

We claim that each $\mu(\mathcal{G}_i^j) \ge 1-\theta^{j/r}$. To see this, we argue by induction on $r-j$. For $r-j=0$ we have $\mu(\mathcal{G}_i^r) \ge 1-\theta$ by assumption. For the induction step, consider any $j \in [r]$ and uniformly random $A \subset B \subset [n]$ with $|A| = j-1$ and $|B| = j$. Given any $A \notin \mathcal{G}_i^{j-1}$ we have $\mathbb{P}(B \notin \mathcal{G}_i^j) \ge \theta^{1/r}$, so $1-\mu(\mathcal{G}_i^j) \ge \theta^{1/r}(1-\mu(\mathcal{G}_i^{j-1}))$. The claim follows.

Next we will construct a cross embedding ϕ of $H_1', \dots, H_{\sigma_2}'$ in $\partial_c\mathcal{G}_1, \dots, \partial_c\mathcal{G}_{\sigma_2}$. We recall that $H_i' = H_i$ for $i \in [\sigma_1]$ and all H_i' are defined on V_1, which is the disjoint union of $V(H_1), \dots, V(H_{\sigma_1})$. We proceed in σ_1 steps, defining ϕ on $V(H_t)$ at step t. When ϕ has been defined on $U_t := \bigcup_{i\le t} V(H_i)$, we say ϕ is *t-good* if $\phi(e \cap U_t) \in \mathcal{G}_i^{|e\cap U_t|}$ for each $i \in [\sigma_2]$ and $e \in \mathcal{G}_i$ with $e \cap U_t \neq \emptyset$.

We note that if ϕ is t-good then $\phi(H_i) \subset \mathcal{G}_i^r = \mathcal{G}_i = \partial_c \mathcal{G}_i$ for all $i \in [t]$ and if ϕ is σ_1-good then $\phi(H_i) \subset \partial_c \mathcal{G}_i$ for all $i \in [\sigma_2]$. As ϕ defined on $U_0 = \emptyset$ is trivially 0-good, it remains to show for any $t \in [\sigma_1]$ that we can extend any $(t-1)$-good ϕ to a t-good embedding.

For clarity of exposition, we start by showing the case $t = 1$. Obtain $\mathcal{H}_1$ from $\mathcal{G}_1$ by removing any edge e such that $f \notin \mathcal{G}_i^{|f|}$ for some $\emptyset \neq f \subset e$ and $i \in [\sigma_2]$ with $V(H_i) \cap V(H_1) \neq \emptyset$. There are at most $r\Delta^2$ such i, so by a union bound and the above claim we have $\mu(\mathcal{H}_1) \geq \mu(\mathcal{G}_1) - r\Delta^2 2^r \theta^{1/r} > n^{-1/K}$. We can assume that G is r-partite, so by Theorem IV.3.6 we can find an embedding ϕ'_1 of $N_1 := \{e \in G : e \cap V(H_1) \neq \emptyset\}$ in $\mathcal{H}_1$. Now $\phi = \phi' \mid_{V(H_1)}$ is 1-good.

Now we consider general $t \in [\sigma_1]$. Obtain $\mathcal{H}_t$ from $(\mathcal{G}_t)^{\emptyset}_{\phi(U_{t-1})}$ by removing any edge e such that $f \notin \mathcal{G}_i^{|f|}$ for some $\emptyset \neq f \setminus \phi(A') \subset e$ where $A \in H_i$ with $V(H_i) \cap V(H_t) \neq \emptyset$ and $A' = A \cap U_{t-1}$. For any such non-empty A', as ϕ is $(t-1)$-good we have $\phi(A') \in \mathcal{G}_i^{|A'|}$, so $\mu((\mathcal{G}_i^j)^{A'}_{A'}) \geq 1 - (j - |A'|)\theta^{1/r}$ for any $|A'| \leq j \leq r$. Thus a union bound gives $\mu(\mathcal{H}_t) \geq \mu(\mathcal{G}_t) - |U_{t-1}|k/n - r\Delta^2 2^r r\theta^{1/r} > n^{-1/K}$. Now as in the case $t = 1$ we obtain a t-good extension by embedding $N_t := \{e \in G : e \cap V(H_t) \neq \emptyset\}$ in $\mathcal{H}_t$ and restricting to $V(H_t)$.

Thus we have constructed a cross embedding ϕ of $H'_1, \dots, H'_{\sigma_2}$ in $\partial_c \mathcal{G}_1, \dots, \partial_c \mathcal{G}_{\sigma_2}$. To complete the proof we extend ϕ to a cross embedding $H_1, \dots, H_\sigma$ in $\mathcal{G}_1, \dots, \mathcal{G}_\sigma$, which requires $\phi(e \setminus V_1) \in (\mathcal{G}_i)^{e \cap V_1}_{e \cap V_1}$ for all $e \in H_i$, $i \in [\sigma_1 + 1, \sigma]$; this is possible by Lemma IV.3.9. □

IV.4.3 Enhanced upgrading

This subsection provides further results on upgrading uncapturability to globalness with enhanced parameters that will be crucial in later proofs. We start by showing that every family has a restriction that is global or large.

Lemma IV.4.4. *Let $b, r \in \mathbb{N}$, $\alpha > 1$ and $\mathcal{F} \subset \binom{[n]}{k}$ with $k \geq br$. Then there is $B \subset [n]$ with $|B| \leq br$ such that if $\mu(\mathcal{F}^B_B) < \alpha^b \mu(\mathcal{F})$ then $\mathcal{F}^B_B$ is $(r, \alpha\mu(\mathcal{F}^B_B))$-global with $\mu(\mathcal{F}^B_B) \geq \alpha^{1_{B \neq \emptyset}} \mu(\mathcal{F})$.*

Proof. We consider $\mathcal{F}_0, \mathcal{F}_1, \dots$, where $\mathcal{F}_0 = \mathcal{F}$, and if $i < b$ and $\mathcal{F}_i$ is not $(r, \alpha\mu(\mathcal{F}_i))$-global then we let $\mathcal{F}_{i+1} = (\mathcal{F}_i)^{B_i}_{B_i}$ so that $|B_i| \leq r$ and $\mu(\mathcal{F}_{i+1}) > \alpha\mu(\mathcal{F}_i)$. When this sequence terminates at some $\mathcal{F}_t$ we let $B = \bigcup_{i \leq t} B_i$. Clearly $\mathcal{F}^B_B = \mathcal{F}_t$ has the required properties. □

By iterating the previous result we obtain the following upgrading lemma.

Lemma IV.4.5. *Suppose $b, r, m \in \mathbb{N}$ and for each $i \in [m]$ that $\alpha_i > 1$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $rb \leq k_i \leq n/2rm\alpha_i$ is (rbm, β_i)-uncapturable with $\alpha_i^b \beta_i > 2rmk_i/n$. Then there are disjoint $B_1, \dots, B_m$ with each $|B_i| \leq rb$ such that, setting $\mathcal{G}_i = (\mathcal{F}_i)^{B_i}_B$ where $B = \bigcup_i B_i$, if $\mu(\mathcal{G}_i) < \alpha_i^b \beta_i/2$ then $\mathcal{G}_i$ is $(r, 4\alpha_i\mu(\mathcal{G}_i))$-global with $\mu(\mathcal{G}_i) > \alpha_i^{1_{B_i \neq \emptyset}} \beta_i/2$.*

Proof. We will choose $B_1, \dots, B_m$ sequentially and define $\mathcal{F}_i^0, \dots, \mathcal{F}_i^m$ for $i \in [m]$ by $\mathcal{F}_i^0 = \mathcal{F}_i$, $\mathcal{F}_j^i = (\mathcal{F}_j^{i-1})^{\emptyset}_{B_i}$ for $j \neq i$ and $\mathcal{F}_i^i = (\mathcal{F}_i^{i-1})^{B_i}_{B_i}$. At step i, we have $\mu(\mathcal{F}_i^{i-1}) \geq \beta_i$ by uncapturability of $\mathcal{F}_i$, so by Lemma IV.4.4 we can choose B_i with $|B_i| \leq rb$ such that if $\mu(\mathcal{F}_i^i) < \alpha_i^b \mu(\mathcal{F}_i^{i-1})$ then $\mathcal{F}_i^i$ is $(r, \alpha\mu(\mathcal{F}_i^i))$-global with $\mu(\mathcal{F}_i^i) \geq \alpha_i^{1_{B_i \neq \emptyset}} \beta_i$. After step m, for any $i \in [m]$ we have $\mathcal{G}_i^m = \mathcal{G}_i = (\mathcal{F}_i)^{B_i}_B$. If $\mu(\mathcal{F}_i^i) \geq \alpha_i^b \mu(\mathcal{F}_i^{i-1})$ then $\mu(\mathcal{G}_i) \geq \alpha_i^b \beta_i - rmk_i/n \geq \alpha_i^b \beta_i/2$. Otherwise, $\mathcal{F}_i^i$ is $(r, \alpha_i\mu(\mathcal{F}_i^i))$-global with $\mu(\mathcal{F}_i^i) \geq \alpha_i^{1_{B_i \neq \emptyset}} \mu(\mathcal{F})$, and $(n/2k_i\alpha_i, \mu(\mathcal{F}_i^i)/2)$-uncapturable by Lemma IV.1.4, so $\mu(\mathcal{G}_i) > \mu(\mathcal{F}_i^i)/2 \geq \alpha_i^{1_{B_i \neq \emptyset}} \beta_i/2$, and $\mathcal{G}_i$ is $(r, 4\alpha_i\mu(\mathcal{G}_i))$-global by Lemma IV.1.2. □

For our final upgrading lemma we apply the previous one twice: the idea is that the globalness from the first application provides the second application with much better uncapturability.

Lemma IV.4.6. *Suppose $b, r, m \in \mathbb{N}$ and for each $i \in [m]$ that $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $rb \leq k_i \leq n/2rmb^2$ is $(2m, \beta_i)$-uncapturable with $\beta_i > 8rmk_i/bn$. Then there are disjoint $B_1, \dots, B_m$ with each $|B_i| \leq rb + 2$ such that, setting $\mathcal{G}_i = (\mathcal{F}_i)^{B_i}_B$ where $B = \bigcup_i B_i$, if $\mu(\mathcal{G}_i) < 2^b \beta_i/8$ then $\mathcal{G}_i$ is $(r, 8\mu(\mathcal{G}_i))$-global with $\mu(\mathcal{G}_i) > 2^{1_{B_i \neq \emptyset}} \beta_i/8$.*

Proof. We start by applying Lemma IV.4.5 with $(b, 1, 2)$ in place of (α_i, r, b). This gives disjoint $S_1, \dots, S_m$ with each $|S_i| \le 2$ such that, setting $\mathcal{H}_i = (\mathcal{F}_i)^{S_i}_S$ where $S = \bigcup_i S_i$, if $\mu(\mathcal{H}_i) < b^2\beta_i/2$ then $\mathcal{H}_i$ is $(1, 4b\mu(\mathcal{H}_i))$-global with $\mu(\mathcal{H}_i) > \beta_i/2$.

We claim that each $\mathcal{H}_i$ is $(rbm, \beta_i/4)$-uncapturable. Indeed, this holds by a union bound if $\mu(\mathcal{H}_i) \ge b^2\beta_i/2$, as then $\mu((\mathcal{H}_i)^{\emptyset}_B) \ge \mu(\mathcal{H}_i) - |J|k_i/n \ge \beta_i/4$ whenever $|J| \le rbm$, as $\beta_i \ge 8rmk_i/bn$. On the other hand, if $\mathcal{H}_i$ is $(1, 4b\mu(\mathcal{H}_i))$-global with $\mu(\mathcal{H}_i) > \beta_i/2$ then $\mathcal{H}_i$ is $(n/2bk_i, \mu(\mathcal{H}_i)/2)$-uncapturable by Lemma IV.1.4, so $(rbm, \beta_i/4)$-uncapturable, as $k_i \le n/2rmb^2$.

Now we can apply Lemma IV.4.5 again to $\mathcal{H}_1, \dots, \mathcal{H}_m$ with $(2, r, b)$ in place of (α_i, r, b). This gives disjoint $S'_1, \dots, S'_m$ with each $|S'_i| \le rb$ such that, setting $\mathcal{G}_i = (\mathcal{H}_i)^{S'_i}_{S'}$ where $S' = \bigcup_i S'_i$, if $\mu(\mathcal{G}_i) < 2^b\beta_i/8$ then $\mathcal{G}_i$ is $(r, 8\mu(\mathcal{G}_i))$-global with $\mu(\mathcal{G}_i) > 2^{1_{S'_i \neq \emptyset}}\beta_i/8$. Thus $B_i = S_i \cup S'_i$ for $i \in [m]$ are as required. □

IV.4.4 Refined capturability for matchings

Here we prove the following sharper version of Lemma IV.2.1, obtaining cross matchings under a much weaker uncapturability condition.

Lemma IV.4.7. *Let $C \gg K \gg d \ge 1$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $k \le k_i \le Kk$ for $i \in [s]$, where $2d \le k \le n/Cs$. Suppose $\mathcal{F}_i$ is $(2dm, (2mk_i/n)^d)$-uncapturable for $i \in [m]$ and $\mu(\mathcal{F}_i) > 12(s + Km\log\frac{n}{mk})k_i/n$ for $i > m$. Then $\mathcal{F}_1, \dots, \mathcal{F}_s$ cross contain a matching.*

Proof. We start by upgrading uncapturability to globalness. We apply Lemma IV.4.5 with $r = 1$, $b = 2d$, $\alpha_i = \sqrt{n/mk_i}$, $\beta_i = (mk_i/n)^d$ noting that each $rb \le k_i \le n/2rm\alpha_i$ and $\alpha_i^b\beta_i = 2^d > 2rmk_i/n$, obtaining $B = \bigcup_i B_i$ with each $|B_i| \le 2d$ such that each $\mathcal{G}_i = (\mathcal{F}_i)^{B_i}_B$ is $(r, 4\alpha_i\mu(\mathcal{G}_i))$-global with $\mu(\mathcal{G}_i) > (2mk_i/n)^d/2$. We note by Lemma IV.1.4 that $\mathcal{G}_i$ is $(n/8\alpha_i k_i, (2mk_i/n)^d/4)$-uncapturable. Now we pass to the biased setting: we let $p_i = k_i/n$ and note that $\mathcal{H}_i = \mathcal{G}_i^{\uparrow}$ is $(n/8\alpha_i k_i, (2mk_i/n)^d/16)$-uncapturable by Lemma IV.2.2.

Now we will apply Lemma IV.2.4.2 to choose $S_1, \dots, S_m$ with each $|S_i| < K\log\frac{n}{mk}$ and define $\mathcal{H}^0_i, \dots, \mathcal{H}^m_i$ for $i \in [s]$ by $\mathcal{H}^0_i = \mathcal{H}_i$, $\mathcal{H}^i_j = (\mathcal{H}^{i-1}_j)^{\emptyset}_{S_i}$ for $j \neq i$ and $\mathcal{H}^i_i = (\mathcal{H}^{i-1}_i)^{S_i}_{S_i}$. At step i, we have $\mu(\mathcal{H}^{i-1}_i) \ge (2mk_i/n)^d/16$ by uncapturability of $\mathcal{H}_i$, as $\sum_{j<i}|S_j| < Km\log\frac{n}{mk}$ and $n/8\alpha_i k_i \ge \frac{1}{8}\sqrt{nm/Kk}$, using $n/mk \ge C \gg K$.

Applying Lemma IV.2.4.2 with $\eta < 1/2d$ and $\sqrt{K}$ in place of K we obtain $S_i \subset [n]$ with $|S_i| \le \sqrt{K}\log\mu(\mathcal{H}^{i-1}_i)^{-1} < K\log\frac{n}{mk}$ and $\mu_{Kp_i}(\mathcal{H}^i_i) \ge \mu^\eta > \sqrt{mp_i}$, so $\mu_{Kp_i}(\mathcal{H}^m_i) \ge \sqrt{mp_i} - |S|Kp_i > 3m(Kp_i)$. For $i > m$, by Lemma IV.2.2 and a union bound we have $\mu_{p_i}(\mathcal{H}^m_i) > \mu(\mathcal{F}_i)/4 - |S|p_i > 3sp_i$. Thus by Lemma IV.2.7 there is a cross matching in $\mathcal{H}^m_1, \dots, \mathcal{H}^m_s$, and so in $\mathcal{F}_1, \dots, \mathcal{F}_s$. □

IV.4.5 Lifted embeddings

We conclude this section by proving Lemma IV.4.2 which completes the proof of Theorem IV.0.2. As mentioned earlier, the proof becomes more complicated as the uniformity k of our family decreases. When it is quite large we can bound the fat shadow using Fairness, but otherwise we must rely on the weaker estimates from Lemma IV.3.4, so there are additional technical challenges, resolved by enhanced upgrading and in one case pausing the shadow embedding for a preliminary lifting step.

Proof of Lemma IV.4.2. Let $C \gg C_1 \gg \theta^{-1} \gg \varepsilon^{-1} \gg r\Delta$ and $C < k < n/Cs$. Let $G, H_1, \dots, H_\sigma$ be as in Setup IV.4.1 with $\sigma_1 \le \theta\sigma$. Let $\mathcal{F}_i \subset \binom{[n]}{k}$ for $i \in [\sigma_1]$ and $\mathcal{F}_i \subset \binom{[n]}{k-1}$ for $i \in [\sigma_1 + 1, \sigma]$. Suppose $\mathcal{F}_i$ is $(C_1\sigma_1, \varepsilon\sigma_1 k/n)$-uncapturable for $i \in [\sigma_1]$, that $\mu(\mathcal{F}_i) \ge 1 - \theta$ for $i \in [\sigma_1 + 1, \sigma_2]$, and $\mu(\mathcal{F}_i) \ge \beta := e^{-k/C_1} + C_1 sk/n$ for $i \in [\sigma_2 + 1, \sigma]$. We need to show that $\mathcal{F}_1, \dots, \mathcal{F}_\sigma$ cross contain $H^+_1, \dots, H^+_\sigma$.

We consider cases according to the size of k. We start with the case $k \ge \sqrt{C_1}\log\frac{n}{\sigma_1}$, for which we will use enhanced upgrading. We apply Lemma IV.4.6 to $\mathcal{F}_1, \dots, \mathcal{F}_{\sigma_1}$ with $m = \sigma_1$, $b = C_1 + \log_2\frac{s}{m}$, each $\beta_i = \varepsilon mk/n$ and $2r$ in place of r, noting that $2rb \le k \le n/2rmb^2$ and $\beta_i > 8rmk/bn$. This gives disjoint $B_1,\dots,B_m$ with each $|B_i| \le 2rb+2$ such that, setting $\mathcal{G}_i = (\mathcal{F}_i)^{B_i}_B$ where $B = \bigcup_i B_i$, if $\mu(\mathcal{G}_i) < 2^b\varepsilon mk/8n$ then $\mathcal{G}_i$ is $(2r, 8\mu(\mathcal{G}_i))$-global with $\mu(\mathcal{G}_i) > \varepsilon mk/8n > m/n \ge e^{-k/\sqrt{C_1}}$. For $i \in [\sigma_1+1,\sigma]$, writing $\mathcal{G}_i = (\mathcal{F}_i)^{\emptyset}_B$, we have $\mu(\mathcal{G}_i) \ge \mu(\mathcal{F}_i) - |B|k/n \ge e^{-k/C_1} + C_1 sk/2n$.

By Fairness (Proposition IV.3.3), with $\sqrt{C_1}$ in place of C, writing $c_i = (1-\varepsilon)\mu(\mathcal{G}_i)$ for $i \in [\sigma]$ we have $\mu(\partial^{r'}_{c_i}\mathcal{G}_i) \ge 1-\varepsilon$ for $r' \in \{r-1, r\}$, so $\partial_{c_1}\mathcal{G}_1,\dots,\partial_{c_\sigma}\mathcal{G}_\sigma$ cross contain a copy $\phi(H_1),\dots,\phi(H_\sigma)$ of $H_1,\dots,H_\sigma$ by Lemma IV.3.9. We write $V' = \operatorname{Im}\phi$ and consider $\mathcal{H}_1,\dots,\mathcal{H}_s$ corresponding to the edges $A_1,\dots,A_s$ of $H_1,\dots,H_\sigma$, where for each edge A_j of H_i with $i \in [\sigma]$ we let $\mathcal{H}_j = (\mathcal{G}_i)^{\phi(A_j)}_{V'}$. To complete the proof of this case it suffices to show that $\mathcal{H}_1,\dots,\mathcal{H}_s$ cross contain a matching.

To do so, we verify the conditions of Lemma IV.2.1. Consider any $A_j \in H_i$. If $i > \sigma_1$ or $i \in [\sigma_1]$ with $\mu(\mathcal{G}_i) \ge 2^b\varepsilon mk/8n > C_1^2 sk/n$ then $\mu(\mathcal{H}_j) \ge c_i - |V'|k/n > C_1 sk/3n$. Now consider $i \in [\sigma_1]$ such that $\mathcal{G}_i$ is $(2r, 8\mu(\mathcal{G}_i))$-global with $\mu(\mathcal{G}_i) > \varepsilon mk/8n$. Then $\mathcal{H}_j$ and $\mathcal{H}'_j = (\mathcal{G}_i)^{\phi(A_j)}_{\phi(A_j)}$ are $(r, 16\mu(\mathcal{G}_i))$-global by Lemma IV.1.2. As $\mu(\mathcal{H}'_j) > c_i = (1-\varepsilon)\mu(\mathcal{G}_i)$, by Lemma IV.1.4 $\mathcal{H}'_j$ is $(n/40k, \mu(\mathcal{H}'_j)/2)$-uncapturable, so $\mu(\mathcal{H}_j) \ge \mu(\mathcal{H}'_j)/2 > \varepsilon mk/20n$, and $\mathcal{H}_j$ is $(n/80k, \mu(\mathcal{H}_j)/2)$-uncapturable again by Lemma IV.1.4. Thus the required conditions hold.

Henceforth we can assume $k < \sqrt{C_1}\log\frac{n}{\sigma_1}$. In this case we upgrade uncapturability to globalness using Lemma IV.1.6 to obtain disjoint $S_1,\dots,S_{\sigma_1}$ with each $|S_i| \le 2r$ such that, setting $\mathcal{G}_i = (\mathcal{F}_i)^{S_i}_S$ where $S = \bigcup_i S_i$, whenever $\mu(\mathcal{G}_i) < \beta$ we have $S_i = \emptyset$ and $\mathcal{G}_i$ is $(2r, 2\beta)$-global with $\mu(\mathcal{G}_i) > \varepsilon\sigma_1 k/n$. For $i > \sigma_1$ we set $\mathcal{G}_i = (\mathcal{F}_i)^{\emptyset}_S$ and note that $\mu(\mathcal{G}_i) \ge \mu(\mathcal{F}_i) - |S|k/n > \beta/2$. As before, for any $i \notin [\sigma_1+1, \sigma_2]$ with $\mu(\mathcal{G}_i) > \beta/2$ Fairness gives $\mu(\partial^{r'}_{c_i}\mathcal{G}_i) \ge 1-\varepsilon$ for $r' \in \{r-1, r\}$, where $c_i = (1-\varepsilon)\mu(\mathcal{G}_i)$. For $i \in [\sigma_1+1, \sigma_2]$ we have the better bound $\mu(\partial^{r'}_{c_i}\mathcal{G}_i) \ge 1-\sqrt{\theta}$ where $c_i = 1-\sqrt{\theta}$ from Lemma IV.3.2. For $i \in I := \{i : \mu(\mathcal{G}_i) < \beta/2\}$ we note that $\mathcal{G}_i$ is G^+-free, as $S_i = \emptyset$, so we can bound the fat shadow by Lemma IV.3.4: we take $\ell = k$, use $(2\varepsilon)^{-1} \gg r\Delta$ in place of C, and write $c_i = \mu(\mathcal{G}_i)/2\binom{k}{r} \ge \mu(\mathcal{G}_i)/2k^r$, to obtain

$$\mu(\partial^r_{c_i}\mathcal{G}_i) \ge \left((\mu(\mathcal{G}_i)/2 - (s/n)^{2k\varepsilon})n/sk^2\right)^{r/(k-1)} \ge z := (\sigma_1/sk^2)^{2r/k} - (s/n)^{r\varepsilon}.$$

Next we consider the case that $k \ge 2C_1\log\frac{s}{\sigma_1}$. Then $z \ge 1-\varepsilon$, so $\partial_{c_1}\mathcal{G}_1,\dots,\partial_{c_\sigma}\mathcal{G}_\sigma$ cross contain a copy $\phi(H_1),\dots,\phi(H_\sigma)$ of $H_1,\dots,H_\sigma$ by Lemma IV.3.9. With notation as in the previous case, it remains to show that $\mathcal{H}_1,\dots,\mathcal{H}_s$ cross contain a matching. To do so, we verify the conditions of Lemma IV.4.7, taking $m = |I|$, $d = 2$ and $K = \varepsilon^{-1}$. Consider any $A_j \in H_i$. If $i \notin I$ then $\mu(\mathcal{H}_j) \ge \beta/3 - |\operatorname{Im}\phi|k/n > 12(s + \varepsilon^{-1}|I|\log\frac{n}{k|I|})k/n$, as $|I|/n \le \sigma_1/n < e^{-k/\sqrt{C_1}}$, so $|I|k/n \cdot \log\frac{n}{k|I|} < k^2 e^{-k/\sqrt{C_1}} < \beta^2$. Now suppose $i \in I$, so that $\mathcal{G}_i$ is $(2r, 2\beta)$-global with $\mu(\mathcal{G}_i) > \varepsilon\sigma_1 k/n$. Then $\mathcal{H}_j$ and $\mathcal{H}'_j = (\mathcal{G}_i)^{\phi(A_j)}_{\phi(A_j)}$ are $(r, 4\beta)$-global by Lemma IV.1.2. As $\mu(\mathcal{H}'_j) > c_i \ge \mu(\mathcal{G}_i)/2k^r$, by Lemma IV.1.4 $\mathcal{H}'_j$ is $(a, \mu(\mathcal{H}'_j)/2)$-uncapturable, where $a = \mu(\mathcal{G}_i)n/8k\beta > \varepsilon\sigma_1/8\beta > rs \ge |\operatorname{Im}\phi|$ as $\sigma_1/s \ge e^{-k/2C_1} \ge \sqrt{\beta}$, since $ks/n < k\Delta\sigma_1/n < \Delta k e^{-k/\sqrt{C_1}}$. Hence $\mu(\mathcal{H}_j) \ge \mu(\mathcal{H}'_j)/2 > \mu(\mathcal{G}_i)/4k^r > 2(2|I|k/n)^2$, and $\mathcal{H}_j$ is $(4|I|, \mu(\mathcal{H}_j)/2)$-uncapturable again by Lemma IV.1.4. Thus the required conditions hold.

It remains to consider the case $k < 2C_1\log\frac{s}{\sigma_1}$. We start by applying IV.4.3 to $(\partial^r_{c_i}\mathcal{G}_i : i \in [\sigma_2])$ with $\theta_0 = \sqrt{\sigma_1/\sigma} \le \sqrt{\theta}$ in place of θ, recalling for $i \in [\sigma_1+1, \sigma_2]$ that $\mu(\partial^r_{c_i}\mathcal{G}_i) \ge 1-\sqrt{\theta} \ge 1-\theta_0$ and $\mu(\partial^r_{c_i}\mathcal{G}_i) \ge 1-\varepsilon$ for $i \in [\sigma_1] \setminus I$, and noting for $i \in I$ that $\mu(\partial^r_{c_i}\mathcal{G}_i) \ge \theta_0^{1/2r} + n^{-\varepsilon} + r\Delta\sigma_1/n$. This gives a cross embedding ϕ of $H'_1,\dots,H'_{\sigma_2}$ in $(\partial_{cc_i}\mathcal{G}_i : i \in [\sigma_2])$, where $c = 1-\theta_0^{1/r}$.

Next we extend $(\phi(H'_i) : i \in [\sigma_1]) = (\phi(H_i) : i \in [\sigma_1])$ to a cross embedding $(\phi(H^+_i) : i \in [\sigma_1])$ in $(\mathcal{G}_i : i \in [\sigma_1])$, by finding a cross matching in $(\mathcal{H}_j : j \in [s_1])$ corresponding to the edges $A_1,\dots,A_{s_1}$ of $H_1,\dots,H_{\sigma_1}$, where for each edge A_j of H_i with $i \in [\sigma_1]$ we let $\mathcal{H}_j = (\mathcal{G}_i)^{\phi(A_j)}_{\operatorname{Im}\phi}$. This is possible by Lemma IV.4.7, which applies similarly to the previous case, where for uncapturability of $\mathcal{H}'_j$ we note that now $|\operatorname{Im}\phi| \le rs_1 \le r\Delta\sigma_1$.

Finally, we extend to a cross embedding $(\phi(H^+_i) : i \in [\sigma])$ in $(\mathcal{G}_i : i \in [\sigma])$ by finding a cross copy of $(A_j \setminus V_1 : s_1 < j \le s)$ in $(\mathcal{H}_j : s_1 < j \le s)$, where for each edge A_j of H_i with $\sigma_1 < i \le \sigma$ we let $\mathcal{H}_j = (\mathcal{G}_i)^{\phi(A_j \cap V_1)}_{\operatorname{Im}\phi}$. This is possible by Lemma IV.1.9, as each $\mu(\mathcal{H}_j) \ge \mu(\mathcal{G}_i) - \Delta\sigma_1 k^2/n > \beta/4$, using $k < 2C_1\log\frac{s}{\sigma_1}$ and $\sigma_1 \le \theta\sigma$. $\square$

Part V

Exact Turán results

This final part of our paper contains our exact results on the Turán numbers of expanded hypergraphs. We prove the Huang–Loh–Sudakov Conjecture on cross containment of matchings in the first section. The second section contains the proof of our Turán result for critical graphs (Theorem I.4.6). We conclude in the third section by proving the Füredi–Jiang–Seiver conjecture on expanded paths; the proof will apply to any graph satisfying a certain generalised criticality condition.

V.1 The Huang–Loh–Sudakov Conjecture

Here we prove Theorem I.4.2, which establishes the Huang–Loh–Sudakov Conjecture. In the first subsection we prove a strong stability version that has independent interest. We then deduce the exact result in the second subsection.

V.1.1 A strong stability result

Here we prove the following strong approximate version of the Huang–Loh–Sudakov conjecture, which will be refined to obtain the exact result in the following subsection.

Theorem V.1.1. *Let $0 < C^{-1} \ll \varepsilon$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $C \le k_i \le n/Cs$ for all $i \in [s]$. If $\mathcal{F}_1, \dots, \mathcal{F}_s$ are cross free of a matching and each $|\mathcal{F}_i| \ge |\mathcal{S}_{n,k_i,s-1}| - (1-\varepsilon)\binom{n-1}{k_i-1}$ then there is $J \in \binom{[n]}{s-1}$ so that $|\mathcal{F}_i \setminus \mathcal{S}_{n,k_i,J}| \le \varepsilon\binom{n-1}{k_i-1}$ for all $i \in [s]$.*

The idea of the proof will be to consider $A = \{a_1, \dots, a_\ell\} \subset [n]$ maximal such that there are distinct $b_1, \dots, b_\ell$ so that all $(\mathcal{F}_{b_i})^{a_i}_{a_i}$ are large. This motivates the setting of the following lemma.

Lemma V.1.2. *Let $0 < C^{-1} \ll \beta \ll \varepsilon \le 1$ and $m, \ell, n, s, k_1, \dots, k_s \in \mathbb{N}$ with $\ell \le m \le s$ and each $k_i \le n/Cs$. Suppose $\mathcal{F}_i \subset \binom{[n]}{k_i}$ and $J_i := \left\{ j \in [n] : \mu\big((\mathcal{F}_i)^j_j\big) \ge \beta \right\}$ for each $i \in [s]$ are such that*

(a) there are distinct $a_1, \dots, a_\ell \in [n]$ with $a_i \in J_i$ for $i \in [\ell]$;

(b) $\mu\big((\mathcal{F}_i)^\emptyset_{J_i}\big) \ge \varepsilon(m - |J_i|)k_i/n$ and $J_i \subset A := \{a_1, \dots, a_\ell\}$ for each $i \in [\ell+1, m]$;

(c) $\mu\big(\mathcal{F}_i\big) \ge Ck_i s/n$ for all $i \in [m+1, s]$.

Then $\mathcal{F}_1, \dots, \mathcal{F}_s$ cross contain a matching.

Proof. It suffices to check the conditions of Lemma IV.2.1 for $\mathcal{G}_1, \dots, \mathcal{G}_s$ defined by $\mathcal{G}_i = (\mathcal{F}_i)^{a_i}_A$ for $i \in [\ell]$ and $\mathcal{G}_i = (\mathcal{F}_i)^\emptyset_A$ otherwise. We do so with $m - \ell$ in place of m and $(\mathcal{G}_i : \ell < i \le m)$ in place of $\mathcal{F}_1, \dots, \mathcal{F}_m$. For $i \in [s] \setminus [m]$ we have $\mu(\mathcal{G}_i) \ge \mu(\mathcal{F}_i) - |A|k/n \ge Ck_i s/2n$. Similarly, for $i \in [\ell]$ we have $\mu(\mathcal{G}_i) \ge \mu((\mathcal{F}_i)^{a_i}_{a_i}) - |A|k/n \ge \beta/2 \ge Ck_i s/2n$. For $i \in [\ell+1, m]$ we note by definition of J_i that $\mathcal{G}_i$ is $(1, 2\beta)$-global with $\mu(\mathcal{G}_i) \ge \mu((\mathcal{F}_i)^\emptyset_{J_i}) - |A \setminus J_i|\beta k/n \ge \varepsilon(m-\ell)k_i/n$, so $(\varepsilon(m-\ell)/4\beta, \varepsilon(m-\ell)k_i/2n)$-uncapturable by Lemma IV.1.4. Thus the required conditions hold. □

We deduce our stability result as follows.

Proof of Theorem V.1.1. Let $0 < C^{-1} \ll \beta \ll \varepsilon \le 1/2$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $k_i \le n/Cs$ for all $i \in [s]$. Let $J_1, \dots, J_s$ be as in Lemma V.1.2. Let $A = \{a_1, \dots, a_\ell\} \subset [n]$ be maximal such that there are distinct $b_1, \dots, b_\ell$ with $a_i \in J_{b_i}$ for all $i \in [\ell]$. Without loss of generality we may assume $b_i = i$ for all $i \in [\ell]$. By maximality, we have $J_i \subset \{a_1, \dots, a_\ell\}$ for all $i \in [\ell+1, s]$.

We may assume $\ell < s$, and that $\mu\big((\mathcal{F}_h)^\emptyset_{J_h}\big) < .1\varepsilon(s - |J_h|)k_h/n$ for some $h \in [\ell+1, s]$, otherwise Lemma V.1.2 provides the required cross matching. Noting that $|\mathcal{S}_{n,k_h,s-1}| - (1-\varepsilon)\binom{n-1}{k_h-1} \le |\mathcal{F}_h| \le$

$|\mathcal{S}_{n,k_h,J_h}| + .1\varepsilon(s - |J_h|)\binom{n-1}{k_h-1}$, we see that $|J_h| = s - 1 = \ell$, $h = s$ and $J_h = A$. Now for each $i \in [s-1]$, as $a_i \in A = J_h$ we can apply the same argument switching the roles of $\mathcal{F}_i$ and $\mathcal{F}_h$ to deduce $\mu\big((\mathcal{F}_i)^\emptyset_{J_i}\big) < .1\varepsilon k_h/n$ and $J_i = A$. The theorem follows. □

V.1.2 The exact result

To complete the proof of the Huang–Loh–Sudakov Conjecture we will upgrade the approximate result of the previous subsection to an exact result via the following bootstrapping lemma (stated in a more general form than needed here as we will also use it for our other Turán results).

Lemma V.1.3. *Let $C \gg \beta^{-1} \gg d \ge 1$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ for all $i \in [s]$ with $\sum_{i=1}^s k_i \le n/C$. Suppose $\mathcal{F}_1, \dots, \mathcal{F}_s$ are cross free of some hypergraph $G = \{e_1, \dots, e_s\}$ with $|e_i| = k_i$ for each $i \in [s]$ and $e_s \cap \bigcup_{i=1}^{s-1} e_i = \emptyset$. If $\sum_{i=1}^{s-1}(1 - \mu(\mathcal{F}_i)) \le \alpha \in (0, \beta)$ then $\mu(\mathcal{F}_s) \le (\alpha k_s/n)^d$.*

Proof. Let $k = n - n/C$ and $\mathcal{G}_s = \mathcal{F}_s^\uparrow \cap \binom{[n]}{k}$. Then $\mathcal{F}_1, \dots, \mathcal{F}_{s-1}, \mathcal{G}_s$ are cross free of G' obtained from G by enlarging e_s to e'_s of size k. Suppose for contradiction that $\mu(\mathcal{F}_s) > (\alpha k_s/n)^d$. Let $t \in [k_s]$ be minimal so that $|\mathcal{F}_s| = (\alpha k_s/n)^d\binom{n}{k_s} \ge \binom{n-t}{k_s-t}$. Then $(\alpha k_s/n)^d < (k_s/n)^{t-1}$, so if $t > 2d$ then $\alpha < (k_s/n)^{t/2d}$. By Kruskal-Katona $|\mathcal{G}_s| \ge \binom{n-t}{k-t}$, so $\mu(\mathcal{G}_s) \ge (1 - 2/C)^t > \sqrt{\alpha}$, as if $t \le 2d$ then $(1 - 2/C)^t > (1 - 2/C)^{2d} > \sqrt{\beta}$ or otherwise $\alpha^{2d/t} < k_s/n \le C^{-1} < (1 - 2/C)^{4d}$. Now we let $\phi : V(G') \to [n]$ be a uniformly random injection. Let E be the event that $\phi(e'_s) \notin \mathcal{G}_s$ or $\phi(e_i) \notin \mathcal{F}_i$ for some $i \in [s-1]$. Then $1 = \mathbb{P}(E) \le 1 - \mu(\mathcal{G}_s) + \sum_{i\in[s-1]}(1 - \mu(\mathcal{F}_i)) < 1 - \sqrt{\alpha} + \alpha$, contradiction. □

Theorem I.4.2 will now follow by combining Theorem V.1.1 and Lemma V.1.3.

Proof of Theorem I.4.2. Let $0 < C \ll \varepsilon \ll 1$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $|\mathcal{F}_i| \ge |\mathcal{S}_{n,k_i,s-1}|$ and $k_i \le n/C$ for all $i \in [s]$, Suppose $\mathcal{F}_1, \dots, \mathcal{F}_s$ have no cross matching. By Theorem V.1.1 there is $J \in \binom{[n]}{s-1}$ such that $\mu\big((\mathcal{F}_i)^\emptyset_J\big) = \varepsilon_i k_i/|V|$ with $V = [n] \setminus J$ and $\varepsilon_i \le \varepsilon$ for all $i \in [s]$. We may assume that ε_s is maximal.

Next we claim that we can list the elements of J as $\mathbf{j} = (j_1, \dots, j_{s-1})$ so that

$$M_{\mathbf{j}} := \sum_{i\in[s-1]} \mu\big((\mathcal{F}_i)^{j_i}_J\big) \ge s - 1 - \varepsilon_s.$$

To see this, we note that $\mathbb{E}_{\mathbf{j}} M_{\mathbf{j}} = \mathbb{E}_{i\in[s-1]} \sum_{j\in J} \mu\big((\mathcal{F}_i)^j_J\big)$ when $\mathbf{j}$ is uniformly random. As each $(\mathcal{F}_i)^I_J \subset (\mathcal{S}_{n,k_i,s-1})^I_J$ whenever $\emptyset \ne I \subset J$ and $\mu(\mathcal{F}_i) \ge \mu(\mathcal{S}_{n,k_i,s-1})$, we have $0 \le \mu(\mathcal{F}_i) - \mu(\mathcal{S}_{n,k_i,s-1}) \le \mu((\mathcal{F}_i)^\emptyset_J) - k|V|^{-1}\sum_{j\in J}(1 - \mu((\mathcal{F}_i)^j_J))$, so $\sum_{j\in J}\mu((\mathcal{F}_i)^j_J)) \ge s - 1 - \varepsilon_s$. The claim follows.

Now let $\mathcal{H}_i = (\mathcal{F}_i)^{j_i}_J \subset \binom{V}{k-1}$ for all $i \in [s-1]$, and $\mathcal{H}_s = (\mathcal{F}_s)^\emptyset_J \subset \binom{V}{k-1}$. Then $\mathcal{H}_1, \dots, \mathcal{H}_s$ have no cross matching, $\sum_{i\in[s-1]}(1 - \mu(\mathcal{H}_i)) \le \varepsilon_s$ and $\mu(\mathcal{H}_s) = \varepsilon_s k_s/|V|$. Therefore $\varepsilon_s = 0$ by Lemma V.1.3 with $d = 1$. By choice of ε_s we deduce $\varepsilon_i = 0$ for all $i \in [s]$. Thus $\mathcal{F}_i = \mathcal{S}_{n,k_i,J}$ for all $i \in [s]$. □

V.2 Critical graphs

In this section we prove Theorem I.4.6, which gives exact Turán results for expanded critical r-graphs of bounded degree. In fact, we will prove the following strong stability version.

Theorem V.2.1. *Let $G \in \mathcal{G}(r, \Delta, s)$ be critical and $C \gg \beta^{-1} \gg dr\Delta$.*
Suppose $\mathcal{F} \subset \binom{[n]}{k}$ with $C \le k \le n/Cs$ is G^+-free and $|\mathcal{F}| \ge |\mathcal{S}_{n,k,\sigma-1}| - \varepsilon\binom{n-1}{k-1}$ with $\varepsilon \in (0, \beta)$.
Then there is $J \in \binom{[n]}{\sigma-1}$ with $|\mathcal{F} \setminus \mathcal{S}_{n,k,J}| \le \varepsilon^d\binom{n-1}{k-1}$.
Furthermore, if $k \le \sqrt{n}$ and $|\mathcal{F}| \ge |\mathcal{S}_{n,k,J}| - \beta\binom{n-r}{k-r}$ then $\mathcal{F} \subset \mathcal{S}_{n,k,J}$.

In the first subsection we will describe the strategy of the proof and complete the proof, assuming a certain bootstrapping lemma that will be proved in the second subsection.

V.2.1 Strategy

Recall that an r-graph G is *critical* if it has an edge e such that $\sigma(G \setminus e) = \tau(G \setminus e) < \tau(G) = \sigma(G)$. Thus we can adopt the following set-up.

Setup V.2.2. Let $G \in \mathcal{G}'(r, s, \Delta)$ be critical. Fix a crosscut S in $G^+(r+1)$ with $|S| = \sigma := \sigma(G)$ and $\{G_x^x : x \in S\} = \{H_i : i \in [\sigma]\}$ with $|H_\sigma| = 1$. Let $I = \{i \in [\sigma - 1] : V(H_i) \cap V(H_\sigma) \neq \emptyset\}$.

The following bootstrapping lemma will be proved in the next subsection. It shows that if we cannot find a cross embedding of $H_1^+, \dots, H_\sigma^+$ as in the above set up, if all but one of the families are nearly complete then the last must be very small.

Lemma V.2.3. *Let $G, H_1, \dots, H_\sigma$ be as in Setup V.2.2. Let $C \gg \beta^{-1} \gg dr\Delta$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $k_i \in [k/2, k]$ for $i \in [\sigma]$, where $C \leq k \leq n/Cs$. Suppose $\mathcal{F}_\sigma$ is G^+-free, $\sum_{i=1}^{\sigma-1}(1 - \mu(\mathcal{F}_i)) \leq \varepsilon \leq \beta$, $\mu(\mathcal{F}_\sigma) \geq \varepsilon^d k/n$ and $1 - \mu(\mathcal{F}_i) \leq \varepsilon_0 := 2\varepsilon/\sigma$ for all $i \in I$. Then $\mathcal{F}_1, \dots, \mathcal{F}_\sigma$ cross contain $H_1^+, \dots, H_\sigma^+$.*

We conclude this subsection by deducing Theorem V.2.1 from Lemma V.2.3.

Proof of Theorem V.2.1. By Theorem IV.0.2 (refined junta approximation) there is $J \in \binom{[n]}{\sigma-1}$ such that $|\mathcal{F} \setminus \mathcal{S}_{n,k,J}| = \delta\binom{n-1}{k-1}$ with $\delta^{-1} \gg dr\Delta$. We write $J = \{j_1, \dots, j_{\sigma-1}\}$, $\mathcal{F}_i = \mathcal{F}_J^{j_i}$ for $i \in [\sigma - 1]$ and $\mathcal{F}_\sigma = \mathcal{F}_J^\emptyset$. Note that $\mathcal{F}_\sigma$ is G^+-free. We may assume $I = [|I|]$ and $|\mathcal{F}_1| \geq \dots \geq |\mathcal{F}_{\sigma-1}|$. Now

$$\begin{aligned}
\mu(\mathcal{F}) &\leq \mu(\mathcal{F}_J^\emptyset) + \mu(\mathcal{S}_{n,k,J}) - \tfrac{k-1}{n-|J|}\sum_{i=1}^{\sigma-1}(1 - \mu(\mathcal{F}_i)) \\
&\leq \delta k/n + \mu(\mathcal{F}) + \varepsilon k/n - \tfrac{k}{2n}\sum_{i=1}^{\sigma-1}(1 - \mu(\mathcal{F}_i)),
\end{aligned}$$

so $\sum_{i=1}^{\sigma-1}(1-\mu(\mathcal{F}_i)) \leq 2(\varepsilon+\delta)$. Now for each $i \in I$ we have $1-\mu(\mathcal{F}_i) \leq 4r\Delta(\varepsilon+\delta)/\sigma$ as if $\sigma \leq 2|I| \leq 2r\Delta$ this follows from $1 - \mu(\mathcal{F}_i) \leq 2(\varepsilon + \delta)$, or otherwise from $1 - \mu(\mathcal{F}_i) \leq \frac{2(\varepsilon+\delta)}{\sigma-|I|}$.

As $\mathcal{F}_1, \dots, \mathcal{F}_\sigma$ are cross free of $H_1^+, \dots, H_\sigma^+$ as in Setup V.2.2, Lemma V.2.3 with $(2r\Delta(\varepsilon+\delta), 2d)$ in place of (ε, d) gives $\delta k/n = \mu(\mathcal{F}_\sigma) < (2r\Delta(\varepsilon+\delta))^{2d}k/n$. As $\varepsilon^{-1}, \delta^{-1} \gg dr\Delta$ we have $((2r\Delta)(\varepsilon+\delta))^{2d} = (2r\Delta)^{2d}\sum_{i=0}^{2d}\binom{2d}{i}\varepsilon^i\delta^{2d-i} < (\varepsilon^d + \delta)/2$, so $\delta < \varepsilon^d$, i.e. $|\mathcal{F}_J^\emptyset| = |\mathcal{F}_\sigma| < \varepsilon^d\binom{n-1}{k-1}$.

Finally, let $k \leq \sqrt{n}$ and suppose for contradiction that $|\mathcal{F}| \geq |\mathcal{S}_{n,k,J}| - \beta\binom{n-r}{k-r}$ but there is some $A \in \mathcal{F} \setminus \mathcal{S}_{n,k,J}$. By the previous statement with $d = 1$ and $\varepsilon = \beta\binom{n-r}{k-r}\binom{n-1}{k-1}^{-1}$ we have $|\mathcal{F}_J^\emptyset| \leq \beta\binom{n-r}{k-r}$, so $|\mathcal{S}_{n,k,J} \setminus \mathcal{F}| \leq 2\beta\binom{n-r}{k-r}$. We fix any $R \in \binom{A}{r}$ and a bijection $\phi : A_s \to R$, where $H_\sigma = \{A_s\}$ and define $\mathcal{G}_1, \dots, \mathcal{G}_{s-1}$ by $\mathcal{G}_j = (\mathcal{F}_i)_A^{\phi(A'_j)}$ whenever A_j is an edge of H_i with $A'_j = A_j \cap A_s$. For each $j \in [s-1]$, writing $r_j = |A'_j| + 1 \in [r]$, we have $\binom{n-k-r_j}{k-r_j} - |\mathcal{G}_j| \leq |\mathcal{S}_{n,k,J} \setminus \mathcal{F}|$, so as $\binom{n-k-r}{k-r} \geq .1\binom{n}{k-r}$ for $k \leq \sqrt{n}$ we have $1 - \mu(\mathcal{G}_j) \leq 20\beta < 1/2$. However, now $\mathcal{G}_1, \dots, \mathcal{G}_{s-1}$ cross contain $A_1 \setminus A_s, \dots, A_{s-1} \setminus A_s$ by Lemma IV.1.9, so we have the required contradiction. $\square$

V.2.2 Bootstrapping

Now we complete the proof of Theorem V.2.1 by proving Lemma V.2.3. The idea is to reduce to the case that the critical edge is disjoint from all other edges, so that we can apply Lemma V.1.3.

Proof of Lemma V.2.3. Let $G, H_1, \dots, H_\sigma$ be as in Setup V.2.2. Let $C \gg \beta^{-1} \gg dr\Delta$ and $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $k_i \in [k/2, k]$ for $i \in [\sigma]$, where $C \leq k \leq n/Cs$. Suppose $\sum_{i=1}^{s-1}(1-\mu(\mathcal{F}_i)) \leq \varepsilon \leq \beta$, $\mu(\mathcal{F}_\sigma) \geq \varepsilon^d k/n$ and $1 - \mu(\mathcal{F}_i) \leq \varepsilon_0 := 2\varepsilon/\sigma$ for all $i \in I$.

We need to show that $\mathcal{F}_1, \dots, \mathcal{F}_\sigma$ cross contain $H_1^+, \dots, H_\sigma^+$. Write $G = \{A_1, \dots, A_s\}$ where $H_\sigma = \{A_s\}$ and $A = A_s \cap \bigcup_{i<s} A_i$. It suffices to find an injection $\phi : A \to [n]$ such that Lemma

V.1.3 provides a cross embedding of $e_1^+, \dots, e_s^+$ in $\mathcal{G}_1, \dots, \mathcal{G}_s$, where for each edge $A_j \in H_i$ we define $e_j = A_j \setminus A_s$ and $\mathcal{G}_j = (\mathcal{F}_i)_{\phi(A)}^{\phi(A \cap A_j)}$. We note that if $A \cap A_j = \emptyset$ then $1 - \mu(\mathcal{G}_j) \le 2(1 - \mu(\mathcal{F}_i))$ for any choice of ϕ. Also, for uniformly random ϕ we have $\mathbb{P}(\mu(\mathcal{G}_j) \ge 1 - \sqrt{\varepsilon_0}) > 1 - \sqrt{\varepsilon_0}$ whenever $i \in I$ by Lemma IV.3.2.

Next suppose $\mu(\mathcal{F}_\sigma) \ge e^{-k\beta}$. Then Fairness (Proposition IV.3.3) gives $\mathbb{P}(\mu(\mathcal{G}_s) \ge \mu(\mathcal{F}_\sigma)/2) > 1/2$. By a union bound we can fix ϕ with $\sum_{i=1}^{s-1}(1 - \mu(\mathcal{G}_i)) \le 2\varepsilon + |I|\sqrt{\varepsilon_0} \le \alpha := 2\Delta\sqrt{\varepsilon}$ and $\mu(\mathcal{G}_s) \ge \mu(\mathcal{F}_\sigma)/2 \ge (\alpha k/n)^{3d}$. Then Lemma V.1.3 applies as required.

It remains to consider the case $\mu(\mathcal{F}_\sigma) < e^{-k\beta}$. We will apply Lemma IV.3.4 to show that we can fix ϕ with $\sum_{i=1}^{s-1}(1 - \mu(\mathcal{G}_i)) \le 2\varepsilon + |I|\sqrt{\varepsilon_0} \le \alpha := 2\Delta\sqrt{\varepsilon}$ as above and $\mu(\mathcal{G}_s) \ge c := \mu(\mathcal{F}_\sigma)/2k^r \ge e^{k\beta}\mu(\mathcal{F}_\sigma) \cdot \mu(\mathcal{F}_\sigma)/2k^r \ge \mu(\mathcal{F}_\sigma)^2 \ge (\alpha k/n)^{6d}$. Again this will suffice by Lemma V.1.3. Lemma IV.3.4 with $\ell = k$ gives $\mathbb{P}(\mu(\mathcal{G}_s) \ge c) \ge (\mu(\mathcal{F}_\sigma)/2)^{r/k} \ge \varepsilon^{1/4} n^{-2r/k}$, so we are done unless $\varepsilon^{1/4} n^{-2r/k} < |I|\sqrt{\varepsilon_0}$, which implies $\sigma^2 n^{-8r/k} < (2\Delta)^4 \varepsilon$. As $\varepsilon \ll \Delta^{-1}$ this implies $k < n^\beta$, say. Furthermore, we can assume $\mathcal{F}_\sigma$ is $(2r, \mu(\mathcal{F}_\sigma)\beta n/sk)$-global, otherwise we can apply the above argument with some $(\mathcal{F}_\sigma)_R^R$ in place of $\mathcal{F}_\sigma$ to get $\mathbb{P}(\mu(\mathcal{G}_s) \ge c) \ge (\mu(\mathcal{F}_\sigma)\beta n/2sk)^{r/k} \ge \varepsilon^{1/4} s^{-2r/k} > |I|\sqrt{\varepsilon_0}$.

Now we claim that $\partial_c^r \mathcal{F}_\sigma$ is G-free. This will suffice to complete the proof, as then Lemma IV.3.4 gives the improved estimate $\mu(\partial_c^r \mathcal{F}_\sigma) \ge (\varepsilon^d/ks)^{2r/k} - (s/n)^\beta > |I|\sqrt{\varepsilon_0}$, using $s \le r\sigma < n^{8r/k}$. To see the claim, we suppose $\phi(G) \subset \partial_c^r \mathcal{F}_\sigma$ and will obtain a contradiction by finding a cross matching in $\mathcal{H}_1, \dots, \mathcal{H}_s$, where for each edge A_j of G we let $\mathcal{H}_j = (\mathcal{F}_\sigma)_{\mathrm{Im}\,\phi}^{\phi(A_j)}$. We verify the conditions of Lemma IV.4.7, with $(s, s, d, 2)$ in place of (s, m, d, K). As $\mathcal{F}_\sigma$ is $(2r, \mu(\mathcal{F}_\sigma)\beta n/sk)$-global, each $\mathcal{H}_j$ is $(r, 2\mu(\mathcal{F}_\sigma)\beta n/sk)$-global by Lemma IV.1.2. Also, $\mathcal{F}_\sigma$ is $(\beta^{-1}s, \mu(\mathcal{F}_\sigma)/2)$-uncapturable by Lemma IV.1.4, so each $\mu(\mathcal{H}_j) \ge \mu(\mathcal{F}_\sigma)/2 \ge \varepsilon^d k/2n$, and each $\mathcal{H}_j$ is $(s/2\beta, \varepsilon^d k/4n)$-uncapturable by Lemma IV.1.4. As $\sigma^2 n^{-8r/k} < (2\Delta)^4\varepsilon$ and $k < n^\beta$ we have $\varepsilon^d k/n > (3sk/n)^d$, and so the conditions of Lemma IV.4.7 hold. But this is a contradiction, as then $\mathcal{H}_1, \dots, \mathcal{H}_s$ cross contain a matching. Therefore $\partial_c^r \mathcal{F}_\sigma$ is G-free, as claimed. □

V.3 The Füredi–Jiang–Seiver Conjecture

In this section we prove the Füredi–Jiang–Seiver Conjecture on the Turán numbers of expanded paths. As previously mentioned, for paths of odd length the conjecture follows from our result on critical graphs (Theorem I.4.6), so it remains to consider paths of even length. We will consider the more general setting of expansions of (normal) graphs (r-graphs with $r = 2$) satisfying the following generalised criticality property. Recall that we denote the crosscut and transversal numbers of an r-graph G by $\sigma(G)$ and $\tau(G)$, and that $\sigma(G) \ge \tau(G)$. Consider any G with $\tau(G) = \sigma(G)$. We say G is *a_1-degree-critical* if (i) $\sigma(G - x) < \sigma(G)$ for some x of degree $|G_x^x| \le a_1$, and (ii) $\tau(G - x) = \tau(G)$ for any x with $|G_x^x| < a_1$. We say G is *a_2-matching-critical* if (i) $\sigma(G \setminus M) < \sigma(G)$ for some matching M with $|M| \le a_2$, and (ii) $\tau(G \setminus M) = \tau(G)$ for any matching M with $|M| < a_2$. We say G is *(a_1, a_2)-critical* if it is both a_1-degree-critical and a_2-matching-critical.

We note that even paths and cycles are $(2, 2)$-critical, and that any G is critical (in the sense defined above) if and only if G is $(a_1, 1)$-critical, where a_1 is the minimum possible degree of any vertex belonging to any minimum size crosscut of G^+. The significance of the generalised definition is that it enables to show that the following natural construction is extremal for the Turán problem for G^+. For any $T \subset [n]$ we write $\mathcal{G}_{n,k}(T) = \{A \in \binom{[n]}{k} : T \subset A\}$ for the family in $\binom{[n]}{k}$ generated by T. For $\mathcal{T} \subset \{0,1\}^n$ we write $\mathcal{G}_{n,k}(\mathcal{T}) = \bigcup_{T \in \mathcal{T}} \mathcal{G}_{n,k}(T)$. We let $\mathcal{F}_{n,k,G} = \mathcal{G}_{n,k}(\mathcal{T})$ where $\mathcal{T}$ is the disjoint union of $\sigma(G) - 1$ singletons and a graph $F_{a_1 a_2}$ with as many edges as possible subject to having no vertex of degree $\ge a_1$ or matching of size $\ge a_2$. Then $\mathcal{F}_{n,k,G}$ is G^+-free by definition of (a, b)-criticality. We will show that it is extremal. When G is a path of even length this will complete the proof of the Füredi–Jiang–Seiver Conjecture.

Theorem V.3.1. *Let $G \in \mathcal{G}(2, \Delta, s)$ be (a_1, a_2)-critical, $C \gg a_2\Delta$ and $C \le k \le n/Cs$. Then $ex(n, G^+(k)) = |\mathcal{F}_{n,k,G}|$.*

Moreover, we will prove the following strong stability version.

Theorem V.3.2. *Let $G \in \mathcal{G}(2,\Delta,s)$ be (a_1,a_2)-critical and $C \gg \beta^{-1} \gg a_2 d\Delta$.*

Suppose $\mathcal{F} \subset \binom{[n]}{k}$ with $C \le k \le n/Cs$ is G^+-free. If $|\mathcal{F}| \ge |\mathcal{S}_{n,k,\sigma-1}|$ then $|\mathcal{F} \setminus \mathcal{G}_{n,k}(\mathcal{T})| \le \beta^{-1}\binom{n-3}{k-3}$ for some $\mathcal{T} = \{\{x\} : x \in J\} \cup F$ where $J \in \binom{[n]}{\sigma-1}$ and $F \subset \binom{[n]\setminus J}{2}$ with $|F| \le |F_{a_1a_2}|$.

Moreover, if $|\mathcal{F}| \ge |\mathcal{F}_{n,k,G}| - \varepsilon\binom{n-2}{k-2}$ with $\varepsilon \in (0,\beta)$ then $\mu(\mathcal{F} \setminus \mathcal{G}) \le (\varepsilon k/n)^d$ for some copy $\mathcal{G}$ of $\mathcal{F}_{n,k,G}$, where if $k \le \sqrt{n}$ then $\mathcal{F} \subset \mathcal{G}$.

Throughout this section we adopt the following set up.

Setup V.3.3. Let $G \in \mathcal{G}'(2,s,\Delta)$ be (a_1,a_2)-critical with $\sigma(G) = \sigma$. Let $\mathcal{B} = \{B_i : i \in [a]\}$ be a r-graph matching with $r \in [2]$, and $\mathcal{B}' = \{B'_i : i \in [a]\} \subset G$, where if $r = 2$ then $a = a_2$ and each $B'_i = B_i$ or if $r = 1$ then $a = a_1$ and each $B'_i = B_i \cup \{x\}$ for some vertex x of degree a. Let $S = \{s_1,\dots,s_{\sigma-1}\}$ be a crosscut in $(G \setminus \mathcal{B}')^+$ and let $H_i = G^{s_i}_{s_i}$ for $i \in [\sigma-1]$. Let $I = \{i \in [\sigma-1] : V(H_i) \cap V(\mathcal{B}) \ne \emptyset\}$.

We prove a bootstrapping lemma in the next subsection and then deduce Theorem V.3.2 in the following subsection.

V.3.1 Bootstrapping

In this subsection we prove the following bootstrapping lemma, which is analogous to Lemma V.2.3, except that rather than concluding that some family is small we conclude that some family is capturable.

Lemma V.3.4. *With notation as in Setup V.3.3, let $C \gg \beta^{-1} \gg ad\Delta$ and $C \le k \le n/Cs$. Let $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $k_i \in [k/2,k]$ for $i \in [\sigma-1]$ and $\mathcal{F}'_i \subset \binom{[n]}{k'_i}$ with $k'_i \in [k/2,k]$ for $i \in [a]$ be such that $\mathcal{F}_1,\dots,\mathcal{F}_{\sigma-1},\mathcal{F}'_1,\dots,\mathcal{F}'_a$ are cross free of $H^+_1,\dots,H^+_{\sigma-1},B^+_1,\dots,B^+_a$. Suppose $\sum_{i=1}^{s-1}(1-\mu(\mathcal{F}_i)) \le \varepsilon \le \beta$ and $1-\mu(\mathcal{F}_i) \le \varepsilon_0 := 2\varepsilon/\sigma$ for all $i \in I$. Then some $\mathcal{F}'_i$ is $(\beta^{-1},\gamma_i+(k/n)^d)$-capturable, where $\gamma_i < \varepsilon^d$, and if $\mathcal{F}'_i$ is G^+-free then $\gamma_i < \varepsilon^d k/n$.*

The proof requires the following lemma which is analogous to Lemma V.1.3.

Lemma V.3.5. *Let $C \gg C' \gg ad$, $\mathcal{F}_i \subset \binom{[n]}{k_i}$ for $i \in [s]$ and $\mathcal{F}'_i \subset \binom{[n]}{k'_i}$ for $i \in [a]$ with $\sum_{i=1}^s k_i + \sum_{i=1}^a k'_i \le n/C$. Suppose $(\mathcal{F}_1,\dots,\mathcal{F}_s,\mathcal{F}'_1,\dots,\mathcal{F}'_a)$ are cross free of $G = (e_1,\dots,e_s,e'_1,\dots,e'_a)$ with each $|e_i| = k_i$, $|e'_i| = k'_i$ and $e \cap e'_i = \emptyset$ for all $i \in [a]$ and $e'_i \ne e \in G$. If $\sum_{i=1}^s(1-\mu(\mathcal{F}_i)) < 1/2$ then some $\mathcal{F}'_i$ is $(C',(k'_i/n)^d)$-capturable.*

Proof. Let $k = n/2a$ and for each $i \in [a]$ let $\mathcal{G}_i = (\mathcal{F}'_i)^\uparrow \cap \binom{[n]}{k}$. Then $(\mathcal{F}_1,\dots,\mathcal{F}_s,\mathcal{G}_1,\dots,\mathcal{G}_a)$ are cross free of G' obtained from G by enlarging each e'_i to e^*_i of size k. Suppose for contradiction that each $\mathcal{F}'_i$ is $(C',(k'_i/n)^d)$-uncapturable. Then an argument of Dinur and Friedgut [15] (apply Russo's Lemma and Friedgut's junta theorem) shows that each $\mu(\mathcal{G}_i) > 1-1/2a$. Consider a uniformly random injection $\phi : V(G') \to [n]$. Let E be the event that some $\phi(e_i) \notin \mathcal{F}_i$ or some $\phi(e^*_i) \notin \mathcal{G}_i$. Then $1 = \mathbb{P}(E) \le \sum_{i\in[s]}(1-\mu(\mathcal{F}_i)) + \sum_{i\in[a]}(1-\mu(\mathcal{G}_i)) < 1/2+1/2$, contradiction. □

Proof of Lemma V.3.4. With notation as in Setup V.3.3, let $C \gg \beta^{-1} \gg b \gg d \gg a\Delta$ and $C \le k \le n/Cs$. Let $\mathcal{F}_i \subset \binom{[n]}{k_i}$ with $k_i \in [k/2,k]$ for $i \in [\sigma-1]$ and $\mathcal{F}'_i \subset \binom{[n]}{k'_i}$ with $k'_i \in [k/2,k]$ for $i \in [a]$ be such that $\mathcal{F}_1,\dots,\mathcal{F}_{\sigma-1},\mathcal{F}'_1,\dots,\mathcal{F}'_a$ are cross free of $H^+_1,\dots,H^+_{\sigma-1},B^+_1,\dots,B^+_a$. Suppose $\sum_{i=1}^{s-1}(1-\mu(\mathcal{F}_i)) \le \varepsilon \le \beta$ and $1-\mu(\mathcal{F}_i) \le \varepsilon_0 := 2\varepsilon/\sigma$ for all $i \in I$. Suppose for contradiction that each $\mathcal{F}'_i$ is $(\beta^{-1},\gamma_i+(k/n)^d)$-uncapturable, where either $\gamma_i \ge \varepsilon^d$ or $\mathcal{F}'_i$ is G^+-free and $\gamma_i \ge \varepsilon^d k/n$.

We start by upgrading uncapturability to globalness. By Lemma IV.4.5 with $(b,4,a)$ in place of (b,r,m) and each $\alpha_i = n/kb$, $\beta_i = \gamma_i + (k/n)^d$, noting that $8b \le k \le n/8a(n/bk)$, $4ba < \beta^{-1}$ and $(n/kb)^b(k/n)^d > n/k \gg 1$, there is a set S' partitioned into $S'_1,\dots,S'_a$ with each $|S'_i| \le 8b$ such that each $\mathcal{G}^0_i := (\mathcal{F}'_i)^{S'_i}_{S'}$ is $(8,4\mu(\mathcal{G}^0_i)n/kb)$-global with $\mu(\mathcal{G}^0_i) > \alpha_i^{1_{S'_i\ne\emptyset}}\beta_i/2$. We have $2\mu(\mathcal{G}^0_i) > \varepsilon^d + (k/n)^d$, unless $\mathcal{F}'_i$ is G^+-free and $S'_i = \emptyset$, in which case $\mathcal{G}^0_i$ is a restriction of $\mathcal{F}'_i$, so is also G^+-free, with $2\mu(\mathcal{G}^0_i) > \varepsilon^d k/n + (k/n)^d$.

Next we define $\mathcal{G}'_i := (\mathcal{F}'_i)^{S_i}_S$ with enhanced globalness, obtaining S partitioned into $S_1, \dots, S_a$ by letting $S_i = S'_i$ if $\mathcal{G}^0_i$ is $(4, \mu(\mathcal{G}^0_i)\beta n/sk)$-global, or otherwise letting $S_i = S'_i \cup R_i$ where $|R_i| \le 4$ and $\mathcal{G}^1_i := (\mathcal{G}^0_i)^{R_i}_{R_i}$ has $\mu(\mathcal{G}^1_i) > \mu(\mathcal{G}^0_i)\beta n/sk$. We also define $\mathcal{G}_i = (\mathcal{F}_i)^{\emptyset}_S$ for $i \in [\sigma - 1]$ and note that each $1 - \mu(\mathcal{G}_i) \le 2(1 - \mu(\mathcal{F}_i))$.

By Lemma IV.1.2, each $\mathcal{G}^1_i$ or $\mathcal{G}'_i$ is $(4, 2\mu(\mathcal{G}^0_i)\beta n/sk)$-global if $R_i = \emptyset$ or $(4, 8\mu(\mathcal{G}^0_i)n/kb)$-global otherwise. By Lemma IV.1.4, each $\mathcal{G}^1_i$ is $(b/8, \mu(\mathcal{G}^1_i)/2)$-uncapturable, so $\mu(\mathcal{G}'_i) > \mu(\mathcal{G}^1_i)/2 \ge \mu(\mathcal{G}^0_i)/2$. Thus $2\beta^{-1}\mu(\mathcal{G}'_i) \ge \gamma'_i + (k/n)^d$, and

(i) $\mathcal{G}'_i$ is $(4, 8\mu(\mathcal{G}'_i)n/kb)$-global with $\gamma'_i \ge \varepsilon^d/s$, or

(ii) $\mathcal{G}'_i$ is G^+-free and $(4, 2\mu(\mathcal{G}'_i)\beta n/sk)$-global with $\gamma'_i \ge \varepsilon^d k/n$.

Indeed, if option (i) does not hold then $\mathcal{G}^0_i$ is G^+-free with $2\mu(\mathcal{G}^0_i) > \varepsilon^d k/n + (k/n)^d$, and also is $(4, \mu(\mathcal{G}^0_i)\beta n/sk)$-global, so $R_i = \emptyset$ and $\mathcal{G}'_i$ is a restriction of $\mathcal{G}^0_i$, so is also G^+-free.

We will show that $\mathcal{G}_1, \dots, \mathcal{G}_{\sigma-1}, \mathcal{G}'_1, \dots, \mathcal{G}'_a$ cross contain $H^+_1, \dots, H^+_{\sigma-1}, B^+_1, \dots, B^+_a$, thus obtaining the required contradiction. It suffices to find an injection $\phi : B \to [n]$, where $B = \bigcup_{i=1}^a B_i$, such that Lemma V.3.5 provides a cross embedding of $e^+_1, \dots, e^+_s$ in $\mathcal{G}_1, \dots, \mathcal{G}_s$, where for each edge $A_j \in H_i$ we define $e_j = A_j \setminus B$ and $\mathcal{H}_j = (\mathcal{G}_i)^{\phi(B \cap A_j)}_{\phi(B)}$, or if $A_j = B_i$ we define $e_j = A_j \setminus B = \emptyset$ and $\mathcal{H}_j = (\mathcal{G}'_i)^{\phi(B_i)}_{\phi(B)}$.

We note that if $B \cap A_j = \emptyset$ then each $1 - \mu(\mathcal{H}_j) \le 2(1 - \mu(\mathcal{G}_i))$ for any ϕ. We consider ϕ obtained by choosing independent uniformly random injections $\phi_i : B_i \to [n]$ for each $i \in [a]$. Then $\mathbb{P}(\phi \text{ is injective}) \ge 1 - 2a^2/n$ and $\mathbb{P}(\mu(\mathcal{H}_j) \ge 1 - \sqrt{\varepsilon_0}) > 1 - 2\sqrt{\varepsilon_0}$ whenever $A_j \in \bigcup_{i \in I} H_i$ by Lemma IV.3.2. We write E_i for the event that $\phi_i(B_i) \in \partial_{c_i}\mathcal{G}'_i$, where $c_i = b^{-.3}\mu(g'_i)$. It suffices to show that conditional on E_i each $\mathcal{H}'_i := (\mathcal{G}'_i)^{\phi_i(B_i)}_{\phi(B)}$ is $(\sqrt{b}, (k/n)^{2d})$-uncapturable, and that $\mathbb{P}(E_i) \ge \varepsilon_0^{1/3a}$.

For uncapturability, we recall that $\mathcal{G}'_i$ is $(4, 8\mu(\mathcal{G}'_i)n/kb)$-global with $2\beta^{-1}\mu(\mathcal{G}'_i) \ge (k/n)^d$. Thus $\mathcal{H}'_i$ and $\mathcal{H}''_i := (\mathcal{G}'_i)^{\phi_i(B_i)}_{\phi_i(B_i)}$ are $(2, 8\mu(\mathcal{G}'_i)n/kb)$-global by Lemma IV.1.2. Conditional on E_i we have $\mu(\mathcal{H}''_i) > c_i$, so $\mathcal{H}''_i$ is $(b^{.7}/16, \mu(\mathcal{H}''_i)/2)$-uncapturable by Lemma IV.1.4. Then $\mu(\mathcal{H}'_i) \ge \mu(\mathcal{H}''_i)/2 \ge b^{-.3}\mu(\mathcal{G}'_i)/4$, so $\mathcal{H}'_i$ is $(b^{.7}/32, \mu(\mathcal{H}'_i)/2)$-uncapturable by Lemma IV.1.4, and so $(\sqrt{b}, (k/n)^{2d})$-uncapturable.

It remains to show $\mathbb{P}(E_i) \ge \varepsilon_0^{1/3a}$. We may assume $\mu(\mathcal{G}'_i) < e^{-k\beta}$, otherwise this holds easily by Fairness (Proposition IV.3.3). As $2\beta^{-1}\mu(\mathcal{G}'_i) \ge (k/n)^d$ this gives $k < n^\beta$. By Lemma IV.3.4 with $\ell = b^{.1}$ we are done unless $\varepsilon_0^{1/3a} > \mathbb{P}(E_i) = \mu(\partial_{c_i}\mathcal{G}'_i) \ge (\mu(\mathcal{G}'_i)/2)^{2/\ell}$, which implies $\gamma'_i + (k/n)^d \le 2\beta^{-1}\mu(\mathcal{G}'_i) < (\varepsilon/s)^{b^{.05}}$. As $\gamma'_i < \varepsilon^d/s$ we have option (ii) above, so $\mathcal{G}'_i$ is G^+-free. As $\varepsilon^d k/n \le \gamma'_i < (\varepsilon/s)^{b^{.05}}$ we also have $s < \varepsilon n^{b^{-.05}}$.

Now we claim that $\partial^2_{c_i}\mathcal{G}'_i$ is G-free. This will suffice to complete the proof, as then Lemma IV.3.4 gives the improved estimate $\mu(\partial^2_{c_i}\mathcal{G}'_i) \ge (\varepsilon^d k/sb + k/n - (s/n)^2)^{b^{-.02}} > (\varepsilon/s)^{b^{-.01}}$. To see the claim, we suppose $\phi'(G) \subset \partial^2_{c_i}\mathcal{G}'_i$ and will obtain a contradiction by finding a cross matching in $\mathcal{A}_1, \dots, \mathcal{A}_s$, where for each edge A_j of G we let $\mathcal{A}_j = (\mathcal{G}'_i)^{\phi'(A_j)}_{\operatorname{Im}\phi'}$. We verify the conditions of Lemma IV.4.7, with $(s, s, d, 2)$ in place of (s, m, d, K). As $\mathcal{G}'_i$ is $(4, 2\mu(\mathcal{G}'_i)\beta n/sk)$-global, each $\mathcal{H}_j$ is $(2, 4\mu(\mathcal{G}'_i)\beta n/sk)$-global by Lemma IV.1.2. Also, $\mathcal{G}'_i$ is $(s/4\beta, \mu(\mathcal{G}'_i)/2)$-uncapturable by Lemma IV.1.4, so each $\mu(\mathcal{H}_j) \ge \mu(\mathcal{G}'_i)/2 \ge \beta\varepsilon^d k/4n$, and each $\mathcal{H}_j$ is $(s/8\beta, \beta\varepsilon^d k/8n)$-uncapturable by Lemma IV.1.4. As $s < \varepsilon n^{b^{-.05}}$ and $k < n^\beta$ we have $\beta\varepsilon^d k/8n > (3sk/n)^d$, so the required conditions hold. □

V.3.2 Strong stability

We conclude with the proof of the main result of this section.

Proof of Theorem V.3.2. Let $G \in \mathcal{G}(2, \Delta, s)$ be (a_1, a_2)-critical and $C \gg \beta^{-1} \gg b \gg d \gg a_2\Delta$. Suppose $\mathcal{F} \subset \binom{[n]}{k}$ with $C \le k \le n/Cs$ is G^+-free and $|\mathcal{F}| \ge |\mathcal{S}_{n,k,\sigma-1}|$.

By Theorem IV.0.2 (refined junta approximation) there is $J \in \binom{[n]}{\sigma-1}$ such that $|\mathcal{F} \setminus \mathcal{S}_{n,k,J}| = \delta\binom{n-1}{k-1}$ with $\delta^{-1} \gg bd\Delta$. We write $J = \{j_1, \dots, j_{\sigma-1}\}$, let $\mathcal{F}_i = \mathcal{F}^{j_i}_J$ for $i \in [\sigma-1]$, say with $|\mathcal{F}_1| \ge \dots \ge |\mathcal{F}_{\sigma-1}|$, and note that $\mathcal{F}^{\emptyset}_J$ is G^+-free. As in the proof of Theorem V.2.1, we have $\sum_{i=1}^{\sigma-1}(1 - \mu(\mathcal{F}_i)) \le 2\delta$, so $1 - \mu(\mathcal{F}_i) \le 4r\Delta\delta/\sigma$ for any $i \le \min\{r\Delta, \sigma - 1\}$.

As G is a_2-matching-critical, we can define $H_1^2, \dots, H_{\sigma-1}^2$, $B_1^2, \dots, B_{a_2}^2$ and I^2 as in Setup V.3.3 with $r = 2$ and $a = a_2$, where we identify I^2 with $[|I^2|]$. Letting $\mathcal{F}'_i = \mathcal{F}_J^\emptyset$ for $i \in [a_2]$, we have $\mathcal{F}_1, \dots, \mathcal{F}_{\sigma-1}, \mathcal{F}'_1, \dots, \mathcal{F}'_{a_2}$ cross free of $(H_1^2)^+, \dots, (H_{\sigma-1}^2)^+, (B_1^2)^+, \dots, (B_{a_2}^2)^+$, so $\mathcal{F}_J^\emptyset$ is $(b, (2\delta)^d k/n + (k/n)^d)$-capturable by Lemma V.3.4. We fix $J' \in \binom{[n]\setminus J}{b}$ so that $\mu(\mathcal{F}_{J\cup J'}^\emptyset) < (2\delta)^d k/n + (k/n)^d$.

As G is a_1-degree-critical, we can define $H_1^1, \dots, H_{\sigma-1}^1$, $B_1^1, \dots, B_a^1$ and I^1 as in Setup V.3.3 with $r = 1$ and $a = a_1$, where we identify I^1 with $[|I^1|]$. For each $x \in J'$, letting $\mathcal{F}'_i = \mathcal{F}_{J\cup\{x\}}^x$ for $i \in [a_1]$, we have $\mathcal{F}_1, \dots, \mathcal{F}_{\sigma-1}, \mathcal{F}'_1, \dots, \mathcal{F}'_{a_1}$ cross free of $(H_1^1)^+, \dots, (H_{\sigma-1}^1)^+, (B_1^1)^+, \dots, (B_{a_2}^1)^+$, so $\mathcal{F}_{J\cup\{x\}}^x$ is $(b, (2\delta)^d k/n + (k/n)^d)$-capturable by Lemma V.3.4. We fix $J_x \in \binom{[n]\setminus(J\cup\{x\})}{b}$ so that $\mu((\mathcal{F}_{J\cup\{x\}}^x)_{J_x}^\emptyset) < (2\delta)^d k/n + (k/n)^d$.

Let $F = \{T \in \binom{[n]\setminus J}{2} : \mu(\mathcal{F}_{T\cup J}^T) > bk/n\}$. Then $F \subset F' := \{xy : x \in J', y \in J_x\}$ and $|F'| \le b^2$. Writing $\mathcal{T} = \{\{x\} : x \in J\} \cup F$, we have $|\mathcal{F} \setminus \mathcal{G}_{n,k}(\mathcal{T})| \le |\mathcal{F}_{J\cup J'}^\emptyset| + \sum_{x\in J'} |\mathcal{F}_{J\cup\{x\}\cup J_x}^x| + \sum_{T\in F'} |\mathcal{F}_{T\cup J}^T|$, so $\mu(\mathcal{F} \setminus \mathcal{G}_{n,k}(\mathcal{T})) \le ((2\delta)^d k/n + (k/n)^d)(1 + bk/n) + (bk/n)^3$. Writing $\mathcal{G} := \mathcal{G}_{n,k}(\mathcal{T})$, as $|\mathcal{F} \setminus \mathcal{G}| \ge |\mathcal{F} \setminus \mathcal{S}_{n,k,J}| - |\mathcal{G}_{n,k}(F)|$ we also have $\mu(\mathcal{F} \setminus \mathcal{G}) \ge \delta k/n - (bk/n)^2$. We deduce $\delta k/n \le (2\delta)^d k/n + 2(bk/n)^2$, so $\delta \le 3bk/n$, giving $|\mathcal{F} \setminus \mathcal{G}| \le 2b^3\binom{n-3}{k-3}$.

To complete the proof of the first statement of the theorem, it remains to show $|F| \le |F_{a_1 a_2}|$. To see this, note that otherwise F contains some $F_0 = (T_i : i \in [a_r])$, where $r = 2$ and F_0 is a matching or $r = 1$ and F_0 is a star. Writing $\mathcal{F}'_i = \mathcal{F}_{J\cup T_i}^{T_i}$, we have $\mathcal{F}_1, \dots, \mathcal{F}_{\sigma-1}, \mathcal{F}'_1, \dots, \mathcal{F}'_{a_2}$ cross free of $(H_1^r)^+, \dots, (H_{\sigma-1}^r)^+, (B_1^r)^+, \dots, (B_{a_r}^r)^+$, so some $\mathcal{F}'_i$ is $(b/2, (k/n)^d)$-capturable by Lemma V.3.4. However, $\mu(\mathcal{F}'_i) > bk/n$ as $T_i \in \mathcal{F}$, so we have a contradiction.

Now suppose $|\mathcal{F}| \ge |\mathcal{F}_{n,k,G}| - \varepsilon\binom{n-2}{k-2}$ with $\varepsilon \in (0, \beta)$. We have

$$\mu(\mathcal{F}) \le \mu(\mathcal{F} \setminus \mathcal{G}) + \mu(\mathcal{G}) - \tfrac{k}{2n}\sum_{i=1}^{\sigma-1}(1 - \mu(\mathcal{F}_i)) - \tfrac{k^2}{2n^2}\sum_{T\in F}(1 - \mu(\mathcal{F}_{J\cup T}^T)),$$

where $\mu(\mathcal{F} \setminus \mathcal{G}) \le 2(bk/n)^3$ and $\mu(\mathcal{G}) \le \mu(\mathcal{F}_{n.k.G}) - (|F_{a_1 a_2}| - |F|)k^2/2n^2 \le \mu(\mathcal{F}) + (|F_{a_1 a_2}| - |F| + 2\varepsilon)k^2/2n^2$. Thus $|F| = |F_{a_1 a_2}|$, so $\mathcal{G} := \mathcal{G}_{n,k}(\mathcal{T})$ is a copy of $\mathcal{F}_{n,k,G}$, and

$$\sum_{i=1}^{\sigma-1}(1 - \mu(\mathcal{F}_i)) + \sum_{T\in F}(1 - \mu(\mathcal{F}_{J\cup T}^T)) \le 3\varepsilon.$$

Next we suppose for contradiction that $\mu(\mathcal{F}\setminus\mathcal{G}) > (\varepsilon k/n)^d$. We fix some $T \in \binom{[n]\setminus J}{2}\setminus F$ with $\mu(\mathcal{F}_{J\cup T}^T) > (\varepsilon k/n)^{d+2}$. By maximality of $F_{a_1 a_2}$ we can fix a matching $T_1, \dots, T_{a_2}$ in F with $T_{a_2} = T$. Writing $\mathcal{F}'_i = \mathcal{F}_{J\cup T_i}^{T_i}$, we have $\mathcal{F}_1, \dots, \mathcal{F}_{\sigma-1}, \mathcal{F}'_1, \dots, \mathcal{F}'_{a_2}$ cross free of $(H_1^2)^+, \dots, (H_{\sigma-1}^2)^+, (B_1^2)^+, \dots, (B_{a_2}^2)^+$. Thus Lemma V.1.3 gives the required contradiction, so $\mu(\mathcal{F} \setminus \mathcal{G}) \le (\varepsilon k/n)^d$, as required.

Finally, let $k \le \sqrt{n}$ and suppose for contradiction that there is some $A \in \mathcal{F} \setminus \mathcal{G}$. From the previous statement we have $|\mathcal{G} \setminus \mathcal{F}| \le 2\beta\binom{n-2}{k-2}$. We fix any $T \in \binom{[n]\setminus J}{2}$ with $T \subset A$, a matching $T_1, \dots, T_{a_2}$ in F with $T_{a_2} = T$, and a bijection $\phi : B_{a_2}^2 \to T$. Writing $A'_j = A_j \cap A_s$ for each edge A_j of G, where $A_s = B_{a_2}^2$, we define $\mathcal{G}_1, \dots, \mathcal{G}_{s-1}$ by $\mathcal{G}_j = (\mathcal{F}_i)_A^{\phi(A'_j)}$ if $A_j \in H_i$ with $i \in [\sigma - 1]$ or $\mathcal{G}_j = (\mathcal{F}_J^\emptyset)_A^{\phi(A'_j)}$ if $A_j = B_i^2$ with $i \in [a_2 - 1]$. For each $j \in [s-1]$, writing $r_j = |A'_j| + 1 \in [2]$, we have $\binom{n-k-r_j}{k-r_j} - |\mathcal{G}_j| \le |\mathcal{G} \setminus \mathcal{F}|$, so as $\binom{n-k-2}{k-2} \ge .1\binom{n}{k-2}$ for $k \le \sqrt{n}$ we have $1 - \mu(\mathcal{G}_j) \le 20\beta < 1/2$. However, now $\mathcal{G}_1, \dots, \mathcal{G}_{s-1}$ cross contain $A_1 \setminus A_s, \dots, A_{s-1} \setminus A_s$ by Lemma IV.1.9, so we have the required contradiction. □

Concluding remarks

We are optimistic that our sharp threshold result in the sparse regime will have many applications in the same vein as the applications of the classical sharp threshold results, e.g. to Percolation [6], Complexity Theory [36], Coding Theory [64], and Ramsey Theory [37].

In particular, it may be possible to estimate the location of thresholds in the spirit of the Kahn-Kalai conjecture [49, Conjecture 2.1] that the threshold probability $p_c(H)$ for finding some graph H in $G(n,p)$ should be within a log factor of its 'expectation threshold' $p_E(H)$ (the probability at which every subgraph H' of H we expect at least one copy of H'). This question is interesting when $|V(H)|$ depends on n, e.g. if H is a bounded degree spanning tree it predicts $p_c(H) = O(n^{-1}\log n)$, which was a longstanding open problem, recently resolved by Montgomery [72].

To obtain similar results from our sharp threshold theorem one needs to show that the property of containing H is not 'local': writing $\mu_p = \mathbb{P}(H \subset G(n,p))$, this means that if we plant any set E of $O(\log \mu_p^{-1})$ edges we still have $\mathbb{P}(H \subset G(n,p) \mid E \subset G(n,p)) \le \mu_p^{O(1)}$. An open problem is to apply this approach to estimate other thresholds that are currently unknown, e.g. the threshold for containing any given H of maximum degree Δ.

Our variant of the Kahn-Kalai conjecture on isoperimetric stability is only effective in the p-biased setting for small p, whereas the corresponding known results [58, 55] for the uniform measure are substantial weaker. This leaves our current state of knowledge in a rather peculiar state, as in many related problems the small p case seems harder than the uniform case! A natural open problem is give a unified approach extending both results for all p.

Another compelling open problem is to generalise Hatami's Theorem to the sparse regime, i.e. to obtain a density increase from $\mu_p(f) = o(1)$ to $\mu_q(f) \ge 1-\varepsilon$ under some pseudorandomness condition on f; we expect that a such result would have profound consequences in Extremal Combinatorics.

Lastly, in relation to our hypergraph Turán results, our notion of generalised criticality seems very restrictive (indeed, we do not know of any examples where it is applicable besides those mentioned above), so it would be interesting to find more refined parameters of expanded hypergraphs that determine the Turán number $\mathsf{ex}(n, G^+(k))$ for more general classes of graphs.

Acknowledgment

We would like to thank Yuval Filmus, Ehud Friedgut, Gil Kalai, Nathan Keller, Guy Kindler, and Muli Safra for various helpful comments and suggestions.

References

[1] Amirali Abdullah and Suresh Venkatasubramanian. A directed isoperimetric inequality with application to bregman near neighbor lower bounds. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, pages 509–518. ACM, 2015.

[2] Dimitris Achlioptas and Ehud Friedgut. A sharp threshold for k-colorability. *Random Structures & Algorithms*, 14(1):63–70, 1999.

[3] Daniel Ahlberg, Erik Broman, Simon Griffiths, and Robert Morris. Noise sensitivity in continuum percolation. *Israel Journal of Mathematics*, 201(2):847–899, 2014.

[4] William Beckner. Inequalities in Fourier analysis. *Annals of Mathematics*, pages 159–182, 1975.

[5] Michael Ben-Or and Nathan Linial. Collective coin flipping. *Randomness and Computation*, 5:91–115, 1990.

[6] Itai Benjamini, Stéphane Boucheron, Gábor Lugosi, and Raphaël Rossignol. Sharp threshold for percolation on expanders. *The Annals of Probability*, 40(1):130–145, 2012.

[7] Itai Benjamini and Jérémie Brieussel. Noise sensitivity of random walks on groups. *arXiv preprint arXiv:1901.03617*, 2019.

[8] Itai Benjamini, Gil Kalai, and Oded Schramm. Noise sensitivity of boolean functions and applications to percolation. *Inst. Hautes Etudes Sci. Publ. Math.*, 90:5–43, 1999.

[9] Béla Bollobás, David E. Daykin, and Paul Erdős. Sets of independent edges of a hypergraph. *Quart. J. Math. Oxford Ser.(2)*, 27(105):25–32, 1976.

[10] Béla Bollobás and Andrew G Thomason. Threshold functions. *Combinatorica*, 7(1):35–38, 1987.

[11] Aline Bonami. Étude des coefficients de Fourier des fonctions de $l^p(g)$. In *Annales de l'institut Fourier*, volume 20(2), pages 335–402, 1970.

[12] Christer Borell. Geometric bounds on the Ornstein-Uhlenbeck velocity process. *Probability Theory and Related Fields*, 70(1):1–13, 1985.

[13] Jean Bourgain and Gil Kalai. Influences of variables and threshold intervals under group symmetries. *Geometric and Functional Analysis*, 7(3):438–461, 1997.

[14] David Conlon, Jacob Fox, and Benny Sudakov. Ramsey numbers of sparse hypergraphs. *Random Structures & Algorithms*, 35(1):1–14, 2009.

[15] Irit Dinur and Ehud Friedgut. Intersecting families are essentially contained in juntas. *Combinatorics, Probability and Computing*, 18(1-2):107–122, 2009.

[16] Irit Dinur, Ehud Friedgut, and Oded Regev. Independent sets in graph powers are almost contained in juntas. *Geometric and Functional Analysis*, 18(1):77–97, 2008.

[17] Irit Dinur and Samuel Safra. On the hardness of approximating minimum vertex cover. *Annals of Mathematics*, pages 439–485, 2005.

[18] Bradley Efron and Charles Stein. The jackknife estimate of variance. *The Annals of Statistics*, pages 586–596, 1981.

[19] David Ellis, Yuval Filmus, and Ehud Friedgut. Triangle-intersecting families of graphs. *Journal of the European Mathematical Society*, 14:841–885, 2012.

[20] David Ellis, Ehud Friedgut, and Haran Pilpel. Intersecting families of permutations. *Journal of the American Mathematical Society*, 24(3):649–682, 2011.

[21] David Ellis, Nathan Keller, and Noam Lifshitz. Stability versions of Erdős–Ko–Rado type theorems, via isoperimetry. *Journal of the European Mathematical Society*, 21:3857–3902, 2019.

[22] David Ellis, Guy Kindler, and Noam Lifshitz. Hypercontractivity for global functions on the bilinear scheme and forbidden intersections. *in preperation*, 2019.

[23] David Ellis and Bhargav Narayanan. On symmetric 3-wise intersecting families. *Proceedings of the American Mathematical Society*, 145(7):2843–2847, 2017.

[24] P. Erdös and T. Gallai. On maximal paths and circuits of graphs. *Acta Math. Acad. Sci. Hungar.*, 10:337–356, 1959.

[25] Paul Erdős. A problem on independent r-tuples. *Ann. Univ. Sci. Budapest*, 8:93–95, 1965.

[26] Paul Erdős, Chao Ko, and Richard Rado. Intersection theorems for systems of finite sets. *The Quarterly Journal of Mathematics*, 12(1):313–320, 1961.

[27] Yuval Filmus, Guy Kindler, and Noam Lifshitz. Hypercontractivity for global functions on the slice. *in preparation*, 2019.

[28] Yuval Filmus, Guy Kindler, Noam Lifshitz, and Dor Minzer. Hypercontractivity for global functions on the symmetric group. *in preparation*, 2019.

[29] Peter Frankl. Improved bounds for Erdős matching conjecture. *Journal of Combinatorial Theory, Series A*, 120(5):1068–1072, 2013.

[30] Peter Frankl and Zoltán Füredi. Exact solution of some Turán-type problems. *Journal of Combinatorial Theory, Series A*, 45(2):226–262, 1987.

[31] Peter Frankl and Andrey Kupavskii. The erdös matching conjecture and concentration inequalities. *arXiv preprint arXiv:1806.08855*, 2018.

[32] Peter Frankl and Andrey Kupavskii. Simple juntas for shifted families. *Discrete Analysis*, 14, 2020.

[33] Peter Frankl, Vojtech Rödl, and Andrzej Ruciński. On the maximum number of edges in a triple system not containing a disjoint family of a given size. *Combinatorics, Probability and Computing*, 21(1-2):141–148, 2012.

[34] Peter Frankl and Norihide Tokushige. Weighted multiply intersecting families. *Studia Scientiarum Mathematicarum Hungarica*, 40(3):287–291, 2003.

[35] Ehud Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):27–35, 1998.

[36] Ehud Friedgut. Sharp thresholds of graph properties, and the k-sat problem (with an appendix by Jean Bourgain). *Journal of the American Mathematical Society*, 12(4):1017–1054, 1999.

[37] Ehud Friedgut, Hiêp Hàn, Yury Person, and Mathias Schacht. A sharp threshold for Van der Waerden's theorem in random subsets. *Discrete Analysis*, 7:19, 2016.

[38] Ehud Friedgut and Gil Kalai. Every monotone graph property has a sharp threshold. *Proceedings of the American mathematical Society*, 124(10):2993–3002, 1996.

[39] Zoltán Füredi. Linear trees in uniform hypergraphs. *European Journal of Combinatorics*, 35:264–272, 2014.

[40] Zoltán Füredi and Tao Jiang. Hypergraph Turán numbers of linear cycles. *Journal of Combinatorial Theory, Series A*, 123(1):252–270, 2014.

[41] Zoltán Füredi, Tao Jiang, and Robert Seiver. Exact solution of the hypergraph Turán problem for k-uniform linear paths. *Combinatorica*, 34(3):299–322, 2014.

[42] N. Fusco, F. Maggi, and A. Pratelli. The sharp quantitative isoperimetric inequality. *Annals of Math.*, 168(3):941–980, 2008.

[43] Spencer Greenberg and Mehryar Mohri. Tight lower bound on the probability of a binomial exceeding its expectation. *Statistics & Probability Letters*, 86:91 – 98, 2014.

[44] L. Gross. Logarithmic Sobolev inequalities. *American J. Math.*, 97:1061–1083, 1975.

[45] Hamed Hatami. A structure theorem for Boolean functions with small total influences. *Annals of Mathematics*, 176(1):509–533, 2012.

[46] Hao Huang, Po-Shen Loh, and Benny Sudakov. The size of a hypergraph and its matching number. *Combinatorics, Probability and Computing*, 21(03):442–450, 2012.

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

[48] Anders Johansson, Jeff Kahn, and Van Vu. Factors in random graphs. *Random Structures & Algorithms*, 33(1):1–28, 2008.

[49] Jeff Kahn and Gil Kalai. Thresholds and expectation thresholds. *Combinatorics, Probability and Computing*, 16(03):495–502, 2007.

[50] Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on Boolean functions. In *Foundations of Computer Science, 1988., 29th Annual Symposium on*, pages 68–80. IEEE, 1988.

[51] Gyula Katona. A theorem of finite sets. In *Classic Papers in Combinatorics*, pages 381–401. Springer, 2009.

[52] Peter Keevash. Hypergraph Turán problems. *Surveys in combinatorics*, 392:83–140, 2011.

[53] Peter Keevash. Hypergraph matchings and designs. *Proceedings of the 2018 ICM*, 2018.

[54] Peter Keevash, Noam Lifshitz, Eoin Long, and Dor Minzer. Forbidden intersections for codes. *in preparation*, 2020.

[55] Peter Keevash and Eoin Long. A stability result for the cube edge isoperimetric inequality. *J. Combin. Theory Ser. A*, 155:360–375, 2018.

[56] Peter Keevash and Eoin Long. Stability for vertex isoperimetry in the cube. *J. Combin. Theory Ser. B*, 145:113–144, 2020.

[57] Nathan Keller and Noam Lifshitz. The junta method for hypergraphs and Chvátal's simplex conjecture. *arXiv preprint arXiv:1707.02643*, 2017.

[58] Nathan Keller and Noam Lifshitz. Approximation of biased Boolean functions of small total influence by DNF's. *Bulletin of the London Mathematical Society*, 50(4):667–679, 2018.

[59] Subhash Khot, Dor Minzer, Dana Moshkovitz, and Muli Safra. Small set expansion in the Johnson graph. In *Electronic Colloquium on Computational Complexity (ECCC)*, 2018.

[60] Subhash Khot, Dor Minzer, and Muli Safra. Pseudorandom sets in grassmann graph have near-perfect expansion. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 592–601. IEEE, 2018.

[61] Alexandr Kostochka, Dhruv Mubayi, and Jacques Verstraëte. Turán problems and shadows I: paths and cycles. *Journal of Combinatorial Theory, Series A*, 129:57–79, 2015.

[62] Alexandr Kostochka, Dhruv Mubayi, and Jacques Verstraëte. Turán problems and shadows II: trees. *Journal of Combinatorial Theory, Series B*, 122:457–478, 2017.

[63] Joseph B Kruskal. The number of simplices in a complex. *Mathematical optimization techniques*, page 251, 1963.

[64] Shrinivas Kudekar, Santhosh Kumar, Marco Mondelli, Henry D. Pfister, Eren Sasoglu, and Rüdiger L. Urbanke. Reed–Muller codes achieve capacity on erasure channels. *IEEE Transactions on Information Theory*, 63(7):4298–4316, 2017.

[65] Noam Lifshitz. Hypergraph removal lemmas via robust sharp threshold theorems. *Discrete Analysis*, 10, 2020.

[66] László Lovász. *Combinatorial Problems and Exercises*. North-Holland, Amsterdam, 1993.

[67] László Lovász and M.D. Plummer. *Matching theory*. AMS Chelsea Publishing, 2009.

[68] Eyal Lubetzky and Jeffrey Steif. Strong noise sensitivity and random graphs. *The Annals of Probability*, 43(6):3239–3278, 2015.

[69] Tomasz Łuczak and Katarzyna Mieczkowska. On Erdős' extremal problem on matchings in hypergraphs. *Journal of Combinatorial Theory, Series A*, 124:178–194, 2014.

[70] G. Margulis. Probabilistic characteristic of graphs with large connectivity. In *Problems Info. Transmission*. Plenum Press, 1977.

[71] Madan Lal Mehta. *Random matrices*. Elsevier, 2004.

[72] Richard Montgomery. Spanning trees in random graphs. *Advances in Mathematics*, 356:106793, 2019.

[73] Elchanan Mossel. Gaussian bounds for noise correlation of functions. *Geometric and Functional Analysis*, 19(6):1713–1756, 2010.

[74] Elchanan Mossel and Joe Neeman. Robust optimality of Gaussian noise stability. *J. Europ. Math. Soc.*, 17(2):433–482, 2015.

[75] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Annals of Mathematics*, pages 295–341, 2010.

[76] Dhruv Mubayi and Jacques Verstraëte. A survey of Turán problems for expansions. In *Recent Trends in Combinatorics*, pages 117–143. Springer, 2016.

[77] Ryan O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, 2014.

[78] Michał Przykucki and Alexander Roberts. Vertex-isoperimetric stability in the hypercube. *J. Combin. Theory Ser. A*, 172:105186, 2020.

[79] Lucio Russo. An approximate zero-one law. *Probability Theory and Related Fields*, 61(1):129–139, 1982.

[80] Oded Schramm and Jeffrey E Steif. Quantitative noise sensitivity and exceptional times for percolation. *Annals of mathematics*, 171(2):619–672, 2010.

[81] Alexander Schrijver. *Combitorial Optimization: Polyhedra and Efficiancy*. Springer-Verlag Berlin Heidelberg, 2003.

[82] Stanislav Smirnov. Critical percolation and conformal invariance. *Proc. ICM*, 2006.

[83] M Talagrand. Approximate 0-1 law. *Ann. Prob*, 22:1576–1587, 1994.