# TOWARDS AN OPTIMAL HYPERGRAPH CONTAINER LEMMA

MARCELO CAMPOS AND WOJCIECH SAMOTIJ

ABSTRACT. The hypergraph container lemma is a powerful tool in probabilistic combinatorics that has found many applications since it was first proved a decade ago. Roughly speaking, it asserts that the family of independent sets of every uniform hypergraph can be covered by a small number of almost-independent sets, called containers. In this article, we formulate and prove two new versions of the lemma that display the following three attractive features. First, they both admit short and simple proofs that have surprising connections to other well-studied topics in probabilistic combinatorics. Second, they use alternative notions of almost-independence in order to describe the containers. Third, they yield improved dependence of the number of containers on the uniformity of the hypergraph, hitting a natural barrier for second-momenttype approaches.

#### 1. INTRODUCTION

The method of hypergraph containers is a powerful and widely-applicable technique in probabilistic combinatorics. The method enables one to control the independent sets of many 'interesting' uniform hypergraphs by exploiting the fact that these sets exhibit a certain subtle clustering phenomenon. The survey [2] provides a gentle introduction to the method, illustrated with several example applications.

The heart of the method is a hypergraph container lemma (HCL for short) – a statement that formalises and quantifies the notion of clustering we alluded to above. A generic HCL asserts that every uniform hypergraph  $\mathcal{H}$  admits a collection  $\mathcal{C} = \mathcal{C}(\mathcal{H})$  of *containers* for the family  $\mathcal{I}(\mathcal{H})$  of independent sets of  $\mathcal{H}$  with the following three properties:

- (i) each  $I \in \mathcal{I}(\mathcal{H})$  is contained in some  $C \in \mathcal{C}$ ;
- (ii) the family C has 'small' cardinality; and
- (iii) each  $C \in \mathcal{C}$  is 'almost independent'.

Note that such a statement is vacuously true when  $\mathcal{H}$  is 1-uniform, as then each  $I \in \mathcal{I}(\mathcal{H})$  is contained in the largest independent set  $\{v \in V(\mathcal{H}) : \{v\} \notin \mathcal{H}\}$ . Although a HCL for 2-uniform hypergraphs is already implicit in the works of Kleitman and Winston [13] from the early 1980s, an explicit statement was formulated and proved only a decade later by Sapozhenko [22]. General HCLs that are applicable to hypergraphs of an arbitrary uniformity were proved much later, independently and simultaneously, by Balogh, Morris, and Samotij [1] and by Saxton and Thomason [23].

The precise meaning of small in (ii) above depends not only on the notion of almost-independence in (iii), but also on the uniformity of the hypergraph. The dependence on the uniformity in the original HCLs of [1, 23] was rather unfavourable. Roughly speaking, the guaranteed upper

This research was supported by: the Israel Science Foundation grant 2110/22; the grant 2019679 from the United States–Israel Binational Science Foundation (BSF) and the United States National Science Foundation (NSF); and the ERC Consolidator Grant 101044123 (RandomHypGra).

bound on log  $|\mathcal{C}(\mathcal{H})|$  for an *r*-uniform hypergraph  $\mathcal{H}$  in [1] and [23] was proportional to  $e^{\Theta(r^2)}$ and *r*!, respectively. To remedy this, Balogh and Samotij [3] proved a more 'efficient' HCL that yielded an upper bound on log  $|\mathcal{C}(\mathcal{H})|$  that was proportional to merely a polynomial in *r*. Let us also mention that Morris, Samotij, and Saxton [18] proved an 'asymmetric' HCL that was fine-tuned to enumeration of sparse graphs not containing an induced copy of a given subgraph, but has since found additional applications [6, 7, 8, 15].

While the proofs of the four HCLs mentioned in the previous paragraph are elementary, they all are rather lengthy and technical (there are much easier arguments that prove the existence of containers in the 2-uniform case, see the survey [21]). This motivated several groups of authors to seek HCLs that admit shorter and simpler proofs. Saxton and Thomason [25] found an easy method for building containers for independent sets in simple hypergraphs. In another work [24], the same authors presented a streamlined proof of their original HCL [23] that yields an additional 'online' property. A few years later, Bernshteyn, Delcourt, Towsner, and Tserunyan [4] formulated a HCL whose statement adapts notions from nonstandard analysis and supplied a compact (spanning approximately four pages), non-algorithmic proof; unfortunately, their HCL still suffers from unfavourable dependence on the uniformity. Finally, Nenadov [19] recently proved a statement that may be interpreted as a probabilistic HCL; while it does not supply a collection of containers for independent sets, it is sufficiently powerful to replace HCL in most of its typical applications.

1.1. Motivation. We embarked on this project with two independent goals in mind. Our first aim was to further improve the dependence of (the logarithm of) the number of containers on the uniformity of the hypergraph. The second aim was to find a simple(r) construction of containers for general uniform hypergraphs that admits a compact proof. Whereas the first goal has been achieved only partially – each HCL formulated and proved in this paper encounters the same 'second-moment barrier' that results in  $O(r^2)$ -type dependence of log  $|\mathcal{C}|$  on the uniformity r, we do suggest a pathway to improving it further. In fact, we present a conjecture that would imply a logarithmic dependence of log  $|\mathcal{C}|$  on r and that is closely connected to the Park–Pham theorem [20]. As for the second goal, we let the reader be the judge.

1.2. Our results. We formulate two new HCLs that are stronger than the efficient HCL of Balogh and the second author and, at the same time, admit short proofs. The two lemmas use different notions of almost-independence that, unlike all previous works, are not expressed in terms of a 'balanced supersaturation' condition; however, both of them are at least as strong as the usual notion (to certify this, we will present short derivations of the efficient HCL from each of our two lemmas). Furthermore, one of our HCLs does not even require the input hypergraph to be uniform.

The statement of our first HCL, Theorem A below, is inspired by the recent developments in the study of threshold phenomena in random sets [9, 20]. In order to state it, we need to introduce some notation. First, given a hypergraph  $\mathcal{G}$ , we write

$$\langle \mathcal{G} \rangle \coloneqq \bigcup_{E \in \mathcal{G}} \{ F \subseteq V(\mathcal{G}) : F \supseteq E \}$$

for the up-set generated by  $\mathcal{G}$ . We say that  $\mathcal{G}$  covers a hypergraph  $\mathcal{H}$  if  $\mathcal{H} \subseteq \langle \mathcal{G} \rangle$ ; in other words,  $\mathcal{G}$  covers  $\mathcal{H}$  is every edge of  $\mathcal{H}$  contains some edge of  $\mathcal{G}$ . Second, for each  $p \in [0, 1]$ , define the

*p*-weight of a hypergraph  $\mathcal{G}$  to be

$$w_p(\mathcal{G}) \coloneqq \sum_{E \in \mathcal{G}} p^{|E|}$$

which is just the expected number of edges of  $\mathcal{G}$  induced by the *p*-random subset<sup>1</sup> of  $V(\mathcal{G})$ .

**Theorem A.** Let  $\mathcal{H}$  be an r-uniform hypergraph with a finite vertex set V. For every  $p \in (0, 1/(8r^2)]$ , there exists a family  $S \subseteq 2^V$  and functions

$$g: \mathcal{I}(\mathcal{H}) \to \mathcal{S}$$
 and  $f: \mathcal{S} \to 2^V$ 

such that:

- (a) For each  $I \in \mathcal{I}(\mathcal{H})$ , we have  $g(I) \subseteq I \subseteq f(g(I))$ .
- (b) Each  $S \in S$  has at most  $8r^2p|V|$  elements.
- (c) For every  $S \in S$ , letting C := f(S), there exists a hypergraph  $\mathcal{G}$  on C with

 $w_p(\mathcal{G}) \leqslant p|C|$ 

that covers  $\mathcal{H}[C]$  and satisfies  $|E| \ge 2$  for all  $E \in \mathcal{G}$ .

Let us point out that the above theorem does not assume anything about the input hypergraph, except that it is uniform. The condition that all edges of the cover have size at least two is necessary to prevent the conclusion from being trivial; indeed, every hypergraph with vertex set C admits a cover (by singletons) of p-weight p|C|. An analogous comment applies to the assumption that  $p \leq 1/(8r^2)$ ; if  $p \geq 1/(8r^2)$ , then one may simply set f(I) = g(I) = I for all  $I \in \mathcal{I}(\mathcal{H})$  and let  $\mathcal{G}$  be the empty hypergraph.

Our second HCL, Theorem B, has a probabilistic flavour; this stems form the fact that its proof views independent sets as samples from a hard-core model. Perhaps surprisingly, it does not assume anything about the hypergraph, not even uniformity. Given an arbitrary set X and a real  $p \in [0, 1]$ , we write  $X_p$  to denote the *p*-random subset of X.

**Theorem B.** Let  $\mathcal{H}$  be a hypergraph with a finite vertex set V. For all reals  $\delta$  and p satisfying  $0 , there exists a family <math>S \subseteq 2^V$  and functions

$$g: \mathcal{I}(\mathcal{H}) \to \mathcal{S}$$
 and  $f: \mathcal{S} \to 2^{V}$ 

such that:

- (a) For each  $I \in \mathcal{I}(\mathcal{H})$ , we have  $g(I) \subseteq I \subseteq f(g(I))$ .
- (b) Each  $S \in S$  has at most  $p|V|/\delta$  elements.
- (c) For every  $S \in S$ , letting  $C \coloneqq f(S)$ ,

$$\mathbb{P}(S \cup C_p \in \mathcal{I}(\mathcal{H})) \ge (1-p)^{\delta |C \setminus S|}$$

The proofs of both Theorems A and B follow the same general strategy that was also used to prove most HCLs to date; it can be traced back to the work of Kleitman and Winston [13]. Namely, we define an algorithm that, given an independent set I as input, builds sets S and Csatisfying  $S \subseteq I \subseteq C$  and, in the proof of Theorem A, also a hypergraph  $\mathcal{G}$  that covers  $\mathcal{H}[C]$ . Crucially, the set C depends on I only through S in the following sense: If for some two inputs

<sup>&</sup>lt;sup>1</sup>The *p*-random subset of a finite set X is the random set formed by independently retaining each element of X with probability p.

4

I, I', the algorithm outputs the same set S, it also returns the same set C. This property allows one to define the functions g and f via the algorithm. The main novelty in the algorithm that underlies the proof of Theorem A is that it considers queries of the form 'Is  $L \subseteq I$ ?' for sets Lthat are not necessarily singletons. The algorithm used in the proof of Theorem B considers only queries of the form 'Does  $v \in I$ ?', but it chooses the vertex v based on its occupancy probability in a certain hard-core model.

1.3. Comparing Theorems A and B. Even though the descriptions of the container sets f(S) in Theorems A and B look rather different, they are in fact essentially equivalent. This is a consequence of the following proposition, which quantifies the relation between the smallest *p*-weight of a cover for a uniform hypergraph  $\mathcal{H}$  and the probability that the *p*-random subset of vertices of  $\mathcal{H}$  is independent.

**Proposition 1.1.** Suppose that  $\mathcal{H}$  is a hypergraph on a finite vertex set V.

(i) For every hypergraph  $\mathcal{G} \subseteq 2^V \setminus \{\emptyset\}$  that covers  $\mathcal{H}$  and all  $p \in (0, 1/2)$ ,

$$\mathbb{P}(V_p \in \mathcal{I}(\mathcal{H})) \geqslant \exp(-2w_p(\mathcal{G}))$$

(ii) If  $\mathcal{H}$  is r-uniform, then, for every  $p \in (0, 1/(4r))$ , there exists a hypergraph  $\mathcal{G} \subseteq 2^V \setminus \{\emptyset\}$ that covers  $\mathcal{H}$  and satisfies

$$\mathbb{P}(V_p \in \mathcal{I}(\mathcal{H})) \leqslant \exp(-w_{p/(4r^2)}(\mathcal{G})/8).$$

We postpone the proof of Proposition 1.1 to Section 2 and now only elaborate on the qualitative equivalence between the notions of almost-independence used in Theorems A and B. First, let  $\mathcal{G}$  be the hypergraph from assertion (c) of Theorem A. The assumption that  $|E| \ge 2$  for all  $E \in \mathcal{G}$  and the fact  $w_p(\mathcal{G}) \le p|C|$  guarantee that, for all  $\delta \in (0, 1)$  and all  $p' \le \delta p$ ,

$$w_{p'}(\mathcal{G}) \leqslant (p'/p)^2 \cdot w_p(\mathcal{G}) \leqslant (p'/p) \cdot p'|C| \leqslant \delta p'|C|.$$

Since  $\mathcal{G}$  covers  $\mathcal{H}[C]$ , it follows from part (i) of Proposition 1.1 that  $\mathbb{P}(V_{p'} \in \mathcal{I}(\mathcal{H}[C])) \geq \exp(-2\delta p'|C|)$ .

Conversely, assume that  $p \leq 1/(4r)$  and let S and C be the sets from assertion (c) of Theorem B. Since  $\mathbb{P}(C_p \in \mathcal{I}(\mathcal{H}[C])) \geq \mathbb{P}(S \cup C_p \in \mathcal{I}(\mathcal{H}))$ , it follows from part (ii) of Proposition 1.1 that  $\mathcal{H}[C]$  admits a cover  $\mathcal{G}$  that satisfies

$$w_{p/(4r^2)}(\mathcal{G}) \leqslant -8\log \mathbb{P}(S \cup C_p \in \mathcal{I}(\mathcal{H})) \leqslant -8\log(1-p)^{\delta|C|} \leqslant 10\delta p|C|$$

1.4. Relation to earlier HCLs. An attractive feature of Theorems A and B is that each of them implies (a slight strengthening of) the 'efficient' HCL of Balogh and Samotij [3, Theorem 1.1], stated here as Theorem C below, which in turn significantly improves the original HCLs proved in [1, 23]. Given a hypergraph  $\mathcal{H}$  with vertex set V and a set  $L \subseteq V$ , we define

$$\deg_{\mathcal{H}} L \coloneqq |\{E \in \mathcal{H} : L \subseteq E\}|.$$

Further, for an integer  $\ell \ge 1$ , we let

$$\Delta_{\ell}(\mathcal{H}) \coloneqq \max\{ \deg_{\mathcal{H}} L : L \subseteq V, |L| = \ell \}.$$

The following statement can be easily derived from both Theorem A as well as Theorem B and Janson's inequality. We present the two derivations in Section 5.

**Theorem C** ([3]). Let  $\mathcal{H}$  be an r-uniform hypergraph with a finite vertex set V. Suppose that  $\tau \in (0,1)$  and  $K \ge r$  are such that, for every  $\ell \in [\![r]\!]$ ,

$$\Delta_{\ell}(\mathcal{H}) \leqslant K \cdot \left(\frac{\tau}{32Kr^2}\right)^{\ell-1} \cdot \frac{e(\mathcal{H})}{|V|}.$$
(1)

There exists a family  $\mathcal{S} \subseteq 2^V$  and functions

$$g: \mathcal{I}(\mathcal{H}) \to \mathcal{S} \qquad and \qquad f: \mathcal{S} \to 2^V$$

such that:

- (a) For each  $I \in \mathcal{I}(\mathcal{H})$ , we have  $g(I) \subseteq I \subseteq f(g(I))$ .
- (b) Each  $S \in S$  has at most  $\tau |V|$  elements.
- (c) For every  $S \in S$ , we have  $|f(S)| \leq (1 1/(2K))|V|$ .

In most applications of the hypergraph container method, one recursively applies a basic HCL, such as Theorem C, to construct a family of containers that are almost independent in the sense that one can no longer prove a balanced supersaturation result in any of the containers. In order to avoid explicitly performing such a recursive procedure, one may instead invoke one of the several 'packaged' HCLs that exist in the literature (see, e.g., [1, Theorem 2.2], [3, Theorem 1.6], or [23, Corollary 3.6]). Both Theorems A and B allow one to deduce such a packaged HCL directly, without the need for recursion. For example, the following statement is a straightforward consequence of Theorem A.

**Theorem D.** Let  $\mathcal{H}$  be an r-uniform hypergraph with a finite vertex set V. For every  $p \in (0, 1/(8r^2)]$ , there exists a family  $S \subseteq 2^V$  and functions

$$g: \mathcal{I}(\mathcal{H}) \to \mathcal{S}$$
 and  $f: \mathcal{S} \to 2^V$ 

such that:

- (a) For each  $I \in \mathcal{I}(\mathcal{H})$ , we have  $g(I) \subseteq I \subseteq f(g(I))$ .
- (b) Each  $S \in S$  has at most  $8r^2p|V|$  elements.
- (c) For every  $S \in S$ , there does not exist a hypergraph  $\mathcal{H}_S \subseteq \mathcal{H}[f(S)]$  that satisfies

$$\Delta_{\ell}(\mathcal{H}_S) < \frac{p^{\ell-1}e(\mathcal{H}_S)}{|f(S)|} \quad \text{for all } \ell \in \{2, \dots, r\}.$$

$$\tag{2}$$

We note that condition (c) in the above theorem is the negation of a balanced supersaturation statement that would enable one to construct a small family of containers for independent sets in  $\mathcal{H}[f(S)]$  using a basic HCL such as Theorem C.

We present the derivation of Theorem D from Theorem A in Section 5 and only mention here that the existence of a hypergraph  $\mathcal{G}$  with small *p*-weight that covers  $\mathcal{H}[f(S)]$  precludes the existence of a hypergraph  $\mathcal{H}_S \subseteq \mathcal{H}[f(S)]$  that satisfies the balancedness condition (2). We further note that the derivation of Theorem D from Theorem A in fact allows one to prove a strengthening of the former, where the assertion (c) is replaced by the stronger assertion that there does not exist a probability distribution on  $\mathcal{H}[f(S)]$  under which the random edge  $\mathbf{R} \in \mathcal{H}[f(S)]$  satisfies

$$\mathbb{P}(L \subseteq \mathbf{R}) < \frac{p^{\ell-1}}{|f(S)|} \quad \text{for all } L \subseteq V \text{ with } 2 \leqslant |L| \leqslant r.$$

Note the striking similarity between this condition and the notion of a *p*-spread measure, introduced by Talagrand [26], which played a key role in the proof [9] of the fractional version of the 'expectation-threshold' conjecture of Kahn and Kalai [12], conjectured by Talagrand [26].

Finally, we remark that one may use Theorem B to derive an alternative version of Theorem D, where assumption (2) is replaced by an analogous sequence of upper bounds on the degrees  $\Delta_1(\mathcal{H}_S), \ldots, \Delta_r(\mathcal{H}_S)$ . The existence of an  $\mathcal{H}_S \subseteq \mathcal{H}[f(S)]$  that satisfies such a modified assumption would imply, via Janson's inequality, an upper bound on the probability that the *p*-random set  $f(S)_p$  is independent in  $\mathcal{H}_S$ , and thus also in  $\subseteq \mathcal{H}[f(S)]$ , that is smaller than the upper bound on this probability asserted by Theorem B.

1.5. Interpolating between Theorems A and B. Before concluding our discussion, we state one more HCL, where the characterisation of almost independence interpolates between those given by Theorems A and B.

**Theorem E.** Let  $\mathcal{H}$  be a hypergraph with a finite vertex set V. For all reals  $\delta$  and p satisfying  $0 , there exists a family <math>S \subseteq 2^V$  and functions

$$g: \mathcal{I}(\mathcal{H}) \to \mathcal{S} \qquad and \qquad f: \mathcal{S} \to 2^{V}$$

such that:

- (a) For each  $I \in \mathcal{I}(\mathcal{H})$ , we have  $g(I) \subseteq I \subseteq f(g(I))$ .
- (b) Each  $S \in S$  has at most  $p|V|/\delta$  elements.
- (c) For every  $S \in S$ , letting C := f(S), there exists a hypergraph  $\mathcal{G}$  with vertex set  $C \setminus S$ and  $|E| \ge 2$  for all  $E \in \mathcal{G}$  that covers  $\mathcal{H}[C]$  and satisfies

$$\mathbb{P}(L \subseteq C_p \mid C_p \in \mathcal{I}(\mathcal{G})) \ge (1-\delta)^{|L|} p^{|L|}$$

for all  $L \in \mathcal{I}(\mathcal{G})$ .

The proof of Theorem E, which we sketch in Section 6, is very similar to the proof of Theorem B, the only difference being that the algorithm that is used here to build the sets S and Cconsiders the more general queries of the form 'Is  $L \subseteq I$ ?', as in the proof of Theorem A.

While it may not be immediately clear, Theorem E simultaneously strengthens both Theorems A and B. We postpone the proofs of these two facts (Theorem  $E \implies$  Theorem A and Theorem  $E \implies$  Theorem B) to Section 6.

1.6. **Organisation.** In Section 2, we set a few notational conventions, state several auxiliary results, and prove Proposition 1.1. In Sections 3 and 4, we prove Theorems A and B, respectively. Section 5 is devoted to derivations of standard HCLs (Theorems C and D) from Theorems A and B. In Section 6, we sketch the proof of Theorem E and show that it implies both Theorems A and B. In Section 7, we present an attractive conjecture that strengthens Proposition 1.1 and discuss some of its implications. Finally, Appendix A contains three proofs of one of our key auxiliary results, Proposition 2.2.

# 2. Preliminaries

2.1. Notation. Given a hypergraph  $\mathcal{H}$  and a positive integer s, we define

$$\mathcal{H}^s \coloneqq \{E \in \mathcal{H} : |E| = s\}.$$

It will be also convenient to introduce the shorthand notations

$$\mathcal{H}^{>1} \coloneqq \bigcup_{s>1} \mathcal{H}^s$$
 and  $\mathcal{H}^{< r} \coloneqq \bigcup_{s < r} \mathcal{H}^s$ 

Finally, for every  $L \subseteq V(\mathcal{H})$ , we denote by  $\partial_L \mathcal{H}$  the link of L in  $\mathcal{H}$ , i.e.,

$$\partial_L \mathcal{H} \coloneqq \{ E \setminus L : L \subseteq E \in \mathcal{H} \};$$

for the sake of brevity, we will write  $\partial_v \mathcal{H}$  in place of  $\partial_{\{v\}} \mathcal{H}$ .

2.2. Auxiliary results. While comparing Theorem B to other HCLs, we will crucially use the following well-known inequality due to Janson [10].

**Theorem 2.1** (Janson's Inequality). Suppose that  $\mathcal{G}$  is a hypergraph on a finite set C. For every  $p \in [0, 1]$ ,

$$\mathbb{P}(C_p \in \mathcal{I}(\mathcal{G})) \leq \exp\left(-\frac{\mu^2}{2\Delta^*}\right),$$

where

$$\mu\coloneqq \sum_{A\in\mathcal{G}}p^{|A|}\qquad and\qquad \Delta^*\coloneqq \sum_{\substack{A,B\in\mathcal{G}\\A\cap B\neq \emptyset}}p^{|A\cup B|}$$

and the second sum ranges over all ordered pairs (A and B are not necessarily distinct).

Our proof of Theorem B, as well as the derivation of this theorem from Theorem E, requires the following probabilistic estimate. Since this estimate follows from standard techniques, we postpone its proof to the appendix (where we in fact present three different proofs).

**Proposition 2.2.** Suppose that C is a finite set and let  $\mathcal{I} \subseteq 2^C$  be a decreasing family of subsets. For every  $p \in (0, 1)$ , we have

$$\log \mathbb{P}(C_p \in \mathcal{I}) \ge \left( |C| - \frac{\mathbb{E}\left[ |C_p| \mid C_p \in \mathcal{I} \right]}{p} \right) \cdot \log(1-p).$$

Finally, we will also employ the following well-known inequality, discovered independently by Bollobás [5], Lubell [16], Meshalkin [17], and Yamamoto [27].

**Proposition 2.3** (LYMB inequality). Suppose that X is a finite set and that  $\mathcal{A} \subseteq 2^X$  is an antichain. Then,

$$\sum_{A \in \mathcal{A}} \binom{|X|}{|A|}^{-1} \leqslant 1.$$

2.3. Proof of Proposition 1.1. The first assertion is a straightforward application of Harris's inequality. Indeed, for every hypergraph  $\mathcal{G} \subseteq 2^V \setminus \{\emptyset\}$  that covers  $\mathcal{H}$ , we have

$$\mathbb{P}(V_p \in \mathcal{I}(\mathcal{H})) \ge \mathbb{P}(V_p \in \mathcal{I}(\mathcal{G})) \ge \prod_{A \in \mathcal{G}} (1 - p^{|A|}) \ge \exp\left(-2\sum_{A \in \mathcal{G}} p^{|A|}\right),$$

where the final inequality holds as  $1 - x \ge e^{-2x}$  for all x < 1/2.

We now prove the second assertion. Suppose that  $\mathcal{H}$  is *r*-uniform and define, for each  $\ell \in [\![r]\!]$ ,  $\lambda_{\ell} := 4^{-\min\{\ell, r-\ell\}}$ . Let  $\mathcal{H}'$  be a maximal subgraph of  $\mathcal{H}$  subject to  $\Delta_{\ell}(\mathcal{H}') \leq \lambda_{\ell} {r \choose \ell}^{-1} p^{\ell-r}$  for all

 $\ell \in [\![r]\!].$  Define  $\mu \coloneqq \sum_{A \in \mathcal{H}'} p^{|A|} = e(\mathcal{H}')p^r$  and

$$\Delta^* \coloneqq \sum_{A \in \mathcal{H}'} p^{|A|} \cdot \sum_{\substack{B \in \mathcal{H}' \\ A \cap B \neq \emptyset}} p^{|B \setminus A|} \leqslant \mu \cdot \sum_{\ell=1}^r \binom{r}{\ell} \Delta_\ell(\mathcal{H}') p^{r-\ell} \leqslant \mu \cdot \sum_{\ell=1}^r \lambda_\ell \leqslant 2\mu.$$

Let  $\mathcal{G}$  be the set of minimal elements in the family

$$\mathcal{G}' \coloneqq \left\{ T \subseteq V : 1 \leqslant |T| \leqslant r \land \deg_{\mathcal{H}'} T = \left\lfloor \lambda_{|T|} \binom{r}{|T|}^{-1} p^{|T|-r} \right\rfloor \right\}$$

and note that  $\mathcal{H}' \subseteq \langle \mathcal{G}' \rangle = \langle \mathcal{G} \rangle$  by maximality of  $\mathcal{H}'$ . Since  $\mathcal{G}$  is an antichain, the LYMB inequality (Proposition 2.3) yields

$$e(\mathcal{H}') \ge \sum_{A \in \mathcal{H}'} \sum_{\substack{T \in \mathcal{G} \\ T \subseteq A}} \binom{r}{|T|}^{-1} = \sum_{T \in \mathcal{G}} \binom{r}{|T|}^{-1} \deg_{\mathcal{H}'} T.$$

Since our upper-bound assumption on p yields

$$\lambda_{\ell} \binom{r}{\ell}^{-1} p^{\ell-r} \ge (4rp)^{\ell-r} \ge 1$$

for every  $\ell \in \llbracket r \rrbracket$ , we may conclude that

$$\mu = e(\mathcal{H}')p^r \ge \frac{1}{2} \sum_{T \in \mathcal{G}} \lambda_{|T|} \binom{r}{|T|}^{-2} p^{|T|} \ge \frac{1}{2} \sum_{T \in \mathcal{G}} \left(\frac{p}{4r^2}\right)^{|T|} = \frac{w_{p/(4r^2)}(\mathcal{G})}{2}.$$

We may now apply Janson's inequality (Theorem 2.1) to conclude that

$$\mathbb{P}(V_p \in \mathcal{I}(\mathcal{H})) \leqslant \mathbb{P}(V_p \in \mathcal{I}(\mathcal{H}')) \leqslant \exp\left(-\frac{\mu^2}{2\Delta^*}\right) \leqslant \exp\left(-\frac{\mu}{4}\right) \leqslant \exp\left(-\frac{w_{p/(4r^2)}(\mathcal{G})}{8}\right),$$

as desired.

#### 3. Proof of Theorem A

3.1. The algorithm. Assume that  $\mathcal{H}$  is an *r*-uniform hypergraph on a finite set *V* and fix some  $p \in (0, 1/(8r^2)]$ . The following algorithm takes as input an independent set  $I \in \mathcal{I}(\mathcal{H})$  and returns sets  $S, C \subseteq V$  and a hypergraph  $\mathcal{G}$  on *V*.

- (1) Let  $\mathcal{H}_0 \coloneqq \mathcal{H}$  and  $S_0 \coloneqq \emptyset$ .
- (2) Do the following for  $i = 0, 1, \ldots$ :
  - (a) Let  $C_i \coloneqq \{v \in V : \{v\} \notin \mathcal{H}_i\}.$
  - (b) If  $w_p(\mathcal{H}_i^{>1}) \leq p|C_i|$ , then set  $J \coloneqq i$  and STOP.
  - (c) Otherwise, let  $s_i$  be the smallest  $s \in \{2, \ldots, r\}$  such that  $w_p(\partial_L \mathcal{H}_i^s) \ge 1/(4r)$  for some nonempty  $L \in 2^V \setminus \mathcal{H}_i$  and let  $L_i$  be an inclusion-maximal set L with this property.<sup>2</sup>
  - (d) If  $L_i \subseteq I$ , then set  $S_{i+1} \coloneqq S_i \cup L_i$  and  $\mathcal{F}_i \coloneqq \partial_{L_i} \mathcal{H}_i^{s_i}$ .
  - (e) Otherwise, if  $L_i \notin I$ , set  $S_{i+1} \coloneqq S_i$  and  $\mathcal{F}_i \coloneqq \{L_i\}$ .
  - (f) Finally, set  $\mathcal{H}_{i+1} \coloneqq (\mathcal{H}_i \setminus \langle \mathcal{F}_i \rangle) \cup \mathcal{F}_i$ .
- (3) Return  $S \coloneqq S_J, C \coloneqq C_J$ , and  $\mathcal{G} \coloneqq \mathcal{H}_J^{>1}$ .

 $<sup>\</sup>overline{^{2}$ It may not be clear at this point that such s and L exist, but we will prove this shortly.

3.2. Well-definedness. We now show that the algorithm described in the previous section is well defined and that it terminates on every input. Our first lemma takes care of the former and shows that the definition of  $s_i$  and  $L_i$  in step (2c) is always valid.

**Lemma 3.1.** If  $w_p(\mathcal{H}_i^{>1}) \ge p|C_i|$ , then there are  $s \in \{2, \ldots, r\}$  and nonempty  $L \in 2^V \setminus \mathcal{H}_i$  with  $w_p(\partial_L \mathcal{H}_i^s) \ge 1/r$ .

*Proof.* Note first that our assumption implies that

$$p|C_i| \leqslant w_p(\mathcal{H}_i^{>1}) = \sum_{s=2}^r w_p(\mathcal{H}_i^s) \leqslant \sum_{s=2}^r s \cdot w_p(\mathcal{H}_i^s) = \sum_{s=2}^r \sum_{v \in C_i} p \cdot w_p(\partial_v \mathcal{H}_i^s)$$

Consequently, there must be some  $v \in C_i$  such that

$$\sum_{s=2}^{r} w_p(\partial_v \mathcal{H}_i^s) \ge 1.$$

Since the fact that  $v \in C_i$  guarantees that  $\{v\} \notin \mathcal{H}_i$ , we may take  $L \coloneqq \{v\}$  and an index s that achieves  $\max_s w_p(\partial_v \mathcal{H}_i^s) \ge 1/r$ .

Next, we show that each hypergraph  $\mathcal{H}_i$  defined in the algorithm is an antichain and that the sequence  $(\langle \mathcal{H}_i \rangle)_i$  is strictly increasing. Note that this implies that the algorithm terminates on every input. Indeed, the fact that  $\mathcal{H}_i$  is an antichain means that it comprises precisely all the minimal elements of  $\langle \mathcal{H}_i \rangle$  and therefore  $\mathcal{H}_0, \mathcal{H}_1, \ldots$  are pairwise distinct.

# **Lemma 3.2.** The hypergraph $\mathcal{H}_i$ is an antichain for every *i*.

Proof. We prove the assertion of the lemma by induction on i. The induction basis holds due to our assumption that  $\mathcal{H}_0 = \mathcal{H}$  is r-uniform. For the induction step, assume that  $\mathcal{H}_i$  is an antichain for some  $i \ge 0$ . Observe first that  $\mathcal{F}_i$  is an antichain as well. Indeed, this is clear when  $\mathcal{F}_i = \{L_i\}$ ; otherwise  $\mathcal{F}_i \subseteq \partial_{L_i} \mathcal{H}_i$  and this is a straightforward consequence of the assumption that  $\mathcal{H}_i$  is an antichain. Suppose that  $\mathcal{H}_{i+1} = (\mathcal{H}_i \setminus \langle \mathcal{F}_i \rangle) \cup \mathcal{F}_i$  contained a pair of sets E, F with  $E \subsetneq F$ . Since both  $\mathcal{H}_i$  and  $\mathcal{F}_i$  are antichains, then either  $E \in \mathcal{F}_i$  and  $F \in \mathcal{H}_i \setminus \langle \mathcal{F}_i \rangle$ , which is clearly impossible, or vice-versa. However, if  $F \in \mathcal{F}_i$ , then  $F \subseteq G$  for some  $G \in \mathcal{H}_i$ . (Indeed, when  $\mathcal{F}_i = \partial_{L_i} \mathcal{H}_i^{s_i}$ , we may take  $G = F \cup L_i$ ; otherwise,  $F = L_i$  and the existence of such a Gfollows from the fact that  $w_p(\partial_{L_i} \mathcal{H}_i) > 0$ .) As a result, if there was  $E \in \mathcal{H}_i$  with  $E \subsetneq F$ , then  $E \subsetneq G$  for some  $G \in \mathcal{H}_i$ , contradicting the inductive assumption that  $\mathcal{H}_i$  is an antichain.  $\Box$ 

**Lemma 3.3.** We have  $\langle \mathcal{H}_i \rangle \subsetneq \langle \mathcal{H}_{i+1} \rangle$  and  $\mathcal{F}_i \cap \langle \mathcal{H}_i \rangle = \emptyset$  for every *i*.

Proof. The definition of  $\mathcal{H}_{i+1}$  implies that  $\langle \mathcal{H}_{i+1} \rangle = \langle \mathcal{H}_i \rangle \cup \langle \mathcal{F}_i \rangle \supseteq \langle \mathcal{H}_i \rangle$ . To complete the proof of the lemma, it is therefore enough to show that  $\mathcal{F}_i \neq \emptyset$  and  $\mathcal{F}_i \cap \langle \mathcal{H}_i \rangle = \emptyset$ . The former statement is clear when  $\mathcal{F}_i = \{L_i\}$ ; otherwise, it follows from the assumption that  $\partial_{L_i} \mathcal{H}_i^{s_i} \neq \emptyset$ . For the latter statement, since  $\mathcal{H}_i$  is an antichain (by Lemma 3.2), it is enough to argue that some element of  $\mathcal{F}_i$  is a proper subset of some element of  $\mathcal{H}_i$ . When  $\mathcal{F}_i = \{L_i\}$ , this is the case since  $L_i \notin \mathcal{H}_i$  and  $\partial_{L_i} \mathcal{H}_i \neq \emptyset$ ; otherwise, if  $\mathcal{F}_i = \partial_{L_i} \mathcal{H}_i^{s_i}$ , this follows as  $L_i \neq \emptyset$  and  $\mathcal{F}_i \neq \emptyset$ .  $\Box$ 

3.3. **Properties.** We now turn to establishing some key properties of the algorithm. We will show that, for every input  $I \in \mathcal{I}(\mathcal{H})$ , we have  $S \subseteq I \subseteq C$ , the set S has at most  $8r^2p|V|$  elements, and the hypergraph  $\mathcal{G}$  (which clearly satisfies  $w_p(\mathcal{G}) \leq p|C|$ , see the stopping condition (2b), and  $|E| \ge 2$  for all  $E \in \mathcal{G}$ ) covers  $\mathcal{H}[C]$ . Further, we will prove that the set C (and the hypergraph  $\mathcal{G}$ ) depends on I only through S, which will allow us to define the functions f and g by g(I) := S and f(g(I)) := C.

**Lemma 3.4.** The hypergraph  $\mathcal{G}$  covers  $\mathcal{H}[C]$ .

*Proof.* Note first that Lemma 3.3 implies that  $\mathcal{H} = \mathcal{H}_0 \subseteq \langle \mathcal{H}_0 \rangle \subseteq \langle \mathcal{H}_J \rangle$ . Let F be an arbitrary edge of  $\mathcal{H}[C]$ . Since  $\mathcal{H}_J$  covers  $\mathcal{H}$ , there must be some  $E \in \mathcal{H}_J$  with  $E \subseteq F$ . However, as  $F \subseteq C = C_J$ , the set E cannot be a singleton and thus  $E \in \mathcal{H}_J^{>1} = \mathcal{G}$ , as required.  $\Box$ 

**Lemma 3.5.** For each *i*, we have  $I \in \mathcal{I}(\mathcal{H}_i)$  and  $S_i \subseteq I \subseteq C_i$ .

Proof. It suffices to show that  $S_i \subseteq I \in \mathcal{I}(\mathcal{H}_i)$  for all i; indeed,  $I \in \mathcal{I}(\mathcal{H}_i)$  implies that  $I \subseteq C_i$ , as  $\{v\} \in \mathcal{H}_i$  for every  $v \in V \setminus C_i$ . We prove this by induction on i. The induction basis follows as  $S_0 = \emptyset$  and  $I \in \mathcal{I}(\mathcal{H}) = I(\mathcal{H}_0)$ . For the induction step, assume that  $S_i \subseteq I \in \mathcal{I}(\mathcal{H}_i)$  for some  $i \ge 0$ . Note first that either  $S_{i+1} = S_i$  or  $S_{i+1} \setminus S_i \subseteq L_i \subseteq I$  and thus  $S_{i+1} \subseteq I$ . Further, since  $\mathcal{H}_{i+1} \subseteq \mathcal{H}_i \cup \mathcal{F}_i$ , it suffices to show that  $I \in \mathcal{I}(\mathcal{F}_i)$ . If  $L_i \not\subseteq I$ , then  $\mathcal{F}_i = \{L_i\}$  and thus  $\mathcal{I}(\mathcal{F}_i)$ comprises all subsets of V not containing  $L_i$  (which includes I). Otherwise, if  $L_i \subseteq I$ , we have  $\mathcal{F}_i \subseteq \partial_{L_i}\mathcal{H}_i$  and thus  $\mathcal{I}(\mathcal{F}_i) \supseteq \mathcal{I}(\partial_{L_i}\mathcal{H}_i)$ ; further,  $\mathcal{I}(\partial_{L_i}\mathcal{H}_i)$  contains all sets  $I' \subseteq V$  such that  $L_i \cup I' \in \mathcal{I}(\mathcal{H}_i)$  (which again includes  $I = L_i \cup I$ ).

**Lemma 3.6.** For every *i*, the hypergraph  $\mathcal{F}_i$  is *u*-uniform for some  $u < s_i$ .

Proof. Suppose first that  $\mathcal{F}_i = \{L_i\}$ . In this case,  $\mathcal{F}_i$  is clearly  $|L_i|$ -uniform and  $|L_i| < s_i$  as  $w_p(\partial_{L_i}\mathcal{H}_i^{s_i}) > 0$  and  $L_i \notin \mathcal{H}_i$ . Suppose now that  $\mathcal{F}_i = \partial_{L_i}\mathcal{H}_i^{s_i}$ . In this case,  $\mathcal{F}_i$  is clearly  $(s_i - |L_i|)$ -uniform and  $s_i - |L_i| < s_i$  as  $L_i$  is nonempty.

**Lemma 3.7.** For every *i* and all *s* and  $L \subseteq V$  with  $1 \leq |L| < s < r$ ,

$$w_p(\partial_L \mathcal{H}_i^s) \leqslant \frac{1}{2r}$$

*Proof.* We prove the lemma by induction on i. The base case i = 0 is trivial, as  $\mathcal{H}$  is r-uniform and thus the hypergraphs  $\mathcal{H}_0^2, \ldots, \mathcal{H}_0^{r-1}$  are all empty. For the induction step, assume that the assertion holds for some  $i \ge 0$  (and all s and L). Since  $\mathcal{H}_{i+1} \subseteq \mathcal{H}_i \cup \mathcal{F}_i$ , we have

$$w_p(\partial_L \mathcal{H}_{i+1}^s) \leqslant w_p(\partial_L \mathcal{H}_i^s) + w_p(\partial_L \mathcal{F}_i^s) \tag{3}$$

for all s and all  $L \subseteq V$ . Further, since the hypergraph  $\mathcal{F}_i$  is u-uniform for some  $u < s_i$  (by Lemma 3.6), we have  $w_p(\partial_L \mathcal{F}_i^s) = 0$  unless s = u. We may therefore assume from now on that s = u, as otherwise the desired inequality (for all  $L \subseteq V$  with  $1 \leq |L| < s$ ) follows from (3) and the inductive assumption. The minimality of  $s_i$  and the fact that  $u < s_i$  imply that  $w_p(\partial_L \mathcal{H}_i^u) < 1/(4r)$ . It thus suffices to show that  $w_p(\partial_L \mathcal{F}_i) < 1/(4r)$  for every  $L \subseteq V$ with |L| < u. Fix some such set L. If  $\mathcal{F}_i = \partial_{L_i} \mathcal{H}_i^{s_i}$ , then the maximality of  $L_i$  implies that  $w_p(\partial_L \mathcal{F}_i) = \mathbbm{1}_{L \cap L_i = \emptyset} \cdot w_p(\partial_{L \cup L_i} \mathcal{H}_i^{s_i}) < 1/(4r)$ . Otherwise, if  $\mathcal{F} = \{L_i\}$ , we have

$$w_p(\partial_L \mathcal{F}_i) = \mathbb{1}_{L \subseteq L_i} \cdot p^{|L_i| - |L|} = \mathbb{1}_{L \subseteq L_i} \cdot p^{u - |L|} \leqslant p \leqslant 1/(4r)$$

as desired.

**Corollary 3.8.** For every *i* and all  $L \in 2^V \setminus \langle \mathcal{H}_i \rangle$ ,

$$w_p(\partial_L \mathcal{H}_i^{< r}) \leqslant \frac{1}{2}.$$

		1
- 14		-1

**Lemma 3.9.** For every  $i \in \{0, ..., J - 1\}$ ,

$$w_p(\mathcal{H}_{i+1}^{< r}) \geqslant w_p(\mathcal{H}_i^{< r}) + \frac{1}{8r} \cdot \mathbb{1}_{L_i \subseteq I}$$

*Proof.* Since the hypergraph  $\mathcal{F}_i$  is *u*-uniform for some  $u < s_i \leq r$  (by Lemma 3.6), it follows from the definition of  $\mathcal{H}_{i+1}$  that

$$w_p(\mathcal{H}_{i+1}^{< r}) - w_p(\mathcal{H}_i^{< r}) \ge \sum_{L \in \mathcal{F}_i} p^{|L|} \cdot \left(1 - w_p(\partial_L \mathcal{H}_i^{< r})\right).$$

Further, since  $\mathcal{F}_i \cap \langle \mathcal{H}_i \rangle = \emptyset$  (by Lemma 3.3), Corollary 3.8 implies that  $w_p(\partial_L \mathcal{H}_i^{< r}) \leq 1/2$  for each  $L \in \mathcal{F}_i$  and thus  $w_p(\mathcal{H}_{i+1}^{< r}) - w_p(\mathcal{H}_i^{< r}) \geq w_p(\mathcal{F}_i)/2$ . The assertion of the lemma follows as  $w_p(\mathcal{F}_i) \geq 0$  and  $w_p(\mathcal{F}_i) = w_p(\partial_{L_i} \mathcal{H}_i^{s_i}) \geq 1/(4r)$  when  $L_i \subseteq I$ .

Lemma 3.10. We have  $|S| \leq 8r^2p|V|$ .

*Proof.* Let *m* denote the number of rounds of the algorithm in which  $L_i \subseteq I$ . Since  $\mathcal{H}_0$  is *r*-uniform, it follows from Lemma 3.9 that

$$w_p(\mathcal{H}_J^{< r}) = w_p(\mathcal{H}_J^{< r}) - w_p(\mathcal{H}_0^{< r}) = \sum_{i=0}^{J-1} \left( w_p(\mathcal{H}_{i+1}^{< r}) - w_p(\mathcal{H}_i^{< r}) \right) \ge \frac{m}{8r}$$

On the other hand, since  $\mathcal{H}_J = \mathcal{H}_J^1 \cup \mathcal{G}$  and  $w_p(\mathcal{G}) \leq p|C_J|$ , we have

$$w_p(\mathcal{H}_J) = w_p(\mathcal{H}_J^1) + w_p(\mathcal{G}) = |V \setminus C_J| \cdot p + w_p(\mathcal{G}) \leq p|V|$$

We conclude that  $m \leq 8r|V|$ . Since  $|S_{i+1}| - |S_i| \leq r$  when  $L_i \subseteq I$  and  $|S_{i+1}| = |S_i|$  otherwise, the assertion of the lemma follows.

**Lemma 3.11.** Suppose that the algorithm with inputs  $I, I' \in \mathcal{I}(\mathcal{H})$  outputs  $(S, C, \mathcal{G})$  and  $(S', C', \mathcal{G}')$ , respectively. If S = S', then  $(C, \mathcal{G}) = (C', \mathcal{G}')$ .

*Proof.* Observe first that the execution of the algorithm on a given input (an independent set of  $\mathcal{H}$ ) depends solely on how the set  $S_{i+1}$  and the hypergraph  $\mathcal{F}_{i+1}$  are defined in each round i (via step (2d) or step (2e)). Thus, if the outputs of the algorithm applied to I and I' are different, then there must be some i for which  $L'_i \subseteq I'$  but  $L_i \notin I$ , or vice-versa. Let i be the *earliest* such round and note that  $\mathcal{H}'_i = \mathcal{H}_i$  and thus also  $L'_i = L_i$ . We may clearly assume that  $L_i = L'_i \subseteq I'$  and  $L_i \notin I$ . However, Lemma 3.5 implies that  $L_i = L'_i \subseteq S'_{i+1} \subseteq S' = S \subseteq I$ , a contradiction.

We finish our discussion with a formal proof of Theorem A.

Proof of Theorem A. For each independent set  $I \in \mathcal{I}(\mathcal{H})$ , set  $g(I) \coloneqq S$  and  $f(S) \coloneqq C$ , where  $(S, C, \mathcal{G})$  is the output of the algorithm with input I. (Recall that Lemmas 3.2 and 3.3 imply that the algorithm terminates on every input.) Further, set  $S \coloneqq \{g(I) : I \in \mathcal{I}(\mathcal{H})\}$ . Lemma 3.11 guarantees that the function f is well-defined. By Lemma 3.5, we have  $g(I) \subseteq I \subseteq f(g(I))$  for every  $I \in \mathcal{I}(\mathcal{H})$  whereas Lemma 3.10 shows that  $|S| \leq 8r^2pn$  for every  $S \in S$ . Finally, writing C = f(g(I)), Lemma 3.4 guarantees that the hypergraph  $\mathcal{G}$  is a cover for  $\mathcal{H}[C]$  with  $w_p(\mathcal{G}) \leq p|C|$ ; the property that  $|E| \geq 2$  for all  $E \in \mathcal{G}$  is straightforward from the definition of  $\mathcal{G}$ . This concludes the proof of the theorem.  $\Box$ 

# 4. Proof of Theorem B

4.1. The algorithm. Assume that  $\mathcal{H}$  is a hypergraph on a finite set V and fix some  $p \in (0, 1]$ . The following algorithm takes as input an independent set  $I \in \mathcal{I}(\mathcal{H})$  and returns sets  $S, C \subseteq V$ .

- (1) Let  $\mathcal{H}_0 \coloneqq \mathcal{H}$  and  $S_0 \coloneqq \emptyset$ .
- (2) Do the following for  $i = 0, 1, \ldots$ :
  - (a) If there exists a vertex  $v \in V \setminus S_i$  such that  $\{v\} \notin \mathcal{H}_i$  and

$$\mathbb{P}(v \in V_p \mid V_p \in \mathcal{I}(\mathcal{H}_i)) < (1 - \delta)p$$

let  $v_i$  be some such vertex; otherwise, set J := i and STOP.

- (b) If  $v_i \in I$ , set  $S_{i+1} \coloneqq S_i \cup \{v_i\}$  and  $\mathcal{H}_{i+1} \coloneqq \mathcal{H}_i \cup \partial_{v_i} \mathcal{H}_i$ .
- (c) Otherwise, if  $v_i \notin I$ , set  $S_{i+1} \coloneqq S_i$  and  $\mathcal{H}_{i+1} \coloneqq \mathcal{H}_i \cup \{\{v_i\}\}\}$ .
- (3) Return  $S \coloneqq S_J$  and  $C \coloneqq \{v \in V : \{v\} \notin \mathcal{H}_J\}.$

4.2. **Properties.** We now establish some key properties of the algorithm. We first claim that the algorithm always terminates with  $J \leq |V|$ . Indeed, for each *i*, the vertex  $v_i$  is either added to  $S_i$  or  $\{v_i\}$  becomes an edge of  $\mathcal{H}_i$ ; in both cases,  $v_i$  is never considered again in step (2a). Further, we claim that the set *C* can be reconstructed with the knowledge of *S* only; indeed, running the algorithm with input *S* instead of *I* results in the same sets *S* and *C*. This allows us to define the functions *f* and *g* by  $g(I) \coloneqq S$  and  $f(g(I)) \coloneqq C$ . It remains to show that, for every input  $I \in \mathcal{I}(\mathcal{H})$ , we have  $S \subseteq I \subseteq C$ , the set *S* has at most  $p|V|/\delta$  elements, and that  $S \cup C_p \in \mathcal{I}(\mathcal{H})$  with probability at least  $(1-p)^{\delta|C|}$ .

**Lemma 4.1.** For each *i*, we have  $S_i \subseteq I \in \mathcal{I}(\mathcal{H}_i)$  and  $I' \in \mathcal{I}(\mathcal{H}_i) \iff S_i \cup I' \in \mathcal{I}(\mathcal{H}_i)$ .

*Proof.* We prove the assertion of the lemma by induction on i. The induction basis follows as  $S_0 = \emptyset$  and  $\mathcal{H} = \mathcal{H}_0$ . For the induction step, assume that the assertion holds for some  $i \ge 0$ . Suppose first that  $v_i \in I$ . In this case,  $S_{i+1} = S_i \cup \{v_i\} \subseteq I$  and  $\mathcal{H}_{i+1} = \mathcal{H}_i \cup \partial_{v_i} \mathcal{H}_i$ , which means that  $\mathcal{I}(\mathcal{H}_{i+1}) = \{I' \subseteq V : \{v_i\} \cup I' \in \mathcal{I}(\mathcal{H}_i)\} \ni I$ . Further

$$I' \in \mathcal{I}(\mathcal{H}_{i+1}) \iff \{v_i\} \cup I' \in \mathcal{I}(\mathcal{H}_i) \iff S_i \cup \{v_i\} \cup I' \in \mathcal{I}(\mathcal{H}_i) \iff S_{i+1} \cup I' \in \mathcal{I}(\mathcal{H}_{i+1}).$$

Suppose now that  $v_i \notin I$ . In this case,  $S_{i+1} = S_i \subseteq I$  and  $\mathcal{H}_{i+1} = \mathcal{H}_i \cup \{\{v_i\}\}$ , which means that  $\mathcal{I}(\mathcal{H}_{i+1}) = \{I' \in \mathcal{I}(\mathcal{H}_i) : v_i \notin I'\} \ni I$ . Further,  $I' \in \mathcal{I}(\mathcal{H}_{i+1}) \iff S_{i+1} \cup I' \in \mathcal{I}(\mathcal{H}_{i+1})$ , as  $S_{i+1} = S_i \not\ni v_i$ .

The above lemma clearly implies that  $S = S_J \subseteq I$ . Further,  $I \in \mathcal{I}(\mathcal{H}_J) \subseteq 2^C$ , as no vertex v satisfying  $\{v\} \in \mathcal{H}_J$  can belong to an independent set of  $\mathcal{H}_J$ . This establishes  $S \subseteq I \subseteq C$ .

**Lemma 4.2.** We have  $\delta |S| \leq p|V|$ .

*Proof.* We claim that the sequence of probabilities  $P_i := \mathbb{P}(V_p \in \mathcal{I}(\mathcal{H}_i))$  satisfies

$$P_{i+1} \leqslant (1-\delta)^{\mathbb{1}_{v_i \in I}} \cdot P_i,$$

for all  $i \in \{0, \ldots, J-1\}$ , which gives

$$(1-p)^{|V|} = \mathbb{P}(V_p = \emptyset) \leqslant P_J \leqslant (1-\delta)^{|\{i:v_i \in I\}|} \cdot P_0 = (1-\delta)^{|S|} \cdot P_0 \leqslant (1-\delta)^{|S|}.$$
 (4)

Indeed, if  $v_i \notin I$ , then the inequality  $P_{i+1} \leq P_i$  follows as  $\mathcal{H}_i \subseteq \mathcal{H}_{i+1}$ , and thus  $\mathcal{I}(\mathcal{H}_{i+1}) \subseteq \mathcal{I}(\mathcal{H}_i)$ . On the other hand, if  $v_i \in I$ , then  $\mathcal{H}_{i+1} = \mathcal{H}_i \cup \partial_{v_i} \mathcal{H}_i$ , and thus  $\mathcal{I}(\mathcal{H}_{i+1}) = \{I' \subseteq V : I' \cup \{v_i\} \in I\}$   $\mathcal{I}(\mathcal{H}_i)$ . Consequently, by the choice of  $v_i$ ,

$$p \cdot P_{i+1} = \mathbb{P}\big(v_i \in V_p \in \mathcal{I}(\mathcal{H}_i)\big) = \mathbb{P}(v_i \in V_p \mid V_p \in \mathcal{I}(\mathcal{H}_i)) \cdot P_i < (1-\delta)p \cdot P_i,$$

which gives the desired inequality. We claim that (4) implies the desired inequality  $\delta |S| \leq p|V|$ . Indeed, if  $\delta |S| > p|V|$ , then  $(1 - \delta)^{|S|} < (1 - \delta)^{p|V|/\delta} \leq (1 - p)^{|V|}$ , where the second inequality follows as  $p \leq \delta$  and the map  $x \mapsto (1 - x)^{1/x}$  is decreasing on (0, 1).

**Lemma 4.3.** We have  $\mathbb{P}(S \cup C_p \in \mathcal{I}(\mathcal{H})) \ge (1-p)^{\delta|C \setminus S|}$ .

*Proof.* Define  $C' \coloneqq C \setminus S$ . Since clearly  $\mathcal{H} = \mathcal{H}_0 \subseteq \mathcal{H}_1 \subseteq \cdots \subseteq \mathcal{H}_J$ , we have

$$\mathbb{P}(S \cup C_p \in \mathcal{I}(\mathcal{H})) \geqslant \mathbb{P}(S \cup C_p \in \mathcal{I}(\mathcal{H}_J)) = \mathbb{P}(S \cup C'_p \in \mathcal{I}(\mathcal{H}_J)).$$

In view of Proposition 2.2, the assertion of the lemma will follow if we prove that

$$\mathbb{E}\big[|C'_p| \mid S \cup C'_p \in \mathcal{I}(\mathcal{H}_J)\big] \ge (1-\delta)p|C'|.$$
(5)

To this end, fix an arbitrary  $v \in C'$ . Since  $\mathcal{I}(\mathcal{H}_J) \subseteq 2^C$ , as we have already observed above, and since  $S \cup V_p \in \mathcal{I}(\mathcal{H}_J) \iff V_p \in \mathcal{I}(\mathcal{H}_J)$ , by Lemma 4.1, we have

$$\mathbb{P}\big(v \in C'_p \mid S \cup C'_p \in \mathcal{I}(\mathcal{H}_J)\big) = \mathbb{P}\big(v \in V_p \mid S \cup V_p \in \mathcal{I}(\mathcal{H}_J)\big) = \mathbb{P}\big(v \in V_p \mid V_p \in \mathcal{I}(\mathcal{H}_J)\big).$$

Finally, since  $v \in C' = C \setminus S$ , the stopping condition (2a) implies that the above probability is at least  $(1 - \delta)p$ . Summing this inequality over all  $v \in C'$  yields (5).

We conclude our discussion with a short derivation of Theorem B.

Proof of Theorem B. For each independent set  $I \in \mathcal{I}(\mathcal{H})$ , set  $g(I) \coloneqq S$  and  $f(S) \coloneqq C$ , where (S, C) is the output of the algorithm with input I. Further, set  $S \coloneqq \{g(I) : I \in \mathcal{I}(\mathcal{H})\}$ . Since the algorithm will perform the exact same operations when we replace the input set I with S, the function f is well-defined. Now, Lemma 4.1 implies assertion (a), Lemma 4.2 proves assertion (b), and Lemma 4.3 establishes assertion (c).

#### 5. Derivations of the standard HCLs

Derivation of Theorem C from Theorem A. We apply Theorem A with  $p := \tau/(8r^2)$  to obtain a family S, functions f and g, and a hypergraph  $\mathcal{G}_S$  for every  $S \in S$ . Assertions (a) and (b) follow immediately from the respective assertions in Theorem A, so we only need to argue that (c) holds as well.

To this end, fix some  $S \in S$ , let C := f(S), and let  $\mathcal{G} := \mathcal{G}_S$  be the hypergraph covering  $\mathcal{H}[C]$ that satisfies  $w_p(\mathcal{G}) \leq p|C| \leq p|V|$  and whose all edges have size at least two. By the assumption of the theorem, for every  $E \in \mathcal{G}$ ,

$$\deg_{\mathcal{H}} E \leqslant K \cdot \left(\frac{p}{4K}\right)^{|E|-1} \cdot \frac{e(\mathcal{H})}{|V|} < \frac{p^{|E|-1}}{2} \cdot \frac{e(\mathcal{H})}{|V|}$$

where the second inequality follows as  $|E| \ge 2$  and  $K \ge r \ge 1$ . The fact that  $\mathcal{G}$  covers  $\mathcal{H}[C]$  implies that

$$e(\mathcal{H}[C]) \leqslant \sum_{E \in \mathcal{G}} \deg_{\mathcal{H}} E < \frac{e(\mathcal{H})}{2p|V|} \cdot \sum_{E \in \mathcal{G}} p^{|E|} = \frac{e(\mathcal{H})}{2p|V|} \cdot w_p(\mathcal{H}[C]) \leqslant \frac{e(\mathcal{H})}{2}.$$

On the other hand,

$$e(\mathcal{H}[C]) \ge e(\mathcal{H}) - |V \setminus C| \cdot \Delta_1(\mathcal{H}) \ge e(\mathcal{H}) - |V \setminus C| \cdot \frac{Ke(\mathcal{H})}{|V|},$$
  
which yields  $|V \setminus C|/|V| > 1/(2K)$  or, equivalently,  $|C| < (1 - 1/(2K))|V|.$ 

*Remark.* A careful examination of the proof allows one to improve the bounds in assumption (1) in Theorem C somewhat further. First, note that we only needed the bounds

$$\Delta_1(\mathcal{H}) \leqslant K \cdot \frac{e(\mathcal{H})}{|V|}$$
 and  $\Delta_\ell(\mathcal{H}) \leqslant \left(\frac{\tau}{32r^2}\right)^{\ell-1} \frac{e(\mathcal{H})}{|V|}$  for  $\ell \in \{2, \dots, r\}$ .

Further, we claim that changing step (2b) in the algorithm from the proof of Theorem A to

(2b') If 
$$w_p(\mathcal{H}_i^{>1}) \leq (\log r/r) \cdot p|C_i|$$
 or  $|C_i| \leq (1 - 1/(2K))|V|$ , then set  $J \coloneqq i$  and STOP

allows us to derive a stronger version of Theorem C, with the upper bound on  $\Delta_{\ell}(\mathcal{H})$  in (1) increased by a factor of roughly  $(r/\log r)^{\ell-1}$ . To see this, notice that the only modifications needed in the proof of Theorem A occur in Lemmas 3.1 and 3.10. The conclusion of Lemma 3.1 still holds since

$$\frac{\log r}{r} \cdot p|C_i| \leqslant w_p(\mathcal{H}_i^{>1}) = \sum_{s=2}^r \sum_{v \in C_i} \frac{p \cdot w_p(\partial_v \mathcal{H}_i^s)}{s}$$

and, consequently, there must be some  $v \in C_i$  such that

$$\sum_{s=2}^{r} \frac{w_p(\partial_v \mathcal{H}_i^s)}{s} \ge \frac{\log r}{r} \ge \sum_{s=2}^{r} \frac{1}{rs}$$

which implies that  $w_p(\partial_v \mathcal{H}_i^s) \ge 1/r$  for some s. On the other hand, the assertion of Lemma 3.10 may now be strengthened. Indeed, since now

$$w_p(\mathcal{H}_J) = |V \setminus C_J| \cdot p + w_p(\mathcal{G}) \leqslant \frac{|V|p}{2K} + \frac{\log r}{r} \cdot |V|p \leqslant \frac{2\log r}{r} \cdot |V|p|$$

we may deduce the stronger bound  $|S| \leq 16(r \log r) \cdot p|V|$ , which allows the conclusion of Theorem C to be reached under the weaker assumption

$$\Delta_1(\mathcal{H}) \leqslant K \cdot \frac{e(\mathcal{H})}{|V|} \quad \text{and} \quad \Delta_\ell(\mathcal{H}) \leqslant \left(\frac{\tau}{2^6 r \log r}\right)^{\ell-1} \frac{e(\mathcal{H})}{|V|} \quad \text{for } \ell \in \{2, \dots, r\}.$$

Derivation of Theorem C from Theorem B. We apply Theorem B with  $p \coloneqq \tau/(16Kr)$  and  $\delta \coloneqq p/\tau = (16Kr)^{-1}$  to obtain a family S and functions f and g. Assertions (a) and (b) follow immediately from the respective assertions in Theorem B, so we only need to argue that (c) holds as well.

Suppose that  $S \in \mathcal{S}$  and let  $C \coloneqq f(S)$ . We may assume that  $e(\mathcal{H}[C]) \ge e(\mathcal{H})/2$ , as otherwise  $|C| \le (1-1/(2K))|V|$  follows from the assumption that  $\Delta_1(\mathcal{H}) \le Ke(\mathcal{H})/|V|$ , as in the previous derivation. Preparing to apply Janson's inequality (Theorem 2.1), we define  $\mu \coloneqq \sum_{A \in \mathcal{H}[C]} p^{|A|} = e(\mathcal{H}[C])p^r$  and

$$\Delta^* \coloneqq \sum_{A \in \mathcal{H}[C]} p^{|A|} \sum_{\substack{B \in \mathcal{H}[C] \\ A \cap B \neq \emptyset}} p^{|B \setminus A|} \leqslant \mu \cdot \sum_{\ell=1}^{r} \binom{r}{\ell} \Delta_{\ell}(\mathcal{H}) p^{r-\ell}.$$

By the assumption of the theorem,

$$\sum_{\ell=1}^{r} \binom{r}{\ell} \Delta_{\ell}(\mathcal{H}) p^{1-\ell} \leqslant \sum_{\ell=1}^{r} r^{\ell} K \left( \frac{\tau}{32Kr^2p} \right)^{\ell-1} \cdot \frac{e(\mathcal{H})}{|V|} = \frac{Kre(\mathcal{H})}{|V|} \cdot \sum_{\ell=1}^{r} 2^{1-\ell},$$

and consequently, by Janson's inequality,

$$\mathbb{P}(C_p \in \mathcal{I}(\mathcal{H})) \leqslant \exp\left(-\frac{\mu^2}{2\Delta^*}\right) \leqslant \exp\left(-\frac{\mu|V|}{4Kre(\mathcal{H})p^{r-1}}\right) \leqslant \exp\left(-\frac{p|V|}{8Kr}\right),$$

where the last inequality holds thanks to our assumption that  $e(\mathcal{H}[C]) \ge e(\mathcal{H})/2$ . On the other hand, the assertion of Theorem B gives

$$\mathbb{P}(C_p \in \mathcal{I}(\mathcal{H})) \ge \mathbb{P}(S \cup C_p \in \mathcal{I}(\mathcal{H})) \ge (1-p)^{\delta|C|} > \exp\left(-2\delta p|V|\right) \ge \exp\left(-\frac{p|V|}{8Kr}\right),$$
  
intradiction.

a contradiction.

Derivation of Theorem D from Theorem A. We apply Theorem A to obtain a family  $\mathcal{S}$ , functions f and g, and a hypergraph  $\mathcal{G}_S$  for every  $S \in \mathcal{S}$ . Assertions (a) and (b) follow immediately from the respective assertions in Theorem A, so we only need to argue that (c) holds as well.

To this end, fix some  $S \in \mathcal{S}$ , let  $C \coloneqq f(S)$ , and let  $\mathcal{G} \coloneqq \mathcal{G}_S$  be the hypergraph covering  $\mathcal{H}[C]$  that satisfies  $w_p(\mathcal{G}) \leq p|C|$  and whose all edges have size at least two. Assume toward contradiction that there is a hypergraph  $\mathcal{H}_S \subseteq \mathcal{H}[C]$  that satisfies (2). Since  $\mathcal{G}$  covers  $\mathcal{H}[C]$ , and thus also  $\mathcal{H}_S$ , we have

$$e(\mathcal{H}_S) \leqslant \sum_{E \in \mathcal{G}} \deg_{\mathcal{H}_S} E < \sum_{E \in \mathcal{G}} \frac{p^{|E|-1}e(\mathcal{H}_S)}{|C|} = w_p(\mathcal{G}) \cdot \frac{e(\mathcal{H}_S)}{p|C|} \leqslant e(\mathcal{H}_S),$$

a contradiction.

#### 6. Interpolating between Theorems A and B

In this section, we sketch the proof of Theorem E, which is a fairly straightforward adaptation of the proof of Theorem B (given in Section 4), and present derivations of Theorems A and B from Theorem E.

6.1. The algorithm. Assume that  $\mathcal{H}$  is a hypergraph on a finite set V and fix some  $p \in (0, 1]$ . The following algorithm takes as input an independent set  $I \in \mathcal{I}(\mathcal{H})$  and returns sets  $S, C \subseteq V$ and a hypergraph  $\mathcal{G}$ .

- (1) Let  $\mathcal{H}_0 \coloneqq \mathcal{H}$  and  $S_0 \coloneqq \emptyset$ .
- (2) Do the following for  $i = 0, 1, \ldots$ :
  - (a) If there exists a set  $L \subseteq V \setminus S_i$  such that  $L \in \mathcal{I}(\mathcal{H}_i)$  and

$$\mathbb{P}(L \subseteq V_p \mid V_p \in \mathcal{I}(\mathcal{H}_i)) < (1-\delta)^{|L|} p^{|L|},$$

let  $L_i$  be a minimal such set; otherwise, set  $J \coloneqq i$  and STOP.

- (b) If  $L_i \subseteq I$ , set  $S_{i+1} \coloneqq S_i \cup L_i$  and  $\mathcal{H}_{i+1} \coloneqq \{E \setminus L_i : E \in \mathcal{H}_i\} \Longrightarrow \partial^*_{L_i}(\mathcal{H}_i)$ .
- (c) Otherwise, if  $L_i \nsubseteq I$ , set  $S_{i+1} \coloneqq S_i$  and  $\mathcal{H}_{i+1} \coloneqq \mathcal{H}_i \cup \{L_i\}$ .

(3) Return  $S \coloneqq S_J, C \coloneqq \{v \in V : \{v\} \notin \mathcal{H}_J\}$  and  $\mathcal{G} \coloneqq \mathcal{H}_J[C]$ .

6.2. Properties. The above algorithm always terminates because no set L is ever considered more than once. Indeed, either  $L_i \subseteq S_{i+1}$ , and thus  $L_i \not\subseteq V \setminus S_j$  for all j > i, or  $L_i \in \mathcal{H}_{i+1}$ , and thus  $L_i \notin \mathcal{I}(\mathcal{H}_j)$  for all j > i. Moreover, the set C and the hypergraph  $\mathcal{G}$  can be reconstructed with the knowledge of S only, as running the algorithm with input S instead of I results in the same sets S and C and hypergraph  $\mathcal{G}$ . This allows us to define  $g(I) \coloneqq S$  and  $f(g(I)) \coloneqq C$ .

It remains to show that, for every input  $I \in \mathcal{I}(\mathcal{H})$ , we have  $S \subseteq I \subseteq C$ , the set S has at most  $p|V|/\delta$  elements, and that  $\mathcal{G}$  is a cover of  $\mathcal{H}[C]$  with the desired property.

The assertion of Lemma 4.1, that is,  $S_i \subseteq I \in \mathcal{I}(\mathcal{H}_i)$  and  $I' \in \mathcal{I}(\mathcal{H}_i) \iff S_i \cup I' \in \mathcal{I}(\mathcal{H}_i)$  for all *i*, remains true, with essentially the same proof. Consequently,  $I \in \mathcal{I}(\mathcal{H}_J)$  and thus  $I \subseteq C$ , as no vertex in  $V \setminus C$  can belong to an independent set of  $\mathcal{H}_J$ . Observe additionally that, for all *i*, we have  $\langle \mathcal{H}_i \rangle \subseteq \langle \mathcal{H}_{i+1} \rangle$  and thus  $\mathcal{I}(\mathcal{H}_{i+1}) \subseteq \mathcal{I}(\mathcal{H}_i)$ . The desired upper bound on |S| can be established by adapting the proof of Lemma 4.2. Indeed, note first that it is enough to show that the sequence  $P_i \coloneqq \mathbb{P}(V_p \in \mathcal{I}(\mathcal{H}_i))$  satisfies the inequality  $P_{i+1} \leq (1-\delta)^{\mathbb{1}_{L_i} \subseteq I \cdot |L_i|} \cdot P_i$  for all *i*. The inequality  $P_{i+1} \leq P_i$  follows as  $\mathcal{I}(\mathcal{H}_{i+1}) \subseteq \mathcal{I}(\mathcal{H}_i)$ . Further, if  $L_i \subseteq I$ , then

$$p^{|L_i|} \cdot \mathbb{P}(V_p \in \mathcal{I}(\mathcal{H}_{i+1})) = \mathbb{P}(L_i \subseteq V_p \in \mathcal{I}(\mathcal{H}_i)) < (1-\delta)^{|L_i|} p^{|L_i|} \cdot \mathbb{P}(V_p \in \mathcal{I}(\mathcal{H}_i)),$$

where the equality holds as  $\mathcal{I}(\mathcal{H}_{i+1}) = \mathcal{I}(\partial_{L_i}^* \mathcal{H}_i) = \{I' \in \mathcal{I}(\mathcal{H}_i) : L_i \cup I' \in \mathcal{I}(\mathcal{H}_i)\}.$ 

Last but not least, we argue that  $\mathcal{G}$  possesses all the claimed properties. First,  $\mathcal{G} = \mathcal{H}_J[C]$ covers  $\mathcal{H}[C]$  since  $\langle \mathcal{H} \rangle = \langle \mathcal{H}_0 \rangle \subseteq \cdots \subseteq \langle \mathcal{H}_J \rangle$ . Second, each  $E \in \mathcal{G}$  satisfies  $|E| \ge 2$  as otherwise we would have  $E \cap C = \emptyset$ , by the definition of C. We also remark that, for all  $L \subseteq C \setminus S$  with  $L \in \mathcal{I}(\mathcal{G})$ , we have

$$\mathbb{P}(L \subseteq C_p \mid C_p \in \mathcal{I}(\mathcal{G})) = \mathbb{P}(L \subseteq V_p \mid V_p \in \mathcal{I}(\mathcal{H}_J)) \ge (1-\delta)^{|L|} p^{|L|},$$

by the terminating condition in the algorithm and the fact that  $\mathcal{I}(\mathcal{H}_J) \subseteq 2^C$ .

Finally, we show that  $E \cap S = \emptyset$  for every  $E \in \mathcal{G}$ . Since  $S = S_J$  and  $\mathcal{G} \subseteq \mathcal{H}_J$ , it is enough to prove the stronger statement that  $E \cap S_i = \emptyset$  for all  $E \in \mathcal{H}_i$  and all *i*. We show this by induction on *i*. The induction basis is vacuously true as  $S_0 = \emptyset$ . Assume that  $i \ge 0$ . If  $L_i \not\subseteq I$ , then  $S_{i+1} = S_i$  and the only edge in  $\mathcal{H}_{i+1} \setminus \mathcal{H}_i$  is  $L_i$ , which is disjoint from  $S_i$ . Else, if  $L_i \subseteq I$ , then  $S_{i+1} \setminus S_i = L_i$  and we replace each  $E \in \mathcal{H}_i$ , which is disjoint from  $S_i$ , with the edge  $E \setminus L_i = E \setminus S_{i+1}$ , which is disjoint from  $S_{i+1}$ .

6.3. **Derivations of Theorems A and B.** We first show that Theorem B follows by a simple application of Proposition 2.2.

Derivation of Theorem B from Theorem E. Since the assumptions of the two theorems are identical, and their assertions differ only in the characterisation of the sets f(S), we just need to show that assertion (c) in Theorem E implies that, for every  $S \in S$ , we have

$$\mathbb{P}(S \cup C_p \in \mathcal{I}(\mathcal{H})) \ge (1-p)^{\delta |C \setminus S|},$$

where C := f(S). Let  $C' := C \setminus S$  and note that the fact that  $\mathcal{G}$  is a cover of  $\mathcal{H}[C]$  implies that

$$\mathbb{P}\big(S \cup C_p \in \mathcal{I}(\mathcal{H})\big) = \mathbb{P}\big(S \cup C'_p \in \mathcal{I}(\mathcal{H})\big) \geqslant \mathbb{P}\big(S \cup C'_p \in \mathcal{I}(\mathcal{G})\big).$$

Further, the fact that  $E \cap S = \emptyset$  for all  $E \in \mathcal{G}$  implies that, for every  $v \in C'$ 

$$\mathbb{P}\big(v \in C'_p \mid S \cup C'_p \in \mathcal{I}(\mathcal{G})\big) = \mathbb{P}\big(v \in C'_p \mid C_p \in \mathcal{I}(\mathcal{G})\big) \ge (1-\delta)p.$$

It follows that

 $\mathbb{E}\big[|C'_p| \mid S \cup C'_p \in \mathcal{I}(\mathcal{G})\big] \ge (1-\delta)p|C'|$ 

and consequently, by Proposition 2.2, that  $\mathbb{P}(S \cup C'_p \in \mathcal{I}(\mathcal{G})) \ge (1-p)^{\delta|C'|}$ .

The derivation of Theorem A from Theorem E is decidedly more complicated. We will show that that the cover  $\mathcal{G}$  of  $\mathcal{H}[C]$  given by Theorem E has the property  $w_p(\mathcal{G}) \leq p|C|$ . We will argue by contradiction, showing that, if  $w_p(\mathcal{G}) > p|C|$ , some maximal set L among those whose degree in  $\mathcal{G}$  is large satisfies  $\mathbb{P}(L \subseteq C_p \mid C_p \in \mathcal{I}(\mathcal{G})) < (1-\delta)^{|L|}p^{|L|}$ ; the proof of this inequality will employ a second-moment argument.

Derivation of Theorem A from Theorem E. Suppose that  $\mathcal{H}$  is an r-uniform hypergraph with vertex set V and let  $p \in (0, 1/(8r^2)]$  be arbitrary. We apply Theorem E with  $\delta \coloneqq 1/(4r)$ and density parameter  $q \coloneqq 2rp \leq 1/(4r) = \delta$  to obtain a family  $S \subseteq 2^V$  of sets satisfying  $|S| \leq q|V|/\delta = 8r^2p|V|$  for each  $S \in S$  and functions  $g: \mathcal{I}(\mathcal{H}) \to S$  and  $f: S \to 2^V$  such that  $g(I) \subseteq I \subseteq f(g(I))$  for all  $I \in \mathcal{I}(\mathcal{H})$ . It suffices to show that, for every  $S \in S$ , there is a hypergraph  $\mathcal{G}$  on V with all edges of size at least two that covers  $\mathcal{H}[C]$  and satisfies  $w_p(\mathcal{G}) \leq p|C|$ , where  $C \coloneqq f(S)$ . To this end, fix some  $S \in S$  and let  $\mathcal{G}$  be the set of all inclusion-minimal elements in the cover given by (c) in Theorem E. As  $\mathcal{G}$  covers  $\mathcal{H}[C]$  and satisfies  $|E| \geq 2$  for all  $E \in \mathcal{G}$ , we only need to show that  $w_p(\mathcal{G}) \leq p|C|$ .

Assume by contradiction that  $w_p(\mathcal{G}) > p|C|$ . We claim that there exists a set  $L \in \mathcal{I}(\mathcal{G})$  and an  $s \in \{2, \ldots, r\}$  such that

$$w_p(\partial_L \mathcal{G}^s) \ge 1/r.$$
 (6)

To this end, note that

$$p|C| < w_p(\mathcal{G}) = \sum_{s=2}^r w_p(\mathcal{G}^s) = \sum_{s=2}^r \frac{1}{s} \sum_{v \in C} p \cdot w_p(\partial_v \mathcal{G}^s) \leqslant p \cdot \sum_{v \in C} \sum_{s=2}^r w_p(\partial_v \mathcal{G}^s),$$

which means that (6) must hold with  $L = \{v\}$  for some  $v \in C$  and  $s \in \{2, \ldots, r\}$ ; note that  $\{v\} \in \mathcal{I}(\mathcal{G})$  for every  $v \in C$  since all edges of  $\mathcal{G}$  have at least two vertices. Let L be an inclusion-maximal set that is independent in  $\mathcal{G}$  and satisfies (6) for some  $s \in \{2, \ldots, r\}$  and let  $\ell := |L| \in [s-1]$ . Such choice of L guarantees that, on the one hand,

$$|\partial_L \mathcal{G}^s| = p^{\ell-s} \cdot w_p(\partial_L \mathcal{G}^s) \geqslant p^{\ell-s}/r$$

and, on the other hand, for every  $t \in [s - \ell - 1]$ ,

$$\Delta_t(\partial_L \mathcal{G}^s) = \max\{|\partial_{L\cup T} \mathcal{G}^s| : T \subseteq V \setminus L \land |T| = t\}$$
  
= max{ $p^{t+\ell-s} \cdot w_p(\partial_{L\cup T} \mathcal{G}^s) : T \subseteq V \setminus L \land |T| = t\} < p^{t+\ell-s}/r,$ 

where the last inequality follows as every set  $L' \supseteq L$  that has fewer than s elements and satisfies  $\partial_{L'} \mathcal{G}^s \neq \emptyset$  is independent in  $\mathcal{G}$  (otherwise, L' would be a proper subset of some edge of  $\mathcal{G}^s$ , contradicting the fact that  $\mathcal{G}$  is an antichain).

Since  $L \in \mathcal{I}(\mathcal{G})$ , assertion (c) of Theorem E yields

$$\mathbb{P}(L \subseteq C_q \mid C_q \in \mathcal{I}(\mathcal{G})) \ge (1-\delta)^{\ell} q^{\ell} \ge (1-\delta\ell)q^{\ell} \ge 3q^{\ell}/4$$

whereas, by Harris's inequality, writing  $\tilde{C}$  for  $C_q$  conditioned on the event  $C_q \in \mathcal{I}(\mathcal{G})$ ,

$$\mathbb{P}(L \subseteq C_q \mid C_q \in \mathcal{I}(\mathcal{G})) = \mathbb{P}(L \subseteq C_q \wedge C_q \in \mathcal{I}(\partial_L \mathcal{G}) \mid C_q \in \mathcal{I}(\mathcal{G}))$$
$$\leqslant q^{\ell} \cdot \mathbb{P}(C_q \in \mathcal{I}(\partial_L \mathcal{G}) \mid C_q \in \mathcal{I}(\mathcal{G})) = q^{\ell} \cdot \mathbb{P}(\tilde{C} \in \mathcal{I}(\partial_L \mathcal{G})),$$

Let  $X := e(\partial_L \mathcal{G}^s[\tilde{C}])$  denote the number of edges that  $\tilde{C}$  induces in  $\partial_L \mathcal{G}^s$ . Using the Paley– Zygmund inequality, we may bound

$$\frac{3}{4} \leqslant \mathbb{P}\big(\tilde{C} \in \mathcal{I}(\partial_L \mathcal{G})\big) \leqslant \mathbb{P}\big(\tilde{C} \in \mathcal{I}(\partial_L \mathcal{G}^s)\big) = \mathbb{P}(X=0) \leqslant 1 - \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}$$

which means that  $\mathbb{E}[X^2] \ge 4\mathbb{E}[X]^2$ .

Since each  $E \in \partial_L \mathcal{G}^s$  is independent in  $\mathcal{G}$  (as  $L \neq \emptyset$  and  $\mathcal{G}$  is an antichain), assertion (c) of Theorem E yields

$$\mathbb{E}[X] = \sum_{E \in \partial_L \mathcal{G}^s} \mathbb{P}(E \subseteq \tilde{C}) \geqslant \sum_{E \in \partial_L \mathcal{G}^s} (1-\delta)^{|E|} q^{|E|} = (1-\delta)^{s-\ell} \cdot \underbrace{|\partial_L \mathcal{G}^s| \cdot q^{s-\ell}}_{\mu}.$$

Further,

$$\mu = |\partial_L \mathcal{G}^s| \cdot q^{s-\ell} = w_p(\partial_L \mathcal{G}^s) \cdot \left(\frac{q}{p}\right)^{s-\ell} \ge \frac{1}{r} \cdot \left(\frac{q}{p}\right)^{s-\ell} \ge 2.$$

On the other hand, by Harris's inequality,

$$\mathbb{E}[X^2] = \sum_{E,E'\in\partial_L\mathcal{G}^s} \mathbb{P}(E\cup E'\subseteq \tilde{C}) \leqslant \sum_{E,E'\in\partial_L\mathcal{G}^s} q^{|E\cup E'|} = \sum_{E\in\partial_L\mathcal{G}^s} q^{|E|} \sum_{E'\in\partial_L\mathcal{G}^s} q^{|E'\setminus E|}$$
$$\leqslant \mu \cdot \left(\mu + 1 + \sum_{t=1}^{s-\ell-1} \binom{s-\ell}{t} \Delta_t(\partial_L\mathcal{G}^s) q^{s-\ell-t}\right).$$

Further, for every  $t \in [\![s - \ell - 1]\!]$ ,

$$\Delta_t(\partial_L \mathcal{G}^s) \cdot q^{s-\ell-t} \leqslant \frac{p^{t+\ell-s} \cdot q^{s-\ell-t}}{r} \leqslant \left(\frac{p}{q}\right)^t \cdot \mu = \frac{\mu}{(2r)^t}$$

and thus

$$\mathbb{E}[X^2] \le \mu + \mu^2 \cdot \sum_{t=0}^{s-\ell-1} \binom{s-\ell}{t} \cdot \frac{1}{(8r)^t} \le \mu + \mu^2 \cdot \left(1 + \frac{1}{2r}\right)^{s-\ell} \le \mu + \frac{3}{2} \cdot \mu^2 \le 2\mu^2.$$

We conclude that

$$\frac{\mathbb{E}[X^2]}{\mathbb{E}[X]^2} \leqslant \frac{2}{(1-\delta)^{s-\ell}} \leqslant \frac{2}{1-\delta r} \leqslant \frac{8}{3}$$

which is a contradiction.

# 7. Concluding Remarks

Comparing the two notions of almost-independence that are used to describe the containers in Theorems A and B led us to proving Proposition 1.1. While part (ii) of this proposition is sufficient for our purposes, we believe that the following much stronger statement is true. We will say that a hypergraph  $\mathcal{H}$  is *r*-bounded, for some positive integer *r*, if  $|E| \leq r$  for all  $E \in \mathcal{H}$ .

**Conjecture 7.1.** There exist an absolute constant C > 0 such that the following holds for all  $p \in [0,1]$  and all  $r \ge 1$ . Every r-bounded hypergraph  $\mathcal{H}$  on a finite vertex set V admits a cover  $\mathcal{G} \subseteq 2^V \setminus \{\emptyset\}$  that satisfies

$$\mathbb{P}(\mathcal{H}[V_p] = \emptyset) \leqslant \exp\left(-w_{p/C}(\mathcal{G}) + C\log r\right).$$

Observe that Conjecture 7.1 implies the Kahn-Kalai Conjecture / Park-Pham Theorem [20, Theorem 1.1]. Indeed, let V be a finite set and let  $\mathcal{F} \subseteq 2^V$  be an arbitrary increasing property. Recall that the *expectation-threshold* of  $\mathcal{F}$  is the number  $q(\mathcal{F})$  defined as follows:

$$q(\mathcal{F}) \coloneqq \max\{q : \exists \mathcal{G} \subseteq 2^V \ \langle \mathcal{G} \rangle \supseteq \mathcal{F} \land w_q(\mathcal{G}) \leqslant 1/2\}.$$

Let  $\mathcal{H}$  be the set of all minimal elements of  $\mathcal{F}$  and let  $r \coloneqq \max\{|E| : E \in \mathcal{H}\} \cup \{2\}$ , so that  $\mathcal{H}$  is an *r*-bounded hypergraph on *V* and  $R \in \mathcal{F}$  if and only if  $\mathcal{H}[R] \neq \emptyset$ . Let  $q \coloneqq q(\mathcal{H})$ , suppose that

 $p \ge Kq \log r$ , and let  $\mathcal{G}$  be the cover of  $\mathcal{H}$  from the assertion of Conjecture 7.1. Observe that

$$w_{p/C}(\mathcal{G}) \ge p/(Cq) \cdot w_q(\mathcal{G}) \ge K/C \cdot \log r \cdot w_q(\mathcal{G}) \ge K/(2C) \cdot \log r.$$

Consequently, setting  $K \coloneqq 2(C^2 + C)$ , we may use Conjecture 7.1 to conclude that

$$\mathbb{P}(V_p \notin \mathcal{F}) = \mathbb{P}(\mathcal{H}[V_p] = \emptyset) \leqslant \exp\left((-K/(2C) + C) \cdot \log r\right) \leqslant 1/r \leqslant 1/2.$$

Furthermore, Conjecture 7.1 implies the celebrated results of Johansson–Kahn–Vu [11] that provide upper bounds on the threshold for appearance of H-factors in the binomial random graph. More precisely, the conjecture implies the following statement, which generalises both [11, Theorem 2.1] and [11, Theorem 2.2] (with an improved lower-bound assumption on p).

**Corollary 7.2** (assuming Conjecture 7.1). For every nonempty graph H, there exists a constant K such that, for all n divisible by  $v_H$  and all p satisfying

$$p \ge K \max_{\emptyset \ne J \subseteq H} \left( n^{1-v_J} \log n \right)^{1/e_J},\tag{7}$$

the random graph  $G_{n,p}$  contains an H-factor with probability at least 1-1/n.

Proof. Let H be a nonempty graph, let m be a large integer, let  $n \coloneqq m \cdot v_H$ , and let  $\mathcal{H}$  denote the family of all copies of the graph  $m \cdot H$  in  $K_n$ . Suppose that  $\mathcal{G} \subseteq 2^{K_n}$  is a cover for  $\mathcal{H}$ . Note that the number of *embeddings* of  $m \cdot H$  into  $K_n$  is n!. For a given  $G \in \mathcal{G}$ , which is a labelled subgraph of  $K_n$ , let  $r_G$  be the proportion of those n! embeddings that are covered by G, so that  $r_G \cdot n!$  is the number of embeddings of  $m \cdot H$  into  $K_n$  that contain G as a subgraph or, equivalently, the number of embeddings of G, viewed as a spanning subgraph of  $K_n$ , into  $m \cdot H$ . Given a  $G \in \mathcal{G}$ , let  $v_G$  be the number of nonisolated vertices of G, let  $\mathcal{C}_G$  denote the collection of connected components of G, and let  $c_G \coloneqq |\mathcal{C}_G|$ ; as usual,  $e_G$  is the number of edges of G. We claim that, for some constant D = D(H),

$$r_G \cdot n! \leq (n)_{c_G} \cdot D^{c_G} \cdot (n - v_G)!$$

Indeed, one can describe an embedding of G into  $m \cdot H$  by first specifying the image of one vertex from every connected component, then the images of the other O(1) vertices in each component (each has only O(1) possible images), and finally the images of the isolated vertices of G. Since  $v_G/v_H \leq c_G \leq v_G/2$ , we have

$$r_G \leqslant \frac{(n)_{c_G} \cdot D^{c_G}}{(n)_{v_G}} \leqslant (D'n)^{c_G - v_G} = \prod_{K \in \mathcal{C}(G)} D'' n^{1 - v_K},$$

where D, D', D'' are constants that depend only on H.

Let C be the constant from the statement of Conjecture 7.1. Since  $\mathcal{G}$  is a cover, the sum of all  $r_G$ , as G ranges over  $\mathcal{G}$ , is at least one. Consequently,

$$w_{p/C}(\mathcal{G}) = \sum_{G \in \mathcal{G}} r_G \cdot \frac{(p/C)^{e_G}}{r_G} \ge \min_{G \in \mathcal{G}} \frac{(p/C)^{e_G}}{r_G} \ge \min_{G \in \mathcal{G}} \prod_{K \in \mathcal{C}(G)} \frac{n^{v_K - 1} p^{e_K}}{C^{e_K} D''}.$$
(8)

It follows from (7) that each term in the product in the right-hand side of (8) is at least  $K/(CD'') \cdot \log n$  and thus  $w_p(\mathcal{G}) \ge K/(CD'') \cdot \log n$  as well (provided that K is sufficiently large). In particular, if K is sufficiently large, then Conjecture 7.1 implies that

$$\mathbb{P}(G_{n,p} \not\supseteq m \cdot H) = \mathbb{P}(\mathcal{H}[V_p] = \emptyset) \leqslant \exp\left(K/(CD'') \cdot \log n - C\log(ne_H/v_H)\right) \leqslant 1/n,$$

provided that K is sufficiently large.

Observe that both the Park–Pham theorem and Corollary 7.2 (with 1 - 1/n replaced by the nonexplicit term 1-o(1)) can also be derived from the following weaker version of Conjecture 7.1, which might be significantly easier to prove.

**Conjecture 7.3.** There exists and absolute constant C > 0 and a function  $f \colon \mathbb{N} \to (0, 1/2)$ satisfying  $\lim_{r\to\infty} f(r) = 0$  such that the following holds for all  $p \in [0, 1]$  and all  $r \ge 1$ . Suppose that  $\mathcal{H}$  is an r-bounded hypergraph on a finite vertex set V whose every cover  $\mathcal{G} \subseteq 2^V \setminus \{\emptyset\}$ satisfies  $w_{p/C}(\mathcal{H}) \ge \log r$ . Then,

$$\mathbb{P}(\mathcal{H}[V_p] = \emptyset) \leqslant f(r).$$

Finally, Conjecture 7.1 implies the following strengthening of Theorem A.

**Theorem F** (assuming Conjecture 7.1). There exists an absolute constant K such that the following holds. Let  $\mathcal{H}$  be an r-bounded hypergraph with a finite vertex set V. For all reals  $\delta$  and p satisfying  $K \log r/(\delta |V|) , there exists a family <math>S \subseteq 2^V$  and functions

$$g: \mathcal{I}(\mathcal{H}) \to \mathcal{S}$$
 and  $f: \mathcal{S} \to 2^{V}$ 

such that:

(a) For each  $I \in \mathcal{I}(\mathcal{H})$ , we have  $g(I) \subseteq I \subseteq f(g(I))$ .

(b) Each  $S \in S$  has at most  $Kp|V|/\delta$  elements.

(c) For every  $S \in S$ , letting  $C \coloneqq f(S)$ , there exists a hypergraph  $\mathcal{G} \subseteq 2^{C \setminus S} \setminus \{\emptyset\}$  with

$$w_p(\mathcal{G}) \leqslant K\delta p|V|$$

that covers  $\mathcal{H}[C]$ .

*Proof.* Denote by D the constant C from the statement of Conjecture 7.1 and let K = 3D. Let S, f, and g be the collection and the functions given by Theorem B with Dp in place of p. Fix some  $S \in S$ , let C := f(S), and let  $\mathcal{G} \subseteq 2^{V \setminus S} \setminus \{\emptyset\}$  be a cover of  $\mathcal{H}[C]$  with smallest Dp-weight. It follows from the assertion of Theorem B and Conjecture 7.1 that

$$(1 - Dp)^{\delta |C \setminus S|} \leq \mathbb{P} \big( S \cup C_{Dp} \in \mathcal{I}(\mathcal{H}) \big) \leq \exp \big( -w_p(\mathcal{G}) + D \log r \big).$$

Taking logarithms of both sides yields the inequality

$$w_p(\mathcal{G}) \leqslant D\log r - \delta |C \setminus S| \log(1 - Dp) \leqslant D\log r + 2D\delta p |C| \leqslant 3D\delta p |V| = K\delta p |V|,$$

as required.

#### References

- J. Balogh, R. Morris, and W. Samotij. Independent sets in hypergraphs. J. Amer. Math. Soc., 28(3):669–709, 2015.
- [2] J. Balogh, R. Morris, and W. Samotij. The method of hypergraph containers. In Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures, pages 3059–3092. World Sci. Publ., Hackensack, NJ, 2018.
- [3] J. Balogh and W. Samotij. An efficient container lemma. Discrete Anal., pages Paper No. 17, 56, 2020.
- [4] A. Bernshteyn, M. Delcourt, H. Towsner, and A. Tserunyan. A short nonalgorithmic proof of the containers theorem for hypergraphs. Proc. Amer. Math. Soc., 147(4):1739–1749, 2019.
- [5] B. Bollobás. On generalized graphs. Acta Math. Acad. Sci. Hungar., 16:447-452, 1965.

- [6] M. Campos. On the number of sets with a given doubling constant. Israel J. Math., 236(2):711–726, 2020.
- [7] M. Campos, M. Collares, R. Morris, N. Morrison, and V. Souza. The typical structure of sets with small sumset. Int. Math. Res. Not. IMRN, 2022(14):11011–11055, 2022.
- [8] M. Campos, M. Coulson, O. Serra, and M. Wötzel. The typical approximate structure of sets with bounded sumset. SIAM J. Discrete Math., 37(3):1386–1418, 2023.
- K. Frankston, J. Kahn, B. Narayanan, and J. Park. Thresholds versus fractional expectation-thresholds. Ann. of Math. (2), 194(2):475–495, 2021.
- [10] S. Janson. Poisson approximation for large deviations. Random Structures Algorithms, 1(2):221–229, 1990.
- [11] A. Johansson, J. Kahn, and V. Vu. Factors in random graphs. Random Structures Algorithms, 33(1):1–28, 2008.
- [12] J. Kahn and G. Kalai. Thresholds and expectation thresholds. Combin. Probab. Comput., 16(3):495–502, 2007.
- [13] D. J. Kleitman and K. J. Winston. On the number of graphs without 4-cycles. Discrete Math., 41(2):167–172, 1982.
- [14] G. Kozma and W. Samotij. Lower tails via relative entropy. Ann. Probab., 51(2):665–698, 2023.
- [15] D. Liu, L. Mattos, and T. Szabó. On the number of sets with small sumset. arXiv:2407.04492.
- [16] D. Lubell. A short proof of Sperner's lemma. J. Combinatorial Theory, 1:299, 1966.
- [17] L. D. Mešalkin. A generalization of Sperner's theorem on the number of subsets of a finite set. Teor. Verojatnost. i Primenen., 8:219–220, 1963.
- [18] R. Morris, W. Samotij, and D. Saxton. An asymmetric container lemma and the structure of graphs with no induced 4-cycle. J. Eur. Math. Soc. (JEMS), 26(5):1655–1711, 2024.
- [19] R. Nenadov. Probabilistic hypergraph containers. Israel J. Math., 261(2):879–897, 2024.
- [20] J. Park and H. T. Pham. A proof of the Kahn-Kalai conjecture. J. Amer. Math. Soc., 37(1):235–243, 2024.
- [21] W. Samotij. Counting independent sets in graphs. European J. Combin., 48:5–18, 2015.
- [22] A. A. Sapozhenko. On the number of independent sets in extenders. Diskret. Mat., 13(1):56–62, 2001.
- [23] D. Saxton and A. Thomason. Hypergraph containers. Invent. Math., 201(3):925–992, 2015.
- [24] D. Saxton and A. Thomason. Online containers for hypergraphs, with applications to linear equations. J. Combin. Theory Ser. B, 121:248–283, 2016.
- [25] D. Saxton and A. Thomason. Simple containers for simple hypergraphs. Combin. Probab. Comput., 25(3):448– 459, 2016.
- [26] M. Talagrand. Are many small sets explicitly small? In STOC'10—Proceedings of the 2010 ACM International Symposium on Theory of Computing, pages 13–35. ACM, New York, 2010.
- [27] K. Yamamoto. Logarithmic order of free distributive lattice. J. Math. Soc. Japan, 6:343–353, 1954.

#### APPENDIX A. THREE PROOFS OF PROPOSITION 2.2

One can see the assertion of the proposition is best-possible by picking an arbitrary subset  $U \subseteq C$  and considering the family  $\mathcal{I} = 2^U$ . Indeed, in this case  $\mathbb{E}[|C_p| \mid C_p \in \mathcal{I}] = \mathbb{E}[|U_p|] = p|U|$  and

$$\log \mathbb{P}(C_p \in \mathcal{I}) = \log \mathbb{P}(C_p \cap (C \setminus U) = \emptyset) = (|C| - |U|) \log(1 - p).$$

We provide three proofs of the proposition. The first proof uses the chain rule for relative entropy, the second proof employs a compression argument combined with the Kruskal–Katona theorem, and the third proof uses a generalisation of the edge-isoperimetric inequality for the hypercube due to Kahn and Kalai [12].

*First proof.* Without loss of generality, we may assume that  $C = \llbracket N \rrbracket$ . Let  $Y = (Y_1, \ldots, Y_N)$  be the indicator function of  $C_p$  conditioned on  $C_p \in \mathcal{I}$  and note that

$$\log \mathbb{P}(C_p \in \mathcal{I}) = -D_{KL}(Y \parallel C_p),$$

where  $D_{KL}(\cdot \| \cdot)$  denotes the Kullback-Leibler divergence between two probability distributions. For any  $J \subseteq \llbracket N \rrbracket$ , write  $I_p((Y_j)_{j \in J})$  in place of  $D_{KL}((Y_j)_{j \in J} \| J_p)$ . With this notational convention,  $D_{KL}(Y \| C_p) = I_p(Y_1, \ldots, Y_N)$ . The first key property of  $I_p$  that we will use is that it obeys the familiar chain rule (see, e.g., [14, Section 4]):

$$I_p(Y_1, \dots, Y_N) = \sum_{i=1}^N I_p(Y_i \mid Y_1, \dots, Y_{i-1}) = \sum_{i=1}^N \mathbb{E}[i_p(\mathbb{E}[Y_i \mid Y_1, \dots, Y_{i-1}])],$$

where we wrote  $I_p(Y_i | Y_1, \ldots, Y_{i-1})$  for the conditional relative entropy and

$$i_p(q) \coloneqq D_{KL}(\operatorname{Ber}(q) || \operatorname{Ber}(p)) = q \log \frac{q}{p} + (1-q) \log \frac{1-q}{1-p}.$$

Since the function  $i_p: [0,1] \to \mathbb{R}$  is convex, we have, for every random variable  $Z \in [0,p]$ ,

$$\mathbb{E}[i_p(Z)] \leqslant (1 - \mathbb{E}[Z]/p) \cdot i_p(0) + (\mathbb{E}[Z]/p) \cdot i_p(p) = (\mathbb{E}[Z]/p - 1)\log(1 - p).$$

Finally, since conditioned on  $Y_1, \ldots, Y_{i-1}$ , the vector  $(Y_i, \ldots, Y_N)$  is the indicator function of  $(\llbracket N \rrbracket \setminus \{1, \ldots, i-1\})_p$  conditioned on belonging to a decreasing family of sets, we have  $\mathbb{E}[Y_i | Y_1, \ldots, Y_{i-1}] \leq p$  almost surely. This implies that

$$I_p(Y_1, \dots, Y_N) \leqslant \sum_{i=1}^N (\mathbb{E}[Y_i]/p - 1) \log(1 - p) = (\mathbb{E}[|Y|] - N) \log(1 - p),$$

as claimed.

Second proof. We are going to prove the following equivalent inequality:

$$\mathbb{E}[|C_p| \mid C_p \in \mathcal{I}] \leqslant \left(|C| - \log_{1-p} \mathbb{P}(C_p \in \mathcal{I})\right) \cdot p.$$
(9)

Let  $\ell = \ell(p) := \log_{1-p} \mathbb{P}(C_p \in \mathcal{I})$ . We first argue that it is enough to establish (9) in the case where  $\ell$  is an integer. To see this, suppose that  $\ell \notin \mathbb{Z}$ . We claim that the (continuous) function  $q \mapsto \ell(q)$  cannot be constant on any open interval I containing p and thus it takes rational values for q arbitrarily close to p. Suppose that this were not true and  $\mathbb{P}(C_q \in \mathcal{I}) = (1-q)^{\ell}$  for all  $q \in I$ . We would then have, for all m,

$$\frac{d^m}{dq^m} \mathbb{P}(C_q \in \mathcal{I}) = \prod_{i=0}^{m-1} (i-\ell) \cdot (1-q)^{\ell-m} \neq 0,$$

contradicting the fact that  $\mathbb{P}(C_q \in \mathcal{I}) = \sum_{I \in \mathcal{I}} q^{|I|} (1-q)^{|C|-|I|}$  is a polynomial in q. Since both sides of (9) are continuous in p, it is enough to prove this inequality when  $\ell$  is rational.

Assume now that  $\ell \in \mathbb{Q}$  and let b be a positive integer such that  $b\ell \in \mathbb{Z}$ . Consider the family  $\mathcal{I}^b \subseteq 2^{C \times \llbracket b \rrbracket}$  defined by

$$\mathcal{I}^b \coloneqq \left\{ I_1 \times \{1\} \cup \dots \cup I_b \times \{b\} : I_1, \dots, I_b \in \mathcal{I} \right\}$$

and observe that  $\mathcal{I}^b$  is decreasing and that  $\mathbb{P}((C \times \llbracket b \rrbracket)_p \in \mathcal{I}^b) = \mathbb{P}(C_p \in \mathcal{I})^b = (1-p)^{b\ell}$ . Invoking (9) with  $\mathcal{I}$  replaced by  $\mathcal{I}^b$ , we conclude that

$$b \cdot \mathbb{E}[|C_p| \mid C_p \in \mathcal{I}] = \mathbb{E}\left[|(C \times \llbracket b \rrbracket)_p| \in \mathcal{I}^b \mid (C \times \llbracket b \rrbracket)_p \in \mathcal{I}^b|\right] \leq (b|C| - b\ell) \cdot p.$$

Assume  $\ell \in \mathbb{Z}$ , let  $\lambda \coloneqq p/(1-p)$ , and let  $Z \coloneqq \sum_{I \in \mathcal{I}} \lambda^{|I|}$ . Our first key observation is that

$$Z = (1-p)^{-|C|} \cdot \mathbb{P}(C_p \in \mathcal{I}) = (1-p)^{\ell - |C|} = (1+\lambda)^{|C|-\ell}$$

Write  $n := |C| - \ell$ , so that  $Z = (1 + \lambda)^n$ . For each  $i \in [n]$ , let  $x_i$  be the unique number in  $\{0\} \cup [i, \infty)$  such that  $\binom{x_i}{i} = |\binom{C}{i} \cap \mathcal{I}|$ . The assumption that  $\mathcal{I}$  is decreasing implies that  $2^I \subseteq \mathcal{I}$  for all  $I \in \mathcal{I}$  and, consequently,  $Z \ge (1 + \lambda)^{|I|}$  for each  $I \in \mathcal{I}$ . Therefore,  $\max_{I \in \mathcal{I}} |I| \le n$  and

$$f(x_1,\ldots,x_n) \coloneqq 1 + \sum_{i=1}^n \binom{x_i}{i} \cdot \lambda^i = \sum_{i=0}^{|C|} \left| \binom{C}{i} \cap \mathcal{I} \right| \cdot \lambda^i = Z.$$

Our second key observation is that

$$g(x_1,\ldots,x_n) \coloneqq \sum_{i=1}^n \binom{x_i}{i} \cdot i\lambda^i = Z \cdot \mathbb{E}[|C_p| \mid C_p \in \mathcal{I}].$$

Since  $\mathcal{I}$  is decreasing, the Kruskal–Katona theorem implies that  $x_1 \ge \cdots \ge x_n$ . In particular, letting

$$\mathcal{X} \coloneqq \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 \ge \dots \ge x_n \text{ and } x_i \in \{0\} \cup (i - 1, \infty) \text{ for all } i \in [[n]]\},\$$

we have

$$\mathbb{E}[|C_p| \mid C_p \in \mathcal{I}] \leqslant Z^{-1} \cdot \max\{g(x_1, \dots, x_n) : (x_1, \dots, x_n) \in \mathcal{X} \land f(x_1, \dots, x_n) = Z\}.$$

Let  $\mathbf{x} = (x_1, \dots, x_n)$  be a vector that achieves the above maximum. We claim that  $x_1 = \dots = x_n$ . Indeed, suppose that  $x_i > x_{i+1}$  for some *i* and let x' be the unique number satisfying  $\max\{x_{i+1}, i\} < x' < x_i$  and

$$d \coloneqq \begin{pmatrix} x_i \\ i \end{pmatrix} - \begin{pmatrix} x' \\ i \end{pmatrix} = \left[ \begin{pmatrix} x' \\ i+1 \end{pmatrix} - \begin{pmatrix} x_{i+1} \\ i+1 \end{pmatrix} \right] \cdot \lambda$$

Let  $\mathbf{x}'$  be the vector obtained from  $\mathbf{x}$  by replacing the *i*th and the (i + 1)th coordinates by x'and note that

$$f(\mathbf{x}') - f(\mathbf{x}) = \left[ \begin{pmatrix} x'\\i \end{pmatrix} - \begin{pmatrix} x_i\\i \end{pmatrix} \right] \cdot \lambda^i + \left[ \begin{pmatrix} x'\\i+1 \end{pmatrix} - \begin{pmatrix} x_{i+1}\\i+1 \end{pmatrix} \right] \cdot \lambda^{i+1} = -d\lambda^i + d\lambda^i = 0$$

and, similarly,

$$g(\mathbf{x}') - g(\mathbf{x}) = (i+1)d\lambda^i - id\lambda^i = d\lambda^i > 0,$$

contradicting the maximality of **x**. Let  $x \in (n-1, \infty)$  be the number such that  $x_1 = \cdots = x_n = x$ . Since

$$(1+\lambda)^n = Z = 1 + \sum_{i=1}^n \binom{x}{i} \cdot \lambda^i = \sum_{i=1}^n \binom{x}{i} \cdot \lambda^i,$$

we have x = n and thus

$$\mathbb{E}[|C_p| \mid C_p \in \mathcal{I}] \leqslant (1+\lambda)^{-n} \cdot \sum_{i=1}^n \binom{n}{i} \cdot i\lambda^i = n \cdot \frac{\lambda}{1+\lambda} = np,$$

as claimed.

Third proof. Let q := 1 - p and let  $f: 2^C \to \{0, 1\}$  be the function defined by f(A) = 1 if and only if  $A^c \in \mathcal{I}$ . Note that f is increasing and that  $\mathbb{E}[f(C_q)] = \mathbb{P}(C_p \in \mathcal{I})$ . Recall that the *total* influence of f with respect to  $C_q$  is the quantity

$$\operatorname{Inf}_{q}(f) \coloneqq \sum_{x \in C} \mathbb{P}\big(f(C_{q} \cup \{x\}) \neq f(C_{q} \setminus \{x\})\big) = \sum_{x \in C} \big(\mathbb{P}(C_{p} \setminus \{x\} \in \mathcal{I}) - \mathbb{P}(C_{p} \cup \{x\} \in \mathcal{I})\big).$$

Observe now that, for every  $x \in C$ ,

$$\mathbb{P}(C_p \setminus \{x\} \in \mathcal{I}) = \mathbb{P}(C_p \in \mathcal{I} \mid x \notin C_p) = \frac{\mathbb{P}(C_p \in \mathcal{I})}{\mathbb{P}(x \notin C_p)} \cdot \mathbb{P}(x \notin C_p \mid C_p \in \mathcal{I})$$

and, similarly,

$$\mathbb{P}(C_p \cup \{x\} \in \mathcal{I}) = \mathbb{P}(C_p \in \mathcal{I} \mid x \in C_p) = \frac{\mathbb{P}(C_p \in \mathcal{I})}{\mathbb{P}(x \in C_p)} \cdot \mathbb{P}(x \in C_p \mid C_p \in \mathcal{I}).$$

Consequently, letting  $p_x \coloneqq \mathbb{P}(x \in C_p \mid C_p \in \mathcal{I})$ , we have

$$\frac{\mathrm{Inf}_q(f)}{\mathbb{P}(C_p \in \mathcal{I})} = \sum_{x \in C} \left( \frac{1 - p_x}{1 - p} - \frac{p_x}{p} \right) = \frac{1}{1 - p} \cdot \sum_{x \in C} \left( 1 - \frac{p_x}{p} \right) = \frac{|C| - \mathbb{E}[|C_p| \mid C_p \in \mathcal{I}]/p}{1 - p}.$$

The desired inequality follows from the edge-isoperimetric inequality for f, which states that

$$q \cdot \operatorname{Inf}_q(f) \ge \mathbb{E}[f(C_q)] \cdot \log_q \mathbb{E}[f(C_q)],$$

see [12, Section 4].

TRINITY COLLEGE, UNIVERSITY OF CAMBRIDGE, CAMBRIDGE CB2 1TQ, UNITED KINGDOM *Email address:* mc2482@cam.ac.uk

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel *Email address:* samotij@tauex.tau.ac.il