The Importance of Being Biased

Irit Dinur* Samuel Safra†

September 10, 2002

Abstract

The Minimum Vertex Cover problem is the problem of, given a graph, finding a smallest set of vertices that touches all edges. We show that it is NP-hard to approximate this problem to within any factor smaller than $10\sqrt{5} - 21 \approx 1.36067$, improving on the previously known hardness result for a $\frac{7}{6}$ factor.

1 Introduction

Many classical combinatorial optimization problems (e.g. finding the maximum independent-set in a given graph, or finding a legal coloring of a given graph with the minimum number of colors) have no efficient algorithm unless $P = NP$. Consequently, a natural next-to-best alternative would be to consider solving their approximation version, the question becoming how close to optimal can a solution be efficiently obtained. This area of research –namely, the complexity of approximation problems– has seen significant progress since the discovery of the PCP theorem [FGL+91, AS92, ALM+92] in the early 1990's. Indeed, the complexity of many approximation problems is by now settled in that a tight bound on the ratio to within which they can be approximated efficiently has been obtained. That is, it has been shown that it is NP-hard to approximate these problems to within a factor even marginally better than the one achieved by the best known polynomial-time algorithm. Therefore, unless P=NP, there is no hope of coming up with an efficient improved approximation scheme.

Hardness results for approximation problems, almost without exception, build on the PCP characterization of NP. The fundamental insight of the PCP characterization of NP is that it is NP-hard to distinguish between SAT formulas that are completely satisfiable, and those that are *extremely* non-satisfiable, or, in particular, that *gap-SAT* is NP-hard.

A gap-SAT instance is a set of local-constraints over values assigned to some variables. These are referred to as *local*-constraints since whether each is satisfied depends only on a constant number of variables. An algorithm for gap-SAT has to distinguish, given such an instance, between the case there is an assignment satisfying all local-constraints, and the case in which no assignment satisfies even a small fraction of those constraints. For cases falling in the in-between gap, where some sizeable fraction, but not all, are satisfied, the algorithm may return an arbitrary result, which attributes to the use of the term *gap*.

* Institute for Advanced Study, Princeton NJ, USA.

† School of Mathematics and School of Computer Science, Tel Aviv University, ISRAEL. Research supported in part by the Fund for Basic Research administered by the Israel Academy of Sciences, and a Binational US-Israeli BSF grant

This problem was proved to be NP-hard [AS92, ALM$^+$92], showing that local constraints can imply global consistency for NP membership proofs. PCP characterizations of NP have served very well in resolving the complexity of approximation problems, leaving only a handful of classical optimization problems with the complexity of their approximation unsettled. One of those problems, and maybe the one that most captures the limitations of current technique for proving hardness of approximation, is Vertex-Cover.

In this paper we try to extend current technique, and show it is hard to approximate Vertex-Cover to within a factor larger than the known $\frac{7}{6}$ [Hås97]. More specifically, it follows by our analysis that:

Corollary 4.3 *Given a graph G, it is NP-hard to approximate the minimum Vertex-Cover to within any factor smaller than* $10\sqrt{5} - 21$. ■

Background. Let us now very briefly describe the background related to PCP, hardness of approximation, and reductions utilizing one to obtain the other.

Vertex-Cover is in fact APX-complete [PY91], namely, it belongs to a class of problems whose hardness of approximation is interrelated. All of the problems in this class (that includes the problem of finding the maximal cut in a given graph, finding the maximal independent-set in a bounded degree graph, etc.) can be efficiently approximated to within some constant factor. Indeed, the minimum vertex-cover can be rather easily approximated to within a factor of 2 as follows: Greedily obtain a maximal matching in the graph, and let the vertex-cover consist of both vertices at the ends of each edge in the matching. The resulting vertex-set covers all the edges and is no more than twice the size of the smallest vertex-cover.

All APX-complete problems have also been shown to be hard to approximate, to within *some* constant factor, via the PCP theorem. This however, is far from providing a tight bound for approximation problems, as the constant factor of approximation whose hardness is thus obtained, is usually quite far from the known upper-bound. For tight bounds one is required to work a little harder, and sometimes devise ingenious reductions and elaborate analysis.

One of the most successful recipes for such reductions, is the scheme of [BGS98, Hås99, Hås97], whose rough sketch is as follows.

The starting point is the parallel repetition lemma of [Raz98], applied to the gap-SAT of [AS92, ALM$^+$92]. This demonstrates that, given a gap-SAT instance Φ, consisting of local-constraints, each over two variables whose range is R; it is NP-hard to decide between (i) Φ is a 'yes' instance and has a satisfying assignment, or (ii) Φ is a 'no' instance and not even an arbitrarily small $\epsilon = \frac{1}{|R|^{O(1)}}$ fraction of the local-constraints of Φ can be satisfied.

The next step calls for applying a version of the Composition technique of [AS92], as proposed in [BGS98], to this specific setting. Namely, one replaces each of the variables with a set of variables representing its encoding, and each local-constraint $\varphi \in \Phi$ with constraints that both verify the consistency of the encoding of φ's variables, and that their encoded values satisfy φ.

This new set of constraints may take different form according to the problem one intends to show hard. The next step of the reduction, if necessary, translates those local-constraints to an instance of the problem at hand, whose solution, even if only an approximated one, implies a satisfying assignment for Φ.

The encoding utilized in that scheme, as proposed in [BGS98], is the *long-code*, the most extensive binary code, whose bits correspond to all possible Boolean functions over the code's domain. The most important merit of the long-code is its *local-testability*: given an arbitrary codeword, probing only two or three of its bits can detect, with high probability, whether or not it is close to a legal-codeword. The long-code's extremely inefficient (doubly exponential) size poses no problem since the range R of values variables of Φ can take is rather small.

Numerous tight bounds for approximation problems, such as Max-3-Sat, Linear Equations, Max-Clique, have been thus obtained. Some of these involve an extensive analysis of consistency tests over long-codes, using Fourier analysis [Hås99, Hås97], showing it suffices, for example, to probe the value of only three bits of the long-codes of x and y to be assured, with high probability, of the consistency within the encoding of x and y as well as the consistency between the two.

Vertex Cover. Nevertheless, where other open questions regarding the hardness of approximation problems have been settled, Vertex-Cover has stood still, leaving its best hardness result at a factor of $\frac{7}{6}$ [Hås97] which is still far from the best known upper bound of $2 - o(1)$ approximation factor [Hal00, BYE85, MS83].

Our analysis herein amends the [BGS98, Hås99] scheme in several places, most notably by introducing a generalized manner by which to probe the bits of a long-code – imposing a non-uniform probability distribution over the long-code bits – a construct which we refer to as the *biased-long-code*. Hence, each bit has a probability attached to it, and, given an assignment A to the bits of a biased–long-code, the weight of A, namely, the fractional size of the set of bits assigned 1 by A, is determined according to the probability distribution. The original long-code is a special case of this construct, in which the distribution over the bits is uniform.

Another important aspect by which our proof differs from the above scheme is in a preprocessing stage preceding the application of the Composition technique. This results in an instance of the Max-Independent-Set problem, much like the one resulting by the reduction of [FGL+91], namely, a graph made of several cliques that are inter-connected. The graph constructed herein, however, satisfies some further specific structural properties, whereby, certain pairs of cliques in this graph, impose a subgraph in which the maximal co-degree of the vertices is extremely small, in particular, one half of the vertices are of co-degree 1, while the other half are of co-degree 2.

Technique. The proof is rather complicated, relying on a two-stage construction with intricate dependency between the two stages.

An object whose analysis is central throughout the course of the proof is the *intersection-graph* over ground set R. This graph consists of a vertex for each subset $F \subseteq R$, and an edge between every pair of disjoint subsets. Our final graph consists of copies of this intersection-graph that are further connected.

An important step in the proof is to find, given a large independent-set in this intersection-graph, a very small set $C \subset R$, that in a sense 'controls' the behavior of the entire independent-set: one can determine whether a subset F belongs to the independent-set or not depending only on the value of $F \cap C$. Thus there is a family $\mathcal{F}_C$ of subsets of C such that with high probability, F is in the independent-set iff $F \cap C \in \mathcal{F}_C$. The probability in the previous statement refers to a certain *biased* distribution over all subsets, as a small set C might not even exist according to the uniform distribution. The set C is referred to as the *core* of the independent set, and $\mathcal{F}_C$ is the *core family*.

We observe that an arbitrary independent set in the intersection-graph corresponds to an intersecting family of subsets. It is then not hard to show that every pair of subsets in $\mathcal{F}_C$ intersect as well. Moreover, venturing into the field of extremal–set-theory, we employ Erdős-Ko-Rado intersection theorems [EKR61, W84, AK97] to show that if the independent-set is sufficiently large then there must be two subsets in $\mathcal{F}_C$ whose intersection consists of *exactly* *one* element. This element plays a key role in analyzing the consistency between pairs of intersection-graphs in our final construction.

The methods by which we find a core come from the field of study of influences of variables on Boolean functions, an area that has seen fruitful progress in recent years [KKL88, BL89, BLS88, BK97, FK96, BKS99]. By associating each subset with a vector in $\{0,1\}^R$, an independent set as above becomes a Boolean function $f : \{0,1\}^R \to \{0,1\}$; and if the independent set is maximal, then the Boolean function is actually monotone.

For $0 < p < 1$, define the p-biased distribution over $\{0,1\}^R$ by setting each bit to 1 independently with probability p. Let $\mu_p(f)$ denote the probability that a given Boolean function f is 1. It is not hard to see that for monotone f, $\mu_p(f)$ increases with p. Moreover, the well-known Russo's Lemma [Mar74, Rus82] (Theorem 5.4) states that the derivative $\frac{d\mu_p(f)}{dp}$, as a function of p, is precisely equal to the *average-sensitivity* of the function f, a quantity that measures how much f changes value upon single-bit flips of its input. Thus, if we allow a slight perturbation in the value of p, and since a bounded continuous function cannot have a large derivative everywhere, Russo's Lemma guarantees that f will have low average-sensitivity. At this point we can apply the insightful Friedgut Lemma [Fri98] (Theorem 5.2) which essentially asserts that if a Boolean function has low average-sensitivity then it is almost entirely determined by only a small number of its variables. These variables make up the core of f that we are seeking.

It is interesting to note that the above combination of Russo's Lemma with the Friedgut Lemma, essentially amounts to the following phenomenon: A monotone Boolean function *always* has a sharp threshold *unless* it depends on a few variables.

Outline. We begin in Section 2 with some preliminaries, introducing the biased long-code. In Section 3 we define our specific variant hIS of the independent-set problem and show it NP-hard, thereby encapsulating all one needs to know – for the purpose of our proof – of the PCP theorem. In section 4 we describe our reduction from an instance of hIS to Vertex-Cover, beginning with a graph G made of m cliques of size r and constructing from it a weighted graph $G_{\mathcal{B}}^{\mathbb{C}}$ whose independent set is either large, in particular roughly p, in case $\mathrm{IS}(G) = m$; or small, that is $< p^\bullet + \varepsilon$, in case $\mathrm{IS}_h(G) \le \epsilon m$. We also immediately show the completeness of the reduction.

Section 5 surveys the necessary combinatorial background for the proof of Soundness; and Section 6 contains the proof of soundness – the major technical proof – showing that if every ϵm vertices in G have an h-clique then $\mathrm{IS}(G_{\mathcal{B}}^{\mathbb{C}}) < p^\bullet + \varepsilon$.

2 Preliminaries

Codes – Long and Biased

A general binary code over a domain R maps an element of R to a sequence of binary bits. The long-code is the the longest possible (without repetition) such binary code, encoding each element of R by a sequence of $2^{|R|}$ binary bits. In this encoding each bit can be identified with an element of $\{0,1\}^R$ and then a codeword is simply the truth table of a Boolean function $\mathcal{F} : \{0,1\}^R \to \{0,1\}$.

The legal codewords of the long-code are specific Boolean functions, called 'dictatorships'. Indeed we define the encoding via the long-code of the element $i \in R$ to be the dictatorship of the ith element $\mathcal{F}_i : \{0,1\}^R \to \{0,1\}$, defined by $\mathcal{F}_i(x_1, \ldots, x_{|R|}) = x_i$. Formally,

Definition 2.1 *The* long-code of R*, is a binary code mapping elements of* R *into the space of all Boolean functions over* $\{0,1\}^R$*:*

$$\mathcal{LC} : R \to \left\{ \mathcal{F} : \{0,1\}^R \to \{0,1\} \right\}.$$

The legal encoding of an element $i \in R$ *is the dictatorship function* $\mathcal{F}_i : \{0,1\}^R \to \{0,1\}$ *defined by* $\mathcal{F}_i(x) = x_i$*.*

Alternative View of a Codeword. Before we continue our discussion, let us consider an alternative and very useful way to view the codewords of the long-code. Each element $x \in \{0,1\}^R$ is naturally identified with a subset of R, denoted usually by $F \subseteq R$. Denoting $\mathcal{P}(R) = \{ F \mid F \subseteq R \}$, a Boolean function becomes $\mathcal{F} : \mathcal{P}(R) \to \{0,1\}$, and is naturally identified with the family of subsets consisting of all F for which $\mathcal{F}(F) = 1$.

In this alternate view, a codeword is some family of subsets $\mathcal{F} \subseteq \mathcal{P}(R)$, and a legal codeword is a 'dictatorship' family $\mathcal{F}_i = \{ F \mid F \ni i \}$.

Long-Code for Hardness Proofs – Linearity-Test. The double exponential length of the long-code (encoding $\log |R|$ bits by $2^{|R|}$ bits) makes it quite inefficient for coding theory, but its advantage over other binary codes lies in its *local-testability* (this size blowup is allowed in our context due to the constant size of R). The amazingly efficient property of this code is that probing even a constant number of bits (chosen at random) of a given codeword can indicate, with good probability, if this codeword is "close to" a legal codeword. Thus, the Long-Code has been utilized in hardness of approximation proofs, to obtain global consistency by testing local-constraints, as can be demonstrated by the canonical example of linearity-testing [BLR93, Hås97]. Given a codeword $\mathcal{F} \subset \mathcal{P}(R)$, consider the following random process: Choose two random subsets $F_1, F_2 \in \mathcal{P}(R)$, and a third subset $H \in \mathcal{P}(R)$ by taking each $i \in R$ to be in H independently with probability ϵ. Now, accept only if an even number (0 or 2) of the three subsets $F_1, F_2, F_1 \Delta F_2 \Delta H$ are in $\mathcal{F}$. If $\mathcal{F}$ is the legal-codeword $\mathcal{F}_i$ of an element $i \in R$, this test accepts with probability $1 - \epsilon$. Moreover, Håstad proved [Hås97], applying Fourier analysis, that if this test accepts with probability $\frac{1}{2} + \epsilon$, then $\mathcal{F}$ is "close to" a small list of legal codewords, i.e. dictatorships. This ties in with the concept of list-decoding employed in coding-theory when, in the absence of one obvious encoded element, the decoding algorithm is required to generate a list of candidates each of which is "near" the given codeword. In Håstad's setting the elements encoded by these dictatorships can be taken as the list-decoding of the codeword $\mathcal{F}$.

The distribution, according to which the subsets F_1, F_2 are chosen, is uniform, implicitly implying that their size is, most probably, roughly $\frac{1}{2} \cdot |R|$. (The third subset, H, is chosen according to a distribution that highlights subsets of size ϵR.)

One may consider other distributions by which to choose these subsets, leading to our generalized version of the long-code, as follows.

The p-Biased Long-Code. Let us consider distributions that highlight subsets of size roughly $p \cdot |R|$. One such natural class of distributions, is the *p–product-distribution* over $\mathcal{P}(R)$, denoted μ_p^R, where, independently for each element $e \in R$, e is in a set with probability p and out of it with probability $1 - p$. More precisely,

Definition 2.2 (μ_p) *Let* $0 < p < 1$*.* μ_p^R *is a distribution over* $\mathcal{P}(R)$ *according to which, every* *subset $F \in \mathcal{P}(R)$ occurs with the following probability:*

$$\mu_p^R(F) \stackrel{def}{=} p^{|F|} \cdot (1-p)^{|R \setminus F|}$$

In some cases, when the set R is clear from the context, we may omit R and refer to $\mu_p^R(F)$ simply as $\mu_p(F)$.

For $p = \frac{1}{2}$, μ_p is simply the uniform distribution. For other values of p, this distribution highlights sets whose cardinality is roughly $p \cdot |R|$, and turns out to be useful especially for $p < \frac{1}{2}$. Let us now introduce the p-biased long-code,

Definition 2.3 (The p-Biased Long-Code $\mathbb{C}$) *The* p-biased long-code *over R, $\langle \mathcal{P}(R), \mu_p^R \rangle$, assigns the distribution μ_p to $\mathcal{P}(R)$.*

We will construct, in the following sections, a weighted graph whose vertices are partitioned into blocks, the vertices in each block corresponding to subsets $F \in \mathcal{P}(R)$ in the long-code of a domain R. An independent-set in this graph would correspond, in each block, to a family $\mathcal{F} \subseteq \mathcal{P}(R)$ satisfying some combinatorial properties, supposedly encoding an element $e \in R$. The weight of an encoding $\mathcal{F}_e = \{F \in \mathcal{P}(R) \mid F \ni e\}$ of an element $e \in R$ in the p-biased–long-code, is $\mu_p^R(\mathcal{F}_e) = p$. This will allow the independent-set in our constructed graph to have large $\geq p - \varepsilon$ weight in case of a 'yes' instance. The hard part will be to construct a graph for which we can show that a 'no' instance has an independent-set of weight no more than roughly $p^\bullet$.

PCP characterization of NP

Our proof relies on the PCP characterization of NP of [Raz98, RS97]. PCP characterizations of NP in general state that given some SAT instance, namely, a set of Boolean-functions $\Phi = \{\varphi_1, ..., \varphi_n\}$, it is NP-hard to distinguish between the case where there is an assignment A to Φ's variables that satisfies all Φ, and the case where any assignment A satisfies at most a small fraction of Φ.

The FGLSS [FGL+91, Kar72] reduction, applied to the PCP characterization of NP of either [Raz98] or [RS97], shows the Independent-Set problem on a co-partite graph to be NP-hard. We introduce a slightly stronger version of that gap-Independent-Set problem – in which in the 'no' case there is not even a sizeable set that does not contain a clique of size h – and show it NP-hard (Section 3). We then proceed to reduce this problem to Vertex-Cover.

Weighted Graphs

Our analysis is more naturally presented over weighted graphs, where the size of a set of vertices is the sum of their weights. Hardness results for these graphs easily translate to hardness for graphs with equal weight.

A *weighted-graph* $G = (V, E, \Lambda)$ is an undirected graph with vertices V and edges E, and a probability distribution Λ over the vertices V. In other words, G is a graph with normalized weights.

Independent-Set. An *independent-set* in G is a set $\mathcal{I} \subseteq V$ such that G restricted to $\mathcal{I}$ is the empty graph. Let us denote by $\mathrm{IS}(G)$ the *maximum*, over all independent-sets $\mathcal{I}$ in G, of $\Lambda(\mathcal{I})$.

Vertex Cover. A *vertex-cover* of G is a set $S \subseteq V$ whose complement $V \setminus S$ is an independent-set. Let us denote by $\overline{\mathrm{IS}}(G)$ the *minimum*, over all vertex-covers S, of $\Lambda(S)$.

It is clearly the case that $\mathrm{IS}(G) + \overline{\mathrm{IS}}(G) = 1$. Hence, computing the one determines the other. When approximation is concerned, however, since $\mathrm{IS}(G)$ may be much smaller than $\overline{\mathrm{IS}}(G)$, a good approximation of $\overline{\mathrm{IS}}(G)$ does not entail a good approximation for $\mathrm{IS}(G)$. Consequently, hardness results for approximating $\mathrm{IS}(G)$ [Hås99, EH00, Kho01, Tre01] do not necessarily imply hardness results for large factors for $\overline{\mathrm{IS}}(G)$. (In particular, it is hard to approximate $\mathrm{IS}(G)$ to within $n^{1-o(1)}$ yet easy to approximate $\overline{\mathrm{IS}}(G)$ to within factor of 2).

3 Co-partite Graphs and h-Clique-Independence

The purpose of this section is to define a gap variant of the Independent-Set problem and prove it is NP-hard. This encapsulates all one needs to know about PCP for our proof, as the rest of the paper starts off with this problem and reduces it to the appropriate gap-Independent-Set problem, one which implies hardness of Vertex-Cover.

First, let us consider the following type of graphs,

Definition 3.1 *An* (m, r)-co-partite graph $G = \langle M \times R, E\rangle$ *is a graph constructed of* $m = |M|$ *cliques each of size* $r = |R|$*, hence the edge set of* G *satisfies*

$$\forall i \in M, j_1, j_2 \in R, (\langle i, j_1\rangle, \langle i, j_2\rangle) \in E$$

This graph is the complement of an m-partite graph, whose parts have r vertices each.

The FGLSS [FGL$^+$91] reduction applied to a gap-SAT problem with constant 'depend' (namely, where each local-constraint depends on a constant number of variables) and arbitrarily small error-probability (see [Raz98, RS97]), results in an (m, r)-co-partite graph G, for which it is NP-hard to distinguish between the case where $\mathrm{IS}(G) = m$ and the case $\mathrm{IS}(G) < \epsilon m$. We define the following generalization:

Definition 3.2 *For any graph* $G = (V, E)$*, define*

$$\mathrm{IS}_h(G) \stackrel{def}{=} \max\{|I| \mid I \subseteq V \text{ contains no clique of size } h\}$$

The gap-h-Clique-Independent-Set Problem *(or* $h\mathrm{IS}(r, \epsilon, h)$ *for short) is as follows:*

Instance: *An* (m, r)*-co-partite graph* G*.*

Problem: *Distinguish between the following two cases:*

- $\mathrm{IS}(G) = m$.
- $\mathrm{IS}_h(G) \le \epsilon m$.

Note that $\mathrm{IS}_2(G) = \mathrm{IS}(G)$ and this becomes the usual gap-Independent-Set problem. Nevertheless, one can show that this problem is still hard, as long as r is large enough compared to h:

Theorem 3.1 *For any* $h, \epsilon > 0$*, the problem* $h\mathrm{IS}(r, \epsilon, h)$ *is NP-hard, as long as* $r \ge (\frac{h}{\epsilon})^c$ *for some constant* c*.*

Proof Sketch: Take a gap-SAT instance Φ, consisting of local-constraints, over variables whose range is R, such that each local-constraint depends on a constant number $D = O(1)$ of the variables. It is NP-hard [Raz98, RS97] to decide whether

Yes: Φ has a satisfying assignment.

No: Not even an arbitrarily small $\epsilon = \frac{1}{|R|^{O(1)}}$ fraction of the local-constraints of Φ can be satisfied.

Next, apply the FGLSS [FGL+91] reduction on this gap-SAT, to obtain an $(m = |\Phi|, r \leq |R|^D)$-co-partite graph G where each clique corresponds to a local-constraint, and its vertices to all possible satisfying assignments to that local-constraint. If there is a set $\mathcal{I}$ in G that contains no clique of size h, there must be an assignment to Φ's variables satisfying $|\mathcal{I}| / h^D$ of the constraints, as one can randomly choose for each variable one out of at most h plausible values.

For a complete proof of this theorem, see Appendix B.

4 Reducing hIS to Vertex-Cover

In this section we present our reduction from $h\mathrm{IS}(r, \varepsilon_0, h)$ to Vertex-Cover by constructing, from any given (m, r)-co-partite graph G, a graph $G_{\mathcal{B}}^{\mathbb{C}}$. Our main theorem is as follows,

Theorem 4.1 *For any $\varepsilon > 0$, and $p < p_{\max} = \frac{3-\sqrt{5}}{2}$, for large enough h, l_{T} and small enough ε_0 (see Definition 4.1 below): Given an (m, r)-co-partite graph $G = (M \times R, E)$, one can construct, in polynomial time, a graph $G_{\mathcal{B}}^{\mathbb{C}}$ so that:*

$$\begin{aligned} \mathrm{IS}(G) = m &\implies \mathrm{IS}(G_{\mathcal{B}}^{\mathbb{C}}) \geq p - \varepsilon \\ \mathrm{IS}_h(G) < \varepsilon_0 \cdot m &\implies \mathrm{IS}(G_{\mathcal{B}}^{\mathbb{C}}) < p^\bullet + \varepsilon = \max(p^2,\, 4p^3 - 3p^4) + \varepsilon \end{aligned}$$

As an immediate corollary we obtain,

Corollary 4.2 (Independent-Set) *Let $p < p_{\max} = \frac{3-\sqrt{5}}{2}$. For any constant $\varepsilon > 0$, given a weighted graph G, it is NP-hard to distinguish between:*

Yes: $\mathrm{IS}(G) > p - \varepsilon$

No: $\mathrm{IS}(G) < p^\bullet + \varepsilon$

In case $p \leq \frac{1}{3}$, $p^\bullet$ reads p^2 and the above asserts that it is NP-hard to distinguish between $\mathcal{I}(G_{\mathcal{B}}^{\mathbb{C}}) \approx p = \frac{1}{3}$ and $\mathcal{I}(G_{\mathcal{B}}^{\mathbb{C}}) \approx p^2 = \frac{1}{9}$ and the gap between the sizes of the minimum vertex cover in the 'yes' and 'no' cases approaches $\frac{1-p^2}{1-p} = 1 + p$, yielding a hardness-of-approximation factor of $\frac{4}{3}$ for Vertex-Cover. In general,

Corollary 4.3 (Vertex Cover) *Given a graph G, it is NP-hard to approximate the minimum Vertex-Cover to within any factor smaller than $10\sqrt{5} - 21 \approx 1.3606$.*

Proof: *(of corollary:)* For $\frac{1}{3} < p < p_{\max}$, $p^\bullet = 4p^3 - 3p^4$, thus it is NP-hard to distinguish between the case $G_{\mathcal{B}}^{\mathbb{C}}$ has a vertex cover of size $1 - p + \epsilon$ and the case $G_{\mathcal{B}}^{\mathbb{C}}$ has a vertex cover of size at least $1 - 4p^3 + 3p^4 - \epsilon$ for any $\epsilon > 0$. Minimum Vertex-Cover is thus shown hard to approximate to within a factor approaching

$$\frac{1 - 4(p_{\max})^3 + 3(p_{\max})^4}{1 - p_{\max}} = 1 + p_{\max} + (p_{\max})^2 - 3(p_{\max})^3 = 10\sqrt{5} - 21 \approx 1.36068...$$

$\blacksquare$

Before we turn to the proof of the main theorem, let us begin by setting the parameters. It is worthwhile to note here that the particular values chosen for these parameters are insignificant. They are merely chosen so as to satisfy some assertions through the course of the proof, nevertheless, most importantly, they are all independent of $r = |R|$. Once the proof has demonstrated that assuming an ε-size independent-set in $G^{\mathbb{L}}_{\mathcal{B}}$, there must be a set of size ε_0 in G that contains no h-clique, one can set r to be large enough so as to imply NP-hardness of $h\mathrm{IS}(r, \varepsilon_0, h)$, which thereby implies NP-hardness for the appropriate gap-Independent-Set problem. This argument is valid due to the fact that none of the parameters of the proof is related to r.

Definition 4.1 (Parameter Setting) *Given $\varepsilon > 0$ and $p < p_{\max}$, let us set the following parameters:*

- *Let $0 < \gamma < p_{\max} - p$ be such that, $(p+\gamma)^{\bullet} - p^{\bullet} < \frac{1}{4}\varepsilon$.*
- *Choosing h: We choose h to accommodate applications of the Friedgut Lemma, a Sunflower Lemma and a pigeon-hole principle. Let $\Gamma(p, \delta, k)$ be the function defined in the Friedgut-Lemma (Theorem 5.2), and let $\Gamma_*(k, d)$ be the function defined in the Sunflower Lemma (Theorem 6.8). Set*

$$h_0 = \sup_{q \in [p, p_{\max}]} \left(\Gamma(q, \tfrac{1}{16}\varepsilon, \tfrac{2}{\gamma}) \right)$$

  and let $\eta = \frac{1}{16 h_0} \cdot p^{5h_0}$, $h_1 = \lceil \frac{2}{\gamma \cdot \eta} \rceil + h_0$, $h_s = 1 + 2^{2h_0} \cdot \sum_{k=0}^{h_0} \binom{h_1}{k}$, and $h = \Gamma_*(h_1, h_s)$.
- *Fix $\varepsilon_0 = \frac{1}{32} \cdot \varepsilon$.*
- *Fix $l_{\mathsf{T}} = \max(4 \ln \frac{2}{\varepsilon},\ (h_1)^2)$.*

Remarks. The value of γ is well defined because the function taking p to $p^{\bullet} = \max(p^2, 4p^3 - 3p^4)$ is a continuous function of p. The supremum $\sup_{q \in [p, p_{\max}]} \left(\Gamma(q, \frac{1}{16}\varepsilon, \frac{2}{\gamma}) \right)$ in the definition of h_0 is bounded, because $\Gamma(q, \frac{1}{16}\varepsilon, \frac{2}{\gamma})$ is a continuous function of q, see Theorem 5.2. Both r and l_{T} remain fixed while the size of the instance $|G|$ increases to infinity, so without loss of generality we can assume that $l_{\mathsf{T}} \cdot r \ll m$.

Proof: *(of main theorem:)* Let us denote the set of vertices of G by $V = M \times R$.

The constructed graph $G^{\mathbb{L}}_{\mathcal{B}}$ will depend on a parameter $l \stackrel{def}{=} 2 l_{\mathsf{T}} \cdot r$.

Consider the family $\mathcal{B}$ of all sets of size l of V:

$$\mathcal{B} = \binom{V}{l} = \{ B \subset V \mid |B| = l \}$$

Let us refer to each such $B \in \mathcal{B}$ as a *block*. The intersection of an independent-set $\mathcal{I}_G \subset V$ in G with any $B \in \mathcal{B}$, $\mathcal{I}_G \cap B$, can take 2^l distinct forms, namely all subsets of B. If $|\mathcal{I}_G| = m$ then expectedly $|\mathcal{I}_G \cap B| = l \cdot \frac{m}{mr} = 2 l_{\mathsf{T}}$ hence for almost all B it is the case that $|\mathcal{I}_G \cap B| > l_{\mathsf{T}}$. Let us consider for each block B its *block-assignments*,

$$R_B \stackrel{def}{=} \left\{ \mathsf{a} \colon B \to \{\mathsf{T}, \mathsf{F}\} \;\middle|\; |\mathsf{a}^{-1}(\mathsf{T})| \geq l_{\mathsf{T}} \right\}.$$

Every *block-assignment* $\mathsf{a} \in R_B$ supposedly corresponds to some independent-set $\mathcal{I}_G$, and assigns T to exactly all vertices of B that are in $\mathcal{I}_G$, that is, where $\mathsf{a}^{-1}(\mathsf{T}) = \mathcal{I}_G \cap B$.

Note that $|R_B|$ is the same for all $B \in \mathcal{B}$, so for $r' = |R_B|$ and $m' = |\mathcal{B}|$, the graph $G_{\mathcal{B}}$ is (m', r')-co-partite. For a block-assignment for B, $\mathsf{a}: B \to \{\mathsf{T}, \mathsf{F}\}$, and any $\hat{B} \subseteq B$, let us denote by $\mathsf{a}|_{\hat{B}}: \hat{B} \to \{\mathsf{T}, \mathsf{F}\}$ the restriction of a to $\hat{B}$, namely, where $\forall v \in \hat{B}$, $\mathsf{a}|_{\hat{B}}(v) = \mathsf{a}(v)$. Given a pair of blocks B_1, B_2 that intersect on $\hat{B} = B_1 \cap B_2$ with $|\hat{B}| = l - 1$, every block-assignment to B_1 is consistent with (i.e. has a non-edge to) at most *two* block-assignments to B_2.

Let us construct the block graph of G, $G_{\mathcal{B}} = (V_{\mathcal{B}}, E_{\mathcal{B}})$

Definition 4.2 *Define the graph $G_{\mathcal{B}} = (V_{\mathcal{B}}, E_{\mathcal{B}})$, with vertices for all block-assignments to every block $B \in \mathcal{B}$,*

$$V_{\mathcal{B}} = \bigcup_{B \in \mathcal{B}} R_B$$

and edges for every pair of block-assignments that are clearly inconsistent,

$$E_{\mathcal{B}} = \bigcup_{\langle v_1, v_2 \rangle \in E,\ \hat{B} \in \binom{V}{l-1}} \left\{ \langle \mathsf{a}_1, \mathsf{a}_2 \rangle \in R_{\hat{B} \cup \{v_1\}} \times R_{\hat{B} \cup \{v_2\}} \,\middle|\, \mathsf{a}_1|_{\hat{B}} \neq \mathsf{a}_2|_{\hat{B}} \quad \textit{or} \quad \mathsf{a}_1(v_1) = \mathsf{a}_2(v_2) = \mathsf{T} \right\}$$

If $\langle \mathsf{a}_1, \mathsf{a}_2 \rangle \in E_{\mathcal{B}}$, we say that the block assignments $\mathsf{a}_1, \mathsf{a}_2 \in V_{\mathcal{B}}$ are *inconsistent.*

Furthermore the (almost perfect) completeness of the reduction from G to $G_{\mathcal{B}}$, can be easily proven:

Proposition 4.4 $\mathrm{IS}(G) = m \quad \Longrightarrow \quad \mathrm{IS}(G_{\mathcal{B}}) \geq m' \cdot (1 - \varepsilon)$.

Proof: Let $\mathcal{I}_G \subset V$ be an independent-set in G, $|\mathcal{I}| = m = \frac{1}{r}|V|$. Let $\mathcal{B}'$ consist of all l-sets $B \in \mathcal{B} = \binom{V}{l}$ that intersect $\mathcal{I}_G$ on at least l_{T} elements $|B \cap \mathcal{I}_G| \geq l_{\mathsf{T}}$. The probability that this does not happen is (see Proposition E.1) $\Pr_{B \in \mathcal{B}}[B \notin \mathcal{B}'] \leq 2e^{-\frac{2l_{\mathsf{T}}}{8}} \leq \varepsilon$. For a block $B \in \mathcal{B}'$, let $\mathsf{a}_B \in R_B$ be the characteristic function of $\mathcal{I}_G \cap B$:

$$\forall v \in B \qquad \mathsf{a}_B(v) \stackrel{def}{=} \begin{cases} \mathsf{T} & v \in \mathcal{I}_G \\ \mathsf{F} & v \notin \mathcal{I}_G \end{cases}$$

The set $\mathcal{I} = \{\mathsf{a}_B \mid B \in \mathcal{B}'\}$ is an independent-set in $G_{\mathcal{B}}$, of size $m' \cdot (1 - \varepsilon)$. ■

The Final Graph

We now define our final graph $G_{\mathcal{B}}^{\mathbb{L}}$, consisting of the same blocks as $G_{\mathcal{B}}$, where the vertices in each block of $G_{\mathcal{B}}^{\mathbb{L}}$ represent the long-code of the vertices of that block in $G_{\mathcal{B}}$.

Vertices and Weights: $G_{\mathcal{B}}^{\mathbb{L}} = \left\langle V_{\mathcal{B}}^{\mathbb{L}}, E_{\mathcal{B}}^{\mathbb{L}}, \Lambda \right\rangle$ has a block of vertices $V_{\mathcal{B}}^{\mathbb{L}}[B]$ for every $B \in \mathcal{B}$, where vertices in each block B correspond to the p-biased–long-code applied to R_B

$$V_{\mathcal{B}}^{\mathbb{L}}[B] = \mathcal{P}(R_B)$$

that is, one vertex for each subset $F \subseteq R_B$ of B's block-assignments. $V_{\mathcal{B}}^{\mathbb{L}}$ consists of one such block of vertices for each $B \in \mathcal{B}$

$$V_{\mathcal{B}}^{\mathbb{L}} = \bigcup_{B \in \mathcal{B}} V_{\mathcal{B}}^{\mathbb{L}}[B]$$

Note that we take the block-assignments to be distinct, hence, subsets of them are distinct, and $V_{\mathcal{B}}^{\mathbb{C}}$ is a disjoint union of $V_{\mathcal{B}}^{\mathbb{C}}[B]$ over all $B \in \mathcal{B}$.

Let Λ_B, for each block $B \in \mathcal{B}$, be the distribution assigning each vertex F, a probability according to μ_p, namely

$$\Lambda_B(F) = \mu_p^{R_B}(F)$$

The block of vertices $V_{\mathcal{B}}^{\mathbb{C}}[B]$ superimposed with Λ_B therefore comprise a p-biased–long-code over R_B (see Section 2).

The probability distribution Λ assigns equal probability to every block: For any $F \in V_{\mathcal{B}}^{\mathbb{C}}[B]$

$$\Lambda(F) \stackrel{def}{=} |\mathcal{B}|^{-1} \cdot \Lambda_B(F)$$

Edges. We have edges between every pair of $F_1 \in V_{\mathcal{B}}^{\mathbb{C}}[B_1]$ and $F_2 \in V_{\mathcal{B}}^{\mathbb{C}}[B_2]$ if in the graph $G_{\mathcal{B}}$ there is a complete bipartite graph between these sets, i.e.

$$E_{\mathcal{B}}^{\mathbb{C}} = \left\{ \langle F_1, F_2 \rangle \in V_{\mathcal{B}}^{\mathbb{C}}[B_1] \times V_{\mathcal{B}}^{\mathbb{C}}[B_2] \;\middle|\; E_{\mathcal{B}} \supseteq F_1 \times F_2 \right\}$$

In particular, there are edges within a block, i.e. when $B_1 = B_2$, iff $F_1 \cap F_2 = \phi$ (formally, this follows from the definition because the vertices of R_B form a clique in $G_{\mathcal{B}}$, and $G_{\mathcal{B}}$ has no self edges).

This completes the construction of the graph $G_{\mathcal{B}}^{\mathbb{C}}$. We have,

Proposition 4.5 *For any fixed $p, l > 0$, the graph $G_{\mathcal{B}}^{\mathbb{C}}$ is polynomial-time constructible given input G.* ■

A simple-to-prove, nevertheless crucial, property of $G_{\mathcal{B}}^{\mathbb{C}}$ is that every independent-set can be monotonely extended,

Proposition 4.6 *Let $\mathcal{I}$ be an independent-set of $G_{\mathcal{B}}^{\mathbb{C}}$: If $F \in \mathcal{I} \cap V_{\mathcal{B}}^{\mathbb{C}}[B]$, and $F \subset F' \in V_{\mathcal{B}}^{\mathbb{C}}[B]$, then $\mathcal{I} \cup \{F'\}$ is also an independent-set.* ■

We conclude this section by proving completeness of the reduction:

Lemma 4.7 (Completeness) $\mathrm{IS}(G) = m \quad \Longrightarrow \quad \mathrm{IS}(G_{\mathcal{B}}^{\mathbb{C}}) \geq p - \varepsilon$.

Proof: By proposition 4.4, if $\mathrm{IS}(G) = m$ then $\mathrm{IS}(G_{\mathcal{B}}) \geq m'(1-\varepsilon)$. In other words, there is an independent-set $\mathcal{I}_{\mathcal{B}} \subset V_{\mathcal{B}}$ of $G_{\mathcal{B}}$ whose size is $|\mathcal{I}_{\mathcal{B}}| \geq m' \cdot (1-\varepsilon)$. Let $\mathcal{I}_0 = \{\{\mathsf{a}\} \mid \mathsf{a} \in \mathcal{I}_{\mathcal{B}}\}$ be the independent-set consisting of all singletons of $\mathcal{I}_{\mathcal{B}}$, and let $\mathcal{I}$ be $\mathcal{I}_0$'s monotone closure. The set $\mathcal{I}$ is also an independent-set due to Proposition 4.6 above. It remains to observe that the weight within each block of the family of all sets containing a fixed $\mathsf{a} \in \mathcal{I}_{\mathcal{B}}$, is p. ■

5 Families of Subsets

Our eventual goal, which will be completed in the next section (Section 6), is to show that an independent-set in $G_{\mathcal{B}}^{\mathbb{C}}$, if large enough, corresponds to a significant subset of G's vertices that is h-clique-free, so G must be a 'yes' instance. Specifically, we will show that an independent-set corresponds, in each block $B \in \mathcal{B}$, to a family $\mathcal{F} \subseteq \mathcal{P}(R_B)$ that can be 'list-decoded' into a small set of permissible values in R_B. We would then distinguish, assuming the independent-set is large enough, one block-assignment per block, for a significant portion of the blocks. We finally translate these block-assignments into a large subset of V that contains no h-clique.

This analysis utilizes two different combinatorial aspects of families of subsets. First, we employ some theorems from the field of influences of variables on Boolean functions, to deduce that a family $\mathcal{F} \subseteq \mathcal{P}(R)$ as above has a *core*, namely, a small set $C \subset R$ of elements of R that are, in a sense, permissible decodings of it.

Next, we turn to extremal set theory to show that if $\mathcal{F}$ is also of large weight according to μ_p, and if it is *intersecting*, it must distinguish, in a specific sense to be defined, *one* element in its core. This element will be important for asserting consistency between blocks, as it will consequently be shown to be consistent with the distinguished elements of other blocks.

Recall that we denote a family of subsets of R by $\mathcal{F} \subset \mathcal{P}(R)$, and one of its member-subsets usually by $F, H \in \mathcal{F}$.

5.1 A Family's Core

Let $\mathcal{F} \subset \mathcal{P}(R)$ be a family of subsets of R. We would be interested in finding when this family is, in a sense, close to an encoding of an element $e \in R$. For starters, we would be satisfied in finding a small set of permissible elements in R, henceforth referred to as a core, such that $\mathcal{F}$ is correlated to the codewords of these values.

A family of subsets $\mathcal{F} \subset \mathcal{P}(R)$ is said to be *determined* by $C \subset R$, if a subset $F \in \mathcal{P}(R)$ is determined to be in or out of $\mathcal{F}$ only according to its intersection with C (no matter whether other elements are in or out of F). Formally, $\mathcal{F}$ is determined by C if,

$$\{F \in \mathcal{P}(R) \mid F \cap C \in \mathcal{F}\} = \mathcal{F}$$

Denote by $\mathcal{F}_1 \sqcup \mathcal{F}_2$ the family consisting of the pairwise union of all subsets of $\mathcal{F}_1$ with all those of $\mathcal{F}_2$,

$$\mathcal{F}_1 \sqcup \mathcal{F}_2 \stackrel{def}{=} \{F_1 \cup F_2 \mid F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2\} .$$

If $C \subset R$ determines $\mathcal{F}$, then there is a family $\mathcal{F}_C \subseteq \mathcal{P}(C)$, such that $\mathcal{F} = \mathcal{F}_C \sqcup \mathcal{P}(R \setminus C)$.

A given family $\mathcal{F}$, is not necessarily determined by any small set C. However, there might be another family $\mathcal{F}'$, that is determined by some small set C, and that approximates $\mathcal{F}$ quite accurately, up to some δ:

Definition 5.1 (Core) *A set $C \subseteq R$ is said to be a (δ, p)-core of the family $\mathcal{F} \subseteq \mathcal{P}(R)$, if there exists a C-determined family $\mathcal{F}' \subseteq \mathcal{P}(R)$ such that $\mu_p(\mathcal{F} \triangle \mathcal{F}') < \delta$.*

As to the family of subsets that best approximates $\mathcal{F}$ on its core, it consists of the subsets $F \in \mathcal{P}(C)$ whose extension to R intersects more than half of $\mathcal{F}$,

$$[\mathcal{F}]_C^{\frac{1}{2}} \stackrel{def}{=} \left\{ F \in \mathcal{P}(C) \;\middle|\; \Pr_{F' \in \mu_p^{R \setminus C}} \left[F \cup F' \in \mathcal{F}\right] > \frac{1}{2} \right\} .$$

Consider the *Core-Family*, defined as the family of all subsets $F \in \mathcal{P}(C)$, for which $\frac{3}{4}$ of their extension to R, i.e. $\frac{3}{4}$ of $\{F' \mid F' \cap C = F\}$, reside in $\mathcal{F}$:

Definition 5.2 (Core-Family) *For a set of elements $C \subset R$, define,*

$$[\mathcal{F}]_C^{\frac{3}{4}} \stackrel{def}{=} \left\{ F \in \mathcal{P}(C) \;\middle|\; \Pr_{F' \in \mu_p^{R \setminus C}} \left[F \cup F' \in \mathcal{F}\right] > \frac{3}{4} \right\}$$

By simple averaging, it turns out that if C is a (δ, p)-core for $\mathcal{F}$, this family approximates $\mathcal{F}$ almost as well as the best family C.

Lemma 5.1 *If C is a (δ, p)-core of $\mathcal{F}$, then $\mu_p^C\left([\mathcal{F}]_C^{\frac{3}{4}}\right) \geq \mu_p^R(\mathcal{F}) - 4\delta$.*

Proof: Let $\mathcal{F}' = [\mathcal{F}]_C^{\frac{1}{2}} \sqcup \mathcal{P}(R \setminus C)$. Since C is a (δ, p)-core of $\mathcal{F}$, $\mu_p^R(\mathcal{F}' \triangle \mathcal{F}) < \delta$. Hence,

$$\mu_p^C\left(\left\{F \in \mathcal{P}(C) \;\middle|\; \Pr_{F' \in \mu_p^{R \setminus C}}\left[F \cup F' \in \mathcal{F}\right] \leq \frac{3}{4}\right\}\right) < 4\delta$$

■

Influence and Sensitivity

Understanding the conditions for a family of subsets to have a small core, has been pursued for some years. This has to do with the probability of every element $e \in R$ to take subsets in or out of $\mathcal{F}$ when flipped, which is referred to as the *influence* of that element. This notion, and its relations with various properties of $\mathcal{F}$, have been the subject of an extensive analysis [BL89, KKL88, Fri98]. Let us now introduce this notion and assert some theorems to be available for good use later.

Assume a family of subsets $\mathcal{F} \subseteq \mathcal{P}(R)$. The *influence* of an element $e \in R$,

$$\mathsf{influence}_p^e(\mathcal{F}) \stackrel{def}{=} \Pr_{F \in \mu_p}[\text{ exactly one of } F \cup \{e\},\ F \setminus \{e\} \text{ is in } \mathcal{F}]$$

The *average sensitivity*[1] of $\mathcal{F}$ with respect to μ_p, denoted $\mathsf{as}_p(\mathcal{F})$, is the sum of the influences of all elements in R,

$$\mathsf{as}_p(\mathcal{F}) \stackrel{def}{=} \sum_{e \in R} \mathsf{influence}_p^e(\mathcal{F})$$

A truly fundamental relation between the average sensitivity of a family $\mathcal{F} \subseteq \mathcal{P}(R)$ and the size of its (δ, p)-core is the following theorem of Friedgut [Fri98]:

Theorem 5.2 (Friedgut) *Let $0 < p < 1$ be some bias, and $\delta > 0$ be any approximation parameter. Consider any family $\mathcal{F} \subset \mathcal{P}(R)$, and let $k = \mathsf{as}_p(\mathcal{F})$. There exists a function $\Gamma(p, \delta, k) \leq (c_p)^{k/\delta}$, where c_p is a constant[2] depending only on p, such that $\mathcal{F}$ has a (δ, p)-core C, with $|C| \leq \Gamma(p, \delta, k)$.* ■

Hence, the number of elements that are necessary in order to approximate $\mathcal{F}$ up to δ depends only on δ and the average sensitivity of $\mathcal{F}$. In particular, if a family $\mathcal{F}$ has low (say, constant) average sensitivity, then it has a (δ, p)-core whose size is merely exponential in $\frac{1}{\delta}$, and is independent of $|R|$.

The next step would be to find sufficient conditions for a family to have low average sensitivity. As it turns out, this is the case with *monotone families* (defined next) assuming we allow some slight shift in p.

[1] The name average-sensitivity is derived from defining the *sensitivity* of a subset $F \in \mathcal{F}$ as the number of elements whose removal from or addition to F takes F in or out of $\mathcal{F}$: $\text{sensitivity}(F) = |\{e \in R \mid \text{exactly one of } F \cup \{e\},\ F \setminus \{e\} \text{ is in } \mathcal{F}\}|$. The above definition is then equivalent to the average, over F, of sensitivity(F).

[2] It follows directly from Friedgut's proof that c_p is bounded by a continuous function of p.

Definition 5.3 (Monotone Family) *A family of subsets $\mathcal{F} \subseteq \mathcal{P}(R)$ is* monotone *if for every $F \in \mathcal{F}$, for all $F' \supset F$, $F' \in \mathcal{F}$.*

Such a family is sometimes called in the literature an 'upset'.

Being monotone restricts a family in certain ways, forcing it, for example, to have relatively more large subsets than it does small subsets. This can be formalized as follows,

Proposition 5.3 *For a monotone family $\mathcal{F} \subseteq \mathcal{P}(R)$, $\mu_p(\mathcal{F})$ is a monotone non-decreasing function of p.*

For a simple proof of this proposition, see Appendix C.

Interestingly, for monotone families, the rate at which μ_p increases with p, is exactly equal to the average-sensitivity:

Theorem 5.4 (Russo-Margulis Identity [Mar74, Rus82]) *Let $\mathcal{F} \subseteq \mathcal{P}(R)$ be a monotone family. Then,*

$$\frac{d\mu_p(\mathcal{F})}{dp} = \mathsf{as}_p(\mathcal{F})$$

■

For a simple proof of this identity, see Appendix C.

This means, that the average sensitivity $\mathsf{as}_p(\mathcal{F})$ of a monotone family $\mathcal{F}$, cannot remain very high for very long. In other words,

Corollary 5.5 *Let $\mathcal{F} \subseteq \mathcal{P}(R)$ be a monotone family, and let $0 \leq p < p + \gamma \leq 1$. There must be some $q \in (p, p + \gamma)$ such that*

$$\mathsf{as}_q(\mathcal{F}) \leq \frac{1}{\gamma}$$

Proof: With the above identity, and a standard application of Lagrange's Mean-Value Theorem, there exists some $q \in (p, p + \gamma)$,

$$\mathsf{as}_q(\mathcal{F}) = \frac{d\mu_q(\mathcal{F})}{dq} = \frac{\mu_{p+\gamma}(\mathcal{F}) - \mu_p(\mathcal{F})}{\gamma} \leq \frac{1}{\gamma}$$

as $\mu_{p+\gamma}(\mathcal{F}) \leq 1$. ■

We have now reached the main point of this discussion. A monotone family $\mathcal{F}$, supposedly representing an encoding with the p-biased long code of an element in R, always has low average sensitivity for some value of $q \in [p, p+\epsilon]$. For this q we can apply the Friedgut Lemma to deduce a small core $C \subset R$, $|C| = O(1)$, for $\mathcal{F}$, on which it is well-approximated according to μ_q. The elements in this core would serve as a set of permissible values, that are the 'list-decoding' of $\mathcal{F}$, in the rest of the proof. That these decoded values indeed represent $\mathcal{F}$, and that consistency of families $\mathcal{F}_1$ and $\mathcal{F}_2$ constitute some form of consistency of their cores C_1 and C_2, is the task we face in the next section.

5.2 Maximal Intersecting Families

We have seen that a monotone family distinguishes a small core of elements, that almost determine it completely. Next, we will show that a monotone family that is of large enough weight, and is also *intersecting*, must exhibit one *distinguished* element in its core. This element is a stricter 'decoding' of the family than is the core, and will consequently serve for establishing consistency between distinct families.

Definition 5.4 *A family $\mathcal{F} \subset \mathcal{P}(R)$ is* t-intersecting, *for $t \geq 1$, if*

$$\forall F_1, F_2 \in \mathcal{F}, \quad |F_1 \cap F_2| \geq t$$

for $t = 1$ such a family is referred to simply as intersecting.

Let us first consider the following natural generalization for a pair of families,

Definition 5.5 (Cross-Intersecting) *Two families $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{P}(R)$ are* cross-intersecting *if for every $F_1 \in \mathcal{F}_1$ and $F_2 \in \mathcal{F}_2$, $F_1 \cap F_2 \neq \phi$.*

Two families cannot be too large and still remain cross-intersecting,

Proposition 5.6 *For any bias parameter $p \leq \frac{1}{2}$, two families of subsets $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{P}(R)$, for which $\mu_p(\mathcal{F}_1) + \mu_p(\mathcal{F}_2) > 1$ are not cross-intersecting.*

Proof: We can assume that $\mathcal{F}_1, \mathcal{F}_2$ are monotone, as their monotone closures must also be cross-intersecting. Since μ_p, for a monotone family, is non-decreasing with respect to p (see Proposition 5.3), it is enough to prove the claim for $p = \frac{1}{2}$. If for all $F \in \mathcal{P}(R)$ contained in both families – that is, so that $F \in \mathcal{F}_1$ and $F \in \mathcal{F}_2$ – it were the case that its complement $F^{\mathsf{c}} = R \setminus F$ would be contained in none of the families – namely, $F^{\mathsf{c}} \notin \mathcal{F}_1, F^{\mathsf{c}} \notin \mathcal{F}_2$ – the sum of sizes would be at most 1. There must therefore be one such pair, F and F^{c}, contained one in $\mathcal{F}_1$ and the other in $\mathcal{F}_2$. ■

It is now easy to prove that if $\mathcal{F}$ is monotone and intersecting, then the same holds for the core-family $[\mathcal{F}]_C^{\frac{3}{4}}$ that is (see Definition 5.2) the threshold approximation of $\mathcal{F}$ on its core C,

Proposition 5.7 *Let $\mathcal{F} \subseteq \mathcal{P}(R)$, and let $C \subseteq R$.*

- *If $\mathcal{F}$ is monotone then $[\mathcal{F}]_C^{\frac{3}{4}}$ is monotone.*
- *If $\mathcal{F}$ is intersecting, and $p \leq \frac{1}{2}$, then $[\mathcal{F}]_C^{\frac{3}{4}}$ is intersecting.*

Proof: The first assertion is immediate. For the second assertion, assume by way of contradiction, a pair of non-intersecting subsets $F_1, F_2 \in [\mathcal{F}]_C^{\frac{3}{4}}$ and observe that the families

$$\{F \in \mathcal{P}(R \setminus C) \mid F \cup F_1 \in \mathcal{F}_1\} \quad \text{and} \quad \{F \in \mathcal{P}(R \setminus C) \mid F \cup F_2 \in \mathcal{F}_2\}$$

both have weight $> \frac{3}{4}$, and by Proposition 5.6, cannot be cross intersecting. ■

An intersecting family whose weight is larger than that of a maximal 2- intersecting family, must contain two subsets that intersect on a unique element $e \in R$. This family behaves in some respects as if it is $\mathcal{F}_e$, a fact that will be critical for our proof.

Definition 5.6 (Distinguished Element) *For a monotone and intersecting family $\mathcal{F} \subseteq \mathcal{P}(R)$, an element $e \in R$ is said to be* distinguished *if there exist $F^\flat, F^\sharp \in \mathcal{F}$ such that*

$$F^\flat \cap F^\sharp = \{e\}$$

Clearly, an intersecting family has a distinguished element if and only if it is not 2-intersecting. We next establish a weight criterion for an intersecting family to have a distinguished element. For each $p < p_{\max}$, define $p^\bullet$ to be

Definition 5.7

$$\forall p < p_{\max}, \qquad p^{\bullet} \stackrel{def}{=} \max(p^2,\, 4p^3 - 3p^4)$$

This maps each p to the size of the maximal 2-intersecting family, according to μ_p. For a proof of such a bound we venture into the field of extremal set theory, where maximal intersecting families have been studied for some time. This study was initiated by Erdős, Ko, and Rado [EKR61], and has seen various extensions and generalizations. The corollary above is a generalization to μ_p of what is known as the Complete Intersection Theorem for finite sets, that was proven by [AK97]. Frankl [Fra78] defined the following families:

$$\mathcal{A}_{i,t} \stackrel{def}{=} \{F \in \mathcal{P}([n]) \mid F \cap [1, t+2i] \geq t+i\}$$

which are easily seen to be t-intersecting for $0 \leq i \leq \frac{n-t}{2}$; and conjectured the following theorem that was finally proven by Ahlswede and Khachatrian [AK97]:

Theorem 5.8 ([AK97]) *Let $\mathcal{F} \subseteq \binom{[n]}{k}$ be t-intersecting. Then,*

$$|\mathcal{F}| \leq \max_{0 \leq i \leq \frac{n-t}{2}} \left| \mathcal{A}_{i,t} \cap \binom{[n]}{k} \right|$$

■

Our analysis requires the extension of this statement to families of subsets that are not restricted to a specific size k, and where $t = 2$. Let us therefore denote $\mathcal{A}_i \stackrel{def}{=} \mathcal{A}_{i,2}$:

Corollary 5.9 *Let $\mathcal{F} \subset \mathcal{P}([n])$ be 2-intersecting. For any $p < \frac{1}{2}$,*

$$\mu_p(\mathcal{F}) \leq \max_i \{\mu_p(\mathcal{A}_i)\}$$

Furthermore, when $p \leq \frac{1}{3}$, this maximum is attained by $\mu_p(\mathcal{A}_0) = p^2$, and for $\frac{1}{3} < p < p_{\max}$ by $\mu_p(\mathcal{A}_1) = 4p^3 - 3p^4$. Having defined $p^{\bullet} = \max(p^2, 4p^3 - 3p^4)$ for every $p < p_{\max}$, we thus have

Corollary 5.10 *If $\mathcal{F} \subset \mathcal{P}(R)$ is 2-intersecting, then $\mu_p(\mathcal{F}) \leq p^{\bullet}$, provided $p < p_{\max}$.*

The proof of these two corollaries can be found in Appendix D.

6 Soundness

This section is the heart, and most technical part, of the proof of correctness, proving the construction is *sound*, that is, that if $G_{\mathcal{B}}^{\mathbb{C}}$ has a large independent-set, then G has a large h-clique–free set.

Lemma 6.1 (Soundness) $\mathrm{IS}(G_{\mathcal{B}}^{\mathbb{C}}) \geq p^{\bullet} + \varepsilon \quad \Longrightarrow \quad \mathrm{IS}_h(G) \geq \varepsilon_0 \cdot m.$

Proof Sketch: Assuming an independent-set $\mathcal{I} \subset V_{\mathcal{B}}^{\mathbb{C}}$ of weight $\Lambda(\mathcal{I}) \geq p^{\bullet} + \varepsilon$, we consider for each block $B \in \mathcal{B}$, its supposed long-code: the family $\mathcal{I}[B] = \mathcal{I} \cap V_{\mathcal{B}}^{\mathbb{C}}[B]$.

The first step (Lemma 6.2) is to find, for a non-negligible fraction of the blocks $\mathcal{B}_q \subseteq \mathcal{B}$, a small core of permissible block-assignments, and in it, one distinguished block-assignment to be used later to form a large h-clique–free set in G. This is done by showing that for every

$B \in \mathcal{B}_q$, $\mathcal{I}[B]$ has both significant weight and low average-sensitivity. This, not necessarily true for p, is asserted for some slightly shifted value $q \in (p, p+\gamma)$. Utilizing the Friedgut Lemma, we deduce the existence of a small core for $\mathcal{I}[B]$. Then, utilizing an Erdős-Ko-Rado-type bound on the maximal size of a 2-intersecting family, we find a distinguished block-assignment for each $B \in \mathcal{B}_q$.

The next step is to focus on one (e.g. random) $l-1$ sub-block $\hat{B} \in \binom{V}{l-1}$, and consider its extensions $\hat{B} \cup \{v\}$ for $v \in V = M \times R$, that represent the initial graph G. The distinguished block-assignments of those blocks that are in $\mathcal{B}_q$ will serve to identify a large set in V.

The final most delicate part of the proof, is Lemma 6.6, asserting that the distinguished block-assignments of the blocks extending $\hat{B}$ must identify an h-clique–free set as long as $\mathcal{I}$ is an independent-set. Indeed, since they all share the same $(l-1)$-sub-block $\hat{B}$, the edge constraints these blocks impose on one another will suffice to conclude the proof. Let us turn then to the proof of Lemma 6.1.

Proof: Let then $\mathcal{I} \subset V_{\mathcal{B}}^{\mathbb{C}}$ be an independent-set of size $\Lambda(\mathcal{I}) \geq p^{\bullet} + \varepsilon$, and denote, for each $B \in \mathcal{B}$,

$$\mathcal{I}[B] \stackrel{def}{=} \mathcal{I} \cap V_{\mathcal{B}}^{\mathbb{C}}[B] .$$

The fractional size of $\mathcal{I}[B]$ within $V_{\mathcal{B}}^{\mathbb{C}}[B]$, according to Λ_B, is $\Lambda_B(\mathcal{I}[B]) = \mu_p(\mathcal{I}[B])$.

Assume w.l.o.g. that $\mathcal{I}$ is maximal,

Observation. $\mathcal{I}[B]$, for any $B \in \mathcal{B}$, is monotone and intersecting.

Proof: It is intersecting, as $G_{\mathcal{B}}^{\mathbb{C}}$ has edges connecting vertices corresponding to non-intersecting subsets, and it is monotone due to maximality (see Proposition 4.6). ■

The first step in our proof is to find for a significant fraction of the blocks, a small core, and in it one distinguished block-assignment. Recall from Definition 5.6, that an element $\mathsf{a} \in C$ would be distinguished for a family $[\mathcal{I}[B]]_C^{\frac{3}{4}} \subseteq \mathcal{P}(C)$ if there are two subsets $F^{\flat}, F^{\sharp} \in [\mathcal{I}[B]]_C^{\frac{3}{4}}$ whose intersection is exactly $F^{\flat} \cap F^{\sharp} = \{\mathsf{a}\}$.

The Friedgut Lemma asserts the existence of a small core only for families with low average-sensitivity. We overcome this by slightly increasing p,

Lemma 6.2 *There exists some $q \in [p, p_{\max})$, and a set of blocks $\mathcal{B}_q \subseteq \mathcal{B}$ whose size is $|\mathcal{B}_q| \geq \frac{1}{4}\varepsilon \cdot |\mathcal{B}|$, such that for all $B \in \mathcal{B}_q$:*

1. *$\mathcal{I}[B]$ has an $(\frac{1}{16}\varepsilon, q)$-core, $\mathsf{Core}[B] \subset R_B$, of size $|\mathsf{Core}[B]| \leq h_0$.*

2. *The core-family $[\mathcal{I}[B]]_{\mathsf{Core}[B]}^{\frac{3}{4}}$ has a distinguished element $\dot{\mathsf{a}}[B] \in \mathsf{Core}[B]$.*

Proof: We will find a value $q \in [p, p_{\max})$ and a set of blocks $\mathcal{B}_q \subseteq \mathcal{B}$ such that for every $B \in \mathcal{B}_q$, $\mathcal{I}[B]$ is of large weight and low average sensitivity, according to μ_q. We will then proceed to show that this implies the above properties. First consider blocks whose intersection with $\mathcal{I}$ has weight not much lower than the expectation,

$$\mathcal{B}' \stackrel{def}{=} \left\{ B \in \mathcal{B} \;\middle|\; \Lambda_B(\mathcal{I}[B]) > p^{\bullet} + \frac{1}{2}\varepsilon \right\}$$

By a simple averaging argument, it follows that $|\mathcal{B}'| \geq \frac{1}{2}\varepsilon \cdot |\mathcal{B}|$, as otherwise

$$\Lambda(\mathcal{I}) \cdot |\mathcal{B}| = \sum_{B \in \mathcal{B}} \Lambda_B(\mathcal{I}[B]) \leq \frac{1}{2}\varepsilon\, |\mathcal{B}| + \sum_{B \notin \mathcal{B}'} \Lambda_B(\mathcal{I}[B]) < \frac{1}{2}\varepsilon\, |\mathcal{B}| + \sum_{B \notin \mathcal{B}'} (p^{\bullet} + \frac{1}{2}\varepsilon) \leq (p^{\bullet} + \varepsilon) \cdot |\mathcal{B}|$$

Since μ_p is non-decreasing with p (see Proposition 5.3), and since the value of γ was chosen so that for every $q \in [p, p+\gamma]$, $p^\bullet + \frac{1}{4}\varepsilon > q^\bullet$, we have for every block $B \in \mathcal{B}'$,

$$\mu_q(\mathcal{I}[B]) \geq \mu_p(\mathcal{I}[B]) > p^\bullet + \frac{1}{2}\varepsilon > q^\bullet + \frac{1}{4}\varepsilon \tag{1}$$

The family $\mathcal{I}[B]$, being monotone, cannot have high average sensitivity according to μ_q for many values of q, so by allowing an increase of at most γ, the set

$$\mathcal{B}_q \stackrel{def}{=} \left\{ B \in \mathcal{B}' \;\middle|\; \mathsf{as}_q(\mathcal{I}[B]) \leq \frac{2}{\gamma} \right\}$$

must be large for some $q \in (p, p+\gamma)$:

Proposition 6.3 *There exists $q \in (p, p+\gamma)$ so that $|\mathcal{B}_q| \geq \frac{1}{4}\varepsilon \cdot |\mathcal{B}|$.*

Proof: Consider the average, within $\mathcal{B}'$, of the size of $\mathcal{I}[B]$ according to μ_q

$$\mu_q[\mathcal{B}'] \stackrel{def}{=} \left|\mathcal{B}'\right|^{-1} \cdot \sum_{B \in \mathcal{B}'} \mu_q(\mathcal{I}[B])$$

and apply a version of Lagrange's Mean-Value Theorem: The derivative of $\mu_q[\mathcal{B}']$ as a function of q is

$$\frac{d\mu_q[\mathcal{B}']}{dq} = \left|\mathcal{B}'\right|^{-1} \cdot \sum_{B \in \mathcal{B}'} \frac{d\mu_q}{dq}(\mathcal{I}[B]) = \left|\mathcal{B}'\right|^{-1} \cdot \sum_{B \in \mathcal{B}'} \mathsf{as}_q(\mathcal{I}[B])$$

where the last equality follows from the Russo-Margulis identity (Lemma 5.4). Therefore, there must be some $q \in (p, p+\gamma)$ for which $\frac{d\mu_q[\mathcal{B}']}{dq} \leq \frac{1}{\gamma}$, as otherwise $\mu_{p+\gamma}[\mathcal{B}'] > 1$ which is impossible. It follows that at least half of the blocks in $\mathcal{B}'$ have $\mathsf{as}_q(\mathcal{I}[B]) \leq \frac{2}{\gamma}$. We have $|\mathcal{B}_q| \geq \frac{1}{2}|\mathcal{B}'| \geq \frac{1}{4}\varepsilon\,|\mathcal{B}|$. ■

Fix then $q \in (p, p+\gamma)$, to be as in the proposition above, so that $|\mathcal{B}_q| \geq \frac{1}{4}\varepsilon \cdot |\mathcal{B}|$. We next show that the properties claimed by the lemma, indeed hold for all blocks in $\mathcal{B}_q$. The first property, namely that $\mathcal{I}[B]$ has an $(\frac{1}{16}\varepsilon, q)$-core, denoted $\mathsf{Core}[B] \subset R_B$, of size $|\mathsf{Core}[B]| \leq h_0$, is immediate from the Friedgut Lemma (see Theorem 5.2), plugging in the average sensitivity of $\mathcal{I}[B]$, and by definition of $h_0 = \sup_{q \in [p, p_{\max}]} \Gamma(q, \frac{1}{16}\,\varepsilon, \frac{2}{\gamma})$, see Definition 4.1.

Having found a core for $\mathcal{I}[B]$, consider the core-family approximating $\mathcal{I}[B]$ on $\mathsf{Core}[B]$, (see Definition 5.2), denoted by

$$\mathcal{CF}_B \stackrel{def}{=} [\mathcal{I}[B]]^{\frac{3}{4}}_{\mathsf{Core}[B]} = \left\{ F \in \mathcal{P}\left(\mathsf{Core}[B]\right) \;\middle|\; \Pr_{F' \in \mu_p^{R \setminus \mathsf{Core}[B]}} \left[F \cup F' \in \mathcal{I}[B]\right] > \frac{3}{4} \right\}$$

By Proposition 5.7, since $\mathcal{I}[B]$ is monotone and intersecting, so is $\mathcal{CF}_B$. Moreover, Corollary 5.1 (a corollary of the Friedgut Lemma) asserts that

$$\mu_q(\mathcal{CF}_B) > \mu_q(\mathcal{I}[B]) - 4 \cdot \frac{\varepsilon}{16} > q^\bullet$$

where the second inequality follows from inequality (1), stating that $\mu_q(I[B]) > q^\bullet + \frac{1}{4}\varepsilon$ for any $B \in \mathcal{B}_q$. We can now utilize the bound on the maximal size of a 2-intersecting family (see Corollary 5.10), to deduce that $\mathcal{CF}_B$ is too large to be 2-intersecting, and must distinguish an element $\dot{\mathsf{a}} \in \mathsf{Core}[B]$, i.e. contain two subsets $F^\sharp, F^\flat \in \mathcal{CF}_B$ that intersect on exactly that block-assignment, $F^\sharp \cap F^\flat = \{\dot{\mathsf{a}}\}$. This completes the proof of Lemma 6.2. ■

Let us now fix q as guaranteed by Lemma 6.2 above. The following implicit definitions appeared in the above proof, and will be used later as well,

Definition 6.1 (Core, Core-Family, Distinguished Block-Assignment) *Let $B \in \mathcal{B}_q$.*

- *B's* core*, denoted $\mathsf{Core}[B] \subset R_B$, is an arbitrary smallest $(\frac{1}{16}\varepsilon, q)$-core of $\mathcal{I}[B]$.*
- *B's* core-family*, is the core-family on B's core (see Definition 5.2), denoted $\mathcal{CF}_B = [\mathcal{I}[B]]^{\frac{3}{4}}_{\mathsf{Core}[B]}$.*
- *B's* distinguished block-assignment*, is an arbitrary distinguished element of $\mathcal{CF}_B$, denoted $\dot{\mathsf{a}}[B] \in \mathsf{Core}[B]$.*

Let us further define for each block $B \in \mathcal{B}_q$, the set of all block-assignments of B that have non-negligible influence (i.e. larger than $\eta = \frac{1}{16h_0} \cdot p^{8h_0}$):

Definition 6.2 (Extended Core) *For $B \in \mathcal{B}$, let the* extended core *of B be*

$$\mathsf{ECore}[B] \stackrel{def}{=} \mathsf{Core}[B] \cup \left\{ \mathsf{a} \in R_B \;\middle|\; \mathsf{influence}^{\mathsf{a}}_q(\mathcal{I}[B]) \geq \eta \right\}$$

The extended core is not much larger than the core, because the total sum of influences of elements in R_B, is bounded for every $B \in \mathcal{B}_q$, by $\mathsf{as}_q(\mathcal{I}[B]) \leq \frac{2}{\gamma}$,

$$|\mathsf{ECore}[B]| \leq h_0 + \frac{\mathsf{as}_q(\mathcal{I}[B])}{\eta} \leq h_0 + \lceil \tfrac{2}{\gamma \cdot \eta} \rceil = h_1$$

The next step in our proof, is to identify for every $(l-1)$-sub-block $\hat{B} \in \binom{V}{l-1}$ a subset $V_{\hat{B}} \subset V$ that is h-clique–free. The members of $V_{\hat{B}}$ will be selected according to the distinguished block-assignments of the blocks extending $\hat{B}$. Analyzing the consistency between the distinguished block-assignments of distinct blocks, is complicated by the fact that families encoding distinct blocks consist of subsets of distinct domains ($R_{B_1} \neq R_{B_2}$ for $B_1 \neq B_2$). Considering only the blocks that extend a specific sub-block $\hat{B} \in \binom{V}{l-1}$, yields a nice 2-to-2 correspondence between their block-assignments. The block-assignments of blocks $B = \hat{B} \cup \{v\}$ are paired according to their restriction to $\hat{B}$, such that all the pairs whose restriction is mapped to the same sub-block-assignment naturally correspond to each other.

It would be undesired to have both block-assignments in a given pair influential in $\mathcal{I}[B]$ for this would mean that the structure of $\mathcal{I}[B]$ is not preserved when reduced to $\hat{B}$. Thus, besides requiring that many of the blocks $\hat{B} \cup \{v\}$ extending $\hat{B}$ reside in $\mathcal{B}_q$, we need them to be *preserved* by $\hat{B}$:

Definition 6.3 (Preservation) *Let $B \in \mathcal{B}$, and let $\hat{B} \subset B$, $|\hat{B}| = l-1$. Let us denote by $\mathsf{a}|_{\hat{B}}$ the restriction to $\hat{B}$ of a block-assignment $\mathsf{a} \in R_B$. We say that $\hat{B}$* preserves *B, if there is no pair of block-assignments $\mathsf{a}_1 \neq \mathsf{a}_2 \in \mathsf{ECore}[B]$ with $\mathsf{a}_1|_{\hat{B}} = \mathsf{a}_2|_{\hat{B}}$.*

It is almost always the case that $\hat{B}$ preserves $\hat{B} \cup \{v\}$:

Proposition 6.4

$$\forall B \in \mathcal{B} \quad |\{ v \in B \mid B \setminus \{v\} \text{ does not preserve } B \}| < \frac{(h_1)^2}{2}.$$

Proof: Each pair of block-assignments $\mathsf{a}_1, \mathsf{a}_2 \in \mathsf{ECore}[B]$ can cause at most one $\hat{B}$ to not preserve B, and for any block $B \in \mathcal{B}_q$, $|\mathsf{ECore}[B]| \leq h_1$; consequently, the number of $\hat{B}$ not preserving B is at most $\binom{h_1}{2} < \frac{(h_1)^2}{2}$. ■

The last step before identifying the required $\hat{B}$ is to note that a distinguished block-assignment for a block $\hat{B} \cup \{v\}$ is useful for constructing an h-clique–free subset in G, if it assigns T to v. Hence, for each $\hat{B}$ we consider the following set $V_{\hat{B}} \subset V$:

Definition 6.4 *Let $V_{\hat{B}} \subseteq V$ be:*

$$V_{\hat{B}} \stackrel{def}{=} \left\{ v \in V \setminus \hat{B} \,\middle|\, B = \hat{B} \cup \{v\} \in \mathcal{B}_q \quad \textit{and} \quad \hat{B} \textit{ preserves } B \quad \textit{and} \quad \dot{\mathsf{a}}[B](v) = \mathsf{T} \right\}$$

It follows from the definition of $V_{\hat{B}}$, that if $v_1, v_2 \in V_{\hat{B}}$ are connected by an edge in G, then the distinguished block-assignments of $B_1 = \hat{B} \cup \{v_1\}$ and $B_2 = \hat{B} \cup \{v_2\}$ are connected by an edge in the graph $G_{\mathcal{B}}$, $\langle \dot{\mathsf{a}}[B_1], \dot{\mathsf{a}}[B_2] \rangle \in E_{\mathcal{B}}$ (see Definition 4.2). Finally, let us identify a sub-block $\hat{B}$, for which $V_{\hat{B}}$ is large:

Proposition 6.5 *There exists $\hat{B} \in \binom{V}{l-1}$, with $\left|V_{\hat{B}}\right| \geq \varepsilon_0 \cdot m$.*

Proof: Observe that

$$\Pr_{\hat{B},\, v \in V \setminus \hat{B}} \left[v \in V_{\hat{B}} \right] \geq \frac{1}{4}\varepsilon \cdot \Pr_{B,\, v \in B} \left[v \in V_{B \setminus \{v\}} \,\middle|\, B \in \mathcal{B}_q \right] \geq \frac{1}{4}\varepsilon \cdot \frac{1}{4r}$$

where the first inequality follows from Proposition 6.3 asserting $\mathcal{B}_q \geq \frac{1}{4}\varepsilon\, |\mathcal{B}|$. The second inequality is a consequence of the fact that for any $\mathsf{a} \in R_B$, there are at least $l_{\mathsf{T}} = \frac{l}{2r}$ elements $v \in B$ with $\mathsf{a}(v) = \mathsf{T}$; and at most $\frac{(h_1)^2}{2}$ $(l-1)$-blocks $\hat{B} \subset B$ not preserving B; hence, conditioned on $B \in \mathcal{B}_q$, the probability of $v \in V_{\hat{B}}$ is at least $\frac{1}{2r} - \frac{(h_1)^2}{2l} \geq \frac{1}{4r}$ as $l \geq 2(h_1)^2 \cdot r$.

This inequality shows that there is at least one $\hat{B}$ for which $\Pr_{v \in V \setminus \hat{B}} \left[v \in V_{\hat{B}} \right] \geq \frac{\varepsilon}{16r}$, hence, $\left|V_{\hat{B}}\right| \geq \frac{1}{16r}\varepsilon \cdot \left| V \setminus \hat{B} \right| \geq \frac{1}{32}\varepsilon \cdot m$, as $\frac{|V \setminus \hat{B}|}{r} > \frac{1}{2}m$, because $|\hat{B}| = l - 1 \ll \frac{1}{2}|V|$, see Definition 4.1. ■

Finally, we establish $\mathrm{IS}_h(G) \geq \varepsilon_0 \cdot m$ by proving,

Lemma 6.6 *The set $V_{\hat{B}}$ contains no clique of size h.*

Proof: (of Lemma 6.6) Assume, by way of contradiction, that there exists a clique over vertices $v_1, \ldots, v_h \in V_{\hat{B}}$. We will show that, for $B_i = \hat{B} \cup \{v_i\}$, the set $\cup_{i \in [h]} \mathcal{I}[B_i]$ is not an independent-set. In fact, we will find two of these blocks, B_{i_1}, B_{i_2}, such that $\mathcal{I}[B_{i_1}] \cup \mathcal{I}[B_{i_2}]$ is not an independent set.

Analyzing consistency between blocks $\hat{B} \cup \{v_i\}$ leads us to consider the mutual sub-block $\hat{B}$, and the sub-block-assignments that are restrictions of block-assignments in R_{B_i} to $\hat{B}$. The $(l-1)$-*block-assignments* of $\hat{B} \in \binom{V}{l-1}$, are defined to be

$$R_{\hat{B}} \stackrel{def}{=} \left\{ \mathsf{a} \colon \hat{B} \to \{\mathsf{T}, \mathsf{F}\} \right\}$$

A block-assignment $\mathsf{a} \in R_{B_i}$ has a natural restriction to $\hat{B}$, denoted $\mathsf{a}|_{\hat{B}} \in R_{\hat{B}}$, where $\forall v \in \hat{B}$, $\mathsf{a}|_{\hat{B}}(v) = \mathsf{a}(v)$.

For the remaining analysis, let us name the three important entities regarding each block B_i, for $i \in [h]$: B_i's distinguished block-assignment, the core of B_i, and the extended core of B_i,

$$\dot{\mathsf{a}}_i \stackrel{def}{=} \dot{\mathsf{a}}[B_i] \qquad C_i \stackrel{def}{=} \mathsf{Core}[B_i] \qquad E_i \stackrel{def}{=} \mathsf{ECore}[B_i]$$

and their natural restrictions to $\hat{B}$ (where the natural restriction of a set is the set comprising the restrictions of its elements),

$$\hat{\mathsf{a}}_i \stackrel{def}{=} \dot{\mathsf{a}}_i|_{\hat{B}} \qquad \hat{C}_i \stackrel{def}{=} C_i|_{\hat{B}} \qquad \hat{E}_i \stackrel{def}{=} E_i|_{\hat{B}}$$

Now, recall the core-family $\mathcal{CF}_{B_i}$, which is the family of subsets, over the core of each B_i, each of which extension is of $\frac{3}{4}$ weight in $\mathcal{I}[B_i]$. For each block B_i, $i \in [h]$, $\dot{\mathsf{a}}_i$ being distinguished implies a pair of subsets

$$F_i^\flat, F_i^\sharp \in \mathcal{CF}_{B_i} \ \text{ so that } \ F_i^\flat \cap F_i^\sharp = \{\dot{\mathsf{a}}_i\}$$

Let their natural restriction to $\hat{B}$ be

$$\hat{F}_i^\flat \stackrel{def}{=} F_i^\flat|_{\hat{B}} \qquad \hat{F}_i^\sharp \stackrel{def}{=} F_i^\sharp|_{\hat{B}}$$

and note that, as $\hat{B}$ preserves every B_i, it follows that, for all $i \in [h]$,

$$\hat{F}_i^\flat \cap \hat{F}_i^\sharp = \{\hat{\mathsf{a}}_i\} \tag{2}$$

Our first goal is to identify two blocks B_{i_1} and B_{i_2} that are extremely inconsistent:

Proposition 6.7 *There exist $i_1 \neq i_2 \in [h]$, such that, denoting $\mathbf{\Delta} = \hat{E}_{i_1} \cap \hat{E}_{i_2}$,*

1. $\hat{C}_{i_1} \cap \mathbf{\Delta} = \hat{C}_{i_2} \cap \mathbf{\Delta}$
2. $\hat{F}_{i_1}^\flat \cap \mathbf{\Delta} = \hat{F}_{i_2}^\flat \cap \mathbf{\Delta}$
3. $\hat{F}_{i_1}^\sharp \cap \mathbf{\Delta} = \hat{F}_{i_2}^\sharp \cap \mathbf{\Delta}$

Proof: Our proof begins by applying the following Sunflower-Lemma over the sets $\hat{E}_i$:

Theorem 6.8 ([ER60]) *There exists some integer function $\Gamma_*(k,d)$ (not depending on $|R|$), such that for any $\mathcal{F} \subset \binom{R}{k}$, if $|\mathcal{F}| \geq \Gamma_*(k,d)$, there are d distinct sets $F_1, \ldots, F_d \in \mathcal{F}$, such that, let $\mathbf{\Delta} \stackrel{def}{=} F_1 \cap \ldots \cap F_d$, the sets $F_i \setminus \mathbf{\Delta}$ are pairwise disjoint.* ■

The sets $F_1, .., F_d$ are called a Sunflower, or a $\mathbf{\Delta}$-system. This statement can easily be extended to families in which each subset is of size *at most* k.

We apply this lemma for $R = R_{\hat{B}}$, and $\mathcal{F} = \{\hat{E}_1, .., \hat{E}_h\}$. Recall (definition 4.1) we have fixed $h > \Gamma_*(h_1, h_s)$, hence Theorem 6.8 implies there exists some $J \subseteq [h]$, $|J| = h_s$, such that

$$\left\{\hat{E}_i \setminus \mathbf{\Delta}\right\}_{i \in J} \text{ are pairwise disjoint for } \mathbf{\Delta} \stackrel{def}{=} \bigcap_{i \in J} \hat{E}_i$$

Consider, for each $i \in J$, the triplet $\left\langle \hat{C}_i \cap \mathbf{\Delta}, \quad \hat{F}_i^\flat \cap \mathbf{\Delta}, \quad \hat{F}_i^\sharp \cap \mathbf{\Delta} \right\rangle$, and note that, since $\hat{F}_i^\flat, \hat{F}_i^\sharp \subseteq \hat{C}_i$ the number of possible triplets is at most

$$\begin{aligned} \left| \left\{ \left\langle \hat{C} \cap \mathbf{\Delta}, \hat{F}^\flat \cap \mathbf{\Delta}, \hat{F}^\sharp \cap \mathbf{\Delta} \right\rangle \;\middle|\; |\hat{C}| \le h_0,\, \hat{F}^\flat, \hat{F}^\sharp \subseteq \hat{C} \right\} \right| &\le \sum_{k=0}^{h_0} \binom{h_1}{k} \cdot 2^{h_0} \cdot 2^{h_0} \\ &< h_s = |J| \end{aligned}$$

(recall we have set (definition 4.1) $h_s = 1 + 2^{2h_0} \cdot \sum_{k=0}^{h_0} \binom{h_1}{k})$). Therefore, by the pigeon-hole principle, there must be some $i_1, i_2 \in J$ for which

$$\left\langle \hat{C}_{i_1} \cap \mathbf{\Delta},\ \hat{F}_{i_1}^\flat \cap \mathbf{\Delta},\ \hat{F}_{i_1}^\sharp \cap \mathbf{\Delta} \right\rangle = \left\langle \hat{C}_{i_2} \cap \mathbf{\Delta},\ \hat{F}_{i_2}^\flat \cap \mathbf{\Delta},\ \hat{F}_{i_2}^\sharp \cap \mathbf{\Delta} \right\rangle$$

■

From now on let us assume w.l.o.g. that $i_1 = 1$, $i_2 = 2$, and continue to denote $\mathbf{\Delta} = \hat{E}_1 \cap \hat{E}_2$. We will arrive at a contradiction by finding an edge between the blocks B_1, B_2, specifically, by finding two extensions, one of $F_1^\flat$ in $\mathcal{I}[B_1]$, and another of $F_2^\sharp$ in $\mathcal{I}[B_2]$, all of whose block-assignments are pairwise inconsistent.

As a first step, let us prove that the block-assignments in $F_1^\flat$ and $F_2^\sharp$ are pairwise inconsistent:

Proposition 6.9 $\left\langle F_1^\flat, F_2^\sharp \right\rangle \in E_{\mathcal{B}}^{\mathbb{C}}$.

Proof: We need to prove that for all $\mathsf{a}_1 \in F_1^\flat, \mathsf{a}_2 \in F_2^\sharp$, $\langle \mathsf{a}_1, \mathsf{a}_2 \rangle \in E_{\mathcal{B}}$. If $\langle \mathsf{a}_1, \mathsf{a}_2 \rangle \notin E_{\mathcal{B}}$, it must be that $\mathsf{a}_1|_{\hat{B}} = \mathsf{a}_2|_{\hat{B}} \in \hat{F}_1^\flat \cap \hat{F}_2^\sharp \subseteq \hat{E}_1 \cap \hat{E}_2 = \mathbf{\Delta}$. B_1 and B_2 are chosen in Proposition 6.7 so that $\hat{F}_1^\flat \cap \mathbf{\Delta} = \hat{F}_2^\flat \cap \mathbf{\Delta}$ and $\hat{F}_1^\sharp \cap \mathbf{\Delta} = \hat{F}_2^\sharp \cap \mathbf{\Delta}$. Consequently $\mathsf{a}_1|_{\hat{B}} = \mathsf{a}_2|_{\hat{B}} \in \hat{F}_1^\flat \cap \hat{F}_1^\sharp \cap \mathbf{\Delta} = \hat{F}_2^\flat \cap \hat{F}_2^\sharp \cap \mathbf{\Delta}$, however (2) asserts that the only block-assignment in these two intersections is the distinguished one, hence $\hat{\mathsf{a}}_1 = \mathsf{a}_1|_{\hat{B}} = \mathsf{a}_2|_{\hat{B}} = \hat{\mathsf{a}}_2$. Since $\hat{B}$ preserves both B_1 and B_2, $\mathsf{a}_1 = \dot{\mathsf{a}}_1$ and $\mathsf{a}_2 = \dot{\mathsf{a}}_2$. However, $\langle \dot{\mathsf{a}}_1, \dot{\mathsf{a}}_2 \rangle \in E_{\mathcal{B}}$ (recall Definition 4.2), as both $\dot{\mathsf{a}}_1, \dot{\mathsf{a}}_2$ assign T to v_1, v_2 respectively and $\langle v_1, v_2 \rangle \in E$. ■

It may well be that $F_1^\flat \notin \mathcal{I}[B_1]$ and $F_2^\sharp \notin \mathcal{I}[B_2]$, thus the proposition above is only a first step towards a contradiction. Nevertheless, we know that $F_1^\flat \in \mathcal{CF}_{B_1} = [\mathcal{I}[B_1]]^{\frac{3}{4}}_{\mathsf{Core}[B_1]}$ means that $\frac{3}{4}$ of $\left\{ F \in \mathcal{P}(R_{B_1}) \mid F \cap \mathsf{Core}[B_1] = F_1^\flat \right\}$ are in $\mathcal{I}[B_1]$; and likewise for $F_2^\sharp$. In what follows, we utilize this large volume of $\frac{3}{4}$ to find extensions of these sets, that are in $\mathcal{I}$, yet are connected by an edge in $E_{\mathcal{B}}^{\mathbb{C}}$.

Let us partition the set of $(l-1)$-block assignments of $R_{\hat{B}}$ into the important ones, which are restrictions of block-assignments in the cores of B_1 or B_2, and the rest,

$$\hat{D} = \hat{C}_1 \cup \hat{C}_2 \quad \text{and} \quad \hat{R} = R_{\hat{B}} \setminus \hat{D}$$

which immediately partitions the block-assignments of R_{B_1} and R_{B_2}, according to whether their restriction falls within $\hat{D}$:

$$D_1 = \left\{ \mathsf{a} \in R_{B_1} \;\middle|\; \mathsf{a}|_{\hat{B}} \in \hat{D} \right\} \quad \text{and} \quad R_1 = R_{B_1} \setminus D_1$$

and similarly for R_{B_2},

$$D_2 = \left\{ \mathsf{a} \in R_{B_2} \;\middle|\; \mathsf{a}|_{\hat{B}} \in \hat{D} \right\} \quad \text{and} \quad R_2 = R_{B_2} \setminus D_2$$

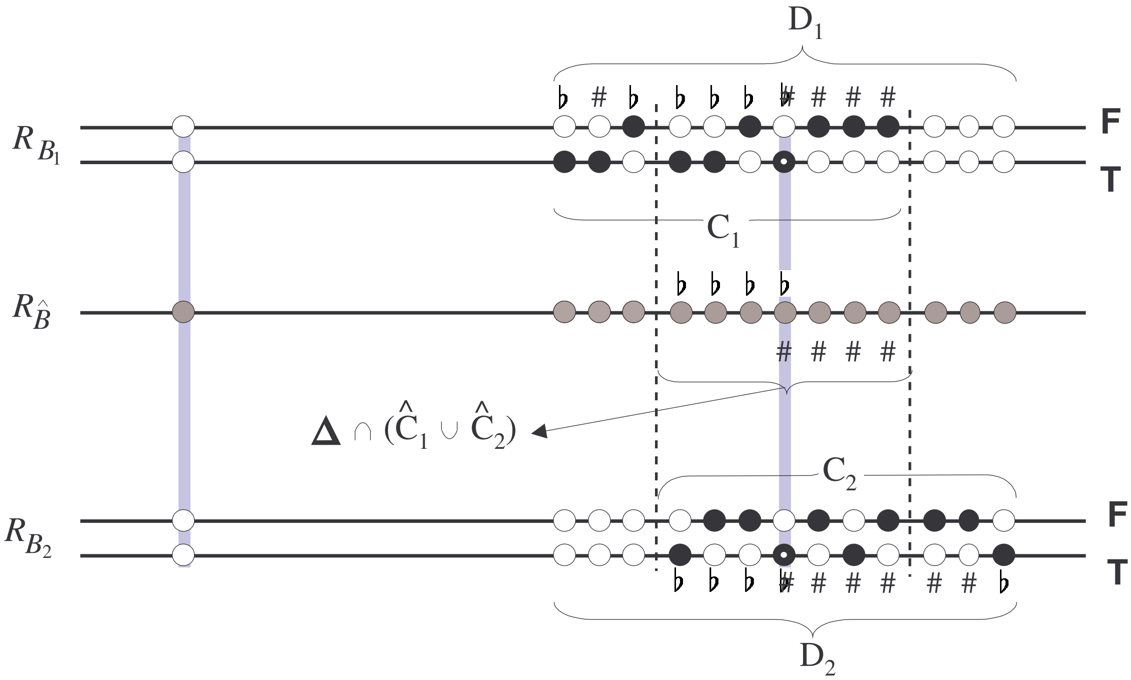

Figure 1: Block Assignments of B_1, B_2 and sub-block-assignments of $\hat{B}$.
R_{B_1} (resp. R_{B_2}) is represented by the two upper (resp. two lower) horizontal lines labelled by T and F to indicate the value assigned to v_1 (resp. v_2) by block-assignments on that line. Each circle represents a single block assignment. On the left a column (highlighted as a light gray vertical line) consists of four block assignments and a sub-block assignment which is their mutual restriction to $\hat{B}$. All block assignments in the same column agree on their restriction to $\hat{B}$, depicted as a gray circle on the middle horizontal line that represents $R_{\hat{B}}$. Two block assignments are consistent *only* if they are in the same column and are not both T. The blackened circles represent members of the core of B_1 and the block-assignments in $F_1^\flat$ and $F_1^\sharp$ are labelled $\flat$ and $\sharp$. The distinguished block-assignment – marked by a white dot – is labelled by both $\flat$ and $\sharp$, and assigns T to v_1. The dashed vertical lines border the intersection of $\hat{C}_1$ with $\hat{C}_2$, which is equal to $\hat{C}_1 \cap \hat{C}_2 = \hat{C}_1 \cap \boldsymbol{\Delta} = \hat{C}_2 \cap \boldsymbol{\Delta}$ and is where the restrictions of $F_1^\sharp, F_1^\flat$ are equal to those of $F_2^\sharp, F_2^\flat$. This also implies that $(\hat{D}_1 \setminus \hat{C}_1) \cap \boldsymbol{\Delta} \subseteq (\hat{D}_1 \setminus \hat{C}_1) \cap \hat{E}_1 = \phi$.

Proposition 6.10 *$|D_1| \leq 4h_0$ and $|D_2| \leq 4h_0$.*

Proof: Simply note that $|D_1|, |D_2| \leq 2|\hat{D}| \leq 2(|\hat{C}_1| + |\hat{C}_2|) \leq 2(|C_1| + |C_2|) = 4h_0$. ■

Notice that $F_1^\flat \in \mathcal{P}(C_1) \subseteq \mathcal{P}(D_1)$ and $F_2^\sharp \in \mathcal{P}(C_2) \subseteq \mathcal{P}(D_2)$, hence it suffices to exhibit two subsets $H_1 \in \mathcal{P}(R_1)$ and $H_2 \in \mathcal{P}(R_2)$ all of whose block-assignments are pairwise-inconsistent, and so that $F_1^\flat \cup H_1 \in \mathcal{I}[B_1]$ and $F_2^\sharp \cup H_2 \in \mathcal{I}[B_2]$.

Let us prove this by showing first that the families of subsets extending $F_1^\flat$ and $F_2^\sharp$ within $\mathcal{I}$ are large; and then proceed to show that this large volume implies the existence of two subsets, H_1 and H_2 as required.

Let us first name these two families of subsets extending $F_1^\flat$ and $F_2^\sharp$ within $\mathcal{I}$:

$$\mathcal{I}_1 = \left\{ F \in \mathcal{P}(R_1) \;\middle|\; (F_1^\flat \cup F) \in \mathcal{I}[B_1] \right\} \quad \text{and} \quad \mathcal{I}_2 = \left\{ F \in \mathcal{P}(R_2) \;\middle|\; (F_2^\sharp \cup F) \in \mathcal{I}[B_2] \right\}$$

and proceed to prove they are large:

Proposition 6.11

$$\mu_q^{R_1}(\mathcal{I}_1) > \frac{1}{2} \qquad \textit{and} \qquad \mu_q^{R_2}(\mathcal{I}_2) > \frac{1}{2}$$

Proof: Let us prove the first case; the second one is proven by a symmetric, but otherwise identical, argument. By definition of $\mathcal{CF}_{B_1} = [\mathcal{I}[B_1]]_{C_1}^{\frac{3}{4}}$, it is the case that

$$\Pr_{F \in \mu_q}\left[F \in \mathcal{I}[B_1] \;\middle|\; F \cap C_1 = F_1^\flat\right] > \frac{3}{4}$$

Note that the only difference between this event and

$$\mu_q^{R_1}(\mathcal{I}_1) = \Pr_{F \in \mu_q}\left[F \in \mathcal{I}[B_1] \;\middle|\; F \cap D_1 = F_1^\flat\right]$$

is the conditioning on F to not contain any block-assignment in $D_1 \setminus C_1$. Simplistically, if the elements in $D_1 \setminus C_1$ have tiny influence, then removing them from a subset does not take it out of $\mathcal{I}[B_1]$. Hence, it suffices to prove that this family, of extensions of $F_1^\flat$ within $\mathcal{I}[B_1]$, is almost independent of the set of block-assignments $D_1 \setminus C_1$, that is, that one can extract a small ($< \frac{1}{4}$) fraction of $\mathcal{I}_1$ and make it completely independent of the block-assignments outside $R_1 \cup C_1$.

Let us first observe that block-assignments in $D_1 \setminus C_1$ indeed have tiny influence,

Proposition 6.12

$$(D_1 \setminus C_1) \cap E_1 = \phi$$

Proof: There are two cases to consider for $\mathsf{a} \in D_1 \setminus C_1$: Either $\mathsf{a}|_{\hat{B}} \in \hat{C}_1$ and in that case, since $\hat{B}$ preserves B_1 and since $\mathsf{a} \notin C_1$, we deduce $\mathsf{a} \notin E_1$; or, $\mathsf{a}|_{\hat{B}} \in \hat{C}_2 \setminus \hat{C}_1$ and since the first condition on B_1 and B_2 in Proposition 6.7 is that $\hat{C}_1 \cap \mathbf{\Delta} = \hat{C}_2 \cap \mathbf{\Delta}$, we deduce $\mathsf{a}|_{\hat{B}} \notin \mathbf{\Delta}$. Now $\mathsf{a}|_{\hat{B}} \in \hat{C}_2 \subseteq \hat{E}_2$, implies $\mathsf{a}|_{\hat{B}} \notin \hat{E}_1$, thus $\mathsf{a} \notin E_1$. ∎

By definition of the extended core E_i (Definition 6.1), it follows that for every $\mathsf{a} \in D_1 \setminus C_1$, $\mathsf{influence}_q^{\mathsf{a}}(\mathcal{I}[B_1]) < \eta$. Since $|D_1 \setminus C_1| < 4h_0$ (Proposition 6.10) we can deduce that $\mathcal{I}[B_1]$ is almost independent of $D_1 \setminus C_1$, utilizing a relatively simple, general property related to influences. Namely, that, given any monotone family of subsets of a domain R, and a set $U \subset R$ of elements of tiny influence, one has to remove only a small fraction of the family to make it completely independent of U, i.e. determined by $R \setminus U$. More accurately, we prove the following simple proposition in Appendix C,

Proposition 6.13 *Let $\mathcal{F} \subset \mathcal{P}(R)$ be monotone, and let $U \subset R$ be such that for all $e \in U$, $\mathsf{influence}_p^e(\mathcal{F}) < \eta$. Let*

$$\mathcal{F}' = \{F \in \mathcal{F} \mid F \setminus U \in \mathcal{F}\}$$

then,

$$\mu_p^R\left(\mathcal{F} \setminus \mathcal{F}'\right) < |U| \cdot \eta \cdot p^{-|U|}$$

Proof: See Appendix C. ∎

Substituting $D_1 \setminus C_1$ for U and $\frac{1}{16h_0} \cdot p^{5h_0}$ for η (see Definition 4.1), this proposition asserts that the weight of the subsets that have to be removed from $\mathcal{I}[B_1]$ to make it independent of $D_1 \setminus C_1$,

$$\mathcal{I}[B_1]' \stackrel{def}{=} \{F \in \mathcal{I}[B_1] \mid (F \setminus (D_1 \setminus C_1)) \notin \mathcal{I}[B_1]\},$$

is bounded by

$$\mu_q^{R_{B_1}}(\mathcal{I}[B_1]') < 4h_0 \cdot \eta \cdot q^{-4h_0} \leq \frac{1}{4} q^{h_0} .$$

Now, even if all $\mathcal{I}[B_1]'$ is concentrated on $F_1^\flat$, since $F_1^\flat$'s weight in $\mathcal{P}(C_1)$ is at least $q^{|C_1|} \geq q^{h_0}$, $\mu_q^{C_1}\left(F_1^\flat\right) \geq q^{h_0}$. It follows that (using $\Pr(A \mid B) \leq \Pr(A)/\Pr(B)$),

$$\Pr_{F \in \mu_q^{R_1}} \left[F \in \mathcal{I}[B_1]' \,\middle|\, F \cap C_1 = F_1^\flat\right] \leq \Pr_{F \in \mu_q^{R_1}} \left[F \in \mathcal{I}[B_1]'\right] \cdot \frac{1}{\mu_q^{C_1}\left(F_1^\flat\right)} < \frac{1}{4}$$

Formally, we write

$$\begin{aligned}
\frac{3}{4} &< \Pr\left[F \in \mathcal{I}[B_1] \mid F \cap C_1 = F_1^\flat\right] = \\
&= \Pr\left[F \in \mathcal{I}[B_1] \setminus \mathcal{I}[B_1]' \,\middle|\, F \cap C_1 = F_1^\flat\right] + \Pr\left[F \in \mathcal{I}[B_1]' \,\middle|\, F \cap C_1 = F_1^\flat\right] \\
&< \Pr\left[F \in \mathcal{I}[B_1] \setminus \mathcal{I}[B_1]' \,\middle|\, F \cap D_1 = F_1^\flat\right] + \frac{1}{4}
\end{aligned}$$

Implying that $\mu_q^{R_1}(\mathcal{I}_1) = \Pr\left[F \in \mathcal{I}[B_1] \mid F \cap D_1 = F_1^\flat\right] > \frac{1}{2}$, and completing the proof of Proposition 6.11. ■

We complete the proof of the Soundness Lemma, by deducing from the large volume of I_1, I_2, the existence of two subsets $H_1 \in I_1$ and $H_2 \in I_2$ so that $\langle H_1, H_2\rangle \in E_{\mathcal{B}}^{\mathbb{C}}$, implying $\left\langle F_1^\flat \cup H_1, F_2^\sharp \cup H_2\right\rangle \in E_{\mathcal{B}}^{\mathbb{C}}$, which is the desired contradiction.

Proposition 6.14 *Let $I_1 \subset \mathcal{P}(R_1), I_2 \subset \mathcal{P}(R_2)$. If $(1-q)^2 \geq q$ and $\mu_q^{R_1}(\mathcal{I}_1) + \mu_q^{R_2}(\mathcal{I}_2) > 1$, there exist $H_1 \in \mathcal{I}_1$ and $H_2 \in \mathcal{I}_2$ such that $\langle H_1, H_2\rangle \in E_{\mathcal{B}}^{\mathbb{C}}$.*

Proof: This proposition is proven by modifying the proof for the case of cross-intersecting families (Proposition 5.6). In that proof, we bounded the size of a pair of cross-intersecting families by pairing each subset with its complement, noting that at $p = \frac{1}{2}$ their weights are equal.

In this case, we focus on the value $q = p_{\max} = \frac{3-\sqrt{5}}{2}$ for which $(1-q)^2 = q$, noting that since $q \leq p_{\max}$, the monotonicity of I_1, I_2 (see Proposition 5.3) yields $\mu_{p_{\max}}(I_1) + \mu_{p_{\max}}(I_2) > 1$. Here let us partition both $\mathcal{P}(R_1)$ and $\mathcal{P}(R_2)$, and define an appropriate 'complement' for each part, rather than for each subset.

Our partition is defined according to a 'representative mapping' mapping each $F \in \mathcal{P}(R_1)$ to a function $\Pi[F_1] : \hat{R} \to \left\{\overline{\mathsf{TF}}, \mathsf{T}\overline{\mathsf{F}}, \mathsf{F}\right\}$ defined as follows:

$$\forall \hat{\mathsf{a}} \in \hat{R}, \quad \Pi[F_1](\hat{\mathsf{a}}) \stackrel{def}{=} \begin{cases} \overline{\mathsf{TF}} & \hat{\mathsf{a}}^{(z_1 \leftarrow \mathsf{T})}, \hat{\mathsf{a}}^{(z_1 \leftarrow \mathsf{F})} \notin F_1 \\ \mathsf{T}\overline{\mathsf{F}} & \hat{\mathsf{a}}^{(z_1 \leftarrow \mathsf{T})} \in F_1,\ \hat{\mathsf{a}}^{(z_1 \leftarrow \mathsf{F})} \notin F_1 \\ \mathsf{F} & \hat{\mathsf{a}}^{(z_1 \leftarrow \mathsf{F})} \in F_1 \end{cases}$$

(symmetrically, we define $\Pi[F_2]$ for each $F_2 \in \mathcal{P}(R_2)$). This mapping is natural when considering the characteristic function of F_1 and asking, for every $\hat{\mathsf{a}} \in \hat{R}$, the value of that function on the two extensions of $\hat{\mathsf{a}}$ in R_1, $\hat{\mathsf{a}}^{(z_1 \leftarrow \mathsf{T})}$ and $\hat{\mathsf{a}}^{(z_1 \leftarrow \mathsf{F})}$.

Additionally, for a function $\Pi = \Pi[F_1]$, $\Pi : \hat{R} \to \left\{\overline{\mathsf{TF}}, \mathsf{T}\overline{\mathsf{F}}, \mathsf{F}\right\}$, let its complement be $\Pi^{\mathsf{c}} : \hat{R} \to \left\{\overline{\mathsf{TF}}, \mathsf{T}\overline{\mathsf{F}}, \mathsf{F}\right\}$ defined as follows:

$$\forall \hat{\mathsf{a}} \in \hat{R}, \quad \Pi^{\mathsf{c}}(\hat{\mathsf{a}}) \stackrel{def}{=} \begin{cases} \overline{\mathsf{TF}} & \Pi(\hat{\mathsf{a}}) = \mathsf{F} \\ \mathsf{T}\overline{\mathsf{F}} & \Pi(\hat{\mathsf{a}}) = \mathsf{T}\overline{\mathsf{F}} \\ \mathsf{F} & \Pi(\hat{\mathsf{a}}) = \overline{\mathsf{TF}} \end{cases}$$

Observe that $\Pi^{\mathsf{cc}} = \Pi$, and that this is indeed a perfect matching of the possible functions $\Pi : \hat{R} \to \left\{\overline{\mathsf{TF}}, \mathsf{T}\overline{\mathsf{F}}, \mathsf{F}\right\}$, and most importantly that $\Pi[H_1] = \Pi^{\mathsf{c}}[H_2]$ implies $\langle H_1, H_2\rangle \in E_{\mathcal{B}}^{\mathbb{C}}$.

Next, observe that for a fixed $\Pi_0 : \hat{R} \to \left\{\overline{\mathsf{TF}}, \mathsf{T}\overline{\mathsf{F}}, \mathsf{F}\right\}$,

$$\Pr_{F_1 \in \mu_q^{R_1}} [\Pi[F_1] = \Pi_0] = \prod_{\hat{\mathsf{a}}:\, \Pi_0(\hat{\mathsf{a}}) = \overline{\mathsf{TF}}} (1-q)^2 \cdot \prod_{\hat{\mathsf{a}}:\, \Pi_0(\hat{\mathsf{a}}) = \mathsf{T}\overline{\mathsf{F}}} q(1-q) \cdot \prod_{\hat{\mathsf{a}}:\, \Pi_0(\hat{\mathsf{a}}) = \mathsf{F}} q$$

Now if $q = p_{\max}$, i.e. $(1-q)^2 = q$, we have $\Pr_F[\Pi[F] = \Pi_0] = \Pr_F[\Pi[F] = \Pi_0^{\mathsf{c}}]$. Since $\mu_q(I_1) + \mu_q(I_2) > 1$, there must be a pair Π, Π^{c} such that

$$\{F_1 \in \mathcal{P}(R_1) \mid \Pi[F_1] = \Pi\} \cap I_1 \neq \phi \quad \text{and} \quad \{F_2 \in \mathcal{P}(R_2) \mid \Pi[F_2] = \Pi^{\mathsf{c}}\} \cap I_2 \neq \phi$$

providing the necessary pair of $H_1 \in I_1, H_2 \in I_2$ with $\langle H_1, H_2\rangle \in E_{\mathcal{B}}^{\mathbb{C}}$. ■

Lemma 6.6 is thereby proved. ■

This completes the proof of the soundness of the construction (Lemma 6.1). ■

The main theorem (Theorem 4.1) is thereby proven as well. ■

7 Tightness

In this section we show our analysis of $G_{\mathcal{B}}^{\mathbb{C}}$ is tight in two respects. First, we show that for any value of p there is always an independent set $\mathcal{I}$ in $G_{\mathcal{B}}^{\mathbb{C}}$ whose size is almost $p^{\bullet}$, regardless of G being a 'yes' or a 'no' instance. Next, we show that if $p > (1-p)^2$ (this happens for $p \geq \frac{3-\sqrt{5}}{2}$), then a large independent set can be formed in $G_{\mathcal{B}}^{\mathbb{C}}$, again, regardless of the size of $\mathrm{IS}(G)$.

The 2-intersecting bound. We will exhibit an appropriate choice of maximal 2-intersecting families for almost all of the blocks $\mathcal{B}$, that constitutes an independent set in $G_{\mathcal{B}}^{\mathbb{C}}$.

Let $V_{red} \cup V_{green} \cup V_{blue} \cup V_{yellow}$ be a partition of V into roughly equal sizes. For every block $B \in \mathcal{B}$, define four special block-assignments, $\mathsf{a}_{red}^B, \mathsf{a}_{green}^B, \mathsf{a}_{blue}^B, \mathsf{a}_{yellow}^B$ defined as being true on their color, and false elsewhere, e.g.

$$\forall v \in B, \quad \mathsf{a}_{red}^B(v) \stackrel{def}{=} \begin{cases} \mathsf{T} & v \in V_{red} \\ \mathsf{F} & otherwise \end{cases}$$

Of course, not all four are defined for every block, as a block-assignment $\mathsf{a} \in R_B$ must contain at least t T's, and there is a negligible fraction of the blocks $\mathcal{B}' \subset \mathcal{B}$ that intersect at least one of $V_{red} \cup V_{green} \cup V_{blue} \cup V_{yellow}$ with less than t values. Neglecting these, we take for each block, the following set of vertices

$$\mathcal{I}[B] = \left\{F \in V[B] \;\middle|\; \left|F \cap \left\{\mathsf{a}_{red}^B, \mathsf{a}_{green}^B, \mathsf{a}_{blue}^B, \mathsf{a}_{yellow}^B\right\}\right| \geq 3\right\}$$

and let $\mathcal{I} \stackrel{def}{=} \bigcup_{B \in \mathcal{B} \setminus \mathcal{B}'} \mathcal{I}[B]$.

Let $\hat{B} \in V^{(l-1)}$, and let $B_1 = \hat{B} \cup \{v_1\}$, and $B_2 = \hat{B} \cup \{v_2\}$. Assume $v_1 \in V_{red}$ (symmetrically for any other color), and observe the following,

1. $\mathsf{a}^{B_1}_{green}, \mathsf{a}^{B_1}_{blue}, \mathsf{a}^{B_1}_{yellow}$ are respectively consistent with $\mathsf{a}^{B_2}_{green}, \mathsf{a}^{B_2}_{blue}, \mathsf{a}^{B_2}_{yellow}$.
2. For any $F_1 \in \mathcal{I}[B_1]$, $\left| F_1 \cap \left\{ \mathsf{a}^{B_1}_{green}, \mathsf{a}^{B_1}_{blue}, \mathsf{a}^{B_1}_{yellow} \right\} \right| \geq 2$, and similarly for $F_2 \in \mathcal{I}[B_2]$, therefore, these vertices are consistent.

Thus, $\mathcal{I}$ is an independent set.

The bound $p < (1-p)^2$. Assume $p > \frac{3-\sqrt{5}}{2}$. We construct an independent set by selecting an arbitrary block assignment for each block, and taking all subsets containing it. By removing a negligible fraction of the vertices (subsets) in each block, we eliminate all edges between blocks.

Consider two blocks $B_1, B_2 \in \mathcal{B}$, such that $B_1 = \hat{B} \cup \{v_1\}, B_2 = \hat{B} \cup \{v_2\}$. Denote by $\hat{R}$ the set of sub-block assignments for $\hat{B}$ that are restrictions of R_{B_1} and of R_{B_2}, and assume for simplicity that every sub-block assignment in $\hat{R}$ has two extensions (to F and to T) in both R_{B_1} and R_{B_2}.

A random subset $F \in_{\mu_p} \mathcal{P}(R_{B_1})$, has expectedly $p \cdot |R_{B_1}|$ block-assignments. Moreover, there are expectedly $(1-p)^2 \cdot |\hat{R}|$ sub-block-assignments in $\hat{R}$ for which $\mathsf{a}^{(v_1 \leftarrow \mathsf{F})}, \mathsf{a}^{(v_1 \leftarrow \mathsf{T})} \notin F$, and expectedly $p \cdot |\hat{R}|$ sub-block-assignments for which $\mathsf{a}^{(v_1 \leftarrow \mathsf{F})} \in F$.

For two vertices $F_1 \in V[B_1]$ and $F_2 \in V[B_2]$ to be inconsistent, one of them must deviate from the expectation, due to the following. Every $\hat{\mathsf{a}} \in \hat{R}$ for which $\mathsf{a}^{(v_1 \leftarrow \mathsf{F})} \in F_1$ must have both $\mathsf{a}^{(v_2 \leftarrow \mathsf{F})}, \mathsf{a}^{(v_2 \leftarrow \mathsf{T})} \notin F_2$. If both F_1, F_2 are near their expectation, there are roughly $(1-p)^2 \cdot |\hat{R}|$ sub-block-assignments in $\hat{R}$ for which $\mathsf{a}^{(v_2 \leftarrow \mathsf{F})}, \mathsf{a}^{(v_2 \leftarrow \mathsf{T})} \notin F_2$, and since $(1-p)^2 < p$, this is not enough to meet the expected $p \cdot |\hat{R}|$ sub-block-assignments for which $\mathsf{a}^{(v_2 \leftarrow \mathsf{F})} \in F_1$.

Standard Chernoff bounds imply that we need to remove only a tiny fraction of the vertices of each block, so as to eliminate all subsets that deviate from the expectation according to at least one sub-block $\hat{B}$.

8 Discussion

The construction presented herein, as is, cannot be utilized to prove stronger hardness results for Vertex-Cover. Nevertheless, an amendment to the structure of the graph $G_{\mathcal{B}}$, on which the biased long-code is applied, might disqualify independent-sets consisting of 2-intersecting families of subsets in the 'no' case where $\mathrm{IS}_h(G) < \epsilon m$, and maybe even allow increasing the bias parameter so that $p = \frac{1}{2} - \epsilon$. An amendment which would allow a proof for the following conjecture:

Conjecture 8.1 *Given a graph G, it is NP-hard to distinguish between*

Yes: $\mathrm{IS}(G) \geq \frac{1}{2} - \epsilon$.

No: $\mathrm{IS}(G) \leq \epsilon$.

for any $\epsilon > 0$.

Thereby showing Vertex-Cover to be hard to approximate to within a factor even slightly smaller than 2.

Let us note that our result also implies, by direct reduction, a hardness of approximation factor of 1.3606 for the 2-CNF clause deletion problem: the problem of finding the minimum weight set of clauses in a 2-CNF formula whose deletion makes the formula satisfiable. The best approximation algorithm for this problem guarantees only a factor of $\log n \log\log n$ [KPRT97].

The framework for proving hardness results suggested herein can be tried on other problems for which the known hardness result does not match the best upper-bound. In essence one has to consider variations on $G_{\mathcal{B}}^{\mathbb{C}}$ and show that if some constraints are satisfied, a sizeable fraction of $G_{\mathcal{B}}$'s blocks must have a core, and in fact a distinguished assignment, that imply global-consistency, i.e. consistency between the blocks.

9 Acknowledgements

We would like to thank Noga Alon for his combinatorial guidance, and Gil Kalai and Ehud Freidgut for highly influential discussions. We also thank Guy Kindler and Amnon Ta-Shma for many constructive remarks on earlier versions of this paper, and Oded Goldreich for inspiring the presentation in its current more combinatorial form. In addition, we would like to thank the brave reading group who had helped us in constructive comments and to convince ourselves of the correctness of the proof: Oded Regev, Robi Krauthgamer, Vera Asodi, Oded Schwartz, Michael Langberg, Dana Moshkovich, Adi Akavia, and Elad Hazan. Thanks also to Sanjeev Arora for pointing out to us the application to the 2CNF deletion problem.

A Weighted vs Unweighted

Given a graph $G = (V, E, \Lambda)$, we construct, for any precision parameter $\varrho > 0$, an unweighted graph $G_\varrho = (V_\varrho, E_\varrho)$ with $\left|\frac{\overline{\mathrm{IS}}(G_\varrho)}{|V_\varrho|} - \overline{\mathrm{IS}}(G)\right| \leq \varrho$, and whose size is polynomial in $|G|$ and $\frac{1}{\varrho}$.

Let $n = |V| \cdot \frac{1}{\varrho}$. We replace each $v \in V$ with $n_v = \lceil n \cdot \Lambda(v) \rfloor$ copies ($\lceil x \rfloor$ denotes the integer nearest x), and set

$$
\begin{aligned}
V_\varrho &\stackrel{def}{=} \{\langle v, i\rangle \mid v \in V,\ 1 \leq i \leq n_v\} \\
E_\varrho &\stackrel{def}{=} \{\{\langle v_1, i_1\rangle, \langle v_2, i_2\rangle\} \mid \{v_1, v_2\} \in E,\ i_1 \in [n_{v_1}],\ i_2 \in [n_{v_2}]\}
\end{aligned}
$$

If $C \subseteq V$ is a vertex cover for G, then $C_\varrho = \bigcup_{v \in C} \{v\} \times [n_v]$ is a vertex cover for G_ϱ. Moreover, every minimal vertex cover $C_\varrho \subseteq V_\varrho$ is of this form. Thus we show $\left|\frac{\overline{\mathrm{IS}}(G_\varrho)}{|V_\varrho|} - \overline{\mathrm{IS}}(G)\right| \leq \varrho$ by the following proposition,

Proposition A.1 *Let $C \subseteq V$, and let $C_\varrho = \bigcup_{v \in C} \{v\} \times [n_v]$. Then $\left|\frac{|C_\varrho|}{|V_\varrho|} - \Lambda(C)\right| \leq \varrho$.*

Proof: For every C, C_ϱ as above,

$$
|C_\varrho| = \sum_{v \in C} n_v = \sum_{v \in C} \lceil n \cdot \Lambda(v) \rfloor = n \cdot \Lambda(C) + \sum_{v \in C} (\lceil n \cdot \Lambda(v) \rfloor - n \cdot \Lambda(v)).
$$

For any v, $|\lceil v \rfloor - v| \leq \frac{1}{2}$, and so

$$
\left|\frac{|C_\varrho|}{n} - \Lambda(C)\right| \leq \frac{1}{2} \frac{|C|}{n} \leq \frac{\varrho}{2} \tag{3}
$$

To complete our proof we need to replace $\frac{|C_\varrho|}{n}$ by $\frac{|C_\varrho|}{|V_\varrho|}$ in (3). Indeed, taking $C = V$ in (3), yields $\left|\frac{|V_\varrho|}{n} - 1\right| \leq \frac{\varrho}{2}$, and multiplying by $\frac{|C_\varrho|}{|V_\varrho|} \leq 1$, we obtain $\left|\frac{|C_\varrho|}{n} - \frac{|C_\varrho|}{|V_\varrho|}\right| \leq \frac{\varrho}{2}$. ■

B Proof of Theorem 3.1

In this section we prove Theorem 3.1 which encapsulates our use of the PCP theorem. PCP characterizations of NP in general state that given some SAT instance, namely, a set of Boolean-functions $\Phi = \{\varphi_1, ..., \varphi_n\}$ over variables W, it is NP-hard to distinguish between 'yes' instances where there is an assignment A to Φ's variables that satisfies all Φ, and 'no' instances where any assignment to A satisfies at most a small fraction of Φ.

Definition B.1 *Let us denote by $\Upsilon(\Phi)$ the maximum, over all assignments to Φ's variables $A : W \to \{0, 1\}$, of the fraction of $\varphi \in \Phi$ satisfied by A, namely*

$$
\Upsilon(\Phi) = \max_A \Pr_{\varphi \in \Phi} [\varphi \textit{ is satisfied by } A]
$$

The basic PCP theorem showing hardness for gap-SAT states that,

Theorem B.1 ([AS92, ALM+92]) *There exists some constant $\beta > 0$ such that given a set $\Phi = \{\varphi_1, .., \varphi_n\}$ of 3-CNF clauses over Boolean variables W (each clause is the OR of exactly 3 variables), it is NP-hard to distinguish between the two cases:*

Yes: Φ *is satisfiable* ($\Upsilon(\Phi) = 1$).

No: $\Upsilon(\Phi) < 1 - \beta$.

■

Let us now turn to the proof of Theorem 3.1.

Theorem 3.1 *For any* $h, \epsilon > 0$*, the problem* $h\text{IS}(r, \epsilon, h)$ *is NP-hard, as long as* $r = (\frac{h}{\epsilon})^c$ *for some constant* c*.*

Proof: Our proof proceeds by reduction from the PCP theorem. Let Φ be as above, and define the parallel repetition version of Φ,

Definition B.2 ($\mathsf{Par}[\Phi, k]$) *Let* $\langle \Phi, W \rangle$ *be a 3-CNF instance, with 3-CNF clauses* Φ *over variables* W*. For any integer* $k > 0$*, let*

$$\mathsf{Par}[\Phi, k] \stackrel{def}{=} \langle \Psi, X, Y \rangle$$

be a SAT instance with Boolean functions Ψ *over two types of variables:* $X \stackrel{def}{=} \Phi^k$ *and* $Y \stackrel{def}{=} W^k$*.*

The range of each variable $x \in X$*, is* $R_X = [7]^k$*, corresponding (by enumerating the 7 satisfying assignments of each 3-CNF clause* $\varphi \in \Phi$*) to the concatenation of the satisfying assignments for* Φ*'s clauses in* x*. The range of each variable* $y \in Y$*, is* $R_Y = [2]^k$*, corresponding to all possible assignments to* W*'s variables in* y*.*

For $y = (w_1, .., w_k)$ *and* $x = (\varphi_1, .., \varphi_k)$*, denote* $y \sqsubseteq x$ *if for all* $i \in [k]$*,* w_i *is a variable in* φ_i*. The Boolean-functions in* Φ *are as follows:*

$$\Psi = \left\{ \psi_{x \to y} \;\middle|\; y \in W^k,\ x \in \Phi^k,\ y \sqsubseteq x \right\}$$

where $\psi_{x \to y}$ *is* T *if the assignment to* y *is the restriction to* y *of the assignment to* x*, and* F *otherwise. Since each test* $\varphi \in \Phi$ *has exactly* 3 *variables, each variable* $x \in X$ *appears in exactly* 3^k *tests in* $\psi_{x \to y} \in \Psi$*.*

Clearly, if $\Upsilon(\Phi) = 1$, then $\Upsilon(\Psi) = 1$. For the converse,

Theorem B.2 (Parallel Repetition, [Raz98]) *There exists some constant* $c > 0$*, such that the following holds. Let* $\langle \Phi, W \rangle$ *be a 3-CNF-instance, and let* $\langle \Psi, X, Y \rangle = \mathsf{Par}[\Phi, k]$*,*

$$\Upsilon(\Psi) \leq \Upsilon(\Phi)^{c \cdot k}.$$

■

Therefore, one may choose k for which $(1 - \beta)^{c \cdot k} \leq \epsilon / h^3$ and $|R_Y|, |R_X| \leq (\frac{\epsilon}{h})^{-O(1)}$, hence it is NP-hard to distinguish whether $\Upsilon(\Psi) = 1$ or $\Upsilon(\Psi) < \epsilon / h^3$.

The FGLSS Construction. We next apply the FGLSS [FGL+91, Kar72] construction to Ψ. Let $\mathcal{G}[\Psi]$ be the (m, r)-co-partite graph, with $m = |X|$ and $r = |R_X|$,

$$\mathcal{G}[\Psi] = \langle V, E \rangle \text{ where } V \stackrel{def}{=} (X \times R_X)$$

that is, where $\mathcal{G}[\Psi]$'s vertices is the set of pairs consisting of a variable x in X and a value $a \in R_X$ for x. For the edge set E of $\mathcal{G}[\Psi]$, let us consider all pairs of vertices whose values cannot possibly correspond to the same satisfying assignment:

$$E = \{\, \{(x_1,a_1),(x_2,a_2)\} \mid \exists y, \psi_{x_1\to y}, \psi_{x_2\to y} \in \Phi,\ \psi_{x_1\to y}(a_1) \neq \psi_{x_2\to y}(a_2)\}$$

Therefore, an independent-set in $\mathcal{G}[\Psi]$ cannot correspond to an inconsistent assignment to Φ.

If Ψ is satisfiable, let $A : X \cup Y \to \{\mathsf{T},\mathsf{F}\}$ be a satisfying assignment for it, and observe that the set $\{\,(x, A(x)) \mid x \in X\} \subset V$ is an independent-set of size $|X| = m$.

Otherwise, let us assume that $\mathcal{I} \subset V$ contains no clique of size h, and that $|\mathcal{I}| > \epsilon \cdot m$, and show that $\Upsilon(\Psi) > \frac{\epsilon}{h^3}$. Let $A_{\mathcal{I}}$ map to each variable a subset of its range, as follows. For every $x \in X$ and $y \in Y$, set

$$\begin{aligned} A_{\mathcal{I}}(x) &\overset{def}{=} \{a \in R_X \mid (x,a) \in \mathcal{I}\} \ \subset R_X \\ A_{\mathcal{I}}(y) &\overset{def}{=} \bigcup_{\psi_{x\to y}\in\Psi} \psi_{x\to y}(A_{\mathcal{I}}(x)) \ \subset R_Y \end{aligned}$$

By the definition of $A_{\mathcal{I}}$, for every x with $A_{\mathcal{I}}(x) \neq \phi$ and for *every* $\psi_{x\to y} \in \Psi$,

$$\psi_{x\to y}(A_{\mathcal{I}}(x)) \cap A_{\mathcal{I}}(y) \neq \phi\,.$$

Denote $X_0 = \{\, x \in X \mid A_{\mathcal{I}}(x) \neq \phi\}$ and observe that since there is an equal number of $\psi_{x\to y} \in \Psi$ for each variable x:

$$\Pr_{\psi_{x\to y}\in\Psi}[\psi_{x\to y}(A_{\mathcal{I}}(x)) \cap A_{\mathcal{I}}(y) \neq \phi] = \Pr_{x\in X}[x \in X_0] = \frac{|X_0|}{|X|} > \frac{1}{h}\cdot\frac{|\mathcal{I}|}{|X|} > \epsilon/h\,.$$

Finally, by picking for each variable $x \in X, y \in Y$ a random assignment

$$\forall x \in X, y \in Y, \qquad a_x \in_R A_{\mathcal{I}}(x), \quad a_y \in_R A_{\mathcal{I}}(y)$$

If $A_{\mathcal{I}}(x) \neq \phi$, the probability that $\psi_{x\to y} \in \Psi$ is satisfied by such a random assignment is at least $\frac{1}{|A_{\mathcal{I}}(x)|}\cdot\frac{1}{|A_{\mathcal{I}}(y)|} > 1/h^2$. Thus the expected number of Boolean functions satisfied by this random assignment is $> \frac{\epsilon}{h^3}\cdot|\Psi|$. Since at least one assignment must meet the expectation, $\Upsilon(\Psi) > \frac{\epsilon}{h^3}$. ■

C Some Propositions about μ_p

Proposition 5.3 *For a monotone family of subsets $\mathcal{F} \subseteq \mathcal{P}(n)$, $q > p \Rightarrow \mu_q(\mathcal{F}) \geq \mu_p(\mathcal{F})$.*

Proof: For a subset $F \in \mathcal{P}([n])$ denote

$$F_{\leq i} \overset{def}{=} F \cap [1,i] \quad \text{and} \quad F_{>i} \overset{def}{=} F \cap [i+1,n]$$

and consider, for $0 \leq i \leq n$, the hybrid distribution, where the first i elements are chosen with bias p and the others are chosen with bias q

$$\mu_{p,i,q}(F) \overset{def}{=} p^{|F_{\leq i}|}\cdot(1-p)^{i-|F_{\leq i}|}\cdot q^{|F_{>i}|}\cdot(1-q)^{n-i-|F_{>i}|}$$

Observe that

$$\forall 0 \leq i \leq n \quad \mu_{p,i,q}(\mathcal{F}) \geq \mu_{p,i+1,q}(\mathcal{F})$$

therefore $\mu_q(\mathcal{F}) = \mu_{p,0,q}(\mathcal{F}) \geq \mu_{p,n,q} = \mu_p(\mathcal{F})$. ■

Theorem 5.4 [Russo-Margulis Identity] *Let $\mathcal{F} \subseteq \mathcal{P}(R)$ be a monotone family. Then,*

$$\frac{d\mu_q(\mathcal{F})}{dq} = \mathsf{as}_q(\mathcal{F})$$

Proof: For a subset $F \in \mathcal{P}(n)$ write

$$\mu_q(F) = \prod_{i\in[n]} \mu_q^i(F), \qquad \text{for} \quad \mu_q^i(F) = \begin{cases} q & i \in F \\ \\ 1-q & i \notin F \end{cases} \tag{4}$$

Observe that

$$\mathsf{influence}_q^i(\mathcal{F}) = \sum_{F\in\mathcal{F}} \left(\frac{d\mu_q^i(F)}{dq} \cdot \prod_{j\neq i} \mu_q^j(F) \right)$$

Differentiating (4) according to q, and summing over all $F \in \mathcal{F}$, we get

$$\frac{d\mu_q(\mathcal{F})}{dq} = \sum_{i\in[n]} \mathsf{influence}_q^i(\mathcal{F}) = \mathsf{as}_q(\mathcal{F})\,.$$

■

We next show that for any monotone family $\mathcal{F} \subset \mathcal{P}(R)$, if $U \subset R$ is a set of elements of tiny influence, one has to remove only a small fraction of $\mathcal{F}$ to make it completely independent of U:

Proposition 6.13 *Let $\mathcal{F} \subset \mathcal{P}(R)$ be monotone, and let $U \subset R$ be such that for all $e \in U$, $\mathsf{influence}_p^e(\mathcal{F}) < \eta$. Let*

$$\mathcal{F}' = \{F \in \mathcal{F} \mid F \setminus U \in \mathcal{F}\}$$

then,

$$\mu_p^R\left(\mathcal{F} \setminus \mathcal{F}'\right) < |U| \cdot \eta \cdot p^{-|U|}$$

Proof: Let

$$\mathcal{F}'' = \{F \in \mathcal{P}(R \setminus U) \;\mid\; F \cup U \in \mathcal{F} \text{ but } F \notin \mathcal{F}\}\,.$$

A set $F \in \mathcal{F}''$ contributes at least $\mu_p^{R\setminus U}(F) \cdot p^{|U|}$ to the influence of at least one element $e \in U$. Since the sum of influences of elements in U is $< |U| \cdot \eta$, we have $\mu_p^{R\setminus U}(\mathcal{F}'') < |U| \cdot \eta \cdot p^{-|U|}$. The proof is complete noting that,

$$\mathcal{F} \setminus \mathcal{F}' \;\subseteq\; \mathcal{F}'' \sqcup \mathcal{P}(U)$$

■

D Erdős-Ko-Rado

In this section we prove a lemma that is a continuous version and follows directly from the complete intersection theorem of Ahlswede and Khachatrian [AK97].

Let us define

$$\mathcal{A}_i \stackrel{def}{=} \{F \in \mathcal{P}([n]) \mid F \cap [1, 2+2i] \geq 2+i\}$$

and prove the following corollary,

Corollary 5.9 *Let $\mathcal{F} \subset \mathcal{P}([n])$ be 2-intersecting. For any $p < \frac{1}{2}$,*

$$\mu_p(\mathcal{F}) \leq \max_i \{\mu_p(\mathcal{A}_i)\}$$

Proof: Denote $\mu = \max_i(\mu_p(\mathcal{A}_i))$. Assuming $\mathcal{F}_0 \subset \mathcal{P}([n_0])$ contradicts the claim, let $a = \mu_p(\mathcal{F}_0) - \mu > 0$. Now consider $\mathcal{F} = \mathcal{F}_0 \sqcup \mathcal{P}([n] \setminus [n_0])$ for $n > n_0$ large enough, to be determined later. Clearly, for any $n \geq n_0$, $\mu_p^{[n]}(\mathcal{F}) = \mu_p^{[n_0]}(\mathcal{F}_0)$, and $\mathcal{F}$ is 2-intersecting. Consider, for $\theta < \frac{1}{2} - p$ to be determined later,

$$S \stackrel{def}{=} \{k \in \mathbb{N} \mid |k - p \cdot n| \leq \theta \cdot n\}$$

and for every $k \in S$, denote by $\mathcal{F}_k = \mathcal{F} \cap \binom{[n]}{k}$. We will show that since most of $\mathcal{F}$'s weight is derived from $\cup_{k \in S} \mathcal{F}_k$, there must be at least one $\mathcal{F}_k$ that contradicts Theorem 5.8. Indeed,

$$\mu + a = \mu_p(\mathcal{F}) = \sum_{k \in S} p^k (1-p)^{n-k} \cdot |\mathcal{F}_k| + o(1)$$

Hence there exists $k \in S$ for which $\frac{|\mathcal{F}_k|}{\binom{[n]}{k}} \geq \mu + \frac{1}{2}a$. We have left to see that $\mu \cdot \binom{n}{k}$ is close enough to $\max_i(|\mathcal{A}_i \cap \binom{[n]}{k}|)$. This follows from usual tail bounds, and is sketched as follows. Subsets in $\binom{[n]}{k}$ for large enough i (depending only on $\frac{k}{n}$ but not on k or n), have roughly $\frac{k}{n} \cdot (2i+2)$ elements in the set $[1, 2i+2]$. Moreover, the subsets in $\mathcal{A}_i$ have at least $i+2$ elements in $[1, 2i+2]$, thus are very few (compared to $\binom{n}{k}$), because $\frac{i+2}{2i+2} > \frac{1}{2} > p + \theta \geq \frac{k}{n}$. In other words, there exists some constant $C_{p+\theta,\mu}$, for which $\left|\mathcal{A}_i \cap \binom{[n]}{k}\right| < \mu \cdot \binom{n}{k}$ for all $i \geq C_{p,\mu}$ as long as $\frac{k}{n} \leq p + \theta$.

Additionally, for every $i < C_{p,\mu}$, taking n to be large enough we have

$$\forall k \in S, \qquad \frac{\left|\mathcal{A}_i \cap \binom{[n]}{k}\right|}{\binom{n}{k}} = \mu_{\frac{k}{n}}(\mathcal{A}_i) + o(1) = \mu_p(\mathcal{A}_i) + o(1) < \mu + o(1)$$

where the first equality follows from a straightforward computation. ■

We have the following corollary,

Corollary 5.10 *Let $\mathcal{F} \subset \mathcal{P}(R)$ be 2-intersecting. For any $q < p_{\max}$, $\mu_q(\mathcal{F}) \leq q^\bullet$.*

Proof: Define a sequence $p_0 < p_1 < \ldots$, by $p_i \stackrel{def}{=} \frac{i}{2i+1}$. We will show that these are the points where the maximum switches from $\mathcal{A}_i$ to $\mathcal{A}_{i+1}$. More accurately, we will show for all $i \geq 0$,

$$\forall p \in (p_i, p_{i+1}] \quad \max_j \{\mu_p(\mathcal{A}_j)\} = \mu_p(\mathcal{A}_i) \tag{5}$$

This, together with Corollary 5.9, implies the corollary, as $p < p_{\max} < 0.4 = p_2$ implies $\mu_p(\mathcal{F}) \leq \max(\mu_p(\mathcal{A}_0), \mu_p(\mathcal{A}_1)) = \max(p^2, 4p^3 - 3p^4) = p^\bullet$.

So we proceed to prove (5). A subset $F \notin \mathcal{A}_i$ must intersect $[1, 2i+2]$ on at most $i+1$ elements. If additionally $F \in \mathcal{A}_{i+1}$ it must then contain $2i+3, 2i+4$. Thus,

$$\mu_p(\mathcal{A}_{i+1} \setminus \mathcal{A}_i) = \binom{2i+2}{i+1} \cdot p^{i+1}(1-p)^{i+1} \cdot p^2$$

Similarly,

$$\mu_p(\mathcal{A}_i \setminus \mathcal{A}_{i+1}) = \binom{2i+2}{i+2} \cdot p^{i+2}(1-p)^i \cdot (1-p)^2$$

Together,

$$\begin{aligned} \mu_p(\mathcal{A}_{i+1}) - \mu_p(\mathcal{A}_i) &= \mu_p(\mathcal{A}_{i+1} \setminus \mathcal{A}_i) - \mu_p(\mathcal{A}_i \setminus \mathcal{A}_{i+1}) \\ &= p^{i+2}(1-p)^{i+1}\binom{2i+2}{i+1}\left(p - (1-p)\frac{i+1}{i+2}\right) \end{aligned}$$

The sign of this difference is determined by $p-(1-p)\frac{i+1}{i+2}$. For a fixed $i \geq 0$, this expression goes from positive to negative passing through zero once at $p = \frac{i+1}{2i+3} = p_{i+1}$. Thus, the sequence $\{\mu_p(\mathcal{A}_j)\}_j$ is maximized at i for $p_i < p \leq p_{i+1}$. (It is increasing when $i \leq \frac{1-3p}{2p-1}$, and decreasing thereafter). ■

E A Chernoff Bound

Proposition E.1 *For any set $I \subset V$ such that $|I| = \frac{1}{r} \cdot |V|$,*

$$\Pr_{B \in \mathcal{B}}[|I \cap B| < l_{\mathsf{T}}] < 2e^{-\frac{2l_{\mathsf{T}}}{8}}$$

Proof: Consider the random variable $\chi_I : V \to \{0,1\}$ taking a 1 iff $v \in I$. We have $\Pr_{v \in V}[\chi_I(v) = 1] = \frac{1}{r}$, and for every $B \in \mathcal{B} = \binom{V}{l}$, $|I \cap B| = \sum_{v \in B} \chi_I(v)$, so the expectation of this is $|B| \cdot \frac{1}{r} = 2l_{\mathsf{T}}$. The standard Chernoff bound directly gives

$$\Pr_{v_1,..,v_l \in V}\left[\sum_{i \in [l]} \chi_I(v_i) < l_{\mathsf{T}} = \frac{1}{2} \cdot l/r\right] < e^{-\frac{l}{8r}}$$

We are almost done, except that the above probability was taken with repetitions, while in our case, for $v_1, \ldots, v_l$ to constitute a block $B \in \mathcal{B}$, they must be l distinct values. In fact, this happens with overwhelming probability and in particular $\geq \frac{1}{2}$, thus we write,

$$\begin{aligned} \Pr_{v_1,\ldots,v_l \in V}\left[\sum_i \chi_I(v_i) < l_{\mathsf{T}} \;\middle|\; |\{v_1, \ldots, v_l\}| = l\right] &\leq \frac{\Pr_{v_1,\ldots,v_l \in V}\left[\sum_i \chi_I(v_i) < l_{\mathsf{T}}\right]}{\Pr_{v_1,\ldots,v_l \in V}\left[|\{v_1, \ldots, v_l\}| = l\right]} \\ &\leq \frac{e^{-\frac{l}{8r}}}{\frac{1}{2}} = 2e^{-\frac{l}{8r}} \end{aligned}$$

■

References

[AK97] Rudolf Ahlswede and Levon H. Khachatrian. The complete intersection theorem for systems of finite sets. *European J. Combin.*, 18(2):125–136, 1997.

[ALM+92] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and intractability of approximation problems. In *Proc. 33rd IEEE Symp. on Foundations of Computer Science*, pages 13–22, 1992.

[AS92] S. Arora and S. Safra. Probabilistic checking of proofs: A new characterization of NP. In *Proc. 33rd IEEE Symp. on Foundations of Computer Science*, pages 2–13, 1992.

[BGS98] Mihir Bellare, Oded Goldreich, and Madhu Sudan. Free bits, PCPs, and nonapproximability—towards tight results. *SIAM Journal on Computing*, 27(3):804–915, June 1998.

[BK97] J. Bourgain and G. Kalai. Influences of variables and threshold intervals under group symmetries. *Geom. Funct. Anal.*, 7(3):438–461, 1997.

[BKS99] Itai Benjamini, Gil Kalai, and Oded Schramm. Noise sensitivity of Boolean functions and applications to percolation. *Inst. Hautes Études Sci. Publ. Math.*, (90):5–43 (2001), 1999.

[BL89] M. Ben-Or and N. Linial. Collective coin flipping. *ADVCR: Advances in Computing Research*, 5, 1989.

[BLR93] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. In *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing (Baltimore, MD, 1990)*, volume 47, pages 549–595, 1993.

[BLS88] M. Ben-Or, N. Linial, and M. Saks. Collective coin flipping and other models of imperfect randomness. In *Combinatorics (Eger, 1987)*, pages 75–112. North-Holland, Amsterdam, 1988.

[BYE85] R. Bar-Yehuda and S. Even. A local-ratio theorem for approximating the weighted vertex cover problem. *Annals of Discrete Mathematics*, 25:27–45, 1985.

[EH00] Lars Engebretsen and Jonas Holmerin. Clique is hard to approximate within $n^{1-o(1)}$. In *Automata, languages and programming (Geneva, 2000)*, pages 2–12. Springer, Berlin, 2000.

[EKR61] P. Erdős, Chao Ko, and R. Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 12:313–320, 1961.

[ER60] P. Erdős and R. Rado. Intersection theorems for systems of sets. *J. London Math. Soc.*, 35:85–90, 1960.

[FGL+91] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy. Approximating clique is almost NP-complete. In *Proc. 32nd IEEE Symp. on Foundations of Computer Science*, pages 2–12, 1991.

[FK96] Ehud Friedgut and Gil Kalai. Every monotone graph property has a sharp threshold. *Proc. Amer. Math. Soc.*, 124(10):2993–3002, 1996.

[Fra78] P. Frankl. The Erdős-Ko-Rado theorem is true for $n = ckt$. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. I*, pages 365–375. North-Holland, Amsterdam, 1978.

[Fri98] Ehud Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1), pages 27–35, 1998.

[Hal00] Eran Halperin. Improved approximation algorithms for the vertex cover problem in graphs and hypergraphs. In *Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 329–337, N.Y., January 9–11 2000. ACM Press.

[Hås97] Johan Håstad. Some optimal inapproximability results. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 1–10, El Paso, Texas, 4–6 May 1997.

[Hås99] Johan Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Math.*, 182(1):105–142, 1999.

[Kar72] R. M. Karp. Reducibility among combinatorial problems. In Miller and Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, 1972.

[Kho01] S. Khot. Improved inapproximability results for Max-Clique, Chromatic Number and Approximate Graph Coloring. In *Proc. 42nd IEEE Symp. on Foundations of Computer Science*, 2001.

[KKL88] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. In IEEE, editor, *29th annual Symposium on Foundations of Computer Science, October 24–26, 1988, White Plains, New York*, pages 68–80. IEEE Computer Society Press, 1988.

[KPRT97] Philip N. Klein, Serge A. Plotkin, Satish Rao, and Eva Tardos. Approximation algorithms for steiner and directed multicuts. *J. Algorithms*, 22(2):241–269, 1997.

[Mar74] G. A. Margulis. Probabilistic characteristics of graphs with large connectivity. *Problemy Peredači Informacii*, 10(2):101–108, 1974.

[MS83] B. Monien and E. Speckenmeyer. Some further approximation algorithms for the vertex cover problem. In G. Ausiello and M. Protasi, editors, *Proceedings of the 8th Colloquium on Trees in Algebra and Programming (CAAP'83)*, volume 159 of *LNCS*, pages 341–349, L'Aquila, Italy, March 1983. Springer.

[PY91] C. Papadimitriou and M. Yannakakis. Optimization, approximation and complexity classes. *Journal of Computer and System Sciences*, 43:425–440, 1991.

[Raz98] Ran Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(3):763–803, June 1998.

[RS97] R. Raz and S. Safra. A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proc. 29th ACM Symp. on Theory of Computing*, pages 475–484, 1997.

[Rus82] Lucio Russo. An approximate zero-one law. *Z. Wahrsch. Verw. Gebiete*, 61(1):129–139, 1982.

[Tre01] L. Trevisan. Non-approximability results for optimization problems on bounded degree instances. In *Proc. 33rd ACM Symp. on Theory of Computing*, 2001.

[W84] Richard M. Wilson. The exact bound in the Erdős-Ko-Rado theorem. *Combinatorica*, 4(2-3), pages 247–257, 1984.