$(1+\Omega(1))$-Approximation to MAX-CUT Requires Linear Space

Michael Kapralov* Sanjeev Khanna† Madhu Sudan‡ Ameya Velingker§

November 1, 2016

Abstract

We consider the problem of estimating the value of MAX-CUT in a graph in the streaming model of computation. We show that there exists a constant $\epsilon_* > 0$ such that any randomized streaming algorithm that computes a $(1+\epsilon_*)$-approximation to MAX-CUT requires $\Omega(n)$ space on an n vertex graph. By contrast, there are algorithms that produce a $(1+\epsilon)$-approximation in space $O(n/\epsilon^2)$ for every $\epsilon > 0$. Our result is the first linear space lower bound for the task of approximating the max cut value and partially answers an open question from the literature [Berb]. The prior state of the art ruled out $(2-\epsilon)$-approximation in $\tilde{O}(\sqrt{n})$ space or $(1+\epsilon)$-approximation in $n^{1-O(\epsilon)}$ space, for any $\epsilon > 0$.

Previous lower bounds for the MAX-CUT problem relied, in essence, on a lower bound on the communication complexity of the following task: Several players are each given some edges of a graph and they wish to determine if the union of these edges is ϵ-close to forming a bipartite graph, using one-way communication. The previous works proved a lower bound of $\Omega(\sqrt{n})$ for this task when $\epsilon = 1/2$, and $n^{1-O(\epsilon)}$ for every $\epsilon > 0$, even when one of the players is given a candidate bipartition of the graph and the graph is promised to be bipartite with respect to this partition or ϵ-far from bipartite. This added information was essential in enabling the previous analyses but also yields a weak bound since, with this extra information, there is an $n^{1-O(\epsilon)}$ communication protocol for this problem. In this work, we give an $\Omega(n)$ lower bound on the communication complexity of the original problem (without the extra information) for $\epsilon = \Omega(1)$ in the three-player setting. Obtaining this $\Omega(n)$ lower bound on the communication complexity is the main technical result in this paper. We achieve it by a delicate choice of distributions on instances as well as a novel use of the convolution theorem from Fourier analysis combined with graph-theoretic considerations to analyze the communication complexity.

*School of Computer and Communication Sciences, EPFL, Lausanne, Switzerland. Email: michael.kapralov@epfl.ch

†Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA 19104. Email: sanjeev@cis.upenn.edu. Supported in part by National Science Foundation grants CCF-1116961, CCF-1552909, CCF-1617851, and IIS-1447470.

‡Harvard John A. Paulson School of Engineering and Applied Sciences, 33 Oxford Street, Cambridge, MA 02138, USA. Email: madhu@cs.harvard.edu. Supported in part by NSF Award CCF 1565641.

§Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213. Email: avelingk@cs.cmu.edu. Supported in part by National Science Foundation grant CCF-0963975.

1 Introduction

In this paper, we consider the space complexity of approximating MAX-CUT in the streaming model of computation. We elaborate on these terms and describe our main result below.

The input to the MAX-CUT problem is an undirected graph, and the goal is to find a bipartition of the vertices of this graph (or a *cut*) that maximizes the number of edges that cross the bipartition. The size of a MAX-CUT on graph G, denoted MAX-CUT(G), is the number of edges that cross the optimal bipartition. An algorithm A is said to produce an α-approximation to the size of the MAX-CUT if for every graph G, the algorithm's output $A(G)$ satisfies MAX-CUT$(G)/\alpha \leq A(G) \leq$ MAX-CUT(G).

In this paper, we study the space complexity of approximating MAX-CUT in the streaming model of computation. The streaming model of computation, formally introduced in the seminal work of [AMS96] and motivated by applications in processing massive datasets, is an extremely well-studied model for designing sublinear space algorithms. For the MAX-CUT problem in this model, the edges of the input graph G are presented as a stream to a (randomized) algorithm, which must output an α-approximation to MAX-CUT(G). The complexity measure is the space complexity, namely, the number of bits of memory used by the streaming algorithm, measured as a function of n, the number of vertices in G.

Our main result is a strong lower bound (optimal to within polylogarithmic factors) on the space required for a strong approximation to the MAX-CUT size. Specifically, we show that there is an $\alpha > 1$ such that every α-approximation algorithm in the streaming model must use $\Omega(n)$ space (see Theorem **??**).

Context and Significance. There are two basic algorithmic results for MAX-CUT in the streaming model: On the one hand, the trivial algorithm that counts the number, say m, of edges in G and outputs $m/2$ is a 2-approximation that uses $O(\log n)$ space. On the other hand, if one has $\tilde{O}(n)$ space[1], one can get an *approximation scheme*, i.e., a $(1+\epsilon)$-approximation algorithm for every $\epsilon > 0$, by building a "cut-sparsifier" [BK96, SS08].

Given just the two algorithms above, it is possible to envision three possible scenarios for improving the approximability of MAX-CUT: (1) Perhaps MAX-CUT has an approximation scheme in polylogarithmic space? (2) Perhaps MAX-CUT admits a space-approximation tradeoff, i.e., for every $\alpha > 1$, there is a $\beta < 1$ such that an α-approximation can be computed in n^β space? (3) Perhaps there is an $\alpha < 2$ and an algorithm using n^β space for some $\beta < 1$ that can compute an α-approximation to MAX-CUT. (Note that the scenarios are nested with (1) $\Rightarrow$ (2) $\Rightarrow$ (3).)

Previous works [KK15, KKS15] have ruled out scenario (1) above, making progress on an open question from [Bera]. In particular, these works have showed that for every $\beta < 1$, there exists $\alpha > 1$ such that a streaming algorithm with space n^β cannot compute an α-approximation to MAX-CUT. The work of [KKS15] also shows that $\beta < 1/2$ and $\alpha < 2$ are not simutaneously achievable. These results still allow for either scenario (2) or (3). Our result achieves the next level of understanding by ruling out scenario (2) as well:

Theorem 1.1 (Main result) *There exists $\epsilon^* > 0$ such that every randomized single-pass streaming algorithm that yields a $(1+\epsilon^*)$-approximation to the MAX-CUT size with probability at least $9/10$ must use $\Omega(n)$ space, where n denotes the number of vertices in the input graph.*

This step has also been suggested as an open problem in the Bertinoro workshop [Berb], though we settle their question only partially since the question suggests a particular approach to proving the lower bound, which we do not follow. Eventually we suspect that even scenario (3) is not achievable, but ruling this out involves more technical challenges. Indeed, one of the hopes of this work is to introduce some techniques that may be useful in the eventual resolution of this problem.

[1] Throughout this paper we use the notation $\tilde{O}(f(n))$ to denote the set $\cup_{c>0} O(f(n)(\log(f(n))^c))$.

Techniques. As with most lower bounds in streaming, ours is obtained by a reduction from a communication complexity problem. However, the communication problem and even communication model in this paper are somewhat new, so we describe our model and then explain why the novelty is necessary and useful.

Roughly, our paper considers a T-player sequential communication game, that we call the **Implicit Hidden Partition Problem**, where player P_i, for $1 \leq i \leq T$, is given a set of edges E_i on vertex set $[n]$, and the players wish to determine whether $\cup_i E_i$ forms a bipartite graph or is ϵ-far from being bipartite. (To be more precise, in our actual game the players also get some "non-edges" F_i and they also need to verify that (most of) the edges of F_i do not cross the bipartition, but we ignore this distinction here since it is not conceptually significant.) The communication is one-way and player P_i is only allowed to broadcast a message based on its own input and broadcast messages from players P_j for $1 \leq j < i$. We show that for $T = 3$ and some $\epsilon > 0$, there is a distribution on inputs for which this task requires $\Omega(n)$ communication.

The communication problems from previous works included an additional player P_0 whose input was a bipartition of the vertices of the graph, and later players needed to verify that the graph was bipartite with respect to this bipartition. The presence of this additional player was essential to previous analyses. These analyses roughly suggested that when the input graph is far from being bipartite, conditioned on not discovering a violating edge, the information of the first i players is effectively dominated by the information of P_0 — i.e., knowledge of the partition subsumes all other knowledge. This suggests a reduction from the T (or $T + 1$) player communication problem to several two-player games involving player P_0 and P_i for $1 \leq i \leq T$, and this two player game can be analyzed as in [GKK$^+$08, VY11]. Implementing this reduction does take technical work, but the intuition works!

For our purposes, the presence of the 0-th player poses an insurmountable obstacle—with this player, there is a $O(\sqrt{n} \cdot \text{poly}(1/\epsilon))$ communication protocol (based on the "birthday paradox") to distinguish bipartite graphs that are ϵ-far from being bipartite! Indeed, one can just send information about the classification of about $\sqrt{n}$ vertices with respect to the bipartition and check how many edges violate the bipartition. Harder communication complexity problems (e.g. the Boolean Hidden Hypermatching Problem of [VY11]—see [KK15, KKS15]) have been considered, leading to stronger $n^{1-O(\epsilon)}$ lower bounds on testing ϵ-closeness to bipartite, but they still use an explicit candidate bipartition and admit $n^{1-O(\epsilon)}$ protocols for any constant ϵ. This forces us to remove the 0-th player, thereby leading to the (in retrospect, more natural) "Implicit Hidden Partition" problem that we introduce explicitly in this paper.

The removal of the 0-th player, however, forces us to introduce new mechanisms to cope with the leakage of information as the protocol evolves. We do so by changing the communication model to allow for some "public inputs" and some "private inputs". All inputs to player i are selected after the transmission of the message of player $i - 1$, and the public input becomes known to all players while the private input is known only to player i. (In our case, the public input is a superset of the edges E_i and the private input is the set E_i.) This separation brings back a little flexibility into our analysis, but the task of bounding the flow of relevant information as the protocol evolves remains challenging and, indeed, we are only able to carry out such an analysis for $T = 3$, by a careful choice of input distributions and paramters.

One major challenge is the task of finding the right set of hard instances for the problem. Natural candidates (for example the one suggested in [Berb]) would involve random bipartite graphs and random graphs; however, the presence of vertices of degree larger than 2 in these graphs poses obstacles to our analysis. So we pick a delicate distribution in which the graph formed by $E_1 \cup E_2$ has no cycles and no vertices of degree > 2 (so $E_1 \cup E_2$ is a union of paths). Of course, this implies that the resulting graph is bipartitite, thereby allowing the final edge set E_3 to come into play. Our final edge set E_3 is chosen to be either a random graph consistent with this bipartition (the **YES** case), or a random sufficiently dense graph (the **NO** case) so that the resulting graph ($E_1 \cup E_2 \cup E_3$) is $\Omega(1)$-far from being bipartite. The choice of parameters is delicate—we need to ensure that the distributions of E_3 in the **YES** and **NO** cases are statistically close while still ensuring that $E_1 \cup E_2 \cup E_3$ is far from bipartite in the case of NO instances.This

combinatorial analysis is carried out in Section 5.

Finally, we are left with the task of actually analyzing the communication protocols aiming to solve the communication problem on the aforementioned distribution. As with previous works [GKK$^+$08], we make use of Fourier analysis. We specifically analyze the set of bipartitions that are consistent with the set of public inputs and messages broadcast thus far and then look at the Fourier coefficients of the indicator function of this set. We employ relatively elementary methods (at least given previous works) to analyze this set after the player P_1 speaks. To analyze the set after player P_2 speaks, we perform some combinatorial analysis involving the special distributions on E_1 and E_2 and then incorporate this combinatorics into the Fourier language, while finally combining the effects of the two steps using the convolution theorem in Fourier analysis. While the use of this theorem is natural in our setting (involving a composition of many messages, that corresponds to a product of various indicator functions), the fact that the convolved coefficients can be subjected to spectral analysis appears somewhat novel, and we hope it will spur further progress on this and other questions.

Related work. The past decade has seen an extensive body of work on understanding the space complexity of fundamental graph problems in the streaming model; see, for instance, the survey by McGregor [McG14]. It is now known that many fundamental problems admit streaming algorithms that only require $\tilde{O}(n)$ space (i.e. they do not need space to load the edge set of the graph into memory) – e.g., sparsifiers [AG09, KL11, AGM12b, KLM$^+$14], spanning trees [AGM12a], matchings [AG11, AG13, GKK12, Kap13, GO12, HRVZ15, Kon15, AKLY15], spanners [AGM12b, KW14]. Very recently it has been shown that it is sometimes possible to approximate the *cost* of the solution without even having enough space to load the *vertex set* of the graph into memory (e.g. [KKS14, EHL$^+$15, CCE$^+$15]). Our work contributes to the study of streaming algorithms by providing a tight impossibility result for non-trivially approximating MAX-CUT value in $o(n)$ space.

Organization. We formally define our communication problem and describe its connection to streaming algorithms for approximating MAX-CUT value in Section 2. We then state the main technical lemmas and prove the main theorem in Section 3. The proof of the main technical lemma of our communication lower bound is given in Section 4, and (an outline of) the gap analysis is given in Section 5.

2 Communication problem and hard distribution

In this section, we introduce a multi-player "sequential" communication problem and state our lower bound for this problem. We first describe the general model in which this problem is presented.

We consider a sequential communication model where T players sequentially receive *public inputs* M_t and *private inputs* w_t, for $t \in [T]$. A problem in this model is specified by an $F(M_1, \ldots, M_T; w_1, \ldots, w_T)$ and the goal of the players is to compute this function. A protocol for this problem Π is specified by a sequence of functions $\Pi = (r_1, \ldots, r_t)$. At stage $t \in [T]$, the t-th player announces its message $a_t = r_t(M_1, \ldots, M_t; a_1, \ldots, a_{t-1}; w_t)$, and the message a_T is defined to be the output of the protocol Π. The complexity of Π, denoted $|\Pi|$, is the maximum length of the messages $\{a_t\}_{t\in[T]}$. We consider the distributional setting, i.e., where the inputs are drawn from some distribution μ and the error of the protocol is the probability that its output does not equal $F(M_1, \ldots, M_T; w_1, \ldots, w_T)$. By Yao's minmax principle, we assume, without loss of generality, that the communication protocol is deterministic. Also, for the remainder of the paper, addition over $\{0,1\}^n$ and matrix multiplication occurs modulo 2.

We now describe the specific communication problem that we consider in this work.

Implicit Hidden Partition (IHP) Problem. The T-player Implicit Hidden Partition problem $\mathbf{IHP}(n)$ for positive integer n is defined as follows: The public inputs are sets of edges, $M_1, \ldots, M_T$, on vertex set

$[n]$, while the private inputs $w_1, \ldots, w_T$ are $\{0,1\}$-colorings of the corresponding sets of edges. The goal is to distinguish the case in which the colorings are *valid* (i.e., there exists a cut such that every edge of $\cup_t M_t$ is colored 1 if and only if it crosses the cut) from the case in which no such cut exists. A convenient representation of the inputs will be to represent the edges M_t as incidence matrices $M_t \in \{0,1\}^{m_t \times n}$ and the coloring by $w_t \in \{0,1\}^{m_t}$, for $t \in [T]$, where m_t denotes the number of edges of M_t. In this representation a coloring $x \in \{0,1\}^n$ is valid if and only if $M_t x = w_t$ for every $t \in [T]$.

In the instances we use, we will set $T = 3$, while M_1 and M_2 will be (incidence matrices of) matchings so that their rows sum to 2 and columns sum to at most 1. Also, $M_3 \in \{0,1\}^{m_3 \times n}$ will be the edge incidence matrix of a suitable cycle-free subgraph of an Erdős-Rényi graph below the threshold for emergence of a giant component.

Distributional Implicit Hidden Partition (DIHP) Problem. In this work, we will actually deal with a *distributional* version of **IHP** with $T = 3$ that we denote **DIHP**. **DIHP** has three parameters: a positive even integer Δ, a positive integer n divisible by Δ, and a real number α with $0 < \alpha < 1$. **DIHP**(n, Δ, α) is defined to be **IHP**(n) on inputs chosen from a distribution $\mathcal{D} = \frac{1}{2}(\mathcal{D}^Y + \mathcal{D}^N)$, where $\mathcal{D}^Y$ and $\mathcal{D}^N$ are defined as follows: In both the distributions $\mathcal{D}^Y$ and $\mathcal{D}^N$, the triples (M_1, M_2, M_3) are chosen identically from a process $\mathcal{P}_{n,\Delta,\alpha}$ that we describe below shortly. In $\mathcal{D}^Y$, the private inputs w_1, w_2, w_3 are chosen by sampling $X^* \in \{0,1\}^n$ uniformly and settting $w_t = M_t X^*$ for $t \in \{1,2,3\}$. Note that the distribution $\mathcal{D}^Y$ is supported on **YES** instances. In the distribution $\mathcal{D}^N$, the w_t's are uniformly random strings chosen independently of each other. As we show later, the distribution $\mathcal{D}^N$ is mostly supported on **NO** instances that are, in fact, far from **YES** instances, where distance is measured in terms of the number of edges that have to be removed in order to produce a valid coloring.

Although the notation M_t denotes an $m_t \times n$ edge incidence matrix, we will often use M_t to denote the corresponding graph as well. However, the sense in which M_t is used will be clear from context. Furthermore, we will use E_t to denote the set of edges specified by M_t.

Edge Sampling Process $\mathcal{P}_{n,\Delta,\alpha}$ We now specify the process $\mathcal{P}_{n,\Delta,\alpha}$, which is used to sample the graphs (edge incidence matrices) M_1, M_2, M_3 in both $\mathcal{D}^Y$ and $\mathcal{D}^N$. The set M_1 is a deterministic perfect matching that matches vertex i to $i + n/2$ for every $i \in [n/2]$. The set M_2 is a also a matching sampled as follows: We sample a permutation $\pi : [n/2] \to [n/2]$ uniformly and then match the vertex $\pi(i)$ to the vertex $\pi(i+1) + n/2$ for every i that is *not* divisible by $\Delta/2$. (Note that by this process, the union of the graphs $M_1 \cup M_2$ is a collection of disjoint paths, each of length $\Delta - 1$.) Finally, we sample M_3 in three steps:

Step 1. We first sample a random graph M_3' from the Erdős-Rényi model with parameter α/n, i.e., every possible edge is included independently with probably α/n.

Step 2. We remove all edges in M_3' that have already been included in $M_1 \cup M_2$ to get a subgraph M_3''.

Step 3. We now consider the connected components of M_3'' and, for every component that contains a cycle, we remove all edges of that component. The resulting subgraph is M_3.

Note that since α is close to 1, the graph M_3 (or M_3' for that matter) is subcritical and most of its components are of constant size. At most a constant number of edges of M_3' appear in $M_1 \cup M_2$ and another small constant appear in cycles. Thus, for all practical purposes, M_3 behaves like M_3'. In particular, as we show later, the fraction of invalidly colored edges in a random coloring of the edges remains nearly the same in $M_1 \cup M_2 \cup M_3$ as in $M_1 \cup M_2 \cup M_3'$.
The following theorem is the main technical contribution of the paper:

Theorem 2.1 *There exist constants $\Delta^* > 0$ and $0 < \alpha^* < 1$ such that for every even integer $\Delta \geq \Delta^*$ and every $\alpha \in (\alpha^*, 1)$, there exists $c > 0$ such that the following holds: For every sufficiently large integer n that*

is divisible by Δ, every protocol Π for **DIHP**(n, Δ, α) *that succeeds with probability at least $2/3$ satisfies $|\Pi| \geq cn$.*

We accompany the above theorem with a reduction from **DIHP** to MAX-CUT:

Theorem 2.2 (Reduction from DIHP to MAX-CUT) *There exist constants $\Delta^* > 0$ and $0 < \alpha^* < 1$ such that for every even integer $\Delta \geq \Delta^*$ and every $\alpha \in (\alpha^*, 1)$, there exists $\epsilon^* > 0$ such that the following holds: If there exists a single-pass streaming $(1+\epsilon^*)$-approximation algorithm for MAX-CUT with space complexity $s(n)$ that succeeds with probability at least 9/10, then there exists a protocol Π for* **DIHP**(n, Δ, α) *with $|\Pi| \leq s(n) + O(\log n)$ that succeeds with probability at least $2/3$.*

Central to both of the above theorems is a combinatorial analysis that establishes that $\mathcal{D}^N$ is supported mostly on **NO** instances and that, furthermore, these instances generate MAX-CUT instances (under the reduction used in Theorem 2.2) whose optimum is bounded away from the total number of edges by a constant fraction. The following definition gives the (simple) reduction which simply outputs the edges of the **DIHP** instance that are labelled 1, and then the lemma establishes the above formally.

Definition 2.3 *Given $\mathcal{I} = (M_1, M_2, M_3; w_1, w_2, w_3)$, the reduction $R(\mathcal{I})$ outputs the stream containing edges of M_1 that are labelled 1 in w_1, followed by the edges of M_2 labelled 1 in w_2, followed by the edges of M_3 labelled 1 in w_3. (Within each M_t, the order of the edges in the stream is arbitrary.)*

Lemma 2.4 *There exist constants $\Delta^* > 0$ and $0 < \alpha^* < 1$ such that for every $\alpha \in (\alpha^*, 1)$ and even integer $\Delta \geq \Delta^*$, there is a constant $\epsilon^* > 0$ for which the following conditions hold for the reduction R from Definition 2.3:*

- **(1)** *If $\mathcal{I} = (M_1, M_2, M_3; w_1, w_2, w_3)$ is sampled from $\mathcal{D}^Y$ of* $\mathbf{DIHP}(n, \Delta, \alpha)$*, then $R(\mathcal{I})$ is a bipartite graph.*
- **(2)** *If $\mathcal{I}$ is sampled from $\mathcal{D}^N$, then with probability at least 95/100, $R(\mathcal{I})$ is a graph on m edges with MAX-CUT value at most $(1 - \epsilon^*)m$.*

Lemma 2.4 is proved in Section 5. Theorem 2.2 is simple to prove using Lemma 2.4. We devote the rest of this section to providing this proof, as well as a proof of Theorem 1.1. The rest of the paper focuses on proving Theorem 2.1.

Reduction from DIHP to MAX-CUT. We now provide a proof of Theorem 2.2.

Proof of Theorem 2.2: Let R be the reduction from Definition 2.3. Let α^* and Δ^* be the constants guaranteed by Lemma 2.4. We fix an $\alpha \in (\alpha^*, 1)$ as well as an even integer $\Delta \geq \Delta^*$. Let $\epsilon^* > 0$ be the constant from Lemma 2.4 for this choice of α and Δ.

It is easy to see that for instances $\mathcal{I}$ sampled from $\mathcal{D}^Y$, the MAX-CUT value of $G = R(\mathcal{I})$ is m, the number of edges of G since G is bipartite. Moreover, by Lemma 2.4, the MAX-CUT value of $R(\mathcal{I})$ for instances $\mathcal{I}$ sampled from $\mathcal{D}^N$ is at most $(1 - \epsilon^*)m$ with probability at least $95/100$.

Now suppose **ALG** is a one-pass streaming algorithm with space complexity $s(n)$ that produces a $(1 - \epsilon^*)$-approximation to the MAX-CUT value with success probability at least $9/10$. Consider the following protocol Π for **DIHP**(n, Δ, α), which makes use of **ALG** as a subroutine: Augment **ALG** with a counter m for the total number of edges presented to it. This takes $O(\log n)$ additional bits of space for simple input graphs on n vertices. Now, for each $t \in \{1, 2, 3\}$, let player t (1.) run (the augmented) **ALG** on the state posted by player $t - 1$ with the stream of edges formed by enumerating all edges in M_t for which the corresponding value in w_t is 1 and, (2.) if $t \in \{1, 2\}$, pass on the resulting state of **ALG** to the next player.

In other words, the players simulate **ALG** on the stream $R(\mathcal{I})$. The last player then takes the ending state of **ALG** and checks whether the output MAX-CUT value of **ALG** is at least $m/(1+\epsilon^*)$. If so, the player outputs **YES**; otherwise, the player outputs **NO** .

It is clear that the aforementioned simulation succeeds on **DIHP**(n, Δ, α) with probability at least $2/3$. Moreover, the amount of communication $|\Pi|$ in Π is at most the amount of space used for our augmented **ALG**. Thus, $|\Pi| \leq s(n) + O(\log n)$, as desired.

■

Given Theorem 2.2 and Theorem 2.1, our main theorem follows easily and the proof is included below for completeness.

Proof of Theorem 1.1: Let α_1^* and Δ_1^* be the constants guaranteed by Theorem 2.1, and let α_2^* and Δ_2^* be the constants of Theorem 2.2. Let Δ be the smallest even integer larger than $\max\{\Delta_1^*, \Delta_2^*\}$ and choose $\alpha \in (\max\{\alpha_1^*, \alpha_2^*\}, 1)$. Let ϵ^* be the constant given by Theorem 2.2 for this choice of α and Δ.

Now, suppose there exists a randomized single-pass streaming algorithm **ALG** that yields a $(1+\epsilon^*)$-approximation to MAX-CUT with probability at least $9/10$. Let $s(n)$ be the amount of space used by **ALG** on input graphs with n nodes. By Theorem 2.2, there is a protocol Π for **DIHP**(n, Δ, α) with $|\Pi| \leq s(n) + O(\log n)$ such that Π succeeds with probability at least $2/3$.

Now, Theorem 2.1 implies that $|\Pi| \geq c'n$ for some constant c'. Hence, $s(n) \geq c'n - O(\log n) \geq cn$ for some constant $c > 0$ and sufficiently large n, which completes the proof. ■

3 Analysis of communication problem via Fourier techniques

In this section, we first review Fourier analysis on the boolean hypercube, then review relevant communication complexity techniques that were developed in prior work [GKK$^+$08], explain why they do not suffice for our result, and give an outline of our approach.

3.1 Fourier analysis on the boolean hypercube

Let $p : \{0,1\}^n \to \mathbb{R}$ be a real valued function defined on the boolean hypercube. We use the following normalization of the Fourier transform:

$$\hat{p}(v) = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} p(x) \cdot (-1)^{x \cdot v}.$$

With this normalization, the inverse transform is given by

$$p(x) = \sum_{v \in \{0,1\}^n} \widehat{p}(v) \cdot (-1)^{x \cdot v}.$$

We will use the relation between multiplication of functions in the time domain and convolution in the frequency domain to analyze the Fourier spectrum of $f_1 \cdot f_2$. With our normalization of the Fourier transform the convolution identity is

$$\widehat{(p \cdot q)}(v) = (\widehat{p} * \widehat{q})(v) = \sum_{x \in \{0,1\}^n} \widehat{p}(x)\widehat{q}(x+v). \tag{1}$$

The main object of our analysis will be the Fourier transform of $h_2 = f_1 \cdot f_2$ (these functions are defined later in Definition 3.2). By (1), we have $\widehat{h}_2 = \widehat{f}_1 * \widehat{f}_2$. This identity will form the basis of our proof. We will also need Parseval's equality, which, with our normalization, takes the form

$$||\widehat{p}||^2 = \sum_{v\in\{0,1\}^n} \widehat{p}(v)^2 = \sum_{v\in\{0,1\}^n} \left(\frac{1}{2^n} \sum_{x\in\{0,1\}^n} p(x)\cdot(-1)^{x\cdot v} \right)^2 = \frac{1}{2^n} \sum_{x\in\{0,1\}^n} p(x)^2 = \frac{1}{2^n}||p||^2. \quad (2)$$

Remark 3.1 *If $f(x) : \{0,1\}^n \to \{0,1\}$ is the indicator of a set $\mathbf{A} \subseteq \{0,1\}^n$, we have $||f||^2 = |\mathbf{A}|$, so that $||\widehat{f}||^2 = \frac{|\mathbf{A}|}{2^n}$.*

3.2 The basic setup

We use the notation $X_{i:j}$ to denote $(X_i, X_{i+1}, \ldots, X_j)$. Recall that the messages posted by the players are denoted by $a_t = r_t(M_{1:t}, a_{1:t-1}, w_t)$, where M_t are public $m_t \times n$ edge incidence matrices and w_t are private inputs to players. We use s to denote the maximum of the bit lengths of messages posted by the players. Our goal is to show that if $s \ll n$, then the total variation distance between the distribution of the publicly shared information (messages a_1, a_2, a_3 and graphs M_1, M_2, M_3) in the **YES** and **NO** cases is small. As we show, this task can be simplified as follows. It suffices to consider the **YES** case only and show that if $s \ll n$, then the distribution of $w_t = M_t X^*$ conditional on the publicly posted content up to time t (namely, $a_1, \ldots, a_{t-1}$ and $M_1, \ldots, M_t$) is close to the uniform distribution in total variation distance for $t = 1, 2, 3$ (recall that w_t is actually uniformly distributed in the **NO** case). Our proof of this fact relies on Fourier analytic techniques for reasoning about the distribution of $M_t X^*$ conditioned on typical communication history.

More specifically, our goal is to show that the total variation distance between the distribution of $(M_{1:3}, a_{1:3})$ for the **YES** and **NO** instances is vanishingly small. It suffices to consider the **YES** case only. Fix $t \in \{1,2,3\}$ and let $X^* \in \{0,1\}^n$ denote a uniform random vector conditioned on the graphs $M_{1:t}$ and messages $a^Y_{1:t-1}$. In Lemma 3.4, we show that it suffices to show that with high probability, for each $t = 1, 2, 3$, the distribution of $M_t X^*$ is close to uniform in $\{0,1\}^{m_t}$ and is, hence, indistinguishable from the **NO** case.

Conditioning on messages posted up to time t makes X^* uniformly random over a certain subset of the binary cube. We will analyze this subset of the hypercube or, rather, the Fourier transform of its indicator function, and show that if communication is small, the distribution of X^* conditional on typical history is such that $M_t X^*$ is close to uniformly random in total variation distance.

We now define notation that lets us reason about the distribution of X^* at each step t. Since we assume that the protocol is deterministic and the prior distribution of X^* is uniform over $\{0,1\}^n$, the distribution of X^* conditioned on the publicly posted content thus far is uniform over some set $\mathbf{B}_t \subseteq \{0,1\}^n$. We prove the desired claim by analyzing the Fourier spectrum of the indicator function of $\mathbf{B}_t$. It turns out to be convenient to represent $\mathbf{B}_t$ as the intersection of simpler subsets $\mathbf{A}_t$ of the hypercube, where each $\mathbf{A}_t$ essentially conveys the information that the t-th player's message gives about X^*. We give formal definitions below.

Definition 3.2 (Sets $\mathbf{A}_t, \mathbf{B}_t$ and their indicator functions f_t, h_t) *Fix $\alpha \in (0,1)$ and integers $n \geq 1$ and $t \in \{1,2,3\}$. Consider a* **YES** *instance $(M_{1:3}, w_{1:3})$ of* **DIHP**(n, Δ, α) *with X^* being the (random) hidden partition (so that $w_t = M_t X^*$). Recall that $a_t = r_t(M_{1:t}, a_{1:t-1}, w_t)$.*

We define $\mathbf{A}_{reduced,t} \subseteq \{0,1\}^{m_t}$ as the set of possible values of $w_t = M_t X^*$ that lead to the message a_t, and we define $\mathbf{A}_t$ to be the set of values of $X^* \in \{0,1\}^n$ that correspond to $\mathbf{A}_{reduced,t}$. Formally, letting $g_t(\cdot) := r_t(M_{1:t}, a_{1:t-1}, \cdot) : \{0,1\}^{m_t} \to \{0,1\}^s$, we define

$$\mathbf{A}_{reduced,t} = g_t^{-1}(a_t) \subseteq \{0,1\}^{m_t} \quad \textit{and} \quad \mathbf{A}_t = \{x \in \{0,1\}^n : M_t x \in \mathbf{A}_{reduced,t}\}. \quad (3)$$

Moreover, for each $t = 1, 2, 3$, let $f_t : \{0,1\}^n \to \{0,1\}$ denote the indicator function of $\mathbf{A}_t$, and let $h_t = f_1 f_2 \cdots f_t$, so that h_t is the indicator of $\mathbf{B}_t := \mathbf{A}_1 \cap \mathbf{A}_2 \cap \ldots \cap \mathbf{A}_t$. We let $\mathbf{B}_0 := \{0,1\}^n$ for convenience.

Our proof of near-uniformity of $M_t X^*$ conditioned on a typical history of communication in **DIHP**(n, Δ, α) is inspired by the work of [GKK$^+$08], which used Fourier analysis to give a communication lower bound on the (explicit) hidden partition problem (where Alice is given X^*, Bob gets (M, w), and Bob needs to check whether $w = MX^*$). In our setting, their results translate to showing that if X^* is uniform in $\mathbf{B} \subseteq \{0,1\}^n$, where $|\mathbf{B}|/2^n \geq 2^{-s}$ with $s = O(\sqrt{n})$, and the indicator function h of $\mathbf{B}$ satisifies

$$\left(\frac{2^n}{|\mathbf{B}|}\right)^2 \sum_{v \in \{0,1\}^n, |v|=2\ell} \widehat{h_t}(v)^2 \leq (4\sqrt{2}s/\ell)^{2\ell} \ \ \forall \ell \in [0:s], \tag{4}$$

where $|v|$ denotes the Hamming weight of v, then the distribution of MX^* is close to uniform for a random sparse graph M (a random matching in [GKK$^+$08]). This translates to a lower bound of $\Omega(\sqrt{n})$ on the communication complexity of the explicit hidden partition problem, but this is too weak for our purposes.

To improve this bound we need to replace the right hand side of the inequality above to a form $(O(s)/\ell)^\ell$ from $(O(s)/\ell)^{2\ell}$. Unfortunately, such an improvement is not possible for the explicit hidden partition problem, which stems from the fact that X^* is known to Alice. In our case, X^* is not known to any player, but we need an analysis that can take advantage of this key fact. We now outline our approach for doing so.

Our first observation is that if the bound in (4) could be strengthened by replacing the exponent on the righthand side with ℓ (i.e., reducing the exponent by a factor of 2), an $\Omega(n)$ lower bound would follow. This observation is formalized in Lemma 3.3, which is stated below and proved formally in section 6.

Lemma 3.3 *Let $\Delta > 0$ be an even integer. Then, for every $0 < \alpha < 1$, there exists a constant $0 < c < 1$ such that for every $\delta \in (n^{-1/10}, c)$, the following conditions hold if n is any sufficiently large multiple of Δ:*

- **(1)** *Let $\mathbf{B} = \mathbf{A}_1$, as defined in Definition 3.2. Then, for every choice of matchings M_1, M_2 sampled according to $\mathcal{P}_{n,\Delta,\alpha}$, the distribution of $M_2 x$ is uniform over $\{0,1\}^{m_2}$ when x is uniformly random in $\mathbf{B}$.*
- **(2)** *Let $\mathbf{B} \subseteq \{0,1\}^n, |\mathbf{B}| = 2^{n-z}$ for $z \leq \delta^4 n$, and let $h : \{0,1\}^n \to \{0,1\}$ be the indicator of $\mathbf{B}$. If $\left(\frac{2^n}{|\mathbf{B}|}\right)^2 \sum_{v:|v|=2\ell} \hat{h}(v)^2 \leq \left(\frac{64\delta^4 n}{\ell}\right)^\ell$ holds for all $\ell \leq \delta^4 n$, then the following conditions hold: Let M_1, M_2, M_3 be sampled according to $\mathcal{P}_{n,\Delta,\alpha}$. Then, with probability at least $1 - O(\delta)$ over the choice of M_3, the total variation distance between the distribution of $M_3 x$, where x is uniformly random in $\mathbf{B}$, and the uniform distribution over $\{0,1\}^{m_3}$ is $O(\delta/\sqrt{1-\alpha})$. In particular, one can take $c = \min\left\{\left(\frac{1-\alpha}{512}\right)^{1/4}, \left(\frac{e^{-\alpha}\log_2(32/(31+\alpha))}{32}\right)^{1/4}\right\}$.*

We note that such a strengthening of (4) is **impossible** for an indicator function $f : \{0,1\}^n \to \{0,1\}$ of an **arbitrary** subset $\mathbf{B} \subseteq \{0,1\}^n$ with $|\mathbf{B}| = 2^{n-z}$, $z \leq \delta^2 n$—a subcube of appropriate size shows that (4) is essentially the best possible bound. Our improvement crucially uses the fact that unlike in the boolean hidden matching problem, in **DIHP**, the players only have **indirect access** to X^* via linear functions $M_t X^*$. In particular, the sets whose indicator functions we analyze are of a special form (see Definition 3.2).

If we could prove that the preconditions of Lemma 3.3 hold w.h.p. for h_2, we would be done by Lemma 3.3. It turns out that one can prove that these preconditions are satisfied for $h_1 = f_1$ rather directly (see Theorem 4.6) using the fact that the compression function g_1 (see Definition 3.2) is applied to the parities of $x_a + x_b, (a, b) \in M_1$. Proving a similar result for the function $h_2 = f_1 \cdot f_2$ is challenging, and

this proof is the main technical contribution of our paper. In order to do that, we need to analyze the Fourier transform $h_2 = f_1 \cdot f_2$, which we do using the convolution identity $\widehat{h_2} = \widehat{f_1} * \widehat{f_2}$. Our main bound on the Fourier transform of $f_1 \cdot f_2$ is stated below.

Lemma 3.4 *There exists $C > 1$ such that for every even integer $\Delta > 2$, $\gamma > n^{-1/5}$ smaller than an absolute constant, and $\alpha \in (0,1)$, the following conditions hold for sufficiently large n divisible by Δ: Let Π be a protocol for* **DIHP**(n, Δ, α) *such that $|\Pi| =: s$, where $s = s(n) = \omega(\sqrt{n})$ and $s(n) \leq \frac{1}{2048C\Delta^2}\gamma^5 n$. Then, there exists an event $\mathcal{E}$ that only depends on X^*, M_1, M_2 and occurs with probability at least $1 - O(\gamma)$ over $\mathcal{P}_{n,\Delta,\alpha}$ and the choice of $X^* \in \{0,1\}^n$ such that, conditioned on $\mathcal{E}$, one has*

(1) $|\mathbf{B}_2|/2^n \geq 2^{-\gamma^4 n}$.

(2) $\left(\frac{2^n}{|\mathbf{B}_2|}\right)^2 \sum_{v \in \{0,1\}^n, |v|=2\ell} \widehat{h}_2(v)^2 \leq (C\Delta^2\gamma^4 n/\ell)^\ell$ *for all $\ell \leq \gamma^4 n$.*

Before we present the proof of Theorem 2.1, we require one simple lemma about total variation distance of two probability distributions, which appears with proof in [KKS15].

Lemma 3.5 (Lemma 5.6 in [KKS15]) *Let (X, Y^1), (X, Y^2) be random variables taking values on a finite sample space $\Omega = \Omega_1 \times \Omega_2$. For any $x \in \Omega_1$, let Y^i_x, $i = 1, 2$ denote the conditional distribution of Y^i given $X = x$. Then,*

$$\|(X, Y^1) - (X, Y^2)\|_{tvd} = \mathbf{E}_X \left[\|Y^1_X - Y^2_X\|_{tvd}\right].$$

Proof of Theorem 2.1:

Suppose $\Delta > 0$ is an even integer and $0 < \alpha < 1$. Then, we choose $\delta \in (0,1)$ as well as $\gamma \in (0,1)$ such that $\gamma < (64/C\Delta^2)^{1/4}\delta$. Moreover, we pick δ and γ to be sufficiently small such they obey the upper bounds in the hypotheses of Lemmas 3.3 and 3.4. Also, assume n is a sufficiently large multiple of Δ (in particular, $n^{-1/10} < \delta$ and $n^{-1/5} < \gamma$) so that δ and γ obey the lower bounds in the hypotheses of Lemmas 3.3 and 3.4. Moreover, assume γ is sufficiently small so that the event $\mathcal{E}$ in Lemma 3.4 occurs with probability greater than $1/2$.

We now assume that Π is a protocol for **DIHP**(n, Δ, α) that uses less than $\frac{1}{2048C\Delta^2}\gamma^5 n$ bits of communication, where $C > 0$ is the constant in Lemma 3.4.

Recall that the first player posts the message $a_1 = r_1(M_1, w_1)$. We now consider the distribution of (M_1, M_2, a_1, w_2). Let D_1^Y and D_1^N be the distributions of (M_1, M_2, a_1, w_2) on **YES** and **NO** instances, respectively. Thus, $D_1^Y = (M_1, M_2, a, p_{M_1,M_2,a})$, where $p_{M_1,M_2,a}$ is the distribution of $M_2 x$ conditional on $r_1(M_1, x) = a$. For any M_1, M_2, a, we let $D^Y_{(M_1,M_2,a)} = p_{M_1,M_2,a}$ and $D^N_{(M_1,M_2,a)} = U_{M_2}$ denote the distribution of w_2 given the message a and edge incidence matrices M_1, M_2 for the **YES** and **NO** instances, respectively. (Here, U_r denotes the uniform distribution on $\{0,1\}^r$.) Moreover, note that the distribution of (M_1, M_2, a_1) is identical in both the **YES** and **NO** cases. Thus, by Lemma 3.5 and part (1) of Lemma 3.3, we have

$$\begin{aligned}\|D_1^Y - D_1^N\|_{tvd} &= \mathbf{E}_{M_1,M_2,a}\left[\|D^Y_{(M_1,M_2,a)} - D^N_{(M_1,M_2,a)}\|\right] \\ &= 0.\end{aligned}$$

Moreoever, since $a_2 = r_2(M_1, M_2, a_1, w_2)$, another simple application of Lemma 3.5 implies that

$$\left\|D_2^Y - D_2^N\right\|_{tvd} = 0,$$

where D_2^Y and D_2^N denote the distributions of (M_1, M_2, a_1, a_2) for the **YES** and **NO** instances, respectively.

Figure 1: Illustration of $P^*(v)$, where $v = \{v_1, \ldots, v_{10}\}$ (marked red). Edges of M_1 are shown as solid lines, edges of M_2 as dashed lines. The set of paths $P(v)$ is the set of edges between the marked nodes. The paths $v_5 - v_6$, $v_7 - v_8$ and $v_9 - v_{10}$ consist only of an edge of M_2, and hence are not grounded. Grounded paths $P^*(v)$ are marked green (paths $v_1 - v_2$ and $v_3 - v_4$).

Now, let $\mathcal{E}$ be the event for the **YES** case that is guaranteed by Lemma 3.4. Recall that $\mathcal{E}$ occurs with probability $1 - O(\gamma)$ over $\mathcal{P}_{n,\Delta,\alpha}$ and the random choice of $X^* \in \{0,1\}^n$. Moreover, Lemma 3.4 implies that for any **YES** instance conditioned on $\mathcal{E}$, we have that for all $\ell \leq \gamma^4 n$,

$$\left(\frac{2^n}{|\mathbf{B}_2|}\right)^2 \sum_{\substack{v \in \{0,1\}^n \\ |v|=2\ell}} \widehat{h}_2(v)^2 \leq \left(\frac{C\Delta^2\gamma^4 n}{\ell}\right)^{\ell} \leq \left(\frac{64\delta^4 n}{\ell}\right)^{\ell}$$

where h_2, $\mathbf{B}_2$ are defined as in Definition 3.2. Thus, letting q_{M_1,M_2,M_3,a_1,a_2} denote the distribution on $M_3 x$ conditioned on $r_1(M_1, x) = a_1$ and $r_2(M_1, M_2, a_1) = a_2$, we see that part (2) of Lemma 3.3 implies that, given the occurrence of $\mathcal{E}$,

$$\|q_{M_1,M_2,M_3,a_1,a_2} - U_{M_3}\|_{tvd} = O(\delta/\sqrt{1-\alpha}). \tag{5}$$

with probability $p \geq 1 - \frac{O(\delta)}{\mathbf{Pr}[\mathcal{E}]} \geq 1 - O(\delta)$ over the choice of M_3 (since γ was chosen small enough to guarantee that $\mathbf{Pr}[\mathcal{E}] \geq 1/2$). Therefore, since $\mathcal{E}$ only depends on X^*, M_1, M_2, Lemma 3.5 and (5) imply that

$$\begin{aligned}
\|D_3^Y - D_3^N\|_{tvd} &= \mathbf{Pr}[\mathcal{E}] \cdot \mathbf{E}_{M_1,M_2,a_1,a_2|\mathcal{E}}\left[\mathbf{E}_{M_3}\left[\|q_{M_1,M_2,M_3,a_1,a_2} - U_{M_3}\|_{tvd}\right]\right] + \mathbf{Pr}[\overline{\mathcal{E}}] \cdot 1 \\
&\leq \mathbf{Pr}[\mathcal{E}] \cdot \mathbf{E}_{M_1,M_2,a_1,a_2|\mathcal{E}}\left[p \cdot O(\delta/\sqrt{1-\alpha}) + (1-p)\cdot 1\right] + \mathbf{Pr}[\overline{\mathcal{E}}] \\
&\leq \mathbf{Pr}[\mathcal{E}](1 - p(1 - O(\delta/\sqrt{1-\alpha}))) + (1 - \mathbf{Pr}[\mathcal{E}]) \\
&\leq 1 - \mathbf{Pr}[\mathcal{E}] \cdot p(1 - O(\delta/\sqrt{1-\alpha})) \\
&\leq 1 - (1 - O(\gamma))(1 - O(\delta))(1 - O(\delta/\sqrt{1-\alpha})) \\
&= O(\gamma) + O(\delta) + O(\delta/\sqrt{1-\alpha}),
\end{aligned}$$

where D_3^Y and D_3^Y denote the distribution of $(M_1, M_2, M_3, a_1, a_2, w_3)$ in the **YES** and **NO** instances, respectively. We choose δ, γ to be small enough so that the above total variation distance is less than $1/3$.

Finally, observe that since $a_3 = r_3(M_1, M_2, M_3, a_1, a_2, w_3)$, the total variation distance of the distributions of $(M_1, M_2, M_3, a_1, a_2, a_3)$ in the **YES** and **NO** cases is also less than $1/3$, which means that Π cannot distinguish the **YES** and **NO** cases with advantage more than $1/6$ over random guessing, i.e., the success probability of Π is less than $2/3$.

Hence, it follows that any algorithm Π for **DIHP** that succeeds with probability at least $2/3$ must use at least cn bits of communication, for $c = \frac{1}{2048C\Delta^2}\gamma^5$. This completes the proof of the claim. ■

4 Proof of main lemma (Lemma 3.4)

The main result of this section is a proof of Lemma 3.4. The main idea behind the proof is to use the convolution identity to express the Fourier transform of h_2 in terms of the Fourier transform of f_1 and f_2. Specifically, for every $v \in \{0,1\}^n$, we have, by the convolution identity,

$$\widehat{h_2}(v) = \widehat{f_1 \cdot f_2}(v) = \sum_{w \in \{0,1\}^n} \widehat{f_1}(w) \cdot \widehat{f_2}(w+v). \tag{6}$$

Besides the convolution identity, we use the structure of the Fourier transform of f_1 and f_2. Specifically, we use the fact that $\widehat{f_1}$ and $\widehat{f_2}$ are supported on edges of M_1 and M_2, respectively (equivalently, they are zero except on the column span of $(M_1; M_2)$). This allows us to classify the terms $\widehat{f_1}(w) \cdot \widehat{f_2}(w+v)$ on the rhs of (6) according to the weight of w and $w+v$. We would like to show that only very few large weight coefficients $\widehat{f_1}(w)$ can contribute to $\widehat{h_2}(v)$ for a low weight v. Note that this is intuitively necessary for the proof, as according to our bounds the ℓ_2^2 mass of coefficients of $\widehat{f_1}$ or $\widehat{f_2}$ grows with the weight level. We prove that a high weight coefficient is unlikely to appear on the rhs of (6) if the coefficient on the lhs is low weight in section 4.1 (see Lemma 4.5). Then in section 4.2, we show how these bounds imply that not too much ℓ_2^2 mass of $\widehat{f_1}$ can be trasferred from high weight levels to low weight levels (see Lemma 4.9). Finally, in section 4.3, we put the developed results together into a proof of Lemma 3.4.

4.1 Useful definitions and basic claims

The following definitions form the basis of our analysis.

Definition 4.1 *Given matchings M_1, M_2 such that $M_1 \cup M_2$ is a union of paths, a vector $v \in \{0,1\}^n$ is called* admissible *with respect to M_1, M_2 if v has an even number of nonzeros on every path in $M_1 \cup M_2$.*

Definition 4.2 (Path decomposition of admissible coefficients) *Given M_1, M_2 such that $M_1 \cup M_2$ is a union of paths, for any $v \in \{0,1\}^n$ admissible wrt M_1, M_2, let $P(v)$ denote the unique set of vertex disjoint paths in $M_1 \cup M_2$ whose endpoints are exactly the nonzeros of v.*

Claim 4.3 *The path decomposition is well defined for any admissible $v \in \{0,1\}^n$.*

Proof: It suffices to show that for any admissible v the set of paths $P(v)$ exists and is unique. Existence follows immediately from definition of admissibility. Uniqueness follows since $M_1 \cup M_2$ is a collection of simple vertex disjoint paths. ■

Definition 4.4 *Given M_1, M_2 such that $M_1 \cup M_2$ is a union of paths, for any $v \in \{0,1\}^n$ admissible wrt M_1, M_2, let $P^*(v) \subseteq P(v)$ denote the set of paths in $P(v)$ that contain at least one edge of M_1. We refer to $P^*(v)$ as the* core *of the path decomposition of v.*

Note that paths in $P(v) \setminus P^*(v)$ are all of length one, i.e. edges of M_2. See Fig. 1 for an illustration.

We will often associate matchings M with the sets of vertices that they match. For example, for $w \in \{0,1\}^n$, we will write $w \subseteq M_1$ to denote the fact that w is a subset of the vertices matched by M_1. We will say that w is supported on edges of M_1 if for every $e = \{u, v\} \in M_1$, one has either $w \cap \{u, v\} = \emptyset$ or $w \cap \{u, v\} = \{u, v\}$. The following claim is crucial to our subsequent analysis:

Lemma 4.5 *For every even integer $\Delta > 2$ and $\alpha \in (0,1)$, if matchings M_1, M_2 are sampled from $\mathcal{P}_{n,\Delta,\alpha}$, then the following conditions hold for every $\ell, k \geq 0$. Conditioned on M_1, for every subset $w \subseteq M_1$ such that $|w| = 2k$, we have the following:*

(1) $\mathbf{Pr}_{M_2}[\exists M' \subseteq M_2 \text{ s.t. } |P^*(w+M')| = \ell \mid M_1] \leq (O(\Delta))^\ell \binom{n/2}{\ell} \binom{n/2}{k}^{-1}$.

(2) *For every* $M' \subseteq M_2$*, one has* $|P^*(w+M')| \geq |w|/\Delta$.

Proof: The second claim follows by recalling that our input distribution on matchings is such that $M_1 \cup M_2$ does not contain cycles, and the largest path length in the graph induced by $M_1 \cup M_2$ is not larger than Δ.

We now prove the first claim. We first upper bound the number of $w \subseteq M_1$ such that $|P^*(w+M')| = \ell$ for some $M' \subseteq M_2$, i.e. the core of $w + M'$ contains ℓ paths. We then show that since the distribution of M_2 is invariant under permutation of edges of M_1, this gives the result.

We now upper bound the number of sets of ℓ paths that each contain at least one edge of M_1, given M_1 and M_2 (we refer to such paths as *grounded*). Given M_1, M_2, in order to select a grounded set of paths, it suffices to first select ℓ edges from M_1, one per path (at most $\binom{n/2}{\ell}$ choices). Then order these edges arbitrarily, and for each $t = 1, \ldots, \ell$,

- choose whether the path starts with an edge of M_1 or an adjacent edge of M_2 (three choices);
- choose a direction to go on the corresponding path in $M_1 \cup M_2$ (at most 2 choices);
- choose a number of steps to go for (at most 2Δ choices).

Putting the bounds above together, we get that for any M_1, M_2, the number of grounded sets of k paths is bounded by $(12\Delta)^\ell \binom{n/2}{\ell}$.

Next, we recall that the matchings M_1, M_2 are generated as follows (our description here is somewhat more detailed than in Section 2, and results in exactly the same distribution; this formulation is more convenient for our analysis):

- Let M_1 be a perfect matching that matches, for each $i = 1, \ldots, n/2$, vertex i to vertex $i + n/2$. Note that edges of M_1 are naturally indexed by $[n/2]$: the i-th edge matches i to $i + n/2$, for $i \in [n/2] = \{1, 2, \ldots, n/2\}$.
- Choose a permutation π of $[n/2] = \{1, 2, \ldots, n/2\}$ uniformly at random. Partition edges of M_1 into $r = n/\Delta$ sets $S_1, \ldots, S_r$ with $\Delta/2$ edges each, where Δ is an even integer that divides n, by letting for each $j = 1, \ldots, n/\Delta$

$$S_j = \left\{ \pi\left(\frac{\Delta}{2} \cdot (j-1) + 1\right), \pi\left(\frac{\Delta}{2} \cdot (j-1) + 2\right), \ldots, \pi\left(\frac{\Delta}{2} \cdot (j-1) + \frac{\Delta}{2}\right) \right\}.$$

- For each $j = 1, \ldots, n/\Delta$, let $M_{2,j}$ match, for each $i = 1, \ldots, \Delta/2 - 1$, the node $\pi\left(\frac{\Delta}{2} \cdot (j-1) + i\right)$ to the node $\pi\left(\frac{\Delta}{2} \cdot (j-1) + i + 1\right) + n/2$. Note that $|M_{2,j}| = \frac{\Delta}{2} - 1$ for each j.

Let $M_1 := \bigcup_{j=1}^r M_{1,j}$ and $M_2 := \bigcup_{j=1}^r M_{2,j}$.

By the derivation above, we have that for any permutation π, the number of grounded sets of k paths in the union $M_1 \cup M_2$ generated by our process is bounded by $(12\Delta)^\ell \binom{n/2}{\ell}$. Denote this set by $\mathcal{P}^\ell(\pi)$ and note that for every P there exists a unique $w \subseteq M_1$ such that $|P^*(w+M')| = \ell$ for some $M' \subseteq M_2$. Specifically, $w = P \cap M_1$ satisfies these constraints. Let $\mathcal{S}(\pi) := \{P \cap M_1 : P \in \mathcal{P}(\pi)\}$. Thus, we have $|\mathcal{S}(\pi)| = (12\Delta)^\ell \binom{n/2}{\ell}$. We now note that $S(\mathrm{id})$ is hence a fixed set of at most $(12\Delta)^\ell \binom{n/2}{\ell}$ subsets of edges. At the same time for every permutation π of $[n/2]$ one has

$$\mathcal{S}(\pi) = \pi^{-1}(\mathcal{S}(\mathrm{id})). \tag{7}$$

Since π is uniformly random, we thus get for every $w \in \{0,1\}^n$ with $|w| = 2k$

$$\begin{aligned}\mathbf{Pr}_\pi[w \in \mathcal{S}(\pi)] &= \mathbf{Pr}_\pi[w \in \pi^{-1}(\mathcal{S}(\mathrm{id}))] = \mathbf{Pr}_\pi[\pi(w) \in \mathcal{S}(\mathrm{id})] \\ &= |\mathcal{S}(\mathrm{id})|/\binom{n/2}{k} = (12\Delta)^\ell \binom{n/2}{\ell}\binom{n/2}{k}^{-1},\end{aligned} \tag{8}$$

where we used the fact that $\pi(w)$ is uniformly random in the set of unordered k-tuples of edges of M_1 when π is uniformly random. This completes the proof. ■

4.2 Bounds on expected transfer of Fourier mass

In this section, we use the convolution identity (6) to bound the contribution of Fourier transforms $\widehat{f_1}$ and $\widehat{f_2}$ to the Fourier transform $\widehat{h_2}$ of $h_2 = f_1 \cdot f_2$ (see Definition 3.2). The main result of this section is Lemma 4.9. The more basic bounds on the Fourier transform of f_1 and f_2 are provided by Theorem 4.6, stated below and proved in section 7. Part **(1)** of the theorem shows that $\widehat{f_1}$ and $\widehat{f_2}$ are supported on edges of matchings M_1 and M_2 respectively, while parts **(2)** and **(3)** use this fact to derive upper bounds of the form $(O(s)/\ell)^\ell$ (i.e. with the improved exponent of ℓ as opposed to 2ℓ that we are looking for) for the amount of mass on weight level ℓ in $\widehat{f_1}$ and $\widehat{f_2}$, respectively.

Theorem 4.6 *Let $M \in \{0,1\}^{m\times n}$ be the incidence matrix of a matching M, where the rows correspond to edges e of M ($M_{eu} = 1$ if e is incident on u and 0 otherwise). Let $g : \{0,1\}^m \to \{0,1\}^s$ for some $s > 0$. Let $a \in \{0,1\}^s$ and let $\mathbf{A}_{reduced} := \{z \in \{0,1\}^m : g(z) = a\}$. Further, let $f : \{0,1\}^n \to \{0,1\}$ denote the indicator of the set*

$$\mathbf{A} := \{x \in \{0,1\}^n : g(Mx) = a\}.$$

Suppose that $|\mathbf{A}| = 2^{n-d}$ for some $d \in [0,n]$.

Then

1. *the only nonzero Fourier coefficients of $\widehat{f}$ are of the form $\widehat{f}(M^T w)$ for some $w \in \{0,1\}^M$;*
2. *for all $\ell \in [0:d]$ and every $Q \subseteq M$*

$$2^{2d} \sum_{\substack{v\in\{0,1\}^n, |v|=2\ell+|Q| \\ v\supseteq Q}} \widehat{f}(v)^2 \le 2^{|Q|}(64d/\ell)^\ell,$$

   where $|Q|$ denotes the number of vertices in Q;

3. $2^{2d}\sum_{v\in\{0,1\}^n} \widehat{f}(v)^2 = 2^d$ *(Parseval's equality).*

The proof of Theorem 4.6 is given in section 7.

Lemma 4.7 *For any $v \in \{0,1\}^n$, one has $(\widehat{f_1 \cdot f_2})(v) = 0$ if v is not admissible with respect to M_1, M_2, and $(\widehat{f_1 \cdot f_2})(v) = \widehat{f_1}(P(v) \cap M_1) \cdot \widehat{f_2}(P(v) \cap M_2)$ otherwise.*

Proof: By the convolution identity (6) we have $(\widehat{f_1 \cdot f_2})(v) = \sum_{x\in\{0,1\}^n} \widehat{f_1}(x)\widehat{f_2}(v+x)$. By Theorem 4.6, **(1)** applied to the sets $\mathbf{A}_i$, messages a_i, functions g_i, $i \in \{1,2\}$ (as per Definition 3.2) we also have that $\widehat{f_1}(x) \neq 0$ only if x is a union of edges of M_1, and $\widehat{f_2}(v+x) \neq 0$ if $v+x$ is a union of edges of M_2. We can thus write $(\widehat{f_1 \cdot f_2})(v) = \sum_{\substack{M_1'\subseteq M_1, M_2'\subseteq M_2 \\ M_1'+M_2'=v}} \widehat{f_1}(M_1')\widehat{f_2}(M_2')$. Since M_1 and M_2 are edge disjoint and $M_1 \cup M_2$ is a union of paths, we have that for every admissible $v \in \{0,1\}^n$, there exists a unique pair $M_1' \subseteq M_1, M_2' \subseteq M_2$ such that $v = M_1' + M_2'$. ■

Lemma 4.8 *For any $w \subseteq M_1$ with $|w| = 2k$, the number of $v \in \{0,1\}^n$ with $|v| = 2\ell$ and $|P^*(v)| = \ell$ such that $v = w + M_2'$ for some $M_2' \subseteq M_2$ is upper bounded by 2^{2k}.*

Proof: For each path $M_1 \cup M_2$, designate one endpoint to be the left endpoint and the other to be the right endpoint arbitrarily. Note that for each path, this fixes an ordering of vertices (left to right). We associate two binary variables with each of the two endpoints of each edge $e \in w$. Denote these binary variables by $L(e)$ and $R(e)$. Then for each $v \in \{0,1\}^n$ and every $e \in w$, we let $L(e) = 1$ if $P(v)$ extends beyond the left endpoint of e, and 0 otherwise. Similarly, $R(e) = 1$ if $P(v)$ extends beyond the right endpoint of e, and 0 otherwise. Note that the collection of variables $\{(L(e), R(e))\}_{e \in w}$ uniquely determines $P(v)$. On the other hand, the number of possible assignments of $L(e), R(e)$ for $e \in w$ is upper bounded by 2^{2k}, proving the lemma. ■

We now state and prove Lemma 4.9. For an event $\mathcal{E}$, we let $\mathbf{I}[\mathcal{E}]$ denote the indicator function of $\mathcal{E}$.

Lemma 4.9 *For every even integer $\Delta > 2$, every $\alpha \in (0,1)$, every $s \leq n/256$, and any protocol Π for* **DIHP**(n, Δ, α)*, the following conditions hold for sufficiently large n. If $f_1, f_2 : \{0,1\}^n \to \{0,1\}$ are indicator functions of $\mathbf{A}_1$ and $\mathbf{A}_2$, respectively, then for every $0 \leq \ell \leq s$, $0 \leq k \leq n/2$, and $w \in \{0,1\}^n$ with $|w| = 2k$, the following conditions hold for every $M_1, \mathbf{A}_1$.*

(1) *If $k \leq \ell$, then*

$$\mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{A}_2|}\right)^2 \cdot \mathbf{I}\left[\frac{|\mathbf{A}_2|}{2^n} \geq 2^{-s}\right] \cdot \sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell}} \widehat{f_2}(w+v)^2 \,\middle|\, M_1, \mathbf{A}_1\right] \leq 4^\ell (O(\Delta))^k (64s/(\ell-k))^{\ell-k}.$$

(2) *If $k \geq \ell$, then*

$$\mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{A}_2|}\right)^2 \cdot \mathbf{I}\left[\frac{|\mathbf{A}_2|}{2^n} \geq 2^{-s}\right] \cdot \sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell}} \widehat{f_2}(w+v)^2 \,\middle|\, M_1, \mathbf{A}_1\right] \leq (O(\Delta))^\ell 8^k \left(\frac{k-\ell}{n/2}\right)^{k-\ell}.$$

Proof: We classify elements v according to the size of the core $P^*(v)$. Let $w = M_1' \subseteq M_1$. For any $v \in \{0,1\}^n, |v| = 2\ell$ admissible wrt M_1, M_2, note that $|P(v)| = \ell$, as every path in $P(v)$ contributes 2 to the weight of v via its two endpoints. Note that $P^*(v) \subseteq P(v)$, so $|P^*(v)|$ is between 0 and ℓ:

$$\sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell}} \widehat{f_2}(w+v)^2 = \sum_{r=0}^{\ell} \sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell \\ |P^*(v)| = r \\ P^*(v) \cap M_1 = w}} \widehat{f_2}(w+v)^2.$$

Since paths in $P(v) \setminus P^*(v)$ are all of length 1 and correspond to edges of M_2, any admissible v can be represented uniquely as $v = v' + x$, where $P^*(v) = P^*(v') = P(v')$ and $x \subseteq M_2$ is supported on edges of M_2 and is disjoint from $P^*(v)$ (see Fig. 1 for an illustration of $P^*(v)$). Substituting this into the rhs of the

equation above, we get

$$\sum_{v\in\{0,1\}^n,|v|=2\ell} \widehat{f_2}(w+v)^2 = \sum_{r=0}^{\ell} \sum_{\substack{v\in\{0,1\}^n \\ |v|=2\ell \\ |P^*(v)|=r}} \widehat{f_2}(w+v)^2$$
$$= \sum_{r=0}^{\ell} \sum_{\substack{v'\in\{0,1\}^n \\ |v'|=2r \\ |P^*(v')|=r \\ P^*(v')\cap M_1=w}} \sum_{\substack{x\subseteq M_2 \\ x\cap P^*(v')=\emptyset \\ |x|=\ell-r}} \widehat{f_2}(w+v'+x)^2 =: Y_1.$$

By Theorem 4.6, **(2)** invoked with $\mathbf{A} = \mathbf{A}_2$, $f = f_2$, $g = g_2$, $M = M_2$, $Q = P^*(v') \cap M_2$, $k = \ell - r$, and $d = \log_2\left(\frac{2^n}{|\mathbf{A}_2|}\right)$, we get

$$\left(\frac{2^n}{|\mathbf{A}_2|}\right)^2 \cdot \mathbf{I}\left[\frac{|\mathbf{A}_2|}{2^n} \geq 2^{-s}\right] \sum_{\substack{x\subseteq M_2 \\ x\cap P^*(v')=\emptyset \\ |x|=\ell-r}} \widehat{f_2}(w+v'+x)^2 \leq 2^{|Q|}(64s/(\ell-r))^{\ell-r}$$
$$\leq 2^{2k}(64s/(\ell-r))^{\ell-r}, \tag{9}$$

where we have used the fact that $|Q| \leq 2\ell$ (the set $P^*(v')$ is a disjoint union of edges of M_1 and M_2 that form paths; the number of edges of M_2 on each such path is no more than a factor of 2 larger than the number of edges of M_1). Putting the bounds above together, and taking expectation over M_2 conditional on

M_1 and $\mathbf{A}_1$, we get that $\mathbf{E}_{M_2}[Y_1]$ is bounded from above by

$$\begin{aligned}
&\mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{A}_2|}\right)^2 \cdot \mathbf{I}\left[\frac{|\mathbf{A}_2|}{2^n} \geq 2^{-s}\right] \cdot \sum_{r=0}^{\min\{k,\ell\}} \sum_{\substack{v' \in \{0,1\}^n \\ |v'|=2r \\ |P^*(v')|=r \\ P^*(v') \cap M_1 = w}} \sum_{\substack{x \subseteq M_2 \\ x \cap P^*(v') = \emptyset \\ |x| = \ell - r}} \widehat{f}_2(w+v'+x)^2 \,\middle|\, M_1, \mathbf{A}_1\right] \\
&= \mathbf{E}_{M_2}\left[\sum_{r=0}^{\min\{k,\ell\}} \sum_{\substack{v' \in \{0,1\}^n \\ |v'|=2r \\ |P^*(v')|=r \\ P^*(v') \cap M_1 = w}} \left(\frac{2^n}{|\mathbf{A}_2|}\right)^2 \cdot \mathbf{I}\left[\frac{|\mathbf{A}_2|}{2^n} \geq 2^{-s}\right] \cdot \sum_{\substack{x \subseteq M_2 \\ x \cap P^*(v') = \emptyset \\ |x| = \ell - r}} \widehat{f}_2(w+v'+x)^2 \,\middle|\, M_1, \mathbf{A}_1\right] \\
&\leq \mathbf{E}_{M_2}\left[2^{2k} \cdot \sum_{r=0}^{\min\{k,\ell\}} (64s/(\ell-r))^{\ell-r} \cdot \sum_{\substack{v' \in \{0,1\}^n \\ |v'|=2r \\ |P^*(v')|=r}} \mathbf{I}[P^*(v') \cap M_1 = w] \,\middle|\, M_1, \mathbf{A}_1\right] \\
&\leq \mathbf{E}_{M_2}\left[2^{2k} \cdot 2^{2k} \cdot \sum_{r=0}^{\min\{k,\ell\}} (64s/(\ell-r))^{\ell-r} \cdot \mathbf{I}[\exists M_2' \subseteq M_2 : |P^*(w+M_2')| = r] \,\middle|\, M_1, \mathbf{A}_1\right] \\
&= 2^{4k} \cdot \sum_{r=0}^{\min\{k,\ell\}} (64s/(\ell-r))^{\ell-r} \cdot \mathbf{Pr}_{M_2}\left[\exists M_2' \subseteq M_2 : |P^*(w+M_2')| = r \,\middle|\, M_1, \mathbf{A}_1\right] =: Y_2,
\end{aligned} \tag{10}$$

where $\mathbf{I}[\mathcal{E}]$ stands for the indicator of event $\mathcal{E}$. We have used (9) to go from the second line to the third, as well as Lemma 4.8 to conclude that $\sum_{\substack{v' \in \{0,1\}^n \\ |v'|=2r \\ |P^*(v')|=r}} \mathbf{I}[P^*(v') \cap M_1 = w] \leq 2^{2k}$ and obtain the fourth line. Note that the summation above is over r between 0 and $\min\{k,\ell\}$. To see that the size of the core $P^*(w+M_2') = P^*(v)$ cannot be larger than 2ℓ, note that each path in the core contributes two distinct endpoints to the weight of v. To see that the size of the core $P^*(w+M_2') = P^*(v)$ cannot be larger than k, note that every path in the core must contain at least one edge in M_2 that belongs to w, and these edges are disjoint.

We have by Lemma 4.5 that

$$\mathbf{Pr}\left[\exists M_2' \subseteq M_2 : |P^*(w+M_2')| = r \mid M_1\right] \leq (O(\Delta))^r \binom{n/2}{r} \binom{n/2}{|w|/2}^{-1}.$$

Substituting this bound into the equation above, we get

$$\begin{aligned} Y_2 &\le 16^k \cdot \sum_{r=0}^{\min\{k,\ell\}} (64s/(\ell-r))^{\ell-r} \cdot \mathbf{Pr}\left[\exists M_2' \subseteq M_2 : |P^*(w+M_2')| = r \mid M_1\right] \\ &\le 16^k \cdot \sum_{r=0}^{\min\{k,\ell\}} (64s/(\ell-r))^{\ell-r} \cdot (O(\Delta))^r \binom{n/2}{r}\binom{n/2}{k}^{-1} =: Y_3. \end{aligned} \tag{11}$$

We now consider two cases, depending on whether $k \le \ell$ or $k \ge \ell$ (the cases overlap, giving us two rather similar bounds for $k = \ell$).

Case 1: $k \le \ell$**.** Using the bound $(n/k)^k \le \binom{n}{k} \le (en/k)^k$ in (11), we obtain

$$\begin{aligned} Y_3 &= 16^k \cdot \sum_{r=0}^{\min\{k,\ell\}} (64s/(\ell-r))^{\ell-r} \cdot (O(\Delta))^r \binom{n/2}{r}\binom{n/2}{k}^{-1} \\ &= 16^k \cdot \sum_{r=0}^{k} (64s/(\ell-r))^{\ell-r} \cdot (O(\Delta))^r \binom{n/2}{r}\binom{n/2}{k}^{-1} \\ &\le 16^k \cdot (64s/(\ell-k))^{\ell-k} \sum_{r=0}^{k} (64s)^{k-r} \left[\frac{(\ell-k)^{\ell-k}}{(\ell-r)^{\ell-r}}\right] \cdot (O(\Delta))^r (en/2r)^r ((n/2)/k)^{-k} \\ &\le 4^\ell \cdot (O(\Delta))^k (64s/(\ell-k))^{\ell-k} \sum_{r=0}^{k} (128s/n)^{k-r} \cdot \frac{(\ell-k)^{\ell-k}(k-r)^{k-r}}{(\ell-r)^{\ell-r}} \cdot \frac{k^k}{r^r(k-r)^{k-r}} \\ &=: Y_4 \end{aligned} \tag{12}$$

We now note that $\frac{a^a b^b}{(a+b)^{a+b}} = \exp(a\ln a + b\ln b - (a+b)\ln(a+b)) \le 1$ for all $a \ge 0, b \ge 0$, by convexity of the function $x \ln x$. Furthermore, for fixed $a+b$, the maximum of $\frac{(a+b)^{a+b}}{a^a b^b}$ is achieved when $a = b$ and equals 2^{a+b}. Applying the first bound with $a = \ell - k, b = k - r$ gives

$$\frac{(\ell-k)^{\ell-k}(k-r)^{k-r}}{(\ell-r)^{\ell-r}} \le 1, \tag{13}$$

and applying the second bound with $a = r, b = k - r$ gives

$$\frac{k^k}{r^r(k-r)^{k-r}} \le 2^k. \tag{14}$$

Substituting these bounds into (12) yields

$$\begin{aligned} Y_4 &= 4^\ell \cdot (O(\Delta))^k (64s/(\ell-k))^{\ell-k} \sum_{r=0}^{k} (128s/n)^{k-r} \cdot \frac{(\ell-k)^{\ell-k}(k-r)^{k-r}}{(\ell-r)^{\ell-r}} \cdot \frac{k^k}{r^r(k-r)^{k-r}} \\ &\le 4^\ell \cdot (O(\Delta))^k (64s/(\ell-k))^{\ell-k} \sum_{r=0}^{k} (128s/n)^{k-r} \cdot \frac{k^k}{r^r(k-r)^{k-r}} \quad \text{(by (13))} \\ &\le 4^\ell \cdot (O(\Delta))^k (64s/(\ell-k))^{\ell-k} \sum_{r=0}^{k} (128s/n)^{k-r} \quad \text{(by (14))} \\ &\le 4^\ell (O(\Delta))^k (64s/(\ell-k))^{\ell-k}. \end{aligned}$$

Substituting this into (11) and then in (10), we get the result for the case $k \leq \ell$ (Case 1).

Case 2: $k \geq \ell$**.** Using the bound $(n/k)^k \leq \binom{n}{k} \leq (en/k)^k$ in (11), we obtain

$$\begin{aligned}
Y_3 &= 16^k \cdot \sum_{r=0}^{\min\{k,\ell\}} (64s/(\ell-r))^{\ell-r} \cdot (O(\Delta))^r \binom{n/2}{r}\binom{n/2}{k}^{-1} \\
&= 16^k \cdot (O(\Delta))^\ell \sum_{r=0}^{\ell} (64s/(\ell-r))^{\ell-r} \cdot \binom{n/2}{r}\binom{n/2}{k}^{-1} \\
&\leq 16^k (O(\Delta))^\ell \sum_{r=0}^{\ell} (64s/(\ell-r))^{\ell-r} (n/2)^{r-k} k^k / r^r \\
&\leq 16^k (O(\Delta))^\ell \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \sum_{r=0}^{\ell} (64s/(\ell-r))^{\ell-r} (n/2)^{r-\ell} \frac{k^k}{r^r (k-\ell)^{k-\ell}} \\
&\leq 16^k (O(\Delta))^\ell \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \sum_{r=0}^{\ell} (128s/n)^{\ell-r} \frac{k^k}{r^r (k-\ell)^{k-\ell} (\ell-r)^{\ell-r}} \\
&\leq 16^k (O(\Delta))^\ell \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \sum_{r=0}^{\ell} (128s/n)^{\ell-r} \frac{k^k \ell^\ell}{r^r (\ell-r)^{\ell-r} (k-\ell)^{k-\ell} \ell^\ell} \\
&\leq 16^k (O(\Delta))^\ell \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \sum_{r=0}^{\ell} (128s/n)^{\ell-r} \frac{\ell^\ell}{r^r (\ell-r)^{\ell-r}} \frac{k^k}{(k-\ell)^{k-\ell} \ell^\ell} \\
&=: Y_5.
\end{aligned}$$

Again, by convexity arguments as in Case 1, we have $\frac{\ell^\ell}{r^r(\ell-r)^{\ell-r}} \frac{k^k}{(k-\ell)^{k-\ell}\ell^\ell} \leq 2^{\ell+k}$. Substituting this in the derivation above, we get

$$\begin{aligned}
Y_5 &= 16^k (O(\Delta))^\ell \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \sum_{r=0}^{\ell} (128s/n)^{\ell-r} \frac{\ell^\ell}{r^r (\ell-r)^{\ell-r}} \frac{k^k}{(k-\ell)^{k-\ell}\ell^\ell} \\
&\leq 16^k (O(\Delta))^\ell 2^{\ell+k} \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \sum_{r=0}^{\ell} (128s/n)^{\ell-r} \\
&\leq 32^k (O(\Delta))^\ell \left(\frac{k-\ell}{n/2}\right)^{k-\ell},
\end{aligned}$$

since $s < n/256$, by assumption of the lemma. ■

4.3 Putting it together

We now present a proof of Lemma 3.4, which we restate here for convenience of the reader:

Lemma 3.4 *There exists $C > 1$ such that for every even integer $\Delta > 2$, $\gamma > n^{-1/5}$ smaller than an absolute constant, and $\alpha \in (0,1)$, the following conditions hold for sufficiently large n divisible by Δ: Let Π be a protocol for* **DIHP**(n, Δ, α) *such that $|\Pi| =: s$, where $s = s(n) = \omega(\sqrt{n})$ and $s(n) \leq \frac{1}{2048C\Delta^2}\gamma^5 n$. Then, there exists an event $\mathcal{E}$ that only depends on X^*, M_1, M_2 and occurs with probability at least $1 - O(\gamma)$ over $\mathcal{P}_{n,\Delta,\alpha}$ and the choice of $X^* \in \{0,1\}^n$ such that, conditioned on $\mathcal{E}$, one has*

(1) $|\mathbf{B}_2|/2^n \geq 2^{-\gamma^4 n}$.

(2) $\left(\frac{2^n}{|\mathbf{B}_2|}\right)^2 \sum_{v\in\{0,1\}^n,|v|=2\ell} \widehat{h}_2(v)^2 \leq (C\Delta^2\gamma^4 n/\ell)^\ell$ *for all* $\ell \leq \gamma^4 n$.

Proof:

We denote

$$\begin{aligned}\mathcal{E}_1 &:= \left\{|\mathbf{A}_1|/2^n \geq 2^{-s-\log_2(2/\gamma)}\right\}\\ \mathcal{E}_2 &:= \left\{|\mathbf{A}_t|/2^n \geq 2^{-s-\log_2(2/\gamma)} \text{ for } t \in \{1,2\}\right\} \quad \text{(note that } \mathcal{E}_2 \subseteq \mathcal{E}_1\text{)}.\end{aligned} \tag{15}$$

We will later show that for every $t \in \{1,2\}$,

$$\mathbf{Pr}[\mathcal{E}_t] \geq 1 - O(\gamma). \tag{16}$$

Note that neither $\mathcal{E}_1$ nor $\mathcal{E}_2$ coincides with the event $\mathcal{E}$—we define $\mathcal{E}$ at the end of the proof as the intersection of $\mathcal{E}_2$ and the success event for an application of Markov's inequality (see Eq. (25) and Eq. (26)).

We prove that if matchings M_1, M_2 are selected according to the random process $\mathcal{P}_{n,\Delta,\alpha}$, then the following conditions hold:

(1) $\frac{|\mathbf{B}_2|}{2^n} = \frac{|\mathbf{A}_1|}{2^n} \cdot \frac{|\mathbf{A}_2|}{2^n}$ for all choices of M_1, M_2, f_1, f_2;

(2) Conditioned on $\mathcal{E}_1$ for all $\ell \in [1, 2s]$,

$$\left(\frac{2^n}{|\mathbf{B}_1|}\right)^2 \sum_{\substack{v\in\{0,1\}^n\\|v|=2\ell}} \widehat{f_1}(v)^2 \leq (128s/\ell)^\ell.$$

(3) Conditioned on $\mathcal{E}_2$ for all $\ell \in [1, 2s]$,

$$\left(\frac{2^n}{|\mathbf{B}_2|}\right)^2 \sum_{\substack{v\in\{0,1\}^n\\|v|=2\ell}} \widehat{f_1 \cdot f_2}(v)^2 \leq (O(\Delta^2/\gamma) \cdot s/\ell)^\ell.$$

We prove claims above, then put them together to get the proof of the lemma. Claims **(1)** and **(2)** are simple, and the proof of the lemma from the claims is simple as well. The bulk of the proof is in **(3)**. We now give the proof of the lemma assuming the claims above.

We now combine **(1)-(3)** to obtain the result of the lemma. Recall that $h_2 = f_1 \cdot f_2$.

First, for $\ell \in [1, 2s]$, we have that **(3)** implies

$$\left(\frac{2^n}{|\mathbf{B}_2|}\right)^2 \sum_{\substack{v\in\{0,1\}^n\\|v|=2\ell}} \widehat{f_1 \cdot f_2}(v)^2 \leq (O(\Delta^2/\gamma)s/\ell)^\ell \leq (C\Delta^2\gamma^4 n/\ell)^\ell$$

for sufficiently large C, since $s \leq \frac{1}{2048C\Delta^2}\gamma^5 n$ by assumption of the lemma.

It remains to show that this bound holds for all $\ell \leq \gamma^4 n$, i.e., we need to consider ℓ in the range $[2s, \gamma^4 n]$. We note that, conditioned on $\mathcal{E}$, one has

$$\frac{2^n}{|\mathbf{B}_2|} = \frac{2^n}{|\mathbf{A}_1|} \cdot \frac{2^n}{|\mathbf{A}_2|} \leq (2^{2s})^2 \leq 2^{4s},$$

where we have combined **(1)** with the fact that conditioned on $\mathcal{E} \subseteq \mathcal{E}_2$, one has $|\mathbf{A}_t|/2^n \geq 2^{-s-\log_2(2/\gamma)} \geq 2^{-2s}$ for every $t \in \{1,2\}$ and sufficiently large n, since $\gamma > n^{-1/5}$ and $s = s(n) = \omega(\sqrt{n})$. Thus, by Theorem 4.6, **(3)** (Parseval's equality), we have

$$\left(\frac{2^n}{|\mathbf{B}_t|}\right)^2 \sum_{\substack{v\in\{0,1\}^n \\ |v|=2\ell}} \widehat{h}_t(v)^2 \leq \left(\frac{2^n}{|\mathbf{B}_t|}\right)^2 \sum_{v\in\{0,1\}^n} \widehat{h}_t(v)^2 \leq \frac{2^n}{|\mathbf{B}_t|} \leq 2^{4s}. \tag{17}$$

for $t \in \{1,2\}$ and all ℓ.

We now show that the rhs above is dominated by $(C\Delta^2\gamma^4 n/\ell)^\ell$ for $\ell \in [2s, \gamma^4 n]$, provided that $C > 0$ is a sufficiently large absolute constant. Indeed, recalling that Δ is a positive integer, we note that as long as $C \geq e$, we have that $(C\Delta^2\gamma^4 n/\ell)^\ell$ is monotonically increasing[2] for $\ell \in [2s, \gamma^4 n]$. Thus, the smallest value is achieved when $\ell = 2s$ and equals

$$(C\Delta^2\gamma^4 n/(2s))^{2s} \geq (4C\Delta^2\gamma^4 n/(\gamma^5 n))^{2s} \geq (4C\Delta^2/\gamma)^{2s} \geq 2^{4s},$$

where we have used the assumption that $s \leq \frac{1}{2048C\Delta^2}\gamma^5 n \leq \frac{1}{8}\gamma^5 n$. This establishes part **(2)** of the lemma statement. Also, note that **(1)** of the lemma statement holds, since

$$\frac{|\mathbf{B}_2|}{2^n} \geq 2^{-4s} \geq 2^{-\gamma^4 n},$$

since $4s \leq 4 \cdot (\gamma^5 n/2048C\Delta^2) \leq \gamma^4 n$. This completes the proof of the lemma assuming claims **(1)**-**(3)** above.

We now prove the claims.

First, we establish Claim **(1)**, which follows from the fact that $M_1 \cup M_2$ does not contain cycles. Indeed, by (6), we have

$$\widehat{h_2}(0) = \sum_{w\in\{0,1\}^n} \widehat{f_1}(w) \cdot \widehat{f_2}(w) = \widehat{f_1}(0^n) \cdot \widehat{f_2}(0^n) + \sum_{w\in\{0,1\}^n \setminus 0^n} \widehat{f_1}(w) \cdot \widehat{f_2}(w),$$

and by Theorem 4.6, **(1)**, all $w \in \{0,1\}^n \setminus 0^n$ such that $\widehat{f_1}(w) \neq 0$ and $\widehat{f_2}(w) \neq 0$ can be perfectly matched by both M_1 and M_2. Let $M_1' \subseteq M_1$ denote the set of edges of M_1 that perfectly match elements of w to each other, and let $M_2' \subseteq M_2$ denote the set of edges of M_2 that perfectly match elements of w to each other. However, this implies that $M_1' \cup M_2'$ must be a union of cycles (note that M_1' and M_2' do not share edges by our construction), which is impossible as $M_1 \cup M_2$ does not contain cycles. Thus, the second term on the rhs of the equation above is zero, and we get

$$\frac{|\mathbf{B}_2|}{2^n} = |\widehat{h_2}(0^n)| = |\widehat{f_1}(0^n)| \cdot |\widehat{f_2}(0^n)| = \frac{|\mathbf{A}_1|}{2^n} \cdot \frac{|\mathbf{A}_2|}{2^n},$$

as desired. This establishes Claim **(1)**.

Let us now concentrate on Claim **(2)**. Note that the claim only applies to $\ell \geq 1$, which will be useful for simplifying calculations somewhat below.

Typical messages. First, note that for each $t \in \{1,2\}$, the function g_t induces a partition $K_1^t, K_2^t, \ldots, K_{2^s}^t$ of $\{0,1\}^{m_t}$, where s is the bit length of the message a_t (recall that we assume wlog that messages are the same length for all t). The number of points in $\{0,1\}^{m_t}$ that belong to sets K_i^t of size less than $\gamma 2^{m_t - s}$ is

[2]Since the function $(ea/b)^b$ is monotone increasing for any $b \in (0, b]$.

bounded by $2^s \cdot \gamma 2^{m_t - s} < \gamma 2^{m_t}$, i.e., at least a $1-\gamma$ fraction of $\{0,1\}^{m_t}$ is contained in large sets K_i^t, whose size is at least $\gamma 2^{m_t - s}$. We call a message m *typical* if $|K_m^t| \geq \gamma 2^{m_t - s}$. Moreover, we say that $a_t = g_t(M_t x)$ is typical if $M_t x$ is typical. We have that $a_t = g_t(z)$ is not typical with probability at most γ if z is uniformly random in $\{0,1\}^{m_t}$. Letting $d := \log_2\left(2^n/|\mathbf{A}_1|\right)$, we now conclude that with probability at least $1 - \gamma/2$ over the choice of $X^* \in \{0,1\}^n$, one has $d \leq s + \log_2(2/\gamma)$. Since $\gamma > n^{-1/5}$ and $s = \omega(\sqrt{n})$ by assumption of the lemma, we have $d \leq s + \log_2(2/\gamma) \leq 2s$ for sufficiently large n. We now invoke Theorem 4.6, **(2)** on the function f_1 with $d \leq s + \log_2(2/\gamma) \leq 2s$, which establishes Claim **(2)**.

Finally, we establish Claim **(3)**. First note that by Lemma 3.3, **(1)** applied to $\mathbf{A}_1$ and M_2, we get that $M_2 X^*$ is uniformly distributed over $\{0,1\}^{m_2}$ when X^* is uniformly distributed over $\mathbf{A}_1$. We thus have that the argument on 'typical' sets from the above paragraph applies even when we condition on M_1 and the first player's message a_1 (equivalently, on the set $\mathbf{A}_1$). Thus, with probability $1 - O(\gamma)$, we have that $\log_2\left(2^n/|\mathbf{A}_2|\right) \leq s + \log_2(2/\gamma)$, which establishes (16).

Thus, assume $\mathcal{E}_2$ holds. Recall that $d = \log_2\left(2^n/|\mathbf{A}_1|\right) \leq s + \log_2(1/\gamma) \leq 2s$. We now claim that for every $k \leq 2s$,

$$\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \sum_{\substack{w \in \{0,1\}^n \\ |w| = 2k}} \widehat{f_1}(v)^2 \leq (128s/k)^k \tag{18}$$

and

$$\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \sum_{\substack{w \in \{0,1\}^n \\ |w| \geq 4s}} \widehat{f_1}(w)^2 \leq 2^{4s}. \tag{19}$$

Indeed, (19) holds by Theorem 4.6, **(3)**:

$$\begin{aligned}\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \sum_{\substack{w \in \{0,1\}^n \\ |w| \geq 4s}} \widehat{f_1}(w)^2 &\leq 2^{2d} \sum_{w \in \{0,1\}^n} \widehat{f_1}(w)^2 \\ &= 2^d \\ &\leq 2^{4s}.\end{aligned}$$

For (18), note that if $k \leq d$, then Theorem 4.6, **(2)** implies that

$$\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \sum_{\substack{w \in \{0,1\}^n \\ |w| = 2k}} \widehat{f_1}(w)^2 \leq (64d/k)^k \leq (128s/k)^k,$$

as desired, while if $d < k \leq 2s$, then Theorem 4.6, **(3)** implies that

$$\begin{aligned}\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \sum_{\substack{w \in \{0,1\}^n \\ |w| = 2k}} \widehat{f_1}(w)^2 &\leq 2^d \\ &\leq (128s/d)^d \\ &\leq (128s/k)^k,\end{aligned}$$

since $(128s/k)^k$ is a monotonically increasing function in k for $k \leq 2s$. This establishes (18).

Next, by Lemma 4.7, we have that for any $v \in \{0,1\}^n$, $\widehat{(f_1 \cdot f_2)}(v) = 0$ if v is not admissible with respect to M_1, M_2, while $\widehat{(f_1 \cdot f_2)}(v) = \widehat{f}_1(P(v) \cap M_1) \cdot \widehat{f}_2(P(v) \cap M_2)$ otherwise. Thus, for any $\ell \geq 0$,

$$\begin{aligned}\sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell}} \widehat{f_1 \cdot f_2}(v)^2 &= \sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell \\ v \text{ admissible wrt } M_1, M_2}} \widehat{f}_1(P(v) \cap M_1)^2 \cdot \widehat{f}_2(P(v) \cap M_2)^2 \\ &= \sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell}} \sum_{w \in \{0,1\}^n} \widehat{f}_1(w)^2 \cdot \widehat{f}_2(w+v)^2.\end{aligned}$$

Note that the second line follows from the first by letting $w := P(v) \cap M_1$ (so that $w + v = (P(v) \cap M_1) + v = P(v) \cap M_2$ and, thus, $\widehat{f}_1(w)^2 \cdot \widehat{f}_2(w+v)^2 = \widehat{f}_1(P(v) \cap M_1)^2 \cdot \widehat{f}_2(P(v) \cap M_2)^2$) as well as noting that there exists at most one $w \in \{0,1\}^n$ such that $\widehat{f}_1(w)^2 \cdot \widehat{f}_2(w+v)^2 \neq 0$ (see the proof of Lemma 4.7).

We now further partition the set of $w \in \{0,1\}^n$ in the inner summation on the rhs above according to weight and obtain

$$\begin{aligned}\sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell}} \widehat{f_1 \cdot f_2}(v)^2 &= \sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell}} \sum_{w \in \{0,1\}^n} \widehat{f}_1(w)^2 \cdot \widehat{f}_2(w+v)^2 \\ &= \sum_{k=0}^{\Delta \cdot \ell} \sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell}} \sum_{\substack{w \in \{0,1\}^n \\ |w| = 2k}} \widehat{f}_1(w)^2 \cdot \widehat{f}_2(w+v)^2 \\ &= \sum_{k=0}^{\Delta \cdot \ell} \sum_{\substack{w \in \{0,1\}^n \\ |w| = 2k}} \widehat{f}_1(w)^2 \cdot \left(\sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell}} \widehat{f}_2(w+v)^2 \right).\end{aligned} \tag{20}$$

Note that we have restricted the summation over k to the range $[0, \Delta \cdot \ell]$ in line 3, as this is justified by Lemma 4.5, **(2)**, which implies that $|P^*(w + M')| \geq |w|/\Delta$ for all $M' \subseteq M_2$, and so, $|v| = |w + (v+w)| = |w + M'| \geq |w|/\Delta$, or $k = |w|/2 \leq \Delta \cdot \ell$ for all v, w such that $\widehat{f}_1(w)\widehat{f}_2(v+w) \neq 0$.

Taking the expectation of (20) with respect to M_2 (conditional on M_1, $\mathbf{A}_1$, and $\mathcal{E}_2$), we obtain

$$\begin{aligned}&\mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{B}_2|} \right)^2 \sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell}} \widehat{f_1 \cdot f_2}(v)^2 \,\middle|\, M_1, \mathbf{A}_1, \mathcal{E}_2 \right] \\ &= \left(\frac{2^n}{|\mathbf{A}_1|} \right)^2 \mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{A}_2|} \right)^2 \sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell}} \widehat{f_1 \cdot f_2}(v)^2 \,\middle|\, M_1, \mathbf{A}_1, \mathcal{E}_2 \right] \\ &\leq \sum_{k=0}^{\Delta \cdot \ell} \left(\frac{2^n}{|\mathbf{A}_1|} \right)^2 \sum_{\substack{w \in \{0,1\}^n \\ |w| = 2k}} \widehat{f}_1(w)^2 \cdot \mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{A}_2|} \right)^2 \sum_{\substack{v \in \{0,1\}^n \\ |v| = 2\ell}} \widehat{f}_2(w+v)^2 \,\middle|\, M_1, \mathbf{A}_1, \mathcal{E}_2 \right].\end{aligned}$$

In what follows, we apply Lemma 4.9 to the inner summation on last line above. In order to reason

about 'typical' messages as defined above, we let

$$\mathbf{I}_1^\star := \mathbf{I}\left[\frac{|\mathbf{A}_1|}{2^n} \geq 2^{-2s}\right] \quad \text{and} \quad \mathbf{I}_2^\star := \mathbf{I}\left[\frac{|\mathbf{A}_2|}{2^n} \geq 2^{-2s}\right].$$

Note that

$$\mathbf{I}_1^\star \geq \mathbf{I}\left[\frac{|\mathbf{A}_1|}{2^n} \geq 2^{-s-\log_2(2/\gamma)}\right] \text{ and } \mathbf{I}_2^\star \geq \mathbf{I}\left[\frac{|\mathbf{A}_2|}{2^n} \geq 2^{-s-\log_2(2/\gamma)}\right]. \tag{21}$$

Specifically, we have for that for any $\ell \leq 2s$,

$$\begin{aligned}
&\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \sum_{k=0}^{\Delta\cdot\ell} \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f_1}(w)^2 \cdot \mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{A}_2|}\right)^2 \sum_{\substack{v\in\{0,1\}^n \\ |v|=2\ell}} \widehat{f_2}(w+v)^2 \middle| M_1, \mathbf{A}_1, \mathcal{E}_2\right] \\
&\quad = \left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \sum_{k=0}^{\Delta\cdot\ell} \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f_1}(w)^2 \cdot \mathbf{E}_{M_2}\left[\mathbf{I}_1^\star \cdot \mathbf{I}_2^\star \left(\frac{2^n}{|\mathbf{A}_2|}\right)^2 \sum_{\substack{v\in\{0,1\}^n \\ |v|=2\ell}} \widehat{f_2}(w+v)^2 \middle| M_1, \mathbf{A}_1, \mathcal{E}_2\right] \\
&\quad \leq \left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \frac{1}{\mathbf{Pr}[\mathcal{E}_2]} \sum_{k=0}^{\Delta\cdot\ell} \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f_1}(w)^2 \cdot \mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{A}_2|}\right)^2 \cdot \mathbf{I}_1^\star \cdot \mathbf{I}_2^\star \cdot \sum_{\substack{v\in\{0,1\}^n \\ |v|=2\ell}} \widehat{f_2}(w+v)^2 \middle| M_1, \mathbf{A}_1\right] \\
&\quad = 2\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \cdot \mathbf{I}_1^\star \cdot \sum_{k=0}^{\Delta\ell} \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f_1}(w)^2 \cdot \mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{A}_2|}\right)^2 \cdot \mathbf{I}_2^\star \cdot \sum_{\substack{v\in\{0,1\}^n \\ |v|=2\ell}} \widehat{f_2}(w+v)^2 \middle| M_1, \mathbf{A}_1\right],
\end{aligned}$$

where we have used (21) to conclude that both $\mathbf{I}_1^\star$ and $\mathbf{I}_2^\star$ equal 1 when $\mathcal{E}_2$ occurs, as well as the fact that $\mathbf{Pr}[\mathcal{E}_2] = 1 - O(\gamma) \geq 1/2$ (when γ is smaller than an absolute constant) by (16) and the fact that $\mathbf{I}_1^\star$ is independent of M_2.

We now apply Lemma 4.9 to the expectation over M_2 in the last line above. Since Lemma 4.9 provides two bounds (one for $\ell \leq k$ and another for $\ell \geq k$), we split the summation into two and apply the respective part of the lemma to each summation. Specifically, we have

$$\begin{aligned}
&2\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \cdot \mathbf{I}_1^\star \cdot \sum_{k=0}^{\Delta\ell} \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f_1}(w)^2 \cdot \mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{A}_2|}\right)^2 \cdot \mathbf{I}_2^\star \cdot \sum_{\substack{v\in\{0,1\}^n \\ |v|=2\ell}} \widehat{f_2}(w+v)^2 \middle| M_1, \mathbf{A}_1\right] \\
&\quad \leq 2\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \cdot \mathbf{I}_1^\star \cdot \sum_{k=0}^{\ell} \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f_1}(w)^2 \cdot \mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{A}_2|}\right)^2 \cdot \mathbf{I}_2^\star \cdot \sum_{\substack{v\in\{0,1\}^n \\ |v|=2\ell}} \widehat{f_2}(w+v)^2 \middle| M_1, \mathbf{A}_1\right] \\
&\qquad + 2\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \cdot \mathbf{I}_1^\star \cdot \sum_{k=\ell+1}^{\Delta\ell} \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f_1}(w)^2 \cdot \mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{A}_2|}\right)^2 \cdot \mathbf{I}_2^\star \cdot \sum_{\substack{v\in\{0,1\}^n \\ |v|=2\ell}} \widehat{f_2}(w+v)^2 \middle| M_1, \mathbf{A}_1\right] \\
&\quad = S_1 + S_2,
\end{aligned}$$

where we let

$$S_1 = 2\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \cdot \mathbf{I}_1^\star \cdot \sum_{k=0}^{\ell} 4^\ell (O(\Delta))^k (64(2s)/(\ell-k))^{\ell-k} \cdot \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f_1}(w)^2$$

$$S_2 = 2\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \cdot \mathbf{I}_1^\star \cdot \sum_{k=\ell+1}^{\Delta\ell} (O(\Delta))^\ell 8^k \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \cdot \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f_1}(w)^2.$$

We now proceed to bound the terms S_1 and S_2 separately.

Bounding S_1. We have

$$\begin{aligned}
S_1 &= 2\left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \cdot \mathbf{I}_1^\star \cdot \sum_{k=0}^{\ell} 4^\ell (O(\Delta))^k (64(2s)/(\ell-k))^{\ell-k} \cdot \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f_1}(w)^2 \\
&= 2\sum_{k=0}^{\ell} 4^\ell (O(\Delta))^k (64(2s)/(\ell-k))^{\ell-k} \cdot \mathbf{I}_1^\star \cdot \left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f_1}(w)^2 \\
&\le 4^{\ell+1} \sum_{k=0}^{\ell} (O(\Delta))^k (64(2s)/(\ell-k))^{\ell-k} \cdot (64(2s)/k)^k \qquad \text{(by Eq. (18))} \\
&= (O(\Delta))^\ell \sum_{k=0}^{\ell} ((128s)/(\ell-k))^{\ell-k} \cdot ((128s)/k)^k \\
&= (O(\Delta))^\ell ((128s)/\ell)^\ell \sum_{k=0}^{\ell} \frac{\ell^\ell}{(\ell-k)^{\ell-k}k^k} \\
&= (O(\Delta))^\ell ((128s)/\ell)^\ell \sum_{k=0}^{\ell} 2^\ell \qquad \left(\text{since } \frac{(a+b)^{a+b}}{a^a b^b} \le 2^{a+b}\right) \text{ for all } a,b>0) \\
&\le (128s/\ell)^\ell (O(\Delta))^\ell \\
&\le (O(\Delta)s/\ell)^\ell.
\end{aligned} \tag{22}$$

Note that we have absorbed the factor of $4^{\ell+1}$ into $(O(\Delta))^\ell$ crucially using the assumption that $\ell > 0$.

Bounding S_2. Observe that

$$S_2 = \mathbf{I}_1^\star \cdot \sum_{k=\ell+1}^{\Delta\cdot\ell} (O(\Delta))^\ell 8^k \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \cdot \left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f_1}(w)^2.$$

We split this summation further into two summations, one over $k \in [\ell+1, 2s]$ and the other over $k \in [2s, \Delta\cdot\ell]$ (assuming that the second range is nonempty).

Case 1: $k \in [\ell+1, 2s]$**.** We have

$$\begin{aligned}
&\sum_{k=\ell+1}^{2s} (O(\Delta))^{\ell} 8^{\ell+(k-\ell)} \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \cdot \mathbf{I}_1^{\star} \cdot \left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f}_1(w)^2 \\
&\quad \le \sum_{k=\ell+1}^{s} (O(\Delta))^{\ell} 8^{\ell+(k-\ell)} \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \cdot (64(2s)/k)^k \qquad \text{(by Eq. (18))} \\
&\quad \le (O(\Delta)s/\ell)^{\ell} \sum_{k=\ell+1}^{s} (2048s/n)^{k-\ell} \frac{(k-\ell)^{k-\ell}\ell^{\ell}}{k^k} \qquad (23) \\
&\quad \le (O(\Delta)s/\ell)^{\ell} \sum_{k=\ell+1}^{s} (2048s/n)^{k-\ell} \qquad \text{(since } a^a b^b/(a+b)^{a+b} \le 1 \text{ for all } a,b>0\text{)} \\
&\quad \le (O(\Delta)s/\ell)^{\ell} \sum_{k=\ell+1}^{s} (2048s/n)^{k-\ell} \qquad \text{(since } s < n/4096 \text{ by assumption)} \\
&\quad \le (O(\Delta)s/\ell)^{\ell}.
\end{aligned}$$

Case 2: $k \in [2s, \Delta\cdot\ell]$**.** Note that increasing the upper limit in the summation to $\Delta \cdot 2s \ge \Delta \cdot \ell$ may only increase the sum since the summands are non-negative. We upper bound the sum of k in this range as follows:

$$\begin{aligned}
&\sum_{k=2s}^{2\Delta\cdot s} (O(\Delta))^{\ell} 8^{\ell+(k-\ell)} \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \cdot \mathbf{I}_1^{\star} \cdot \left(\frac{2^n}{|\mathbf{A}_1|}\right)^2 \sum_{\substack{w\in\{0,1\}^n \\ |w|=2k}} \widehat{f}_1(w)^2 \\
&\quad \le \sum_{k=2s}^{2\Delta\cdot s} (O(\Delta))^{\ell} 8^{\ell+(k-\ell)} \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \cdot 2^{4s} \qquad \text{(by Eq. (19))} \\
&\quad \le \sum_{k=2s}^{2\Delta\cdot s} (O(\Delta))^{\ell} 8^{\ell+(k-\ell)} \left(\frac{k-\ell}{n/2}\right)^{k-\ell} \cdot (8\Delta s/k)^k,
\end{aligned}$$

where we have used the fact that

$$\left(\frac{8\Delta s}{k}\right)^k \ge \left(\frac{8\Delta s}{2s}\right)^{2s} \ge (4\Delta)^{2s} \ge 2^{4s}.$$

We now upper bound the expression on the last line above as follows:

$$\begin{aligned}
&\sum_{k=2s}^{2\Delta\cdot s}(O(\Delta))^{\ell}8^{\ell+(k-\ell)}\left(\frac{k-\ell}{n/2}\right)^{k-\ell}\cdot(8\Delta s/k)^k \\
&\quad\le (O(\Delta^2)s/\ell)^{\ell}\sum_{k=2s}^{2\Delta\cdot s}(O(\Delta)s/n)^{k-\ell}\frac{(k-\ell)^{k-\ell}\ell^{\ell}}{k^k} \\
&\quad\le (O(\Delta^2)s/\ell)^{\ell}\sum_{k=2s}^{\infty}(O(\Delta)s/n)^{k-\ell} \qquad \text{(since } a^ab^b/(a+b)^{a+b}\le 1 \text{ for all } a,b>0\text{)} \\
&\quad\le (O(\Delta^2)s/\ell)^{\ell}\sum_{k=\ell+1}^{\infty}(O(\Delta)s/n)^{k-\ell} \\
&\quad\le (O(\Delta^2)s/\ell)^{\ell}
\end{aligned} \tag{24}$$

Putting Eq. (22), Eq. (23) and Eq. (24) together, we get that for every $0<\ell\le 2s$,

$$S_1+S_2\le (O(\Delta)s/\ell)^{\ell}+(O(\Delta)s/\ell)^{\ell}+(O(\Delta^2)s/\ell)^{\ell}=(O(\Delta^2)s/\ell)^{\ell},$$

where we again used the assumption that $\ell>0$ to absorb a constant factor into the $O(\Delta)$ term. Substituting this bound in the derivations above, we note that for every $0<\ell\le 2s$,

$$\mathbf{E}_{M_2}\left[\left(\frac{2^n}{|\mathbf{B}_2|}\right)^2\sum_{\substack{v\in\{0,1\}^n\\|v|=2\ell}}\widehat{f_1\cdot f_2}(v)^2\,\middle|\,M_1,\mathbf{A}_1,\mathcal{E}_2\right]=(O(\Delta^2)s/\ell)^{\ell}.$$

Thus, Markov's inequality implies that with probability at least $1-O(\gamma)$, one has that for every $0<\ell\le 2s$, there exists an absolute constant $K>0$ such that

$$\mathbf{Pr}_{M_2}\left[\left(\frac{2^n}{|\mathbf{B}_2|}\right)^2\sum_{\substack{v\in\{0,1\}^n\\|v|=2\ell}}\widehat{f_1\cdot f_2}(v)^2>(K(\Delta^2/\gamma)s/\ell)^{\ell}\,\middle|\,M_1,\mathbf{A}_1,\mathcal{E}_2\right]\le\gamma^{\ell}. \tag{25}$$

Therefore, by a union bound over $0<\ell\le 2s$,

$$\begin{aligned}
&\mathbf{Pr}_{M_2}\left[\left(\frac{2^n}{|\mathbf{B}_2|}\right)^2\sum_{\substack{v\in\{0,1\}^n\\|v|=2\ell}}\widehat{f_1\cdot f_2}(v)^2>(K(\Delta^2/\gamma)s/\ell)^{\ell}\ \text{ for some }\ \ell\in[1,2s]\,\middle|\,M_1,\mathbf{A}_1,\mathcal{E}_2\right] \\
&\le\sum_{\ell\ge 1}\gamma^{\ell}=O(\gamma),
\end{aligned} \tag{26}$$

since γ is bounded from above by an absolute constant. We now define the event $\mathcal{E}$ (promised by the lemma) as the intersection of $\mathcal{E}_2$ and the success event for the application of Markov's inequality above. This completes the proof of Claim **(3)**, as desired. ■

5 Gap analysis (proof of Lemma 2.4)

In this section, we first give informal intuition about the existence of a MAX-CUT value gap in our instance and then give the formal proof. Recall that in the **YES** case, the MAX-CUT value is exactly the number of edges in the graph, as the graph is bipartite. The involved part of the argument consists of showing that our input graph is $\Omega(1)$-far from bipartite in the **NO** case, with high probability. Recall that our input MAX-CUT instance in the **NO** case is the union $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}$, where $\tilde{M}_1, \tilde{M}_2, \tilde{G}$ are obtained by first generating (M_1, M_2, G) from the distribution $\mathcal{P}_{n,\Delta,\alpha}$ and then keeping each edge independently with probability $1/2$. Note that the graph $\tilde{G}$ generated in this way is distributed as $\mathcal{G}_{n,\frac{1}{2}(1-\eta)/n}$ for a sufficiently small constant $\eta > 0$, i.e., it is slightly below half of the threshold for emergence of a giant component.

At a high level, the proof that $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}$ is $\Omega(1)$-far from bipartite proceeds by showing that $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}$ contains an $\Omega(n)$ size connected component (giant component) and then showing that this component is robust with respect to removal of a small positive fraction of the edges of the graph. To see why a giant component exists in our graph, it is useful to first see why it *does not exist* in the graph $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{M}_3$, where $\tilde{M}_1, \tilde{M}_2$ are (nearly)-perfect matchings subsampled at rate $1/2$ as above, and $\tilde{M}_3$ is a yet another perfect matching subsampled at rate $1/2$. Note that the number of edges in this graph is very close to the number of edges in our graph $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}$, as the expected degree of every vertex in $\tilde{G}$ is $\frac{1}{2}(1-\eta)$, just like in $\tilde{M}_3$. There is a subtle conditioning issue that precludes a giant component in $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{M}_3$. The reason is that $M_1 \cup M_2 \cup M_3$ is (almost) a 3-regular graph, and neighborhoods (of small sets) in it expand by less than a factor of 2 (see Fig. 2, left panel), so subsampling at rate $1/2$ pushes the process slightly below the critical limit and destroys the growth. This is because a vertex can only be incident to at most one edge of $\tilde{M}_i, i = 1, 2, 3$, as $\tilde{M}_i$ are matchings. We replace $\tilde{M}_3$ with an Erdős-Rényi graph, so that vertices can occasionally have degree more than 1 in it – and this pushes the process over the critical limit, leading to a giant component! This is illustrated in Fig. 2 (right panel; note the extra dashed edge on the right). In what follows we formally prove that the graph that we get in the **NO** case is $\Omega(1)$-far from bipartite.

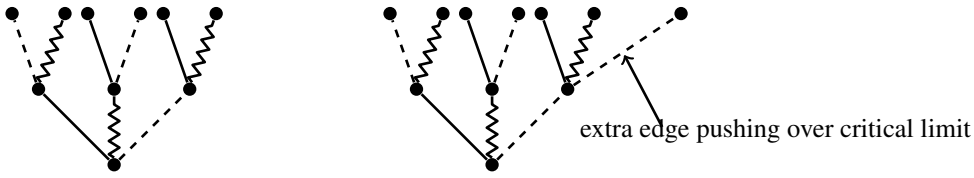

Figure 2: Illustration of neighborhood growth in a random 3-regular graph $M_1 \cup M_2 \cup M_3$ (left) and our instance $M_1 \cup M_2 \cup G_{n,(1-\eta)/n}$.

Distributions on Graphs. We will use the following two related distributions on n-vertex graphs. For a parameter $p \in (0, 1)$, we let $\mathcal{G}_{n,p}$ denote the Erdős-Rényi distribution with edge probability p. For every integer m, we let $\mathcal{G}_{n,m}$ denote the distribution on (multi)graphs with m edges, where a graph is selected by choosing m edges $e_t = (u_t, v_t), t = 1, \ldots, m$ independently and uniformly at random.

Recall that the process for generating our input random graph instance is as follows (see Section 2). We restate the process here for convenience of the reader.

Edge Sampling Process $\mathcal{P}_{n,\Delta,\alpha}$. Recall the process $\mathcal{P}_{n,\Delta,\alpha}$, which is used to sample the graphs (edge incidence matrices) M_1, M_2, M_3 in **DIHP**(n, Δ, α). We first describe how to generate M_1, M_2 and then describe how to generate M_3.

Sampling the matchings M_1, M_2. First, we generate the matchings M_1, M_2 as follows:

- Let M_1 be a perfect matching that matches, for each $i = 1, \ldots, n/2$, vertex i to vertex $i + n/2$. Note that edges of M_1 are naturally indexed by $[n/2]$: the i-th edge matches i to $i + n/2$, for $i \in [n/2] = \{1, 2, \ldots, n/2\}$.
- Choose a permutation π of $[n/2] = \{1, 2, \ldots, n/2\}$ uniformly at random. Partition the edges of M_1 into $r = n/\Delta$ sets $S_1, \ldots, S_r$ with $\Delta/2$ edges each by letting

$$S_j = \left\{ \pi\left(\frac{\Delta}{2} \cdot (j-1) + 1\right), \pi\left(\frac{\Delta}{2} \cdot (j-1) + 2\right), \ldots, \pi\left(\frac{\Delta}{2} \cdot (j-1) + \frac{\Delta}{2}\right) \right\}$$

  for each $j = 1, \ldots, n/\Delta$.
- For each $j = 1, \ldots, n/\Delta$ let $M_{2,j}$ match, for each $i = 1, \ldots, \Delta/2 - 1$, the node $\pi(\frac{\Delta}{2} \cdot (j-1) + i)$ to the node $\pi(\frac{\Delta}{2} \cdot (j-1) + i + 1) + n/2$. Note that $|M_{2,j}| = \frac{\Delta}{2} - 1$ for all j.

Let $M_2 := \bigcup_{j=1}^{r} M_{2,j}$. Note that $M_1 \cup M_2$ is a union of n/Δ disjoint paths of length $\Delta - 1$.

Sampling M_3. Having sampled M_1 and M_2, we now sample the graph M_3 in three steps:

(1) First, we sample an intermediate graph M_3', which is taken to be an Erdős-Rényi graph obtained by including every edge between vertices in $[n]$ independently with probability α/n (recall that $\alpha < 1$; we will choose α to be close to 1).

(2) Next, we form a graph M_3'' by removing any edges of M_3' that are already in M_1 or M_2. This serves the purpose of ensuring that our hard distribution is supported on simple graphs only. Note that the number of edges excluded from M_3' is at most a constant with high probability. This means, as we show below, that this slight change in the distribution does not affect analysis of the gap between MAX-CUT value in the **YES** and **NO** instances.

(3) Finally, we consider the connected components of the resulting graph M_3''. We remove all edges of each component that contains a cycle. We call the resulting graph M_3. Note that M_3 is guaranteed to contain no cycles.

Note that our input graph instance uses only the Erdős-Rényi distribution $\mathcal{G}_{n,p}$. The distribution $\mathcal{G}_{n,m}$ is used only in the proof. The main technical result of this section is the following lemma. As we show below, it leads directly to a proof of Lemma 2.4.

Lemma 5.1 *There exists $\eta^* > 0$ such that for every $\Delta > 10^4$ and every $\eta \in (0, \eta^*)$, $c > 0$, there exists $\delta > 0$ such that the following conditions hold for sufficiently large n: If M_1, M_2 are generated according to the process $\mathcal{P}_{n,\Delta,1-\eta}$, $\tilde{M}_1$ (resp. $\tilde{M}_2$) is obtained by sampling edges of M_1 (resp. M_2) independently with probability $1/2$, $\tilde{G}_1 \sim \mathcal{G}_{n,\frac{1}{2}(1-\eta)/n}$, and $\tilde{G}_2 \sim \mathcal{G}_{n,cn}$, then $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1 \cup \tilde{G}_2$ is δ-far from being bipartite with probability at least $97/100$.*

We now give a proof of Lemma 2.4, assuming Lemma 5.1. We restate Lemma 2.4 here for convenience of the reader:

Lemma 2.4*There exist constants $\Delta^* > 0$ and $0 < \alpha^* < 1$ such that for every $\alpha \in (\alpha^*, 1)$ and even integer $\Delta \geq \Delta^*$, there is a constant $\epsilon^* > 0$ for which the following conditions hold for the reduction R from Definition 2.3:*

(1) *If $\mathcal{I} = (M_1, M_2, M_3; w_1, w_2, w_3)$ is sampled from $\mathcal{D}^Y$ of $\mathbf{DIHP}(n, \Delta, \alpha)$, then $R(\mathcal{I})$ is a bipartite graph.*

(2) *If $\mathcal{I}$ is sampled from $\mathcal{D}^N$, then with probability at least 95/100, $R(\mathcal{I})$ is a graph on m edges with MAX-CUT value at most $(1 - \epsilon^*)m$.*

The proof uses the following claim:

Claim 5.2 *For any integer $m > 1$ and integers $1 \leq a \leq b \leq m$, if $X \subseteq [m]$ is a uniformly random subset of $[m]$ of size a and $Y \subseteq [m]$ is a uniformly random subset of $[m]$ of size b, then there exists a coupling between X and Y such that $X \subseteq Y$ with probability* 1.

Proof: To construct the coupling, it suffices to first sample $X \subseteq [m]$ of size a, then sample a uniformly random subset of $[m] \setminus X$ of size $b - a$ and let $Y := X \cup Y'$. To see that Y is uniformly random among all subsets of $[m]$ of size b, it suffices to note that for every pair $S, S' \subseteq [m], |S| = |S'| = b$ one has

$$\mathbf{Pr}[Y = S] = \mathbf{Pr}[X \subseteq S] \cdot \mathbf{Pr}[Y' = S \setminus X | X] = \mathbf{Pr}[X \subseteq S'] \cdot \mathbf{Pr}[Y' = S' \setminus X | X] = \mathbf{Pr}[Y = S'],$$

as the distributions of X and Y' are invariant under permutations of $[m]$ and $[m] \setminus X$ respectively. ■

Proof of Lemma 2.4:

In the **YES** case, the value of MAX-CUT is exactly m, the number of edges in the input graph, as the graph is bipartite by construction (the sides of the bipartition are given by $X^* \in \{0,1\}^n$).

We now show that in the **NO** case, the value of MAX-CUT is $(1-\Omega(1))m$. The proof proceeds over two steps. In the **first step**, we show that if $\tilde{G}_1 \sim \mathcal{G}_{n,\frac{1}{2}(1-\eta)/n}$, $\tilde{G}_2 \sim \mathcal{G}_{n,cn}$, $\tilde{G} \sim \mathcal{G}_{n,\alpha/(2n)}$, and $\alpha^* \geq 1-\eta+6c$, then $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}$ stochastically dominates $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1 \cup \tilde{G}_2$ on all but a vanishingly small fraction of the probability space. Note that the former is the distribution that Lemma 5.1 reasons about, and the latter is very close to our target distribution. Indeed, the distribution of $\tilde{M}_3$ is only different from the distribution of $\tilde{G}$ in at most $O(\log^2 n)$ edges with extremely high probability.

In the **second step**, we apply Lemma 5.1 to obtain the result.

Step 1. We first show that for every $c \in (0,1)$, we have that with probability $1 - o(1)$ over the choice of the edges of $\tilde{G}_2$, **(1)** the graph $\tilde{G}_1 \cup \tilde{G}_2$ has no edges of multiplicity higher than two, and **(2)** the total number of edges of multiplicity 2 is at most $O(\log n)$. Note that by the union bound, it suffices to show that both **(1)** and **(2)** individually occur with probability $1 - o(1)$. We prove this below.

Proof of (1) w.h.p. Note that an edge appears with multiplicity higher than two if either **(a)** at least three copies of the edge appear in $\tilde{G}_2$, or **(b)** one copy of the edge appears in $\tilde{G}_1$, while at least two copies appear in $\tilde{G}_2$. Note that by the union bound, **(a)** occurs with probability at most

$$\binom{cn}{3} \cdot \left(\frac{1}{\binom{n}{2}}\right)^2 \leq \frac{3c^3}{n},$$

since there are $\binom{cn}{3}$ possible triples of distinct edge indices (i_1, i_2, i_3) in $\tilde{G}_2$, and the probability that all three edges in a triple are copies of each other is $(1/\binom{n}{2})^2$.

Meanwhile, note that by the Chernoff bound, $\tilde{G}_1$ has at most $2n$ edges with probability $1 - o(1)$. Conditioned on this event, (2) occurs with probability at most

$$2n \cdot \binom{cn}{2} \cdot \frac{1}{\binom{n}{2}^2} \leq \frac{16c^2}{n}.$$

It follows that the graph $\tilde{G}_1 \cup \tilde{G}_2$ has no edges of multiplicity higher than two with probability $1 - o(1)$.

Proof of (2) w.h.p. Note that any edge that appears with multiplicity two must either **(a)** have both copies in $\tilde{G}_2$, or **(b)** have one copy in $\tilde{G}_1$ and one copy in $\tilde{G}_2$.

First, consider T, the number of edges of multiplicity two that obey **(a)**. Let $T_1, T_2, \ldots, T_{\binom{n}{2}}$ be indicator random variables such that T_i indicates whether edge i is sampled at least twice in $\tilde{G}_2$. Note that $T \leq T_1 + T_2 + \cdots + T_{\binom{n}{2}}$. Moreover, for every i, the probability that $T_i = 1$ is at most $\binom{cn}{2} \cdot (1/\binom{n}{2})^2$. Thus,

$$\mathbf{E}[T] \leq \mathbf{E}[T_1] + \mathbf{E}[T_2] + \cdots + \mathbf{E}[T_{\binom{n}{2}}] \leq \binom{n}{2} \cdot \binom{cn}{2} \cdot \left(\frac{1}{\binom{n}{2}}\right)^2 \leq 2c^2.$$

Moreover, since $T_1, T_2, \ldots, T_{\binom{n}{2}}$ are negatively associated, it follows from the Chernoff bound that $T = O(\log n)$ with probability $1 - n^{-\Omega(1)} = 1 - o(1)$.

Next, consider edges of multiplicity two that obey **(b)**. Note that the probability that any specific edge in $\tilde{G}_1$ also appears in $\tilde{G}_2$ is

$$1 - \left(1 - \frac{1}{\binom{n}{2}}\right)^{cn} \leq 1 - e^{-4c/n} \leq \frac{4c}{n}.$$

Recall that, by the Chernoff bound, $\tilde{G}_1$ has at most $2n$ edges with probability $1 - o(1)$. Conditioned on this event, the expectation of T', the number of edges of $\tilde{G}_1$ that also occur in $\tilde{G}_2$, is

$$\mathbf{E}[T'] \leq 2n \cdot \frac{4c}{n} = 8c.$$

Now, T' is the sum of indicator random variables for each edge of $\tilde{G}_1$ (which indicate whether the corresponding edge also appears in $\tilde{G}_2$). Since these variables are negatively associated, one can apply the Chernoff bound to T' in order to deduce that $T' = O(\log n)$ with probability $1 - n^{-\Omega(1)} = 1 - o(1)$, as desired.

Thus, ignoring the multiplicities of edges of $\tilde{G}_1 \cup \tilde{G}_2$ leads to an additive error of no more than $O(\log n)$ in any cut, and this is what we do in what follows—denote as $\tilde{G}^*$ the graph obtained from $\tilde{G}_1 \cup \tilde{G}_2$ by disregarding multiplicities. The expected number of edges in $\tilde{G}_1$ is $\frac{1-\eta}{4}(n-1)$. Thus, it follows by standard concentration inequalities, along with the aforementioned argument about edge multiplicities, that the number of edges in $\tilde{G}^*$ is at most $(\frac{1-\eta}{4} + c + o(1))n$ with probability $1 - o(1)$. Furthermore, conditioned on the number of edges in $\tilde{G}^*$ being equal to t for some t, the edge set of $\tilde{G}^*$ is a uniformly random set of edges of size t in $\binom{[n]}{2}$.

Similarly, it follows by concentration inequalities that the number of edges in $\tilde{G}$ is at least $(\frac{\alpha}{4} - o(1))n$ with probability $1 - o(1)$. Furthermore, conditioned on the number of edges being equal to t, the set of edges is a uniformly random set of size t in $\binom{[n]}{2}$.

Claim 5.2 therefore implies that the graph $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}$ stochastically dominates the graph $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1 \cup \tilde{G}_2$ with respect to inclusion as long as $\frac{\alpha}{4} > \frac{1-\eta}{4} + c + C'$ for an absolute constant $C' > 0$. We now let $\alpha^* = 1 - \eta^*/2$. For every $\alpha \in (\alpha^*, 1)$ we let $\eta = \eta^*$ and let $c = \eta^*/12$, so that

$$\frac{\alpha}{4} - \left(\frac{1-\eta}{4} + c\right) \geq \frac{1 - \eta^*/2}{4} - \left(\frac{1-\eta^*}{4} + \frac{\eta^*}{12}\right) \geq \frac{\eta^*}{4} - \frac{\eta^*}{8} - \frac{\eta^*}{12} = \frac{\eta^*}{24} = \Omega(1),$$

as required.

Step 2. Now, by Lemma 5.1 invoked with $\eta = \eta^*$ and $c = \eta^*/12$, as above, we get that the graph $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1 \cup \tilde{G}_2$ is $\Omega(\delta)$-far from bipartite for some $\delta = \Omega_{\eta^*}(1)$ with probability at least $97/100$. Since

the graph $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}$ stochastically dominates $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1 \cup \tilde{G}_2$ by **step 1** and both graphs contain $\Theta(n)$ edges with high probability, the graph $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}$ is also $O(\delta)$-far from being bipartite for some $\delta > 0$ with probability $1 - o(1)$.

Finally, it remains to note that the actual graph $\tilde{M}_3$ generated in the **NO** case only differs from $\tilde{G}$ in $O(\log^2 n)$ edges with probability $1 - o(1)$. It suffices to consider the edges that are removed from M_3' in order to produce M_3 in $\mathcal{P}_{n,\Delta,\alpha}$ and show that this number is $O(\log^2 n)$ with probability $1 - o(1)$. Recall that edges are removed in two stages. Consider the first stage, in which edges of M_3' that are already in M_1 or M_2 are removed in order to form M_3''. Since there are are $n - n/\Delta$ edges in $M_1 \cup M_2$, we have that the expected number of edges M_3' that are removed is $(n - n/\Delta) \cdot \alpha/n = \alpha(1 - 1/\Delta) \leq 1$. Thus, by the Chernoff bound, the number of removed edges is $O(\log n)$ with probability $1 - n^{-\Omega(1)}$.

Next, consider the second stage, in which edges are removed from M_3'' to form M_3. In order to bound the number of such edges, we use the following facts, which appear in [Dur06]:

Fact 5.3 (follows from Theorem 2.3.1 in [Dur06]) *Suppose $\lambda < 1$. Then, all connected components of $\mathcal{G}_{n,\lambda/n}$ are of size $O(\log n)$ with probability $1 - o(1)$.*

Fact 5.4 (follows from Theorem 2.6.1 in [Dur06]) *Suppose $\lambda < 1$, and let $A < \infty$ be a constant. Then, consider all connected components of $\mathcal{G}_{n,\lambda/n}$ with at most $A \log n$ vertices. With probability $1 - o(1)$, there are no complex components (i.e., components whose number of edges is at least 2 more than the number of vertices).*

Fact 5.5 (follows from Corollary 2.6.6 in [Dur06]) *Suppose $\lambda < 1$. Then, the expected number of unicyclic components (i.e., components whose number of edges equals the number of vertices) in $\mathcal{G}_{n,\lambda/n}$ is at most a constant $c = c(\lambda)$ (independent of n).*

Now, recall that M_3 is formed from M_3'' by removing all connected components that have cycles. Since such components are either unicyclic or complex, the above facts as well as the fact that the edges of M_3'' are a subset of the edges of M_3' imply that, with probability $1 - o(1)$, (1) there are no complex components in M_3'', (2) the number of unicyclic components in M_3'' is $O(\log n)$ (by Fact 5.5 and Markov's inequality), and (3) all connected components of M_3'' have size $O(\log n)$. Thus, it follows that $O(\log^2 n)$ edges are removed from M_3''.

Putting the bounds above together, we get that there exists a $1 + \epsilon_* = 1 + \Omega(1)$ gap between **YES** and **NO** instances with probability at least $97/100 + o(1) \geq 95/100$ for sufficiently large n.

■

The rest of the section is devoted to proving Lemma 5.1. The proof proceeds in two steps:

Step 1. We start by showing that the graph $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1$, where $\tilde{G}_1 \sim \mathcal{G}_{n,\frac{1}{2}(1-\eta)/n}$, contains a giant component of size $\Omega(n)$ with probability at least $99/100$ if $\eta > 0$ is smaller than an absolute constant. This proof is given in Section 5.1. The main result of that section is Lemma 5.6.

Step 2. We then condition on the existence of a giant component in $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1$ and show that the addition of a $\tilde{G}_2 \sim \mathcal{G}_{n,c_2 n}$ to the graph makes the union $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1$ δ-far from bipartite for some constant $\delta > 0$, as long as $c_2 > 0$ is a positive constant (so δ depends on c_2). This argument as well as a proof of Lemma 5.1 are provided in section 5.2.

5.1 Existence of a giant component

The goal of this subsection is to establish the lemma below:

Lemma 5.6 *There exists $\eta^* > 0$ such that for every $\Delta \geq 10^4$, $\eta \in (0, \eta^*)$ there exists $C_0 \geq 1$ such that the following conditions hold for sufficiently large n. If (M_1, M_2, G) is generated according to the process $\mathcal{P}_{n,\Delta,1-\eta}$, and $\tilde{M}_1, \tilde{M}_2, \tilde{G}$ are generated from M_1, M_2, G by sampling edges independently with probability $1/2$, then with probability at least $99/100$ the graph $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}$ contains a connected component of size at least n/C_0.*

The Component Growing Process: We will analyze the following process for growing a giant component. We will maintain at all times a partition of the vertex set V into two sets D and U where D denotes the set of vertices that have already been *discovered*, and U denotes the set of *undiscovered* vertices. We start with an arbitrary seed vertex, say s, and grow a component by iteratively including vertices that are reachable from s. Specifically, in the ith iteration, we start with a set A_i of the active vertices; initially, $A_0 = \{s\}$. Let $B_i \subseteq U$ denote the set of vertices in U that are connected to some vertex in A_i via edges in $\tilde{G}$. We next include all vertices reachable from vertices in B_i using edges in $\tilde{M}_1$ and $\tilde{M}_2$. Let C_i denote this set of vertices. We now set $A_{i+1} = B_i \cup C_i$, and add *all* vertices in the **blocks** of vertices in $B_i \cup C_i$ to the set D of discovered vertices; recall that the vertices are partitioned into blocks of size Δ. Finally, we set $U = V \setminus D$.

We say that an iteration i *succeeds* if $|A_{i+1}| \geq \frac{28}{25}|A_i|$, and it *fails* otherwise. We terminate this growth process if either an iteration fails or the size of the component has reached n/C_0. In the latter case, we are done, but if the growth process terminates due to failure at some iteration, we start the entire component growth process again starting with an arbitrary new seed vertex in U.

Overview of the Analysis: We will show that with probability at least $99/100$, one of the invocations of the component growth process described above reaches a connected component of size n/C_0. Let us first focus on the analysis of a single invocation of the component growth process. It is clear that we will obtain a connected component of size at least n/C_0 after $\Theta(\log n)$ successive iterations without any failure. The heart of the proof is to show the following lemma which bounds the probability of failure in any iteration to be exponentially small in the number of active vertices.

Lemma 5.7 *The probability that an iteration i succeeds is at least $1 - e^{-|A_i|/K}$ for some absolute constant $K > 0$ whenever $|U| \geq 9n/10$ at the start of the iteration.*

We now complete the analysis assuming the above lemma. First note that Lemma 5.7 implies that there is a positive probability p_0 that the component growth process succeeds for the first $\Theta(\log K)$ iterations (since K is a constant), allowing the growth process to reach an iteration j with $|A_j| \geq 10K$. Once the size of the active set A_j exceeds $10K$, the probability that any subsequent iteration fails while $|U| \geq 9n/10$, can now be bounded by

$$\sum_{\ell \geq 0} e^{-\frac{\left(\frac{28}{25}\right)^\ell |A_j|}{K}} \leq \sum_{\ell \geq 0} e^{-10\left(\frac{28}{25}\right)^\ell} \leq e^{-10} \sum_{\ell \geq 0} e^{\left(-\frac{28}{25}\right)^\ell} = \frac{e^{-10}}{1 - e^{-\frac{28}{25}}} \leq 10^{-3}.$$

Thus any single invocation of the component growth process finishes with a component of size at least n/C_0 with probability at least $p_0 \times 10^{-3}$, provided we satisfy the condition $|U| \geq 9n/10$ during the growth process. Note that any failed invocation of the component growth process removes at most $(n/C_0)\Delta$ vertices from U. Since we start with $|U| = n$, we can invoke the component growth process at least $\Gamma = \frac{(n/10)}{(n/C_0)\Delta} = \frac{C_0}{10\Delta}$ times before $|U|$ falls below $9n/10$. The probability that none of the first Γ invocations succeeds in growing a component of size at least n/C_0 is at most

$$\left(1 - \frac{p_0}{10^3}\right)^{\Gamma} \leq e^{-\frac{p_0\Gamma}{10^3}} = e^{-\frac{p_0 C_0}{10^4\Delta}} \leq \frac{1}{100},$$

provided we choose C_0 to be any constant greater than $5 \times (10^4\Delta)/p_0$. This completes the overview of our analysis assuming Lemma 5.7.

In the remainder of this section, we focus on establishing Lemma 5.7. Let $p = \frac{1}{2}(1-\eta)/n$ denote the probability of edge realization in the graph $\tilde{G}$. We will assume that η is chosen to be a fixed constant smaller than 10^{-2}, ensuring that $p \geq \frac{49}{100n}$. Our proof relies on the following two claims. We say that a vertex v is *well-placed* in a block P if v is at least distance 10 away from either end-point of P.

Claim 5.8 *Let $A \subseteq V \setminus U$ be an arbitrary subset of vertices of size at most $n/(10^3\Delta)$, and let $B \subseteq U$ denote the set of vertices that are adjacent to at least one vertex in A via edges in $\tilde{G}$. Furthermore, let $B' \subseteq B$ be any maximal subset of well-placed vertices such that B' contains at most one vertex from any block in U. Then for any $\Delta \geq 10^4$, whenever $|U| \geq 9n/10$,*

$$\mathbf{Pr}\left[|B'| > \frac{2|A|}{5}\right] \geq 1 - e^{-|A|/K_1},$$

for some positive constant K_1.

Proof: For any vertex $v \in U$, let q denote the probability that v is adjacent to one of the vertices in A via edges of $\tilde{G}$. Then

$$\begin{aligned}
q &= 1 - (1-p)^{|A|} \geq 1 - e^{-p|A|} \\
&\geq p|A| - \frac{p^2|A|^2}{2} \qquad (\text{as } e^{-x} \leq 1 - x + \frac{x^2}{2} \quad \forall x \geq 0) \\
&\geq p|A| - \frac{p|A|}{100} \geq \frac{99p|A|}{100} \qquad (\text{since } p \leq \frac{1}{2n} \text{ and } |A| \leq \frac{n}{10^3\Delta}).
\end{aligned}$$

Let $P_1, P_2, ..., P_k$ denote the blocks inside U where $k = |U|/\Delta \geq (9n)/(10\Delta)$. Let $Y = Y_1 + Y_2 + ... + Y_k$ where Y_i is a $0/1$-random variable that takes value 1 iff B contains a well-placed vertex v in P_i. Clearly, $|B'| = Y$, and it suffices to analyze the variable Y. Now

$$\begin{aligned}
\mathbf{Pr}[Y_i = 1] &= 1 - (1-q)^{\Delta - 20} \geq 1 - e^{-q(\Delta-20)} \\
&\geq q(\Delta - 20) - \frac{q^2(\Delta-20)^2}{2} \qquad (\text{as } e^{-x} \leq 1 - x + \frac{x^2}{2} \quad \forall x \geq 0) \\
&\geq q(\Delta - 20) - \frac{q(\Delta-20)}{2} \cdot (q(\Delta - 20)) \\
&\geq q(\Delta - 20) - \frac{q(\Delta-20)}{2} \cdot (p|A| \cdot (\Delta - 20)) \qquad (\text{since } q \leq p|A|) \\
&\geq q(\Delta - 20) - \frac{q(\Delta-20)}{2} \cdot \left(\frac{1}{2000}\right) \qquad (\text{since } p \leq \frac{1}{2n} \text{ and } |A| \leq \frac{n}{10^3\Delta}) \\
&\geq \frac{99q\Delta}{100} \qquad (\text{since } \Delta \geq 10^4)
\end{aligned}$$

Thus

$$\begin{aligned}\mathbf{E}[Y] &\geq k \cdot \frac{99q\Delta}{100} \geq \frac{9n}{10\Delta} \cdot \frac{99q\Delta}{100} \\ &\geq \frac{891n}{1000} \cdot q \geq \frac{891n}{1000} \cdot \frac{99p|A|}{100} \\ &\geq \frac{43}{100}|A|.\end{aligned}$$

An application of the standard Chernoff bound now suffices to conclude that $\mathbf{Pr}\left[Y > \frac{2|A|}{5}\right] \geq 1 - e^{-|A|/K_1}$ for some positive constant K_1.

■

Claim 5.9 *Suppose we are given a set B' of ℓ well-placed vertices in ℓ distinct blocks, say $P_1, P_2, ..., P_\ell$. Let C be the set of new vertices that are reachable in $\cup_{i=1}^{\ell} P_i$ from B' when we sample the edges in each P_i with probability $1/2$. Then for any $\Delta \geq 10^4$, we have*

$$\mathbf{Pr}\left[|C| \geq \frac{9\ell}{5}\right] \geq 1 - e^{-\ell/K_2},$$

for some positive constant K_2.

Proof: To analyze the number of vertices reachable in $\cup_{i=1}^{\ell} P_i$ from B' when we sample the edges in each P_i with probability $1/2$, it suffices to analyze the sum of truncated geometric random variables of the following form. Let $Z_1, Z_2, ..., Z_{2\ell}$ be identically distributed independent random variables where each Z_i indicates the number of successive heads seen when a fair coin is tossed 10 times. It is easy to see that the distribution of $|C|$ stochastically dominates the variable $Z = \sum_{i=1}^{2\ell} Z_i$. In what follows, we will show that

$$\mathbf{Pr}\left[|C| \geq \frac{9\ell}{5}\right] \geq 1 - e^{-\ell/K_2},$$

for some positive constant K_2.

To show this, we view each variable Z_i as sum of $0/1$ random variables, $Z_i^{(1)}, Z_i^{(2)}, ..., Z_i^{(10)}$ where the variable $Z_i^{(j)}$ is 1 iff $Z_i \geq j$. Thus $\mathbf{Pr}[Z_i^{(j)} = 1] = 2^{-j}$, and moreover, for any j, the variables $Z_1^{(j)}, Z_2^{(j)}, ..., Z_{2\ell}^{(j)}$ are independent and identically distributed. Let $Z^{(j)} = \sum_{i=1}^{2\ell} Z_i^{(j)}$. Since $E[Z^{(j)}] = (2\ell)/2^j$, an application of the standard Chernoff bound implies that

$$\mathbf{Pr}\left[Z^{(j)} < \frac{99}{100} \times \frac{2\ell}{2^j}\right] \leq e^{-\Omega(\ell/2^j)}.$$

By taking the union bound over all $j \in [1..10]$, we conclude that with probability at least $1 - e^{-\Omega(\ell/2^{10})}$, we have $Z^{(j)} \geq .99 \times \frac{2\ell}{2^j}$ for all $j \in [1..10]$. Hence

$$\mathbf{Pr}\left[Z \geq \frac{99}{100} \times 2\ell\left(\frac{1}{2} + \frac{1}{2^2} + ... + \frac{1}{2^{10}}\right)\right] \geq 1 - e^{-\Omega(\ell/2^{10})}.$$

Since $\frac{99}{100} \times 2\ell\left(\frac{1}{2} + \frac{1}{2^2} + ... + \frac{1}{2^{10}}\right) = \frac{99}{100} \times 2\ell\left(1 - \frac{1}{2^{10}}\right) \geq \frac{9\ell}{5}$, it follows that

$$\mathbf{Pr}\left[Z \geq \frac{9\ell}{5}\right] \leq e^{-\ell/K_2},$$

for a suitably large constant K_2 as desired. ■

Proof of Lemma 5.7: We now complete the proof of Lemma 5.7 using the above claims. Consider any iteration i in the component growth process such that $|U| \geq 9n/10$. Let A_i be the set of active vertices at the start of the iteration, and suppose that $|A_i| \leq n/(10^3\Delta)$ (otherwise, we already have a component of desired size). Let $B_i \subseteq U$ denote the set of vertices in U that are adjacent to some vertex in A_i via edges in $\tilde{G}$. Furthermore, let $B_i' \subseteq B_i$ be any maximal subset of well-placed vertices such that B' contains at most one vertex from any block in U. Then by Claim 5.8, the size of the set B_i' is at least $(2|A_i|)/5$ with probability at least $1 - e^{-|A_i|/K_1}$. Let us denote this event by $\mathcal{E}_1$. Now let C_i denote the set of vertices that are reachable from vertices in B_i' in the blocks containing them when we sample the edges in each block with probability $1/2$. By Claim 5.9, we have that $|C_i| \geq \frac{9|B_i'|}{5}$ with probability at least $e^{-|B_i'|/K_2}$.

Thus assuming that each of the events $\mathcal{E}_1$ and $\mathcal{E}_2$ occur, we have

$$|A_{i+1}| = |B_i| + |C_i| \geq |B_i'| + \frac{9|B_i'|}{5} = \frac{14|B_i'|}{5} \geq \frac{28}{25}|A_i|.$$

Finally, we observe that

$$\begin{aligned}
\mathbf{Pr}[\mathcal{E}_1 \wedge \mathcal{E}_2] &= \mathbf{Pr}[\mathcal{E}_1]\mathbf{Pr}[\mathcal{E}_2 \mid \mathcal{E}_1] \\
&\geq (1 - e^{-|A_i|/K_1})(1 - e^{-|B_i'|/K_2}) \\
&\geq (1 - e^{-|A_i|/K_1})(1 - e^{-2|A_i|/5K_2}) \\
&\geq 1 - e^{-|A_i|/K},
\end{aligned}$$

for a suitably large constant K as desired.

5.2 Distance to bipartiteness

The goal of this section is to establish Lemma 5.1, which we restate here for convenience of the reader:

Lemma 5.1 *(Restated) There exists $\eta^* > 0$ such that for every $\Delta > 10^4$ and every $\eta \in (0, \eta^*), c > 0$ there exists $\delta > 0$ such that the following conditions hold for sufficiently large n. If M_1, M_2 are generated according to the process $\mathcal{P}_{n,\Delta,1-\eta}$, $\tilde{M}_1, \tilde{M}_2$ obtained by sampling edges of M_1 (resp. M_2) independently with probability $1/2$, $\tilde{G}_1 \sim \mathcal{G}_{n,\frac{1}{2}(1-\eta)/n}$ and $\tilde{G}_2 \sim \mathcal{G}_{n,cn}$, then $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1 \cup \tilde{G}_2$ is δ-far from being bipartite with probability at least $97/100$.*

We start with an overview of the analysis. Since Lemma 5.6 (see Section 5.1) guarantees the existence of a giant component of $\Omega(n)$ size in $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1$, it suffices to argue that the addition of a random graph $\tilde{G}_2 \sim \mathcal{G}_{n,cn}$ to the giant component in $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1$ results in a graph that is $\Omega(1)$-far from bipartite with high probability. Our analysis in this section is completely oblivious to the process that generates the giant component: we will prove that the addition of G_2 to any graph with an $\Omega(n)$ size connected component that additionally exhibits certain regularity of vertex degrees makes this graph $\Omega(1)$-far from bipartite with high probability. Specifically, we show that the addition of a random graph $G_2 \sim \mathcal{G}_{n,cn}$ to any tree T of size n/C_0 for a constant $C_0 > 1$ with balanced vertex degrees (see Definition 5.10) results in a graph that is δ-far from bipartite for a constant δ with high probability. The proof proceeds over three steps.

Step 1 We show that if δ is small enough, the removal of δn edges from a tree T on at least n/C_0 vertices most of whose vertices have small degree results in a forest (denoted by F) that can be partitioned into a large number of rather large connected components (see Lemma 5.13 below). The formal definition of what it means for the tree to consist mostly of nodes of bounded degree is given in Definition 5.10. It is easy to see that the giant component in $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1$ satisfies these conditions: we show in Claim 5.12 that G_1 satisfies these conditions (by a simple Chernoff bound), and addition of matchings $\tilde{M}_1, \tilde{M}_2$ only leads to a slight change in parameters.

Step 2 After partitioning the forest F into connected components of size at least r, we contract these components into supernodes (denote the resulting graph by H). We then show that for any $c > 0$, if r is substantially larger than c, the effect of this is that the addition of the graph $\tilde{G} \sim \mathcal{G}_{n,cn}$ to H essentially amounts to sampling an Erdős-Rényi graph **well above the threshold for emergence of a giant component** on the supernodes H and hence makes $H \cup \tilde{G}$ non-bipartite with extremely high probability. This argument proceeds in two steps. We first write $\tilde{G} = \tilde{G}_1 \cup G_2$, where $\tilde{G}_1, \tilde{G}_2 \sim \mathcal{G}_{n,(c/2)n}$ are sampled independently. We then show by a simple union bound over all sufficiently large cuts in H that $H \cup \tilde{G}_1$ contains a linear size connected component with extremely high probability. This connected component has a unique bipartition, which the second graph $\tilde{G}_2$ destroys with high probability. The details of the proof are given in Lemma 5.16.

Step 3 Finally, the proof of Lemma 5.1 is obtained by a union bound over at most $O(\delta)n$ edges removed from G_2 and the result of Lemma 5.16.

Definition 5.10 *We say that a graph $G = (V, E)$ has (D, τ)-bounded degrees if there exists a subset $V^* \subseteq V$ of vertices such that*

1. *vertex degrees in $G^* = (V^*, E \cap (V^* \times V^*))$ are upper bounded by D;*
2. *$|E \cap (V^* \times V^*)| \geq |E| - \tau|V|$.*

Theorem 5.11 (Chernoff bound) *If $X_1, \ldots, X_n$ are independent $0/1$ random variables, $X = \sum_{i=1}^{n} X_i$ and $\mu = \mathbf{E}[X]$, then for any $\delta > 1$ one has $\mathbf{Pr}[X > (1+\delta)\mu] \leq e^{-\delta\mu/3}$.*

Claim 5.12 (Degree based pruning) *Let $G = (V, E)$ be a graph sampled from $\mathcal{G}_{n,p}$ distribution with $p < \frac{1}{n}$. Then for every $\lambda > 4$ with probability at least $99/100$ the graph G is $(\lambda, 100e^{-\lambda/6})$-bounded (as per Definition 5.10).*

Proof:

We need to argue the existence of a set $V^* \subseteq V$ such that

1. vertex degrees in $G^* = (V^*, E \cap (V^* \times V^*))$ are upper bounded by λ;
2. $|E \cap (V^* \times V^*))| \geq |E| - (100 \cdot e^{-\lambda/6})|V|$.

We let $V^* := \{u \in V : \deg_E(u) \leq \lambda\}$ and show that V^* satisfies both conditions above with probability at least $99/100$. First for each $u \in V$ define

$$X_u := \begin{cases} 1 & \text{if } \deg_E(u) > \lambda \\ 0 & \text{o.w.} \end{cases}$$

We now bound the expected number of edges incident on vertices with degrees larger than λ in G. For each fixed $u \in V$ one has $\deg_E(u) = \sum_{w \in V \setminus \{u\}} Y_w$, where Y_w is a $0/1$ Bernoulli random variable with on probability p. For every $z \in V$ one has

$$\begin{aligned} \mathbf{Pr}[(u,z) \in E \text{ and } \deg_E(u) > \lambda] &= \mathbf{Pr}[(u,z) \in E] \cdot \mathbf{Pr}[\deg_E(u) > \lambda | (u,z) \in E] \\ &\leq \frac{p}{n} \cdot \mathbf{Pr}\left[\sum_{w \in V \setminus \{u,z\}} Y_w > \lambda - 1\right] \\ &\leq \frac{p}{n} \cdot \mathbf{Pr}\left[\sum_{w \in V \setminus \{u,z\}} Y_w > \lambda/2\right] \quad (\text{since } \lambda > 4) \end{aligned}$$

One has $\mathbf{E}[\sum_{w \in V \setminus \{u,z\}} Y_w] \leq p(n-1) =: \mu$, and hence by Theorem 5.11 (Chernoff bound)

$$\mathbf{Pr}\left[\sum_{w \in V \setminus \{u,z\}} Y_w \geq (1+\delta)\mu\right] \leq e^{-\delta\mu/3}.$$

for every $\delta \geq 1$. We set $\delta := \lambda/\mu - 1$. Since $\lambda > 4$ and $\mu = p(n-1) < 1$ by assumption, we have $\delta = (\lambda/2)/\mu - 1 \geq \lambda/(4\mu) > 1$. This gives

$$\mathbf{Pr}[\sum_{w \in V \setminus \{u,z\}} Y_w > \lambda/2] \leq e^{-(\lambda/(2\mu))\mu/3} = e^{-\lambda/6},$$

and hence

$$\mathbf{Pr}[(u,z) \in E \text{ and } \deg_E(u) > \lambda] \leq \frac{1}{n} e^{-\lambda/6}.$$

We have, using the analysis above

$$\mathbf{E}[|\{(u,z) \in E : \deg_E(u) > \lambda\}|] \leq n e^{-\lambda/6}.$$

By Markov's inequality we thus have

$$\mathbf{Pr}[|V \setminus V^*| > 100 \cdot n e^{-\lambda/6}] \leq 1/100,$$

as required. ■

Claim 5.13 *For any forest $F = (V, E_F)$ with vertex degrees bounded by $D \geq 1$ and any $r \geq 1$, if all connected components in F have size at least r, then there exists a partitioning of $V = V_0 \cup V_1 \cup \ldots \cup V_K$ such that*

1. *$|V_j| \geq r$ for all $j \in [1:K]$;*
2. *$K \geq |E_F|/(2rD)$.*

Proof: We prove the bound for the case when F is a tree, and the desired result then follows by applying the bound to every tree in F.

Consider a tree $T = (V_T, E_T)$ with at least r nodes. Consider the following iterative procedure. Start by letting $T^0 \leftarrow T$, and letting $q \leftarrow 0$. Then for every $q \geq 0$, repeat the following until T^q contains fewer than r nodes. Root T^q arbitrarily, and let $u_q \in V_{T^q}$ be the furthest node from the root of T^q whose subtree $T^q_{u_q}$ contains at least r nodes. Remove the subtree $T^q_{u_q}$ from T^q, denote the remaining tree by T^{q+1} and repeat. This is formalized as Algorithm 1 below, where we denote the number of iterations that the loop runs for by Q. Note that a choice of the node u_q always exists, since the root itself satisfies the condition. Also note that the maximum degree in T^q is upper bounded by D for all q (T^q is a subtree of the original tree T^0, which satisfies this condition by assumption of the lemma). Also note that since u_q is the furthest node whose subtree contains at least r nodes, it must be that

$$|T^q_{u_q}| \leq 1 + \sum_{c \in T_{u_q} : c \text{ child of } u_q} |T_c| \leq 1 + D \cdot \max_{c \in T_{u_q} : c \text{ child of } u_q} |T_c| \leq (r-1)D + 1 \leq rD,$$

where we used the fact that u_q's children are strictly further from the root than u_q (and hence their subtrees contain fewer than r nodes) and the fact that maximum degree in T^q is upper bounded by D.

To prove the result of the lemma for trees, we note that

$$|V_T| \leq \sum_{q=0}^{Q-1} |T^q_{u_q}| + |T^Q| \leq \sum_{q=0}^{Q-1} |T^q_{u_q}| + r - 1$$
$$\leq 2 \sum_{q=0}^{Q-1} |T^q_{u_q}| \qquad \text{(since } |T^q_{u_q}| \geq r \text{ for all } q \in [0 : Q-1] \text{ and } Q \geq 1)$$

Since $|T^q_{u_q}| \leq rD$ for every q, and $Q \geq 1$ since $|T^0| \geq r$, we thus get $Q - 1 \geq |V_T|/(2rD)$, and hence $Q \geq |V_T|/(2rD) + 1 \geq |E_T|/(2rD)$.

Algorithm 1 Partitioning a tree into components of large size

```
1:  procedure PARTITIONTREE(T, r)
2:      T^0 ← T, q ← 0
3:      C ← ∅                                  ▷ Initialize collection of components to empty
4:      while |T^q| ≥ r do
5:          R ← root of T (arbitrarily chosen)
6:          u_q ← furthest node from R such that |T^q_u| ≥ r   ▷ Note that |T^q_R| ≥ r by loop condition, so u exists
7:          T^{q+1} ← T^q \ T^q_u
8:          C ← C ∪ {T^q_{u_q}}                ▷ Add subtree to collection of components
9:          q ← q + 1
10:     end while
11:     Q ← q
12:     return C
13: end procedure
```

■

Lemma 5.14 *For every $C_0 > 2$, $r \geq 1$, $\delta < 1/(8C_0 r)$, and every $\lambda \geq 3\log(1600C_0 r)$ the following conditions hold for sufficiently large n.*

For every tree $T = (V_T, E_T)$ with $V_T \geq V$, $|V_T| \geq n/C_0$ such that E_T is $(2\lambda, 100e^{-\lambda/6})$-bounded, for every $E^* \subseteq E_T$ with $|E^*| \leq \delta n$ there exists $E^{**} \subseteq E_T$ such that the forest $\tilde{F} := (V_T, E_F \setminus (E^* \cup E^{**}))$ consists of at least $n/(8C_0 \cdot \lambda r)$ components of size $\geq r$ each.

Proof: Since the tree T is $(2\lambda, 100e^{-\lambda/6})$ bounded by assumption, there exists a subset V^* of vertices in F such that vertex degrees in $F \cap (V^* \times V^*)$ are bounded by 2λ and $|E_F \cap (V^* \times V)| \geq |E_F| - 100e^{-\lambda/6}n$. Let

$$E'_F := (E_F \setminus E^*) \cap (V^* \times V^*).$$

Note that since $E'_F \subseteq E_F$, vertex degrees in E'_F are bounded by 2λ.

By assumption of the lemma we have $|E_T| = |V_T| - 1 \geq n/C_0 - 1$ and $|E^*| \leq \delta n$, so

$$|E'_F| \geq n/C_0 - 1 - (\delta n + 100e^{-\lambda/6}n) = N - 2C_0(\delta N + 100e^{-\lambda/6}N) = N(1 - 2C_0\delta + 200C_0 e^{-\lambda/6}),$$

where we let $N := n/C_0$ to simplify notation.

We now would like to invoke Lemma 5.13 on the set E'_F. Before we apply the lemma, however, we need to remove components of size below r from E'_F. Let E^{**} denote the set of edges of E'_F that belong to connected components of size at most r. We have

$$|E^{**}| \leq r \cdot 2C_0(\delta N + 100e^{-\lambda/6}N),$$

as **(1)** every such component contains an edge from E^* or $E_T \setminus (V^* \times V^*)$ and **(2)** every such component contains at most $r-1$ edges. Using the assumption that $\delta < 1/(8C_0r)$ and $\lambda \geq 3\log(1600C_0r)$ made in the claim, we get

$$r \cdot 2C_0(\delta N + 100e^{-\lambda/6}N) \leq (2C_0r(1/(8C_0r) + 1/(8C_0r)))N \leq (1/4 + 1/4)N \leq \frac{1}{2}N. \tag{27}$$

Let $E''_F := E_T \setminus (E^* \cup E^{**})$ denote the set of edges in E'_F that belong to a connected component of size $\geq r$ inE'_F, and note that $|E''_F| \geq N/2$ by (27).

We now apply Claim 5.13 to E''_F with parameter r. Since all components in E''_F have size $\geq r$ by construction and degrees are upper bounded by 2λ, we get that E''_F can be partitioned into $\geq (N/2)/(4\lambda r)$ components of size at least r each. Since $N \geq n/C_0$, this gives the result of the lemma. ■

Claim 5.15 *For every graph $H = (V_H, E_H), N = |V_H|$, if every cut in H with at least $N/3$ vertices on each side is non-empty, then the graph H contains a connected component of size at least $N/3$.*

Proof: The proof is by contradiction. We show that if all connected components in H are of size less than $N/3$, then there exists an empty cut in H with at least $N/3$ vertices on each side.

Let connected component sizes be $s_1 \leq \ldots \leq s_K$, where $1 \leq K \leq N$ is the number of connected components in H. Let k be the smallest such that $\sum_{j=1}^{k} s_j \geq N/3$. Since $\sum_{j=1}^{k-1} s_j < N/3$ by definition of k, and $s_j < N/3$ for all $j = 1, \ldots, K$, we have $\sum_{j=1}^{k} s_j = \sum_{j=1}^{k-1} s_j + s_k < N/3 + N/3 < 2N/3$. But in this case the graph contains a cut with at least $N/3$ vertices on one side which is empty. Indeed, take all of components $[1:k]$ on one side, and other vertices in H on the other side. The number of vertices on one size of this cutis $\sum_{j=1}^{k} s_j \in [N/3, 2N/3)$, leading to a contradiction. This completes the proof. ■

Lemma 5.16 *For every $C_0 > 1$ and every $c > 0$ there exists $r \geq 1$ and $\lambda \geq 3\log(1600C_0r)$ such that for every forest F with at least $n/(8 \cdot C_0 \cdot \lambda r)$ components of size $\geq r$ each, the graph $F \cup G'$ with $G' \sim \mathcal{G}_{n,cn}$ is not bipartite with probability at least $1 - \exp\left(-\Omega(n \cdot c/(C_0 \cdot \lambda r)^2)\right)$.*

Proof: Let V_H denote the set of nodes obtained from F by contracting each connected component into a **supernode**, and let N denote the number of resulting supernodes (i.e. $|V_H| = N$). We have $N \geq n/(8 \cdot C_0 \cdot \lambda r)$ by assumption of the lemma. It is also convenient two write $G' \sim \mathcal{G}_{n,cn}$ as $G' = G'_1 \cup G'_2$, where $G'_1, G'_2 \sim \mathcal{G}_{n,cn/2}$.

The proof proceeds in two steps. In **step 1** we show that every cut in H with at least $N/3$ nodes on each side is nonempty with very high probability, which implies that H contains a connected component of size at least $N/3$. In **step 2** we show that this implies that addition of the $\mathcal{G}_{n,cn}$ graph turns H into a non-bipartite graph with high probability, as long as the parameter r is chosen large enough (depending on c). We then show that such a choice indeed exists for every C_0, proving the claim.

Step 1. First note that for every pair of supernodes a, b we have that the expected number of edges in G'_1 between a and b is $\geq r^2c/(2n)$. The smallest number of edge slots going across a cut with at least $N/3$ supernodes on each side is $(N/3)(2N/3)$. Since each edge slot between a pair of supernodes corresponds

to at least r^2 edge slots in the original graph $G_1' \sim \mathcal{G}_{n,cn/2}$, we get that a fixed such cut is empty with probability at most

$$\left(1 - \frac{(2/9)N^2 \cdot r^2}{n^2}\right)^{cn/2} \leq e^{-\frac{(2/9)N^2 \cdot r^2}{n^2} \cdot cn/2} \qquad \text{(since } 1 - x \leq e^{-x} \text{ for } x \geq 0\text{)}$$
$$= e^{-(c/2)\frac{(2/9)N^2 \cdot r^2}{n}}.$$

Taking a union bound over at most 2^N cuts, we get using the relation $n \leq 8 \cdot C_0 N \lambda r$

$$\begin{aligned} 2^N \cdot e^{-(c/2)\frac{(2/9)N^2 \cdot r^2}{n}} &= 2^N \cdot \exp\left(-N \cdot \frac{(c/2)(2/9)N \cdot r^2}{n}\right) \\ &\leq 2^N \cdot \exp\left(-N \cdot \frac{(c/2)(2/9)(n/(8 \cdot C_0 \cdot \lambda r)) \cdot r^2}{n}\right) \qquad \text{(since } N \geq n/(8C \cdot C_0 \cdot \lambda r)\text{)} \\ &= 2^N \cdot \exp\left(-N \cdot \frac{r(c/2)(2/9)}{8 \cdot C_0 \cdot \lambda}\right) \\ &= 2^N \cdot \exp\left(-N \cdot \frac{rc}{72 \cdot C_0 \cdot \lambda}\right) \\ &\leq (2/e)^N \end{aligned} \tag{28}$$

as long as $r \geq \frac{72 \cdot C_0 \lambda}{c}$.

We now exhibit a setting of r and λ that satisfies this and the constraints of the lemma. Specifically, we need to show that there exists a setting of integer r and $\lambda > 2$ such that

$$r \geq \frac{72 \cdot C_0 \lambda}{c} \quad \text{and} \quad \lambda \geq 3\log(1600 C_0 r).$$

Equivalently (after exponentiating both sides of the second inequality above), we need to ensure that

$$r \geq \frac{72 \cdot C_0 \lambda}{c} \quad \text{and} \quad r \leq \frac{1}{1600} C_0^{-1} e^{\lambda/6}.$$

For every $C_0 > 1$ and $c > 0$ we let λ be a sufficiently large constant so that

$$1 \leq \frac{72 \cdot C_0 \lambda}{c} \text{ and } \frac{72 \cdot C_0 \lambda}{c} + 1 \leq \frac{1}{1600} C_0^{-1} e^{\lambda/6}.$$

Such a value of λ exists since $\frac{1}{1600} C_0^{-1} e^{\lambda/6}$ is grows asymptotically faster with λ than $\frac{72 \cdot C_0 \lambda}{c}$. Setting λ sufficiently large so that the interval $[\frac{72 \cdot C_0 \lambda}{c}, \frac{1}{1600} C_0^{-1} e^{\lambda/6}]$ contains an integer and letting r equal this integer satisfies all the constraints above.

Finally, it remains to recast the upper bound on failure probability from (28) in terms of n. We have, using the assumption that $N \geq n/(8 \cdot C_0 \cdot \lambda r)$

$$(2/e)^N \leq (2/e)^{n/(8 \cdot C_0 \cdot \lambda r)}$$

as required.

Step 2. By **Step 1** with probability at least $1 - (2/e)^{n/(8 \cdot C_0 \cdot \lambda r)}$ over the choice of G_1' no cut in $H \cup G_1'$ with at least $N/3$ supernodes on each side in H is empty. By Claim 5.15 this also implies that $H \cup G_1'$ contains a connected component with $\geq N/3$ supernodes. Denote the success event by $\mathcal{E}_1$. Let $C \subseteq [n]$

denote the set of nodes in this component (i.e. expanding the supernodes of H to the nodes that they represent). This connected component has a unique bipartition. Denote the sides of the bipartition by $A, B \subseteq [n]$. Note that $|A| + |B| \geq r \cdot N/3$, as each of the supernodes contains at least r nodes. We thus have $\min\{|A|, |B|\} \geq rN/6$. We now show that with overwhelming probability at least one of the edges of G_2' connects two nodes belonging to the same side of the bipartition $A \cup B$ of the large connected component in $H \cup G_1'$, thereby making $H \cup G_1' \cup G_2'$ non-bipartite.

Recall that $G_2' \sim \mathcal{G}_{n,(c/2)n}$ is obtained by selecting $(c/2)n$ edges uniformly at random from $\binom{[n]}{2}$. We thus have that a single such edge has both endpoints on the larger side of the bipartition with probability at least

$$(rN/6)^2 / \binom{n}{2} \geq 2(rN/6)^2/n^2$$

as long as $n \geq 2$. The probability that none of the $(c/2)n$ sampled edges of G_2' have both endpoints in the larger side of the bipartition is upper bounded by

$$\begin{aligned}
(1 - 2(rN/6)^2/n^2)^{(c/2)n} &\leq \exp(-2(rN/6)^2/n^2 \cdot (c/2)n) \\
&\leq \exp(-c(rN/6)^2/n) \qquad \text{(since } N \geq n/(8 \cdot C_0 \cdot \lambda r)\text{)} \\
&\leq \exp(-(c/(C_0 \cdot 48 \cdot \lambda)^2)n)
\end{aligned}$$

We thus get that, conditioned on $\mathcal{E}_1$, the graph $H \cup G_1' \cup G_2'$ is bipartite with probability at most $\exp(-(c/(C_0 \cdot 48 \cdot \lambda)^2)n)$. Hence, $H \cup G_1' \cup G_2'$ is not bipartite conditioned on $\mathcal{E}_1 \cap \mathcal{E}_2$, which satisfies

$$\begin{aligned}
\mathbf{Pr}[\mathcal{E}_1 \cap \mathcal{E}_2] &\geq 1 - \mathbf{Pr}_{G_1'}[\bar{\mathcal{E}}_1] - \mathbf{Pr}_{G_2'}[\bar{\mathcal{E}}_2 | \mathcal{E}_1] \\
&\geq 1 - (2/e)^{n/(8 \cdot C_0 \cdot \lambda r)} - \exp\left(-\Omega(n \cdot c/(C_0 \cdot \lambda)^2)\right) \\
&\geq 1 - \exp\left(-\Omega(n \cdot c/(C_0 \cdot \lambda r)^2)\right)
\end{aligned}$$

as required.

■

We now prove Lemma 5.1, the main result of this section:

Proof of Lemma 5.1: Let η^* and C_0 denote the constants whose existence is guaranteed by Lemma 5.6. Denote the graph $\tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1$ by $H = (V, E_H)$. By Lemma 5.6 with probability at least $99/100$ the graph H contains a giant component $V_{gc} \subseteq V$ of size at least n/C_0 (denote the success event by $\mathcal{E}_{gc}$). We condition on $\mathcal{E}_{gc}$ in what follows. Let $T := (V, E_T)$, where E_T is an arbitrarily chosen spanning tree of V_{gc} in E_H.

We prove:

> **(*)** There exists $\delta > 0$ that depends only on c and C_0 such that for any subset $E^* \subseteq E_T$ with $|E^*| \leq \delta n$ the following conditions hold for $F := (V, E_F), E_F = E_T \setminus E_1^*$, then the graph $F \cup \tilde{G}_2$ is δ-far from bipartite with probability at least $98/100$ over the choice of $\tilde{G}_2 \sim \mathcal{G}_{n,cn}, \tilde{G}_2 = (V, E_2)$.

Claim **(*)** implies the required result after a union bound over the failure event from **(*)** and $\bar{\mathcal{E}}_{gc}$, leading to $97/100$ success probability overall.

We prove **(*)** using a union bound over all choices of $E^{**} \subseteq E_2$. Formally, recall that $\tilde{G}_2$ contains $m = cn$ edges $e_1, \ldots, e_m$, where each e_i is independently chosen from $\binom{V}{2}$ (we say that the corresponding edge has index i). For each $J \subseteq \{1, 2, \ldots, m\}$ we let $\tilde{G}_2(J) = (V, E_2 \setminus E(J))$ denote the random graph obtained from $\tilde{G}_2$ by removing edges with indices in J. Our proof proceeds by a union bound over the choices of $J \subseteq [cn]$, as we describe next.

For any fixed $J \subseteq \{1, 2, \ldots, m\}$ the graph $\tilde{G}_2(V, E_2 \setminus J)$ is distributed as $\mathcal{G}_{n,m-|J|}$. We assume that $\delta < c/2$, so that $\mathcal{G}_{n,m-|J|}$ stochastically dominates $\mathcal{G}_{n,(c/2)n}$. Let $\tilde{G}_2(J) := G_2(V, E_2 \setminus E(J))$. We now show that $H \cup \tilde{G}_2(J)$ is non-bipartite with extremely high probability for any fixed $J \subseteq [cn]$, and then apply a union bound over all $J \subseteq [cn]$ to conclude the result.

Step 1. First note that by Claim 5.12 one has for any $\lambda > 4$ the graph $\tilde{G}_1$ is $(\lambda, 100e^{-\lambda/6})$-bounded with probability at least $99/100$. Since $H = \tilde{M}_1 \cup \tilde{M}_2 \cup \tilde{G}_1$, where $\tilde{M}_1$ and $\tilde{M}_2$ are matchings, and $\lambda \geq 4$, we thus have that H is $(2\lambda, 100e^{-\lambda/6})$ bounded with probability at least $99/100$. Denote this event by $\mathcal{E}_{bounded}$.

Step 2. Now by Lemma 5.14 we have for every $r \geq 1$, $\lambda \geq 3\log(1600C_0 r)$ and $\delta < 1/(8C_0 r)$ for every $E_2^* \subseteq E_H$ with $|E_2^*| \leq \delta n$ there exists $E_2^{**} \subseteq E_T$ such that $\tilde{F} := (V_T, E_F \setminus (E_2^* \cup E_2^{**}))$ consists of at least $n/(8C_0 \cdot \lambda r)$ components of size $\geq r$ each.

Step 3. Then by Lemma 5.16 there exists $r \geq 1$ and $\lambda \geq 3\log(1600C_0 r)$ such that the graph $\tilde{F} \cup G'$ with $G' \sim \mathcal{G}_{n,(c/2)n}$ is not bipartite with probability at least $1 - e^{-\Omega((c/(C_0\lambda r)^2)n}$.

Steps 1-3 above show that for any fixed $J \subseteq \{1, 2, \ldots, m\}$ the graph $H \cup \tilde{G}_2(J)$ is non-bipartite with probability at least $1 - e^{-\Omega((c/(C_0\lambda r)^2)n}$. To obtain the final result, we take a union bound over possible choices of the set $J \subseteq \{1, 2, \ldots, m\}$. The number of such choices is bounded by $\binom{cn}{\delta n} \leq (e \cdot c/\delta)^{\delta n} = e^{(\delta \ln(ec/\delta))n}$. Using the fact that $\delta \ln(ec/\delta)$ is increasing in δ for $\delta \in (0, 1/10)$, we can choose δ to be a sufficiently small constant so that

$$e^{(\delta \ln(ec/\delta))n} \cdot e^{-\Omega((c/(C_0\lambda r)^2)n} < 1/100.$$

Note that λ and r we chosen as functions of C_0 only, so δ depends only on c and C_0, as required by (*).

We have shown that conditioned on $\mathcal{E}_{gc}$ the probability of $H \cup \tilde{G}_2$ being δ-close to bipartite is upper bounded by $1/100 + \mathbf{Pr}[\bar{\mathcal{E}}_{bounded}] \leq 2/100$. This proves **(*)** and completes the proof.

■

6 Proof of Lemma 3.3 (distance to uniformity)

In this section, we prove Lemma 3.3, which we restate below.

Lemma 3.3 *Let $\Delta > 0$ be an even integer. Then, for every $0 < \alpha < 1$, there exists a constant $0 < c < 1$ such that for every $\delta \in (n^{-1/10}, c)$, the following conditions hold if n is any sufficiently large multiple of Δ:*

- **(1)** *Let $\mathbf{B} = \mathbf{A}_1$, as defined in Definition 3.2. Then, for every choice of matchings M_1, M_2 sampled according to $\mathcal{P}_{n,\Delta,\alpha}$, the distribution of $M_2 x$ is uniform over $\{0,1\}^{m_2}$ when x is uniformly random in $\mathbf{B}$.*
- **(2)** *Let $\mathbf{B} \subseteq \{0,1\}^n, |\mathbf{B}| = 2^{n-z}$ for $z \leq \delta^4 n$, and let $h : \{0,1\}^n \to \{0,1\}$ be the indicator of $\mathbf{B}$. If $\left(\frac{2^n}{|\mathbf{B}|}\right)^2 \sum_{v:|v|=2\ell} \hat{h}(v)^2 \leq \left(\frac{64\delta^4 n}{\ell}\right)^{\ell}$ holds for all $\ell \leq \delta^4 n$, then the following conditions hold: Let M_1, M_2, M_3 be sampled according to $\mathcal{P}_{n,\Delta,\alpha}$. Then, with probability at least $1 - O(\delta)$ over the choice of M_3, the total variation distance between the distribution of $M_3 x$, where x is uniformly random in $\mathbf{B}$, and the uniform distribution over $\{0,1\}^{m_3}$ is $O(\delta/\sqrt{1-\alpha})$. In particular, one can take $c = \min\left\{\left(\frac{1-\alpha}{512}\right)^{1/4}, \left(\frac{e^{-\alpha}\log_2(32/(31+\alpha))}{32}\right)^{1/4}\right\}$.*

In order to establish Lemma 3.3, we will require the following lemmas, which we prove later.

Lemma 6.1 *Let $n \geq 2$, and let $\Delta > 0$ be any even integer such that Δ divides n. Then, for every fixed constant $0 < \alpha < 1$, the following statement holds: Suppose M_1 and M_2 are sampled according to $\mathcal{P}_{n,\Delta,\alpha}$. If $v \in \{0,1\}^n$ such that $v \neq 0^n$ and $\widehat{h_1}(v) \neq 0$, then*

$$|\{s \in \{0,1\}^{m_2} : v = M_2^T s\}| = 0.$$

Lemma 6.2 *Let $n \geq 2$, and let $\Delta > 0$ be any even integer such that Δ divides n. Then, for every fixed constant $0 < \alpha < 1$, the following statement holds: Assume that M_1 and M_2 have been sampled according to $\mathcal{P}_{n,\Delta,\alpha}$. Then, if $v \in \{0,1\}^n$ has even weight ℓ, we have*

$$\mathbf{E}_{M_3}\left[|s \in \{0,1\}^{m_3} : s \neq 0^{m_3}, v = M_3^T s|\right] \leq 2^{\ell}(\ell/2)!(\alpha/(1-\alpha)n)^{\ell/2}.$$

Note that the lemma appears almost exactly in [KKS15], except that instead of sampling the Erdős-Rényi graph $\mathcal{G}_{n,\alpha/n}$, we sample all individual edges except those in M_1, M_2 with probability α/n. However, leaving out the edges of M_1, M_2 can only cause the expectation on the left-hand side to decrease. Hence, the proof is similar.

Fact 6.3 *Let $n \geq 2$, and let $\Delta > 0$ be any even integer such that Δ divides n. Then, for every fixed constant $0 < \alpha < 1$, the following statement holds: Let M_3 be sampled according to $\mathcal{P}_{n,\Delta,\alpha}$. Then, for every $v \in \{0,1\}^n$, there exists at most one $s \in \{0,1\}^{m_3}$ such that $M_3^T s = v$.*

In light of the above fact, it is also clear that there exists an s such that $M_3^T s = v$ if and only if the intersection of the support of v with each component of M_3 has even size.

Lemma 6.4 *Let $\Delta > 0$ be any even integer. Then, for every choice of fixed constants $0 < \alpha, \beta, c < 1$, the following statement holds for $b = c\log_2(1/(1-\beta))/2$ and sufficiently large $n \geq 2$ divisible by Δ: Let M_3 be sampled according to $\mathcal{P}_{n,\Delta,\alpha}$, and let S be the random variable denoting the set of all vertices u such that $\{u\}$ is a connected component of M_3. Also, let $V \subseteq \{1,2,\ldots,n\}$ be a uniformly random set of nodes such that $|V| = \ell$, where $\ell \geq \beta n$. Then, if $|S| \geq cn$, we have $\mathbf{Pr}_V[V \cap S = \emptyset] \leq 2^{-bn}$.*

Lemma 6.5 *Let $\Delta > 0$ be any even integer. Then, for every choice of fixed constants $0 < \alpha, \beta < 1$, the following statement holds for $b = e^{-\alpha}\log_2(1/(1-\beta))/16$ and sufficiently large $n \geq 2$ divisible by Δ: Let M_3 be sampled according to $\mathcal{P}_{n,\Delta,\alpha}$. There exists an event $\mathcal{E}$ (depending on M_3) with $\mathbf{Pr}_{M_3}[\mathcal{E}] \geq 1 - 6n^{-1/3}$ such that for every $v \in \{0,1\}^n$ with $|v| \geq \beta n$,*

$$\mathbf{E}_{M_3}\left[|\{s \in \{0,1\}^{m_3} : v = M_3^T s\}| \mid \mathcal{E}\right] \leq 2^{-bn}.$$

Now, we prove Lemma 3.3 assuming the validity of the aforementioned lemmas and facts.

Proof of Lemma 3.3: Let $\beta = (1-\alpha)/32$, and let $b = e^{-\alpha}\log_2(1/(1-\beta))/16$ be the constant guaranteed by Lemma 6.5 for our choice of α, β. We will choose $c = \min\{((1-\alpha)/512)^{1/4}, (b/2)^{1/4}\}$. For the remainder of the proof, we assume that $\delta \in (n^{-1/10}, c)$.

Now, for any $z \in \{0,1\}^m$ and $m \times n$ edge incidence matrix M of a graph on n vertices, we let

$$p_M(z) = \frac{|\{x \in \mathbf{B} : Mx = z\}|}{|\mathbf{B}|}.$$

Later in the proof, we will instantiate M as M_2 and M_3. Note that $p_M(z)$ is a function of $\mathbf{B}$. We will supress this dependence in what follows to simplify notation. This will not cause any ambiguity since $\mathbf{B}$ is fixed as a typical large set arising from Alice's partition. We would like to prove that $p_M(z)$ is close to uniform. We

will do that by bounding the Fourier mass in positive weight coefficients of $p_M(z)$. By the same calculation as in [GKK$^+$08] (Lemma 10), we have

$$\begin{aligned}
\widehat{p_M}(s) &= \frac{1}{2^m} \sum_{z \in \{0,1\}^m} p_M(z)(-1)^{z \cdot s} \\
&= \frac{1}{|\mathbf{B}|2^m} \left(|\{x \in \mathbf{B} : (Mx) \cdot s = 0\}| - |\{x \in \mathbf{B} : (Mx) \cdot s = 1\}| \right) \\
&= \frac{1}{|\mathbf{B}|2^m} \left(|\{x \in \mathbf{B} : x \cdot (M^T s) = 0\}| - |\{x \in \mathbf{B} : x \cdot (M^T s) = 1\}| \right) \\
&= \frac{1}{|\mathbf{B}|2^m} \sum_{x \in \{0,1\}^n} h(x) \cdot (-1)^{x \cdot (M^T s)} \\
&= \frac{2^n}{|\mathbf{B}|2^m} \widehat{h}(M^T s),
\end{aligned}$$

where h is the indicator function of $\mathbf{B}$, and

$$\begin{aligned}
||p_M - U_r||_{tvd}^2 &\leq 2^m ||p_M - U_r||_2^2 \\
&= 2^{2m} \sum_{\substack{s \in \{0,1\}^m \\ s \neq 0}} \widehat{p_M}(s)^2 \\
&= \frac{2^{2n}}{|\mathbf{B}|^2} \sum_{\substack{s \in \{0,1\}^m \\ s \neq 0}} \widehat{h}(M^T s)^2 \qquad (29) \\
&= \frac{2^{2n}}{|\mathbf{B}|^2} \sum_{v \in \{0,1\}^n} \widehat{h}(v)^2 \cdot |\{s \in \{0,1\}^m : s \neq 0, v = M^T s\}|.
\end{aligned}$$

Here, the first transition in (29) holds by Cauchy-Schwarz, the subsequent equality is a result of Parseval's equality, and U_r is the uniform distribution over $\{0,1\}^m$.

Now, let us prove part (1) of the lemma statement. We fix a perfect matching M_1. Recall that we are interested in the distribution of $M_2 x$, where x is uniformly random in $\mathbf{B}$. Then, by (29) and Lemma 6.1, we have

$$\begin{aligned}
\|p_{M_2} - U_r\|_{tvd}^2 &\leq \frac{2^{2n}}{|\mathbf{B}|^2} \sum_{\substack{s \in \{0,1\}^{m_2} \\ s \neq 0^{m_2}}} \widehat{h_1}(M_2^T s)^2 \\
&= \frac{2^{2n}}{|\mathbf{B}|^2} \sum_{\substack{v \in \{0,1\}^n \\ v \neq 0^n}} \widehat{h_1}(v)^2 \cdot |\{s \in \{0,1\}^{m_2} : v = M_2^T s\}| \qquad (30) \\
&= 0,
\end{aligned}$$

which proves the claim.

Next, we prove part (2) of the statement of Lemma 3.3. Suppose that M_1 and M_2 have already been sampled according to $\mathcal{P}_{n,\Delta,\alpha}$. Then, let $\mathcal{E}$ be the event guaranteed by Lemma 6.5. Note that $\mathbf{Pr}[\mathcal{E}] \geq 1 - 6n^{-1/3} \geq 1 - 6\delta^2$.

By (29), we have

$$\begin{aligned}\mathbf{E}_{M_3}\left[\|p_{M_3}-U_r\|_{tvd}^2 \mid \mathcal{E}\right] &\le \frac{2^{2n}}{|\mathbf{B}|^2}\mathbf{E}_{M_3}\left[\sum_{\substack{s\in\{0,1\}^{m_3}\\ s\neq 0^{m_3}}} \widehat{h_2}(M_3^T s)^2 \;\middle|\; \mathcal{E}\right] \\ &= \frac{2^{2n}}{|\mathbf{B}|^2}\sum_{v\in\{0,1\}^n} \widehat{h_2}(v)^2 \cdot \mathbf{E}_{M_3}\left[|\{s\in\{0,1\}^{m_3} : s\neq 0^{m_3}, v = M_3^T s\}| \mid \mathcal{E}\right] \\ &= \frac{2^{2n}}{|\mathbf{B}|^2}\sum_{\substack{0\le\ell\le n\\ \text{even } \ell}}\sum_{\substack{v\in\{0,1\}^n\\ |v|=\ell}} \widehat{h_2}(v)^2 \cdot \mathbf{E}_{M_3}\left[|\{s\in\{0,1\}^{m_3} : s\neq 0^{m_3}, v = M_3^T s\}| \,|\mathcal{E}\right].\end{aligned} \tag{31}$$

Note that in the above sum, ℓ is restricted to be even, since the fact that all rows of M_3 have even weight implies that any v in the row space of M_3 must also have even weight.

Now, we split (31) into three sums S_1, S_2, S_3 over different ranges of ℓ. In particular, S_1, S_2, and S_3 will be the sums over the ranges $\ell \in [0, \delta^4 n]$, $\ell \in (\delta^4 n, \beta n)$, and $\ell \in [\beta n, n]$, respectively.

Bounding S_1. First, note that by Lemma 6.2, we have

$$\begin{aligned}S_1 &= \frac{2^{2n}}{|\mathbf{B}|^2}\sum_{\substack{0\le\ell\le\delta^4 n\\ \text{even } \ell}}\sum_{\substack{v\in\{0,1\}^n\\ |v|=\ell}} \widehat{h_2}(v)^2 \cdot \mathbf{E}_{M_3}\left[|\{s\in\{0,1\}^{m_3} : s\neq 0^{m_3}, v = M_3^T s\}| \mid \mathcal{E}\right] \\ &\le \frac{2^{2n}}{|\mathbf{B}|^2}\sum_{\substack{0\le\ell\le\delta^4 n\\ \text{even } \ell}} \frac{1}{\mathbf{Pr}[\mathcal{E}]}\cdot 2^\ell(\ell/2)!\left(\frac{\alpha}{(1-\alpha)n}\right)^{\ell/2}\sum_{\substack{v\in\{0,1\}^n\\ |v|=\ell}} \widehat{h_2}(v)^2 \\ &\le \sum_{\substack{0\le\ell\le\delta^4 n\\ \text{even } \ell}} 2\cdot 2^\ell(\ell/2)!\left(\frac{\alpha}{(1-\alpha)n}\right)^{\ell/2}\left(\frac{64\delta^4 n}{\ell/2}\right)^{\ell/2} \\ &\le 2\cdot\sum_{\substack{0\le\ell\le\delta^4 n\\ \text{even } \ell}}\left(\frac{256\alpha\delta^4}{1-\alpha}\right)^{\ell/2} \\ &= O\left(\delta^4/(1-\alpha)\right),\end{aligned} \tag{32}$$

since $\delta < c < ((1-\alpha)/512)^{1/4}$.

Bounding S_2. Next, we bound S_2 as follows, again using Lemma 6.2:

$$\begin{aligned}
S_2 &= \frac{2^{2n}}{|\mathbf{B}|^2} \sum_{\substack{\delta^4 n < \ell < \beta n \\ \text{even } \ell}} \sum_{\substack{v \in \{0,1\}^n \\ |v| = \ell}} \widehat{h_2}(v)^2 \cdot \mathbf{E}_{M_3} \left[|\{s \in \{0,1\}^{m_3} : s \neq 0^{m_3}, v = M_3^T s\}| \mid \mathcal{E} \right] \\
&\leq \frac{2^{2n}}{|\mathbf{B}|^2} \cdot \|\widehat{h_2}\|^2 \cdot \max_{\substack{\delta^4 n < \ell < \beta n \\ \text{even } \ell}} \left\{ \frac{1}{\mathbf{Pr}[\mathcal{E}]} \cdot 2^\ell (\ell/2)! \left(\frac{\alpha}{(1-\alpha)n} \right)^{\ell/2} \right\} \\
&= \frac{2^n}{|\mathbf{B}|} \max_{\substack{\delta^4 n < \ell < \beta n \\ \text{even } \ell}} \left\{ 2 \cdot (2\alpha\ell/(1-\alpha)n)^{\ell/2} \right\} \\
&\leq 2^{\delta^4 n} \cdot \max_{\substack{\delta^4 n < \ell < \beta n \\ \text{even } \ell}} \left\{ (4\alpha\beta/(1-\alpha))^{\ell/2} \right\} \\
&\leq 2^{\delta^4 n} \cdot 8^{-\delta^4 n/2} \\
&= 2^{-\Omega(\delta^4 n)}, \qquad (33)
\end{aligned}$$

since $4\alpha\beta/(1-\alpha) = \alpha/8 < 1/8$.

Bounding S_3. Finally, by Lemma 6.5, we can bound S_3 as follows:

$$\begin{aligned}
S_3 &= \frac{2^{2n}}{|\mathbf{B}|^2} \sum_{\substack{\beta n \leq \ell \leq n \\ \text{even } \ell}} \sum_{\substack{v \in \{0,1\}^n \\ |v| = \ell}} \widehat{h_2}(v)^2 \cdot \mathbf{E}_{M_3} \left[|\{s \in \{0,1\}^{m} : v = M_3^T s\}| \mid \mathcal{E} \right] \\
&\leq \frac{2^{2n}}{|\mathbf{B}|^2} \sum_{\substack{\beta n \leq \ell \leq n \\ \text{even } \ell}} \sum_{\substack{v \in \{0,1\}^n \\ |v| = \ell}} \widehat{h_2}(v)^2 \cdot 2^{-bn} \\
&\leq \frac{2^{2n}}{|\mathbf{B}|^2} \cdot 2^{-bn} \sum_{\substack{\beta n \leq \ell \leq n \\ \text{even } \ell}} \sum_{\substack{v \in \{0,1\}^n \\ |v| = \ell}} \widehat{h_2}(v)^2 \\
&\leq \frac{2^{2n}}{|\mathbf{B}|^2} \cdot 2^{-bn} \cdot \frac{|\mathbf{B}|}{2^n} \\
&\leq 2^{\delta^4 n - bn} \\
&\leq 2^{-\Omega(\delta^4 n)}, \qquad (34)
\end{aligned}$$

where the final step uses the fact that $b > 2\delta^4$.

Now, we can combine (31), (32), (33), and (34) to obtain

$$\begin{aligned}
\mathbf{E}_{M_3} \left[\|p_{M_3} - U_r\|_{tvd}^2 \mid \mathcal{E} \right] &\leq S_1 + S_2 + S_3 \\
&= O(\delta^4/(1-\alpha)) + 2^{-\Omega(\delta^4 n)} + 2^{-\Omega(\delta^4 n)} \\
&= O(\delta^4/(1-\alpha)),
\end{aligned}$$

where we use the fact that $\delta > n^{-1/10}$.

Finally, recall that $\mathbf{Pr}[\mathcal{E}] \geq 1 - 6n^{-1/3} \geq 1 - 6\delta^2$. Since $\|p_{M_3} - U_r\|_{tvd}$ is always at most 1, we have that

$$\begin{aligned}\mathbf{E}_{M_3}\left[\|p_{M_3} - U_r\|_{tvd}^2\right] &\leq \mathbf{E}_{M_3}\left[\|p_{M_3} - U_r\|_{tvd}^2 \mid \mathcal{E}\right] \cdot \mathbf{Pr}[\mathcal{E}] + \mathbf{Pr}[\bar{\mathcal{E}}] \\ &\leq \mathbf{E}_{M_3}\left[\|p_{M_3} - U_r\|_{tvd}^2 \mid \mathcal{E}\right] + O(\delta) \\ &\leq O(\delta^4/(1-\alpha)) + O(\delta^2) \\ &= O(\delta^2/(1-\alpha)).\end{aligned}$$

Thus,

$$\begin{aligned}\mathbf{E}_{M_3}\left[\|p_{M_3} - U_r\|_{tvd}\right] &\leq \sqrt{\mathbf{E}_{M_3}\left[\|p_{M_3} - U_r\|_{tvd}^2\right]} \\ &= O(\delta/\sqrt{1-\alpha}),\end{aligned}$$

as desired. ■

We now prove the supporting lemmas and facts:

Proof of Lemma 6.1: Note that by part (1) of Theorem 4.6, $\widehat{h_1}(v) \neq 0$ implies that v is supported on edges of M_1. Now, consider any $v \neq 0^n$ for which $\widehat{h_1}(v) \neq 0$. Then by part (2) of Lemma 4.5, we have that for every $w \in \{0,1\}^n$ supported on the edges of M_2, $|v + w| > 0$. Thus, v is not in the column space of M_2^T, which proves the claim. ■

Proof of Lemma 6.2: Note that any s satisfying $M_3^T s = v$ must consist of a union of edge-disjoint paths connecting endpoints in the support of v and cycles. Since M_3 does not have any cycles, by definition, it follows that such an s must simply be a union of paths.

Let s be the union of paths $P_1, P_2, \ldots, P_{\ell/2}$ connecting nonzero coordinates of v. Fix a pairing of the ℓ nonzero coordinates of v. Then, for any single path P_i, we have $\mathbf{Pr}[P_i \subseteq M_3] = (\alpha/n)^q$, where q is the length of P_i. Thus, by a union bound over all path lengths $q \geq 1$ and all paths connecting the $(2i-1)$-st nonzero coordinate of v to the $2i$-th nonzero coordinate, we have

$$\mathbf{Pr}[P_i \subseteq M_3] \leq \sum_{q \geq 1} n^{q-1} \cdot (\alpha/n)^q \leq \frac{\alpha}{(1-\alpha)n}.$$

Since $P_1, P_2, \ldots, P_{\ell/2}$ are edge disjoint, we have

$$\mathbf{Pr}[P_i \subseteq M_3 \text{ for all } i = 1, \ldots, \ell/2] \leq \prod_{i=1}^{\ell/2} \mathbf{Pr}[P_i \subseteq M_3] \leq (\alpha/(1-\alpha)n)^{\ell/2}.$$

Finally, since there are at most $\frac{\ell!}{2^{\ell/2}(\ell/2)!}$ ways to pair up the ℓ nonzero coordinates of v, it follows that

$$\begin{aligned}\mathbf{E}_{M_3}\left[|s \in \{0,1\}^m : s \neq 0^m, v = M_3^T s|\right] &\leq \frac{\ell!}{2^{\ell/2}(\ell/2)!} \cdot (\alpha/(1-\alpha)n)^{\ell/2} \\ &\leq 2^{\ell}(\ell/2)!(\alpha/(1-\alpha)n)^{\ell/2},\end{aligned}$$

as desired. ■

Proof of Fact 6.3: Any s satisfying $M_3^T s = v$ must correspond to an edge-disjoint union of paths connecting pairs of nodes in the support of v along with cycles. Since M_3 is, by design, guaranteed to contain no cycles,

we have that such an s must consist of an edge-disjoint union of paths. Moreover, since each connected component of M_3 is a tree, there exists a unique (if at all) selection of edges in each component that can be contained in s. Thus, the desired claim follows. ■

Proof of Lemma 6.4: Consider a subset W of the n vertices chosen as follows: For each vertex, we independently include it in W with probability $p = \ell/n$ and exclude it with probability $1 - p$.

It is apparent that

$$\mathbf{Pr}_W[W \cap S = \emptyset] \leq (1-p)^{|S|} \leq (1-p)^{cn} \leq (1-\beta)^{cn}.$$

Now, we pass from the sampling process for W to V. Note that V is precisely the random variable obtained by conditioning W on the event $|W| = \ell$. Moreover,

$$\begin{aligned}\mathbf{Pr}_W[|W| = \ell] &= \binom{n}{\ell} p^{\ell}(1-p)^{n-\ell} \\ &= \binom{n}{pn} p^{pn}(1-p)^{(1-p)n} \\ &\geq \frac{1}{n+1}.\end{aligned}$$

Therefore,

$$\begin{aligned}(1-\beta)^{cn} &\geq \mathbf{Pr}_W[W \cap S = \emptyset] \\ &\geq \mathbf{Pr}_W[|W| = \ell] \cdot \mathbf{Pr}_W[W \cap S = \emptyset \mid |W| = \ell] \\ &\geq \frac{1}{n+1} \cdot \mathbf{Pr}_V[V \cap S = \emptyset],\end{aligned}$$

which implies that for sufficiently large n, $\mathbf{Pr}_V[V \cap S = \emptyset] \leq (n+1)(1-\beta)^{cn} \leq 2^{-bn}$ for $b = c\log_2(1/(1-\beta))/2$, as desired. ■

Proof of Lemma 6.5: Recall that M_3 is formed by first sampling M_3' and removing edges. Let T_1' be the random variable equal to the number of connected components of M_3' that consist of a single vertex. Then, let $\mathcal{E} = \mathcal{E}(\alpha)$ denote the event that $T_1' \geq cn$, where $c = e^{-\alpha}/8$ is a constant depending on α.

We now show that for sufficiently large n, event $\mathcal{E}$ occurs with high probability, namely,

$$\mathbf{Pr}[\mathcal{E}] \geq 1 - 6n^{-1/3}. \tag{35}$$

The proof follows the proof of the more general Theorem 2.6.3 in [Dur06], but we reproduce it here with our choice of parameters for the sake of completeness.

First, let us calculate the expected value of T_1'. Observe that the probability that any given vertex lies in a connected component by itself is $\left(1 - \frac{\alpha}{n}\right)^{n-1}$. Thus,

$$\mathbf{E}_{M_3}[T_1'] = n\left(1 - \frac{\alpha}{n}\right)^{n-1},$$

This implies that $\lim_{n\to\infty} \mathbf{E}_{M_3}[T_1']/n = e^{-\alpha}$.

Now, we establish concentration to the mean. Let us count the number of ordered pairs (u_1, u_2) of distinct vertices u_1, u_2 that each lie in a connected component by themselves. The expected number of such pairs is

$$\begin{aligned}n\left(1 - \frac{\alpha}{n}\right)^{n-1} \cdot (n-1)\left(1 - \frac{\alpha}{n}\right)^{n-2} &\leq (\mathbf{E}_{M_3}[T_1'])^2 \left(1 - \frac{\alpha}{n}\right)^{-1} \\ &\leq (\mathbf{E}_{M_3}[T_1'])^2 e^{\alpha/n}.\end{aligned}$$

Thus, we can now compute the variance of T_1':

$$\begin{aligned}\mathrm{var}(T_1') = \mathbf{E}_{M_3}[T_1'^2] - \mathbf{E}_{M_3}[T_1']^2 &= \mathbf{E}_{M_3}[T_1'(T_1'-1)] + \mathbf{E}_{M_3}[T_1'] - \mathbf{E}_{M_3}[T_1']^2 \\ &\leq \left(e^{\alpha/n} - 1\right)\mathbf{E}_{M_3}[T_1']^2 + \mathbf{E}_{M_3}[T_1'].\end{aligned}$$

It now follows from Chebyshev's Inequality that

$$\begin{aligned}\mathbf{Pr}_{M_3}\left[\left|T_1' - \mathbf{E}_{M_3}[T_1']\right| \geq n^{2/3}\right] &\leq \frac{\mathrm{var}(T_1')}{(n^{2/3})^2} \\ &\leq \frac{\left(e^{\alpha/n}-1\right)\mathbf{E}_{M_3}[T_1']^2 + \mathbf{E}_{M_3}[T_1']}{n^{4/3}} \\ &\leq \frac{(2\alpha/n)\cdot(2e^{-\alpha}n)^2 + 2e^{-\alpha}n}{n^{4/3}} \\ &\leq 6n^{-1/3}.\end{aligned}$$

for sufficiently large n. Moreover, if $|T_1' - \mathbf{E}_{M_3}[T_1']| < n^{2/3}$, then

$$\begin{aligned}T_1' &> \mathbf{E}_{M_3}[T_1'] - n^{2/3} \\ &> (e^{-\alpha}/2)n - n^{2/3} \\ &> cn\end{aligned}$$

for sufficiently large n. This establishes (35).

Let S' denote the set of vertices u such that $\{u\}$ is a connected component of M_3'. Similarly, we define S to be the set of vertices u such that $\{u\}$ is a connected component of M_3. Note that by Fact 6.3, if $|v| = \ell \geq \beta n$, then

$$\begin{aligned}\mathbf{E}_{M_3}\left[\left|\{s \in \{0,1\}^m : v = M_3^T s\}\right| \mid \mathcal{E}\right] &\leq \mathbf{Pr}_{M_3}[\mathrm{supp}(v) \cap S = \emptyset \mid \mathcal{E}] \\ &\leq \mathbf{Pr}_{M_3'}[\mathrm{supp}(v) \cap S' = \emptyset \mid \mathcal{E}], \qquad (36)\end{aligned}$$

where the second inequality follows from the fact that $S' \subseteq S$. Since the distribution of S' is invariant with respect to permutations of the vertices, it follows from symmetry that if $v, v' \in \{0,1\}^n$ with $|v| = |v'|$, then

$$\mathbf{Pr}_{M_3'}[\mathrm{supp}(v) \cap S' = \emptyset \mid \mathcal{E}] = \mathbf{Pr}_{M_3'}[\mathrm{supp}(v') \cap S' = \emptyset \mid \mathcal{E}]$$

Thus, letting V be a uniformly random set of ℓ vertices, we see that Lemma 6.4 implies

$$\mathbf{Pr}_{M_3'}[\mathrm{supp}(v) \cap S' = \emptyset \mid \mathcal{E}] = \mathbf{Pr}_{M_3',V}[V \cap S' = \emptyset \mid \mathcal{E}] \leq 2^{-bn}$$

for $b = c\log_2(1/(1-\beta))/2 = e^{-\alpha}\log_2(1/(1-\beta))/16$ and sufficiently large n. Combining this with (36) yields

$$\mathbf{E}_{M_3}\left[\left|\{s \in \{0,1\}^m : v = M_3^T s\}\right| \mid \mathcal{E}\right] \leq 2^{-bn},$$

as desired. ■

7 Basic bounds on Fourier mass

In this section we prove some useful properties of Fourier coefficients of boolean functions that arise in our analysis. We start by recalling notation, and then proceed to the proofs. As before, we denote the

message posted on the board by player $t, t \in \{0,1,2\}$ by m_t, where $m_0 = 0$ for convenience. Similarly, the matchings $M_t, t = 1,2$ are made available to all players. This means that for each t the t-th player chooses a function $g_t : \{0,1\}^{M_t} \to \{0,1\}^s$ that depends on the messages $m_1, \ldots, m_{t-1}$ as well as the matchings $M_1, \ldots, M_t$. Then the player computes $m_t = g_t(M_t x)$. The space $\{0,1\}^{M_t}$ will be referred to as the *reduced space* (as opposed to the space $\{0,1\}^n$ from which a partition x is drawn uniformly at random).

Let $\mathbf{A}_t \subseteq \{0,1\}^{M_t}, t = 1, \ldots, T$ be typical messages in the reduced space for round t. Let

$$\mathbf{B}_t = \{x \in \{0,1\}^n : g_t(M_t x) \in \mathbf{A}_t\}$$

be the typical messages in the full space for rounds $t = 1,2$. Let $f_t : \{0,1\}^n \to \{0,1\}$ denote indicator functions of the sets $\mathbf{A}_t$. Note that $f_1 \cdot f_2$ is the indicator of $\mathbf{B}_2 = \mathbf{A}_1 \cap \mathbf{A}_2$ (the indicator of $\mathbf{B}_1 = \mathbf{A}_1$ is f_1).

The bounds that we prove in this section can be summarized as follows. First, we show that the Fourier transform of f_t is supported only on sets of vertices that can be perfectly matched by $M_t, t \in \{1,2\}$. Second, we show that the ℓ_2 mass cannot be to concentrated in every subcube in the Fourier space. The bounds are summarized formally in Theorem 4.6. The first bound is proved in Lemma 7.4, and the second bound is proved in Lemma 7.3.

We will use

Lemma 7.1 ([KKL88]) *Let f be a function $f : \{0,1\}^n \to \{-1,0,1\}$. Let $\mathbf{A} = \{x \in \{0,1\}^n : f(x) \neq 0\}$. Let $|s|$ denote the Hamming weight of $s \in \{0,1\}^n$. Then for every $\delta \in [0,1]$*

$$\sum_{s \in \{0,1\}^n} \delta^{|s|} \hat{f}(s)^2 \leq \left(\frac{|\mathbf{A}|}{2^n}\right)^{\frac{2}{1+\delta}}.$$

Our main result in this section is

Theorem 4.6 (Restated) *Let $M \in \{0,1\}^{m \times n}$ be the incidence matrix of a matching M, where the rows correspond to edges e of M ($M_{eu} = 1$ if e is incident on u and 0 otherwise). Let $g : \{0,1\}^m \to \{0,1\}^s$ for some $s > 0$. Let $a \in \{0,1\}^s$ and let $\mathbf{A}_{reduced} := \{z \in \{0,1\}^m : g(z) = a\}$. Further, let $f : \{0,1\}^n \to \{0,1\}$ denote the indicator of the set*

$$\mathbf{A} := \{x \in \{0,1\}^n : g(Mx) = a\}.$$

Suppose that $|\mathbf{A}| = 2^{n-d}$ for some $d \in [0,n]$.

Then

1. *the only nonzero Fourier coefficients of $\widehat{f}$ are of the form $\widehat{f}(M^T w)$ for some $w \in \{0,1\}^M$;*
2. *for all $\ell \in [0:d]$ and every $Q \subseteq M$*

$$2^{2d} \sum_{\substack{v \in \{0,1\}^n, |v| = 2\ell + |Q| \\ v \supseteq Q}} \widehat{f}(v)^2 \leq 2^{|Q|} (64d/\ell)^\ell,$$

   where $|Q|$ denotes the number of vertices in Q;

3. $2^{2d} \sum_{v \in \{0,1\}^n} \widehat{f}(v)^2 = 2^d$ *(Parseval's equality).*

Remark 7.2 *Note that Theorem 4.6 has two parameters related to the size of the set $\mathbf{A}$: s and d. The first parameter is the dimensionality of the binary cube that the function g maps to. The second parameter is d gives the size of the set $\mathbf{A}_{full}$. For a 'typical' set $\mathbf{A}_{full}$ we expect that $\mathbf{A}_{full}$ occupies a 2^{-s} fraction of the hypercube 2^n, i.e. $d = s \pm O(1)$.*

The proof of Theorem 4.6 relies on Lemma 7.3 (stated below), whose proof is given in section 7.2. The lemma provides a natural extension of the KKL-based bounds used in previous works. It shows that ℓ_2 mass of an indicator function of a large subset of the boolean cube cannot be too concentrated on low weight coefficients of subcubes:

Lemma 7.3 *Let $\mathbf{A} \subseteq \{0,1\}^n$ be a set such that $|\mathbf{A}|/2^n = 2^{-d}$. Let $f : \{0,1\}^n \to \{0,1\}$ be the indicator of $\mathbf{A}$. Then for every set $S \subseteq [n]$ and every $k \in \{1, 2, \ldots, d\}$ one has*

$$2^{2d} \sum_{v \in \{0,1\}^n, wt(v)=k+|S|, S \subseteq v} \hat{f}(v)^2 \leq 2^{2|S|}(64d/k)^k.$$

We note that the case $|S| = \emptyset$ corresponds to Lemma 6 in [GKK$^+$08]. Lemma 6 in [GKK$^+$08] shows that (appropriately normalized) Fourier transform of the indicator function of a large subset of the boolean cube cannot have too much mass on the low weight coefficients. Lemma 7.3 shows that the normalized Fourier mass cannot be too concentrated on low weight Fourier coefficients inside a fixed subcube.

We will also need Lemma 7.4, stated below:

Lemma 7.4 *Let $M \in \{0,1\}^{m \times n}$ be the incidence matrix of a matching M, where the rows correspond to edges e of M ($M_{eu} = 1$ if e is incident on u and 0 otherwise). Let $g : \{0,1\}^m \to \{0,1\}^s$ for some $s > 0$. Let $a \in \{0,1\}^s$ and let $q : \{0,1\}^m \to \{0,1\}$ be the indicator of the set $\mathbf{A}_{reduced} := \{z \in \{0,1\}^m : g(z) = a\}$. Further, let $f : \{0,1\}^n \to \{0,1\}$ denote the indicator of the set*

$$\mathbf{A}_{full} := \{x \in \{0,1\}^n : g(Mx) = a\}.$$

Then for every $v \in \{0,1\}^n$

$$\hat{f}(v) = \begin{cases} 0, & \text{if } v \text{ cannot be perfectly matched via edges of } M \\ \hat{q}(w), & w \text{ the perfect matching of } v \text{ using edges of } M \text{ o.w.} \end{cases} \quad (37)$$

Furthermore, the perfect matching of v, when it exists, is unique. The second condition above is equivalent to the existence of $w \in \{0,1\}^m = \{0,1\}^M$ such that $v = M^T w$. Thus, Fourier coefficients of f are indexed by sets of edges of M. Note that nonzero weight k coefficients of $\hat{g}$ are in one to one correspondence with nonzero weight $2k$ coefficients of $\hat{f}$, i.e. the only nonzero Fourier coefficients of $\widehat{f}$ are of the form $\widehat{f}(M^T w) = \widehat{g}(w)$ for some $w \in \{0,1\}^M$.

Given these two lemmas, the proof of Theorem 4.6 follows:

Proof of Theorem 4.6: By Lemma 7.4 nonzero Fourier coefficients of f are of the form $\widehat{f}(M^T r) = \widehat{g}(r), r \in \{0,1\}^M$, and in particular $|M^T r| = 2|r|$. By Lemma 7.3 applied to $q : \{0,1\}^m \to \{0,1\}$ with S as the set of edges of M that perfectly match all nodes in Q we have

$$2^{2d} \sum_{v \in \{0,1\}^n, wt(v)=k+|S|, S \subseteq v} \hat{q}(v)^2 \leq 2^{2|S|}(64d/k)^k$$

for every $k \leq d$. Putting these two facts together and using the fact that $|Q| = 2|S|$ gives

$$2^{2d} \sum_{v \in \{0,1\}^n, wt(v)=2k+|Q|, Q \subseteq v} \hat{f}(v)^2 \leq 2^{|Q|}(64d/k)^k$$

for every $k \leq d$, as required. ■

In the remainder of this section we prove Lemma 7.4 in section 7.1, then prove Lemma 7.3 in section 7.2.

7.1 Properties of Fourier transform of f_t

First, we prove

Lemma 7.4 *Let $M \in \{0,1\}^{m\times n}$ be the incidence matrix of a matching M, where the rows correspond to edges e of M ($M_{eu} = 1$ if e is incident on u and 0 otherwise). Let $g : \{0,1\}^m \to \{0,1\}^s$ for some $s > 0$. Let $a \in \{0,1\}^s$ and let $q : \{0,1\}^m \to \{0,1\}$ be the indicator of the set $\mathbf{A}_{reduced} := \{z \in \{0,1\}^m : g(z) = a\}$. Further, let $f : \{0,1\}^n \to \{0,1\}$ denote the indicator of the set*

$$\mathbf{A}_{full} := \{x \in \{0,1\}^n : g(Mx) = a\}.$$

Then for every $v \in \{0,1\}^n$

$$\hat{f}(v) = \begin{cases} 0, & \text{if } v \text{ cannot be perfectly matched via edges of } M \\ \hat{q}(w), & w \text{ the perfect matching of } v \text{ using edges of } M \text{ o.w.} \end{cases} \tag{38}$$

Furthermore, the perfect matching of v, when it exists, is unique. The second condition above is equivalent to the existence of $w \in \{0,1\}^m = \{0,1\}^M$ such that $v = M^T w$. Thus, Fourier coefficients of f are indexed by sets of edges of M. Note that nonzero weight k coefficients of $\hat{g}$ are in one to one correspondence with nonzero weight $2k$ coefficients of $\hat{f}$, i.e. the only nonzero Fourier coefficients of $\widehat{f}$ are of the form $\widehat{f}(M^T w) = \widehat{g}(w)$ for some $w \in \{0,1\}^M$.

Proof: We compute the Fourier transform of $g(x)$. For $z \in \{0,1\}^m$ let $x(z)$ be defined by setting, for each edge $(u,v) \in M$,

$$x(z)_u = z_e \text{ and } x(z)_v = 0$$

and $x(z)_w = 0$ if w is not matched by M. Note that $x(z)$ is a particular solution of $Mx = z$. Note that the set of solutions is given by

$$\{x(z) + Ns : s \in \{0,1\}^{n-m}\}, \tag{39}$$

where N is a basis for the nullspace of M. Without loss of generality suppose that M contains the edges $(2i-1, 2i), i = 1, \ldots, m$. Then the matrix $N \in \{0,1\}^{n\times(n-m)}$ may be taken as

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \ldots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \ldots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 1 \end{pmatrix},$$

where the bottom right submatrix is an $(n-m)\times(n-m)$ identity and the top left is M^T.

The Fourier transform of f at $v \in \{0,1\}^n$ is given by

$$\begin{aligned}\hat{f}(v) &= \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x) \cdot (-1)^{x \cdot v} \\ &= \frac{1}{2^n} \sum_{z \in \mathbf{A}_{reduced}} \sum_{s \in \{0,1\}^{n-m}} (-1)^{(x(z)+Ns) \cdot v} \\ &= \frac{1}{2^n} \sum_{z \in \mathbf{A}_{reduced}} (-1)^{x(z) \cdot v} \sum_{s \in \{0,1\}^{n-m}} (-1)^{(v^T N) \cdot s}\end{aligned}$$

First note that

$$\sum_{s \in \{0,1\}^{n-m}} (-1)^{(v^T N) \cdot s} = \mathbf{1}_{v^T N = 0} \cdot 2^{n-m},$$

so $\hat{f}(v) = 0$ unless $v^T N = 0$. Note that all such v are of the form $v = M^T r$ for some $r \in \{0,1\}^m$.

Thus,

$$\hat{f}(M^T r) = \frac{2^{n-m}}{2^n} \sum_{z \in \mathbf{A}_{reduced}} (-1)^{x(z) \cdot M^T r} = \frac{2^{n-m}}{2^n} \sum_{z \in \mathbf{A}_{reduced}} (-1)^{z \cdot r} = \hat{q}(r)$$

and $\hat{f}(v) = 0$ for all v not of the form $M^T r$. Note that we use the fact that $x(z) \cdot M^T r = z \cdot r$ for all z and r.

Note that Fourier coefficients of f only have even weight and weight k Fourier coefficients of g are in direct correspondence with weight $k/2$ coefficients of f (since $|M^T r| = 2|r|$ for all $r \in \{0,1\}^m$). ■

7.2 KKL on subcubes

Proof of Lemma 7.3:

We use the notation $[n]$ for the set of elements $\{1, 2, \ldots, n\}$.For a vector $x \in \{0,1\}^n$ we write supp(x) to denote the set of nonzeros in x. For a vector $x \in \{0,1\}^n$ and a set $S \subseteq [n]$ we write $x_S \in \{0,1\}^S$ to denote the restriction of x to coordinates in S.

Let f denote the indicator of $\mathbf{A}$, $|\mathbf{A}|/2^n = 2^{-d}$. For $\alpha \in \{0,1\}^n$ with supp$(\alpha) \subseteq S$ let

$$\mathbf{B}_\alpha := \{x \in \mathbf{A} : x_S = \alpha_S\},$$

and let $g_\alpha(x)$ denote the indicator of $\mathbf{B}_\alpha$. Note that

$$g_\alpha(x) = f(x) \cdot \mathbf{1}_{x_S = \alpha_S}(x), \tag{40}$$

where $\mathbf{1}_{x_S = \alpha_S}$ is the indicator of the set $\{x \in \{0,1\}^n : x_S = \alpha_S\}$.

The proof proceeds as follows. We first relate energy of the Fourier transform of g_α for a random α to the energy of the Fourier transform of f (see Eq. (46)). This relation shows that the expected energy of g_α contributed by weight k Fourier coefficients lower bounds the sum of energies of Fourier coefficients of f that have weight $k + |S|$ and contain the set S (this is the rhs of the inequality that we would like to prove). By an averaging argument there exists α^* such that the sum of Fourier coefficients of g_{α^*} provides the same lower bound (see Eq. (47)). We then apply KKL to g_{α^*}, obtaining the required bound (see Eq. (48)).

By Eq. (40) together with Eq. (1) the Fourier transform of g_α for all $v \in \{0,1\}^n$ equals

$$\hat{g}_\alpha(v) = \sum_{w \in \{0,1\}^n} \hat{f}(v+w) \widehat{\mathbf{1}_{x_S = \alpha_S}}(w) \tag{41}$$

One has, for every $w \in \{0,1\}^n$,

$$\widehat{\mathbf{1}_{x_S=\alpha_S}}(w) = \frac{1}{2^n} \sum_{x\in\{0,1\}^n, x_S=\alpha_S} (-1)^{x\cdot w} = \mathbf{1}_{\text{supp}(w)\subseteq S} \cdot (-1)^{\alpha\cdot w} 2^{-|S|} \tag{42}$$

Thus,

$$\hat{g}_\alpha(v) = 2^{-|S|} \sum_{w\in\{0,1\}^n, \text{supp}(w)\subseteq S} \hat{f}(v+w)(-1)^{\alpha\cdot w} \tag{43}$$

We now define a useful distribution over vectors in $\{0,1\}^n$. We say that a vector $\alpha \in \{0,1\}^n$ is sampled from $\mathcal{D}_S$ if α_S is uniformly random in $\{0,1\}^S$ and all other entries of α are zero. Note that for every $x \in \{0,1\}^n$ one has

$$\mathbf{E}_{\alpha\sim\mathcal{D}_S}[(-1)^{\alpha\cdot x}] = \begin{cases} 1 & \text{if } x_S = 0 \\ 0 & \text{o.w.} \end{cases} \tag{44}$$

We have for every $v \in \{0,1\}^n$

$$\begin{aligned}
\mathbf{E}_{\alpha\sim\mathcal{D}_S}[\hat{g}_\alpha(v)^2] &= \mathbf{E}_\alpha\left[\left(2^{-|S|} \sum_{w\in\{0,1\}^n, \text{supp}(w)\subseteq S} \hat{f}(v+w)(-1)^{\alpha\cdot w}\right)^2\right] \\
&= 2^{-2|S|}\mathbf{E}_{\alpha\sim\mathcal{D}_S}\left[\sum_{\substack{w,w'\in\{0,1\}^n, \\ \text{supp}(w)\subseteq S, \text{supp}(w')\subseteq S}} \hat{f}(v+w)\hat{f}(v+w')(-1)^{\alpha\cdot(w+w')}\right] \\
&= 2^{-2|S|} \sum_{\substack{w,w'\in\{0,1\}^n, \\ \text{supp}(w)\subseteq S, \text{supp}(w')\subseteq S}} \hat{f}(v+w)\hat{f}(v+w')\mathbf{E}_{\alpha\sim\mathcal{D}_S}\left[(-1)^{\alpha\cdot(w+w')}\right] \\
&= 2^{-2|S|} \sum_{w\in\{0,1\}^n, \text{supp}(w)\subseteq S} \hat{f}(v+w)^2 \quad \text{(by (44) applied to } w+w')
\end{aligned} \tag{45}$$

In particular, for every $k \geq 1$

$$\begin{aligned}
\mathbf{E}_{\alpha\sim\mathcal{D}_S}\Big[\sum_{\substack{v\in\{0,1\}^n:\text{supp}(v)\cap S=\emptyset, \\ wt(v)=k}} \hat{g}_\alpha(v)^2\Big] &= 2^{-2|S|} \sum_{\substack{v\in\{0,1\}^n:\text{supp}(v)\cap S=\emptyset, \\ wt(v)=k}} \sum_{\substack{w\in\{0,1\}^n, \\ \text{supp}(w)\subseteq S}} \hat{f}(v+w)^2 \\
&\geq 2^{-2|S|} \sum_{\substack{v\in\{0,1\}^n:\text{supp}(v)\cap S=\emptyset, \\ wt(v)=k}} \sum_{\substack{w\in\{0,1\}^n, \text{supp}(w)\subseteq S, \\ w_S=\mathbf{1}_S}} \hat{f}(v+w)^2 \\
&= 2^{-2|S|} \sum_{\substack{v\in\{0,1\}^n: S\subseteq\text{supp}(v), \\ wt(v)=k+|S|}} \hat{f}(v)^2
\end{aligned} \tag{46}$$

By Eq. (46) there exists $\alpha^* \in \{0,1\}^n, \text{supp}(\alpha^*) \subseteq S$ such that

$$\sum_{v:wt(v)=k} \hat{g}_{\alpha^*}(v)^2 \geq 2^{-2|S|} \sum_{v\in\{0,1\}^n, S\subseteq v:wt(v)=k+|S|} \hat{f}(v)^2 \tag{47}$$

Recall that $g_{\alpha^*}(x)$ is the indicator of $\mathbf{B}_{\alpha^*} \subseteq \{0,1\}^n$. By Lemma 7.1 we have for all $\delta \in [0,1]$

$$\sum_{v\in\{0,1\}^n} \delta^{|v|}\hat{g}_{\alpha^*}(v)^2 \leq (|\mathbf{B}_{\alpha^*}|/2^n)^{\frac{2}{1+\delta}}.$$

Thus, for every $k \geq 1$ we have

$$\sum_{v \in \{0,1\}^n : wt(v)=k} \hat{g}_{\alpha^*}(v)^2 \leq \delta^{-k}(|\mathbf{B}_{\alpha^*}|/2^n)^{\frac{2}{1+\delta}} \leq \delta^{-k}(|\mathbf{A}|/2^n)^{\frac{2}{1+\delta}},$$

where we used the fact that $|\mathbf{B}_\alpha| \leq |\mathbf{A}|$ for all α.

Putting this together with Eq. (47) yields

$$2^{-2|S|} \sum_{\substack{v \in \{0,1\}^n, S \subseteq v \\ wt(v)=k+|S|}} \hat{f}(v)^2 \leq \delta^{-k}(|\mathbf{A}|/2^n)^{\frac{2}{1+\delta}}. \tag{48}$$

We now set $\delta = k/(2d)$. This is valid, i.e. $\delta \in (0,1)$, since $k \leq d$ by assumption of the lemma. Simplifying the rhs of Eq. (48) with this setting of δ, we get

$$\delta^{-k}(|\mathbf{A}|/2^n)^{\frac{2}{1+\delta}} = \left(\frac{k}{2d}\right)^{-k} 2^{-\frac{2d}{1+k/(2d)}} = 2^{-2d}\left(\frac{k}{2d}\right)^{-k} 2^{-2d\left(\frac{1}{1+k/(2d)}-1\right)}$$

Since $\frac{1}{1+k/(2d)} - 1 = -\frac{k/(2d)}{1+k/(2d)} \geq -k/(2d)$, we conclude that the rhs is bounded by

$$2^{-2d}(k/(2d))^{-k}2^k = 2^{-2d}(4d/k)^k \leq 2^{-2d}(64d/k)^k.$$

Substituting this into Eq. (48), we get

$$2^{2d} \sum_{v \in \{0,1\}^n, S \subseteq v : wt(v)=k+|S|} \hat{f}(v)^2 \leq 2^{-2|S|}(64d/k)^k$$

as required.

■

References

[AG09] K. Ahn and S. Guha. Graph sparsification in the semi-streaming model. *ICALP*, pages 328–338, 2009.

[AG11] K. Ahn and S. Guha. Linear programming in the semi-streaming model with application to the maximum matching problem. *ICALP*, pages 526–538, 2011.

[AG13] K. Ahn and S. Guha. Access to data and number of iterations: Dual primal algorithms for maximum matching under resource constraints. *CoRR*, abs/1307.4359, 2013.

[AGM12a] Kook Jin Ahn, Sudipto Guha, and Andrew McGregor. Analyzing graph structure via linear measurements. *SODA*, pages 459–467, 2012.

[AGM12b] Kook Jin Ahn, Sudipto Guha, and Andrew McGregor. Graph sketching: Sparsification, spanners, and subgraphs. *PODS*, 2012.

[AKLY15] Sepehr Assadi, Sanjeev Khanna, Yang Li, and Grigory Yaroslavtsev. Tight bounds for linear sketches of approximate matchings. *CoRR*, 2015.

[AMS96] Noga Alon, Yossi Matias, and Mario Szegedy. The space complexity of approximating the frequency moments. In *STOC*, pages 20–29, 1996.

[Bera] Bertinoro workshop 2011, problem 45, http://sublinear.info/45.

[Berb] Bertinoro workshop 2014, problem 67, http://sublinear.info/67.

[BK96] András A. Benczúr and David R. Karger. Approximating *s-t* minimum cuts in $\tilde{O}(n^2)$ time. *Proceedings of the 28th annual ACM symposium on Theory of computing*, pages 47–55, 1996.

[CCE+15] Rajesh Hemant Chitnis, Graham Cormode, Hossein Esfandiari, MohammadTaghi Hajiaghayi, Andrew McGregor, Morteza Monemizadeh, and Sofya Vorotnikova. Kernelization via sampling with applications to dynamic graph streams. *CoRR*, abs/1505.01731, 2015.

[Dur06] Rick Durrett. *Random Graph Dynamics (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press, New York, NY, USA, 2006.

[EHL+15] Hossein Esfandiari, Mohammad Taghi Hajiaghayi, Vahid Liaghat, Morteza Monemizadeh, and Krzysztof Onak. Streaming algorithms for estimating the matching size in planar graphs and beyond. *SODA*, 2015.

[GKK+08] Dmitry Gavinsky, Julia Kempe, Iordanis Kerenidis, Ran Raz, and Ronald de Wolf. Exponential separation for one-way quantum communication complexity, with applications to cryptography. *SIAM J. Comput.*, 38(5):1695–1708, 2008.

[GKK12] A. Goel, M. Kapralov, and S. Khanna. On the communication and streaming complexity of maximum bipartite matching. *SODA*, 2012.

[GO12] Venkatesan Guruswami and Krzysztof Onak. Superlinear lower bounds for multipass graph processing. *CCC*, 2012.

[HRVZ15] Zengfeng Huang, Božidar Radunović, Milan Vojnović, and Qin Zhang. Communication complexity of approximate maximum matching in distributed graph data. *STACS*, 2015.

[Kap13] Michael Kapralov. Better bounds for matchings in the streaming model. *SODA*, 2013.

[KK15] Dmitry Kogan and Robert Krauthgamer. Sketching cuts in graphs and hypergraphs. *ITCS*, 2015.

[KKL88] Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on boolean functions (extended abstract). *29th Annual Symposium on Foundations of Computer Science, White Plains, New York, USA, 24-26 October 1988*, pages 68–80, 1988.

[KKS14] Michael Kapralov, Sanjeev Khanna, and Madhu Sudan. Approximating matching size from random streams. In *25th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2014.

[KKS15] Michael Kapralov, Sanjeev Khanna, and Madhu Sudan. Streaming lower bounds for approximating MAX-CUT. In *26th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2015.

[KL11] Jonathan A. Kelner and Alex Levin. Spectral sparsification in the semi-streaming setting. *STACS*, pages 440–451, 2011.

[KLM+14] Michael Kapralov, Yin Tat Lee, Cameron Musco, Christopher Musco, and Aaron Sidford. Single pass spectral sparsification in dynamic streams. *FOCS*, 2014.

[Kon15] Christian Konrad. Maximum matching in turnstile streams. *CoRR*, abs/1505.01460, 2015.

[KW14] Michael Kapralov and David Woodruff. Spanners and sparsifiers in dynamic streams. *PODC*, 2014.

[McG14] Andrew McGregor. Graph stream algorithms: a survey. *SIGMOD Record*, 43(1):9–20, 2014.

[SS08] D.A. Spielman and N. Srivastava. Graph sparsification by effective resistances. *STOC*, pages 563–568, 2008.

[VY11] Elad Verbin and Wei Yu. The streaming complexity of cycle counting, sorting by reversals, and other problems. *SODA*, pages 11–25, 2011.