FRIENDLY BISECTIONS OF RANDOM GRAPHS

ASAF FERBER, MATTHEW KWAN, BHARGAV NARAYANAN, ASHWIN SAH,
AND MEHTAAB SAWHNEY

ABSTRACT. Resolving a conjecture of Füredi from 1988, we prove that with high probability, the random graph $\mathbb{G}(n, 1/2)$ admits a friendly bisection of its vertex set, i.e., a partition of its vertex set into two parts whose sizes differ by at most one in which $n - o(n)$ vertices have more neighbours in their own part as across. Our proof is constructive, and in the process, we develop a new method to study stochastic processes driven by degree information in random graphs; this involves combining enumeration techniques with an abstract second moment argument.

1. INTRODUCTION

In a cut of a graph, i.e., a partition of its vertex set into two parts, we call a vertex *friendly* if it has more neighbours in its own part than across, and *unfriendly* otherwise. Questions about finding friendly and unfriendly partitions of graphs, i.e., partitions in which all (or almost all) the vertices are friendly or unfriendly, have been investigated in various contexts: in combinatorics, on account of their inherent interest [5, 10, 19, 26, 30, 34, 36], in computer science, as 'local' analogues of important NP-complete partitioning problems [4, 13], in probability and statistical physics, owing to their connections to spin glasses [1, 18, 20, 32], and in logic and set theory [2, 31]; this list is merely a representative sample (and by no means exhaustive) since such partitions have been studied extremely broadly. On the other hand, when it comes to finding friendly or unfriendly *bisections*, i.e., partitions into two parts whose sizes differ by at most one, much less is known. Our aim here is to prove an old and well-known conjecture about random graphs due to Füredi [16] from 1988. This problem has gained some notoriety over the years, in part due to its inclusion in Green's list of 100 open problems [21, Problem 91]. Our main result is as follows.

Theorem 1.1. *With high probability, an Erdős–Rényi random graph $G \sim \mathbb{G}(n, 1/2)$ admits a bisection in which $n - o(n)$ vertices are friendly.*

Background and motivation. To place Füredi's conjecture and its resolution here in context, we recall some background and some simple observations. It is a classical fact that every graph admits a partition in which every vertex is unfriendly, as evidenced by any maximum cut. On the other hand, it is also well-known that not every graph admits a partition in which every vertex is friendly, though a general result of Stiebitz [34] ensures that one can always find a partition in which every vertex is 'almost

friendly', and it is easy to deduce from this that $\mathbb{G}(n, 1/2)$ has, with high probability, a partition into two sets of size $n/2 \pm o(n)$ in which $n - o(n)$ vertices are friendly. When it comes to bisections, however, essentially nothing along these lines was previously known (though, see for example [5, 10] for some interesting conjectures).

It is worth mentioning an elementary heuristic for why Theorem 1.1 should plausibly hold. In a random graph $\mathbb{G}(n, 1/2)$, if we fix a bisection $A \cup B$ and a vertex v, and consider the numbers $\deg_A(v)$ and $\deg_B(v)$ of v's neighbours in A and B respectively, then $\deg_A(v)$ and $\deg_B(v)$ are independent and have essentially the same distribution (one is distributed as $\text{Binomial}(n/2, 1/2)$, and the other as $\text{Binomial}(n/2 - 1, 1/2)$). Thus, the probability that a particular vertex is friendly with respect to a particular bisection is about $1/2$. For different vertices v and w, it seems plausible that the events that v is friendly and that w is friendly should be positively correlated, so the probability that at least $(1-\varepsilon)n$ vertices are friendly with respect to $A \cup B$ should be at least about $2^{-(1-\varepsilon)n}$. Since there are $\binom{n}{n/2} = 2^{n-o(n)}$ choices of $A \cup B$, it seems reasonable to expect that there should be at least one bisection satisfying the conclusion of Theorem 1.1.

It seems to be very challenging to make the above line of reasoning rigorous. However, there are a few simple ways in which we can prove the following weakening of Theorem 1.1: with high probability, $G \sim \mathbb{G}(n, 1/2)$ admits a bisection $A \cup B$ in which $0.6n$ vertices are friendly. One such argument is as follows. It is easy to show that in a *random* bisection, typically about half of the vertices in G are friendly. For some small constant $\varepsilon > 0$, we can then take the εn unfriendliest vertices on each side and swap them, and it would appear that one can use this idea to construct a bisection where at least say $0.6n$ vertices are friendly. An alternative iterative argument to establish the same result is as follows: partition the vertices into pairs, go through the pairs one-by-one, at each step revealing the edges between the current pair $\{x, y\}$ and all previous pairs, and decide whether to put x in A and y in B or to put x in B and y in A depending on which choice would make $\{x, y\}$ 'as friendly as possible'. Finally, one could also prove such a weakened bound by considering a max-cut of the *complement* $\overline{G}$ of G, and then randomly 'rebalancing' it into a bisection.

Degree-driven stochastic processes. Although Theorem 1.1 is specifically about friendly bisections of random graphs, the approach we adopt to prove this result is rather general, and it may be that the more important point of this work is its contribution to methodology. Concretely, we develop a method that appears suitable for analysing many different types of stochastic processes on random graphs driven primarily by degree information; for example, the fourth and fifth authors [29] use modifications of these techniques to settle various conjectures of Tran and Vu [37] concerning majority dynamics on random graphs. Below, we outline how our approach allows us to prove Theorem 1.1.

We adopt a constructive approach that yields an efficient algorithm to find the bisection promised by Theorem 1.1. To motivate our approach, it is instructive to consider the following basic algorithm, motivated by the classical large-cut-finding algorithm: starting with any bisection $A \cup B$ of a graph G, repeatedly check whether there are vertices $v \in A$ and $w \in B$ such that $\deg_B(v) > \deg_A(v)$ and $\deg_A(w) > \deg_B(w)$, and if so, swap v and w. Such a swap must decrease the size (i.e., the number of crossing

edges) of the bisection, so this algorithm must terminate.[1] Of course, if we are unlucky, it might happen that when the algorithm terminates, all the vertices in A are friendly, while very few of the vertices in B are friendly, so the resulting bisection may be very far from satisfying the conclusion of Theorem 1.1. However, it seems plausible that such an outcome is rather unusual: if G is sampled from $\mathbb{G}(n, 1/2)$, then one might expect this algorithm (interpreted as a random process) to typically follow a predictable trajectory, and in particular, the number of friendly vertices in A and in B to stay roughly the same for most of the duration of the algorithm.

This is a promising starting point, especially due to the fact that we do not actually need to fully understand the typical trajectory of the process. Indeed, we only need to show that at each step k, the number of friendly vertices in A concentrates around its expected value. By symmetry (assuming for the moment that n is even), the number of friendly vertices in B would then concentrate around that same expected value, so the numbers of friendly vertices in A and B would never get 'too imbalanced'. However, it is far from obvious how to actually establish concentration. Roughly speaking, the main issue is that in order to execute even the first step of the algorithm, we have to inspect every vertex of our graph, meaning that there is seemingly 'no remaining randomness' for the second step. This is in contrast with most other random graph processes in the literature (such as H-free or H-removal processes, as in [7, 8, 15] for example), where each individual step is defined in terms of a random choice.

There are two ideas that allow us to salvage enough randomness to establish the desired concentration. First, instead of swapping vertices one at a time, we shall instead swap a sizeable 'batch' of vertices between A and B in each step; this is strongly reminiscent of the influential 'nibbling' idea introduced by Rödl [28]. We will be able to use discrepancy properties of random graphs to show that, in a typical outcome of the random graph $\mathbb{G}(n, 1/2)$, when we have a bisection $A \cup B$ in which many vertices in A and in B are unfriendly, swapping a large number of the 'unfriendliest' vertices in A and in B dramatically decreases the size of the bisection. That is to say, it should only take a few steps, (about $\exp(1/\varepsilon)$, in fact) to reach a bisection in which one of the two parts has $(1 - \varepsilon)n/2$ friendly vertices. This makes the problem of establishing concentration more tractable, since we now only need to do this for a large constant number of steps. Our second main observation is that in order to execute a step of our algorithm, we only need to know the degrees $\deg_A(v)$ and $\deg_B(v)$ for each vertex v at that stage (and not any other information about the graph). Thus, instead of revealing the whole graph to study the first step, we may simply reveal the required degree information, meaning that our random graph is now conditionally a degree-constrained random graph. We then have the randomness of this degree-constrained random graph with which to show concentration at the next step, for which we again only need to (dynamically) reveal some more degree information, and so on.

The above observations leave us with the task of demonstrating concentration in some (families of) degree-constrained random graphs. In order to study these degree-constrained random graphs, we have at our disposal powerful enumeration theorems due to McKay and Wormald [27], and extensions by Canfield, Greenhill, and

---

[1]Strictly speaking this is not always true, for subtle off-by-one reasons (it is true in the 'opposite' algorithm where we repeatedly swap pairs $v \in A$ and $w \in B$ with $\deg_B(v) < \deg_A(v)$ and $\deg_A(w) < \deg_B(w)$, which would be relevant if we were searching for an *unfriendly* bisection). This subtlety will not be relevant to the present paper.

McKay [12], which give very precise asymptotic formulae for the number of graphs with specified degree information. In principle, this allows one to write down explicit formulae for essentially all relevant probabilities, from which one could attempt to compute the typical trajectory of the process. However, the necessary computations are formidable, and in particular, the various densities under consideration do not appear to have closed-form expressions past the first few iterations.

Our approach to circumventing these issues brings us to the heart of the matter: we develop an abstract second-moment argument with which one can establish concentration of various statistics at a given step, using only *stability* and *anti-concentration* information about the outcomes of previous steps. In particular, this enables us to establish concentration without actually knowing the trajectory of the process. This is superficially reminiscent of martingale arguments establishing concentration around the mean without any knowledge of the location of the mean itself (see [3, Section 7.2]), but the inputs to such arguments, typically Lipschitz-like behaviour of the random variables of interest, are rather different from the inputs to our argument. As mentioned earlier, the methods in our argument are quite general, and we anticipate that a broad range of similar stochastic processes will now become amenable to analysis.

Further directions. The most obvious direction for improvement is to quantify or remove the $o(n)$ term in Theorem 1.1. With appropriate quantification, our proof allows for the $o(n)$ in Theorem 1.1 to be replaced by $O(n/\log\log\log\log\log n)$; however, computer experiments seem to indicate that with iterative swapping processes of the type considered in this paper, one can quite rapidly reach a partition in which only a tiny number of vertices are unfriendly. In fact, we cannot rule out the possibility that $\mathbb{G}(n, 1/2)$ typically admits a bisection where *all* the vertices are friendly. It is not implausible that such bisections exist, but are nonetheless computationally difficult to find, a situation that would be somewhat reminiscent of the largest clique problem in random graphs (see [17], for example).

Notation. Our graph-theoretic notation is for the most part standard; see [9] for terms not defined here. In a graph G, we write $\deg(v)$ for the degree of a vertex $v \in V(G)$, and $N(v)$ for its neighbourhood; also, for a subset $U \subseteq V(G)$, we write $\deg_U(v)$ for the number of neighbours of v in U, i.e., for the size of $N(v) \cap U$. We write $\mathbb{G}(n, p)$ for the Erdős–Rényi random graph on n vertices with edge density p.

Our use of asymptotic notation is mostly standard as well. We say that an event occurs with high probability if it holds with probability $1 - o(1)$ as some parameter (usually n, unless we specify otherwise) tends to infinity. Constants suppressed by asymptotic notation may be absolute, or might depend on other fixed parameters; we shall spell out the latter situation explicitly whenever there might be cause for confusion. To lighten notation, we write $f = g \pm h$ for $|f - g| \leq h$. We maintain this convention with asymptotic notation as well, so $f = g \pm n^{-\Omega(1)}$ for example is taken to mean $|f - g| = n^{-\Omega(1)}$. We also adopt the following non-standard bit of notation: as a parameter n tends to infinity, we write $f \simeq h$ if $f = (1 \pm n^{-\Omega(1)})h$. Finally, following a common abuse, we omit floors and ceilings wherever they are not crucial.

Finally, we write A^c to denote the complement of a set A (with respect to some ground set that will be clear from context).

Organisation. This paper is organised as follows. In Section 2, we describe the swapping process that allows us to prove Theorem 1.1, and also give the deduction of our main result from a few key lemmas. In Section 3, we dispose of the more routine of these lemmas. The beef of our argument is in Section 4, where we must work rather hard to establish the key concentration properties of our swapping process.

2. PROOF OVERVIEW

In this section we make some initial observations, then describe a random swapping process that underlies our argument and state some facts about this process (with proofs to follow later). We then show how to deduce Theorem 1.1 from these facts.

Given a bisection $A \cup B$ of a graph, the *friendliness* $\Delta_{A,B}(v)$ of a vertex v is the difference between the number of its neighbours on its own side and the number of its neighbours on the other side. We say a vertex is *friendly* if its friendliness is positive, and otherwise, we say it is *unfriendly*. The *total friendliness* $\Delta_{A,B}$ of the bisection $A \cup B$ is then given by

$$\Delta_{A,B} = \sum_{v \in V(G)} \Delta_{A,B}(v).$$

We also make a simple observation that allows us to restrict our attention to random graphs of even order (which in turn allows us to somewhat simplify the presentation). A simple union bound (similar to calculations we will see in Section 3) shows that with high probability, in *any* partition of the vertex set of $\mathbb{G}(n, 1/2)$, at most $10n/\log n$ vertices have friendliness 1, i.e., have exactly one more neighbour on their own side than across, or vice versa. Consequently, it clearly suffices to establish Theorem 1.1 for $\mathbb{G}(n, 1/2)$ when n is even; indeed, when n is odd, we may delete an arbitrary vertex from the random graph, apply Theorem 1.1 to the result, and add back the deleted vertex to either part to get the desired bisection. Therefore, all graphs under consideration will be of even order unless explicitly specified otherwise, and we shall not belabour this point any further.

Lemma 2.1 shows that for a typical outcome of the random graph $\mathbb{G}(n, 1/2)$, there is a window of length $O(n^{3/2})$ within which the total friendliness of any bisection lies.

Lemma 2.1. *There is a $\gamma > 0$ such that for a random graph $G \sim \mathbb{G}(n, 1/2)$, with high probability, every bisection $A \cup B$ of G has $|\Delta_{A,B}| < \gamma n^{3/2}$.*

Next, we shall define a simple random 'swap' operation that modifies a bisection with the aim of making it more friendly.

Definition 2.2. Given a bisection $A \cup B$ of an n-vertex graph G and $0 < \alpha < 1/2$, the α-*swap* of $A \cup B$ is the random bisection obtained by the following procedure. First, we take the subset $A' \subseteq A$ of the $\lfloor \alpha n \rfloor$ most unfriendly vertices in A, and the subset $B' \subseteq B$ of the $\lfloor \alpha n \rfloor$ most unfriendly vertices in B (breaking ties according to some *a priori* fixed ordering of the vertex set), and swap A' and B'. At this stage, the parts of the resulting bisection are then $(A \setminus A') \cup B'$ and $(B \setminus B') \cup A'$. Next, we make a uniformly random choice of $\lfloor \alpha^4 n \rfloor$ vertices on both of these sides, and swap these subsets.

We remark that the second (random) swap in the α-swap procedure is not actually necessary for the proof of Theorem 1.1, but the analysis later in the paper would become substantially more involved without it.

Lemma 2.3 shows that in a typical outcome of the random graph $\mathbb{G}(n, 1/2)$, for every bisection $A \cup B$, either our swapping operation always increases the total friendliness by $\Omega(n^{3/2})$, or almost all the vertices in one of the parts (either A or B) are already friendly. Here there are two sources of randomness (the random graph and the random swap); we emphasise that we fix a particular outcome of the random graph, and then consider a random swap in the setting of that particular graph.

Lemma 2.3. *For every fixed $\varepsilon > 0$, there are $\alpha \in (0, \varepsilon)$ and $\beta > 0$ for which a random graph $G \sim \mathbb{G}(n, 1/2)$ has, with high probability, the following property. In any bisection $A \cup B$ of G in which at least εn vertices are unfriendly in each of A and B, the random bisection $A_1 \cup B_1$ obtained from an α-swap of $A \cup B$ always satisfies*

$$\Delta_{A_1, B_1} \geq \Delta_{A, B} + \beta n^{3/2}.$$

Finally, Lemma 2.4 establishes concentration properties for bisections obtained by iterating our swapping operation.

Lemma 2.4. *Fix $\varepsilon > \alpha > 0$, $k \in \mathbb{N}$, and an arbitrary bisection $A \cup B$ of the vertex set of $\mathbb{G}(n, 1/2)$. For a random graph $G \sim \mathbb{G}(n, 1/2)$, let $A_k \cup B_k$ be the bisection obtained by performing k iterations of the α-swap procedure starting from $A \cup B$. Writing X and Y respectively for the number of unfriendly vertices in A_k and B_k, we have with high probability that $|X - Y| = o(n)$.*

With these facts in hand, we may now easily deduce Theorem 1.1.

Proof of Theorem 1.1. For any fixed $\varepsilon > 0$, we shall show that $G \sim \mathbb{G}(n, 1/2)$ with high probability has a bisection in which at most $2\varepsilon n + o(n)$ vertices are unfriendly.

Say $V(G) = \{1, \ldots, n\}$, define the bisection $A_0 \cup B_0$ by $A_0 = \{1, \ldots, n/2\}$ and $B_0 = \{n/2 + 1, \ldots, n\}$. Let γ be as in Lemma 2.1 and β as in Lemma 2.3 applied to ε. Set $K = \lceil 2\gamma/\beta \rceil + 1$, and let

$$A_1 \cup B_1, \ A_2 \cup B_2, \ldots, A_K \cup B_K$$

be the sequence of bisections arising from K iterations of the α-swap procedure starting from $A_0 \cup B_0$.

Say that a bisection $A \cup B$ is ε-*good* if there are at most εn unfriendly vertices in A or at most εn unfriendly vertices in B. Now, the following properties hold with high probability, by Lemmas 2.1, 2.3 and 2.4.

(1) There is an interval of length at most $2\gamma n^{3/2}$ such that the total friendliness of every bisection of G lies in this interval.
(2) For every $0 \leq k \leq K-1$, either $A_k \cup B_k$ is ε-good, or $\Delta_{A_{k+1}, B_{k+1}} \geq \Delta_{A_k, B_k} + \beta n^{3/2}$.
(3) For every $1 \leq k \leq K$, the numbers of unfriendly vertices in A_k and in B_k differ by $o(n)$.

Fix outcomes of G and $A_1 \cup B_1, A_2 \cup B_2, \ldots, A_K \cup B_K$ satisfying all these properties. Now, by property (1), it is not possible for the total friendliness to increase by $\beta n^{3/2}$ in each of the K iterations. So, by property (2), there must be some k for which $A_k \cup B_k$ is ε-good, meaning that there are at most εn unfriendly vertices in A_k or at most εn unfriendly vertices in B_k. The third property (3) now ensures that there are at most $2\varepsilon n + o(n)$ unfriendly vertices in total at this stage. The bisection $A_k \cup B_k$ has the properties we desire, proving the result. □

2.1. **Overview of the proofs of the key lemmas.** We now briefly discuss the proofs of Lemmas 2.1, 2.3 and 2.4. First, Lemma 2.1 is proved via a Chernoff bound and a simple union bound over all possible bisections. Second, Lemma 2.3 is also proved by a union bound: we show that that no bisection of the graph has many vertices with friendliness very close to zero, so that there is always some reasonably large gain from swapping unfriendly vertices; here, one must also control the (small) amount of additional unfriendliness potentially introduced between pairs of swapped vertices.

The proof of Lemma 2.4 is by far the most technical ingredient in the proof. At a high level, one runs the iterated swap algorithm on a random graph $G \sim \mathbb{G}(n, 1/2)$, at each step revealing only that information about G (namely, degrees into certain parts) which is necessary to determine the outcome of the α-swap procedure. So, at every step, we need to study a degree-constrained random graph model; this is accomplished using graph enumeration techniques in the style of McKay–Wormald [27]. One can track the fraction of vertices that live in prescribed parts at prescribed times inductively, showing via the second moment method in our degree-constrained random graph model that the numbers of different types of vertices are concentrated. However, several obstacles arise naturally due to the presence of complicated conditional distributions, and the need for all of the different 'well-conditioned' degree-constrained models (based on different revelations) to converge to a single distribution of degrees. The totality of what must be tracked to implement this argument is contained in Proposition 4.3.

In particular, we note that the first part of the proof (Lemmas 2.1 and 2.3) and the second part of the proof (Lemma 2.4) are essentially logically independent, and the analysis here can be extended to a variety of similar algorithms based on degree sequences. One can think of the first part as providing a monovariant to the graph process analysed in the second part, guaranteeing that the graph partition 'gets better' over time and converges to a friendly distribution of degrees rather than to an abstract (iterated) optimiser of some associated variational problem.

3. SWAPPING DECREMENT

In this section we prove Lemmas 2.1 and 2.3. To start with, we need some simple facts about centered binomial distributions. The first is a Chernoff bound (see [22, Theorem 2.1], for example) and the second follows from either Stirling's approximation or the Erdős–Littlewood–Offord theorem (see [35, Corollary 7.4]).

Theorem 3.1. *For $N \in \mathbb{N}$, let $X_1, \ldots, X_N$ be independent Rademacher random variables (satisfying $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = 1/2$), and let $X = X_1 + \cdots + X_N$.*

(1) *For all $t \geq 0$, we have $\mathbb{P}(|X| \geq t) \leq 2e^{-t^2/(2N)}$.*

(2) *For all $t \geq 1$ and all $x \in \mathbb{R}$, we have $\mathbb{P}(|X - x| \leq t) \leq \sqrt{2}t/\sqrt{N}$.*

The proof of Lemma 2.1 is extremely simple, being a routine application of the union bound.

Proof of Lemma 2.1. There are $\binom{n}{n/2} \leq 2^n$ bisections in total. For each such bisection $A \cup B$, we claim that $\Delta_{A,B} + n/2$ has a centered binomial distribution to which Theorem 3.1 applies (with $N = \binom{n}{2}$). Indeed, for each pair of distinct vertices $e = \{x, y\}$, let $Y_e = 1$ if e is an edge in G, and $Y_e = -1$ if e is not an edge in G. Then, let $X_e = Y_e$ if $e \subset A$ or $e \subset B$, and $X_e = -Y_e$ otherwise, and note that $\Delta_{A,B} + n/2 = \sum_e X_e$.

So, for sufficiently large γ, Theorem 3.1(1) yields

$$\mathbb{P}\left(|\Delta_{A,B}| \geq \gamma n^{3/2}\right) \leq 2\exp\left(-\frac{(\gamma n^{3/2} - n/2)^2}{2\binom{n}{2}}\right) = o(2^{-n}),$$

so the desired result follows from the union bound.   □

Lemma 2.3 is also proved by the union bound, but for this, we will first need to prove some auxiliary lemmas.

Lemma 3.2. *For any $\eta \in (0, 1/4)$, a random graph $G \sim \mathbb{G}(n, 1/2)$ with high probability has the property that for every bisection $A \cup B$ of G, we have $|\Delta_{A,B}(v)| \geq 4^{-1/\eta}\sqrt{n}$ for all but at most ηn vertices $v \in A$, and for all but at most ηn vertices $v \in B$.*

Proof. For each bisection $A \cup B$, if we condition on an outcome of $G[A]$, then the random variables $\{\Delta_{A,B}(v) : v \in A\}$ become mutually independent. Conditionally, for each $v \in A$, the random variable $2(\deg_B(v) - n/4) = 2(-\Delta_{A,B}(v) + \deg_A(v) - n/4)$ has a centered binomial distribution to which Theorem 3.1 applies (with $N = n$). Therefore by Theorem 3.1(2),

$$\mathbb{P}\left(|\Delta_{A,B}(v)| \leq 4^{-1/\eta}\sqrt{n}\right) \leq (2\sqrt{2} \cdot 4^{-1/\eta}\sqrt{n})/\sqrt{n-1} \leq 4 \cdot 4^{-1/\eta}$$

for large n. It follows that the probability that the property in the statement of the lemma does not hold is at most

$$2 \cdot 2^n \binom{n/2}{\eta n}(4 \cdot 4^{-1/\eta})^{\eta n} \leq 2^{3n/2 + 2\eta n}4^{-n} = o(1). \qquad \square$$

Lemma 3.3. *For any sufficiently small fixed $\alpha > 0$, a random graph $G \sim \mathbb{G}(n, 1/2)$ with high probability has the property that for every bisection $A \cup B$ of G and every pair of subsets $A' \subseteq A$ and $B' \subseteq B$ each of size αn, we have*

$$|\Delta_{A',B'}| \leq \alpha^{4/3}n^{3/2},$$

where we view $A' \cup B'$ as a bisection of the induced subgraph $G[A' \cup B']$.

Proof. Note that the event does not depend on A, B, only on A', B'. For subsets A' and B' as in the statement of the lemma, the random variable $\Delta_{A',B'} + \alpha n$ has a centered binomial distribution to which Theorem 3.1 applies (with $N = \binom{2\alpha n}{2}$). We then have

$$\mathbb{P}\left(|\Delta_{A',B'}| \geq \alpha^{4/3}n^{3/2}\right) \leq 2\exp\left(-\frac{(\alpha^{4/3}n^{3/2} - \alpha n)^2}{2\binom{2\alpha n}{2}}\right) \leq 2\exp(-\alpha^{2/3}n/8) = o\left(\binom{n}{\alpha n}^{-2}\right),$$

where in the second inequality we have used that n is sufficiently large with respect to α and in the final inequality we have used that $\binom{n}{\alpha n} \leq (e/\alpha)^{\alpha n}$. The desired result follows from a union bound over all choices of A' and B'.   □

Lemma 3.4. *For any sufficiently small fixed $\delta > 0$, a random graph $G \sim \mathbb{G}(n, 1/2)$ with high probability has the following property. For every bisection $A \cup B$, and every pair of subsets $A' \subseteq A, B' \subseteq B$ each of size δn, if we swap A' and B' to obtain a bisection $A_1 \cup B_1$ with $A_1 = (A \setminus A') \cup B'$ and $B_1 = (B \setminus B') \cup A'$, then we have*

$$|\Delta_{A_1,B_1} - \Delta_{A,B}| \leq \delta^{1/3}n^{3/2}.$$

Proof. For each bisection $A \cup B$ and subsets A' and B' as in the lemma statement, the random variable $\Delta_{A_1,B_1} - \Delta_{A,B}$ has a centered binomial distribution to which Theorem 3.1 applies (with $N = 4(n/2 - \delta n)\delta n$). We then have

$$\mathbb{P}\left(|\Delta_{A_1,B_1} - \Delta_{A,B}| \geq \delta^{1/3} n^{3/2}\right) \leq 2\exp\left(-\frac{(\delta^{1/3} n^{3/2})^2}{4(n/2 - \delta n)\delta n}\right) = o\left(2^{-n}\binom{n/2}{\delta n}^{-2}\right),$$

so the desired result follows once again from the union bound. $\qquad\square$

We are now ready to prove Lemma 2.3.

Proof of Lemma 2.3. Let $\eta < \varepsilon/2$ be small enough for Lemma 3.2 to hold. Let $\alpha \in (0, \varepsilon/2)$ be small enough so that Lemma 3.3 holds and Lemma 3.4 holds for $\delta = \alpha^4$, and also $\alpha \leq 4^{-3/\eta}$. Now assume that the properties in Lemmas 3.2 to 3.4 all hold for G with these parameters, which occurs with high probability.

Now, consider an arbitrary bisection $A \cup B$ where at least εn vertices in A are unfriendly and at least εn vertices in B are unfriendly. Let A' be the subset of the αn most unfriendly vertices in A, and let $B' \subseteq B$ be the subset of the αn most unfriendly vertices in B. By assumption, at least εn vertices in A are unfriendly, so at least $(\varepsilon - \alpha)n \geq \eta n$ vertices in A are unfriendly but not as unfriendly as the vertices in A'. By Lemma 3.2 we deduce that for all $v \in A'$ we have $\Delta_{A,B}(v) \leq -4^{-1/\eta}\sqrt{n}$. Similarly, for all $v \in B'$ we have $\Delta_{A,B}(v) \leq -4^{-1/\eta}\sqrt{n}$.

Next, let $A'' = (A \setminus A') \cup B'$ and $B'' = (B \setminus B') \cup A'$ be the parts resulting from the first step in an α-swap. We know that $|\Delta_{A',B'}| \leq \alpha^{4/3} n^{3/2}$ by Lemma 3.3, so we have

$$\Delta_{A'',B''} = \Delta_{A,B} - 4\sum_{v \in A' \cup B'} \Delta_{A,B}(v) + 4\Delta_{A',B'}$$

$$\geq \Delta_{A,B} + 4(2\alpha n)(4^{-1/\eta}\sqrt{n}) - 4\alpha^{4/3} n^{3/2} \geq \Delta_{A,B} + 4\alpha 4^{-1/\eta} n^{3/2}.$$

Finally, by the guarantee in Lemma 3.4, we note that the final random swap in the definition of the α-swap procedure changes the friendliness of the bisection $A'' \cup B''$ by at most

$$\delta^{1/3} n^{3/2} = \alpha^{4/3} n^{3/2} \leq \alpha 4^{-1/\eta} n^{3/2}$$

in passing to the final bisection $A_1 \cup B_1$. It follows that we have the desired result with $\beta = 3\alpha 4^{-1/\eta}$. $\qquad\square$

4. CONCENTRATION OF THE ITERATED SWAPPING PROCESS

In this section we prove Lemma 2.4. In fact, it will follow from the more technical Proposition 4.3, which we shall shortly state and prove by induction. To get started, we need some definitions.

First, we introduce some notation to handle empirical distributions. Given a sequence $(a_i : i \in I)$, the *uniform measure* $\widehat{\mathcal{L}}$ on this sequence is the probability distribution of a_j where j is chosen uniformly from I. When the sequence $(a_i : i \in I)$ is itself random — for example, comprised of jointly random vectors — we emphasise that the associated uniform measure $\widehat{\mathcal{L}}$ is itself a random object, i.e., each realisation of the random sequence $(a_i : i \in I)$ gives rise to an associated uniform measure on this realisation.

We now define some empirical degree distributions associated with our iterated swapping process.

Definition 4.1. Given a graph G on the vertex set $\{1, \dots, n\}$, we consider the iterated swapping process in which we start with the bisection $A_0 \cup B_0$, where $A_0 = \{1, \dots, n/2\}$ and $B_0 = \{n/2 + 1, \dots, n\}$, and perform k repeated α-swaps to yield a sequence $(A_t \cup B_t)_{t=0}^k$ of bisections. For a binary sequence $x = (x_t)_{t=1}^{k+1} \in \{0,1\}^{k+1}$, let V_x be the set of vertices that are in part A_t at those times t with $x_{t-1} = 0$, and in part B_t at those times t with $x_{t-1} = 1$ for $1 \leq t \leq k+1$. For a binary sequence $x \in \{0,1\}^{k+1}$, let $\widehat{\mathcal{L}}_x$ be the uniform measure on the sequence of degree vectors

$$\left(\left((\deg_{V_y}(v) - |V_y|/2)/\sqrt{n} \right)_{y \in \{0,1\}^{k+1}} : v \in V_x \right).$$

Next, we recall the definition of multidimensional Kolmogorov distance on $\mathbb{R}^d$.

Definition 4.2. Let $\mathcal{L}$ and $\mathcal{L}'$ be probability distributions on $\mathbb{R}^d$. We define the *Kolmogorov distance* $d_K(\mathcal{L}, \mathcal{L}')$ between $\mathcal{L}$ and $\mathcal{L}'$ to be the supremum of $|\mathcal{L}(A) - \mathcal{L}'(A)|$ over all sets A of the form $(-\infty, a_1] \times \cdots \times (-\infty, a_d]$, where $a_1, \dots, a_d \in \mathbb{R}$.

Note that the Kolmogorov distance controls the probability of lying in any half-open box: indeed, for any such box $B = (b_1, c_1] \times \cdots \times (b_d, c_d]$, we can use the inclusion-exclusion principle to express $\mathcal{L}(B)$ as a signed sum of 2^d probabilities of the form $\mathcal{L}((-\infty, a_1] \times \cdots \times (-\infty, a_d])$, so $|\mathcal{L}(B) - \mathcal{L}'(B)| \leq 2^d \, d_K(\mathcal{L}, \mathcal{L}')$.

The promised generalisation of Lemma 2.4 is now as follows.

Proposition 4.3. *Fix $\alpha \in (0, 1/4)$ and $k \in \mathbb{N}$. There are $c_{\alpha,k}, C_{\alpha,k} > 0$ such that for each $x \in \{0,1\}^{k+1}$ there are*

(1) *a 2^{k+1}-dimensional probability distribution $\mathcal{L}_x$ (i.e., a probability distribution on $\mathbb{R}^{2^{k+1}}$), and*

(2) *a real number $\pi_x \geq \alpha^{4k}/2$,*

both of which may depend on α and n, such that the following holds. For $G \sim \mathbb{G}(n, 1/2)$, consider a sequence of k iterated α-swaps, and for $x \in \{0,1\}^{k+1}$, let V_x and $\widehat{\mathcal{L}}_x$ be as in Definition 4.1. Then, with high probability, all of the following hold.

(A1) *For each $x \in \{0,1\}^{k+1}$, we have*

$$\left| |V_x| - \pi_x n \right| \leq n^{1 - c_{\alpha,k}}.$$

(A2) *For each $x \in \{0,1\}^{k+1}$, we have*

$$d_K(\widehat{\mathcal{L}}_x, \mathcal{L}_x) \leq n^{-c_{\alpha,k}}.$$

(A3) *For each vertex $v \in V(G)$ and each $x \in \{0,1\}^{k+1}$, we have*

$$\left| \deg_{V_x}(v) - |V_x|/2 \right| < C_{\alpha,k} \sqrt{n \log n}.$$

(A4) *For each $x \in \{0,1\}^{k+1}$, and each box $B = \prod_{y \in \{0,1\}^{k+1}} (a_y, b_y]$ with side lengths $b_y - a_y = n^{-c_{\alpha,k}}$ (and, therefore, $\mathrm{vol}(B) = (n^{-c_{\alpha,k}})^{2^{k+1}}$) we have*

$$\mathcal{L}_x(B) \leq \mathrm{vol}(B) \exp(C_{\alpha,k} \sqrt{\log n}).$$

Again, we emphasise that we treat α and k as fixed constants for the purpose of the 'with high probability' statement in Proposition 4.3; in particular, Proposition 4.3 only holds if n grows sufficiently fast (with respect to α and k).

Before discussing the proof of Proposition 4.3, we explain how it implies Lemma 2.4. The key observation is that **(A1)** to **(A4)** essentially allow us to read off, from the distributions $\mathcal{L}_x$, arbitrary information about degree statistics (and, in particular, the number of friendly vertices in each part). Lemma 4.4 will be used for this purpose: we can estimate the probability mass in a half-space by approximating that half-space by boxes and using **(A2)** to **(A4)**.

Lemma 4.4. *Suppose that G is such that* **(A2)** *to* **(A4)** *are satisfied, and let $H \subseteq \mathbb{R}^{\{0,1\}^{k+1}}$ be any closed half-space (i.e., a region bounded by a hyperplane). Then for any $x \in \{0,1\}^{k+1}$, we have $\widehat{\mathcal{L}}_x(H) = \mathcal{L}_x(H) + o(1)$.*

We defer the proof of Lemma 4.4 (in a slightly stronger form, see Lemma 4.6) to Section 4.2; we now deduce Lemma 2.4 from Proposition 4.3 and Lemma 4.4.

Proof of Lemma 2.4. Let $A_k \cup B_k$ be the bisection resulting from k iterations of the α-swap process. Recall that in the statement of Lemma 2.4, the random variables X and Y are the numbers of unfriendly vertices in A_k and B_k. It suffices to prove that there is some value N (potentially depending on all of α, k, n) such that $X = N + o(n)$ with high probability. Indeed, by the symmetry of the process with respect to A and B, it would follow that $Y = N + o(n)$ with high probability as well, implying that $|X - Y| = o(n)$ with high probability, as desired.

To this end, for $i \in \{0, 1\}$, let $S_i = \{x \in \{0,1\}^{k+1} : x_{k+1} = i\}$ and note that a vertex $v \in A_k$ is unfriendly if and only if

$$\sum_{y \in S_0} \deg_{V_y}(v) - \sum_{y \in S_1} \deg_{V_y}(v) = \sum_{y \in S_0} \left(\deg_{V_y}(v) - |V_y|/2\right) - \sum_{y \in S_1} \left(\deg_{V_y}(v) - |V_y|/2\right) \leq 0.$$

So, defining the affine half-space

$$H = \left\{ d \in \mathbb{R}^{\{0,1\}^{k+1}} : \sum_{y \in S_0} d_y - \sum_{y \in S_1} d_y \leq 0 \right\},$$

we have $X = \sum_{x \in S_0} |V_x| \widehat{\mathcal{L}}_x(H)$. By Proposition 4.3 and Lemma 4.4, with high probability we have $X = n \sum_{x \in S_0} \pi_x \mathcal{L}_x(H) + o(n)$, as desired.   □

We will prove Proposition 4.3 by induction on k. In its full generality, our argument will rely on a second moment computation that utilises results of McKay–Wormald [27] and Canfield–Greenhill–McKay [12] about enumerating graphs with specified vertex-degrees. Since the argument is rather technical, we shall proceed slowly, first illustrating the base case before jumping into the meat of the argument.

4.1. **The base case.** In this subsection we prove Proposition 4.3 for $k = 0$. This entails some explicit calculations in the random graph $\mathbb{G}(n, 1/2)$; the inductive step can be seen as a 'relativised' version of this argument, with the randomness coming from a well-conditioned random graph with specified degree information rather than $\mathbb{G}(n, 1/2)$.

Recall that we need to prove that the four properties in **(A1)** to **(A4)** each hold with high probability. The most interesting of these properties is **(A2)**, which will be established using Lemma 4.5.

Lemma 4.5. *Fix $c \in (0, 1)$ and $d \in \mathbb{N}$. Let $(\vec{d}(v))_{v \in V}$ be a sequence of n discrete jointly random vectors in $\mathbb{R}^d$, and let $\mathcal{L}$ be the (fixed) distribution on $\mathbb{R}^d$ defined by choosing*

v uniformly at random from V and then sampling from $\vec{d}(v)$. Suppose that for a box $Q = (-q, q]^d$ with $q \geq 1$, the following conditions hold:

(1) for each $\vec{s}, \vec{t} \in Q$ and each pair of distinct $u, v \in V$, we have

$$\mathbb{P}(\vec{d}(u) = \vec{t} \text{ and } \vec{d}(v) = \vec{s}) = (1 \pm n^{-c})\mathbb{P}(\vec{d}(u) = \vec{t})\mathbb{P}(\vec{d}(v) = \vec{s}),$$

(2) $\mathcal{L}(Q^c) \leq n^{-c}$, and

(3) for each box $B \subseteq Q$ with side lengths at least n^{-c}, we have $\mathcal{L}(B) \leq q \operatorname{vol}(B)$.

For a given realisation of the random sequence $(\vec{d}(v))_{v \in V}$, let $\widehat{\mathcal{L}}$ be the (random) distribution on $\mathbb{R}^d$ which is the uniform measure on this realisation. With probability at least $1 - O(q^d n^{-c/8})$ over the randomness of $(\vec{d}(v))_{v \in V}$, we have $d_K(\mathcal{L}, \widehat{\mathcal{L}}) = O(q^d n^{-c/(8d)})$.

In applications, $\vec{d}(v)$ will be a list of degrees from v to a number of other fixed subsets, and $(\vec{d}(v))_{v \in V}$ will be the random ensemble of these lists. Lemma 4.5 roughly states that given decorrelation between these degree statistics, and (for technical reasons) a tail bound and anti-concentration, the empirical degree distribution of V is very likely to concentrate around an explicit distribution.

Here, we again reiterate that the constants suppressed by the asymptotic notation in Lemma 4.5 are allowed to depend on the fixed parameters c and d.

Proof of Lemma 4.5. For any $v \in V$, and any box B, let $\mathcal{E}_{v,B}$ be the event that $\vec{d}(v)$ lies in B, so that $n\widehat{\mathcal{L}}(B)$ is the number of $v \in V$ such that $\mathcal{E}_{v,B}$ holds. For $u, v \in V$ and boxes $B, B' \subseteq Q$, we can sum the bound in (1) over all the points $\vec{t} \in B$ and $\vec{s} \in B'$ to see that

$$\mathbb{P}(\mathcal{E}_{u,B} \cap \mathcal{E}_{v,B'}) = \mathbb{P}(\mathcal{E}_{u,B})\mathbb{P}(\mathcal{E}_{v,B'}) \pm n^{-c}.$$

It follows that $\operatorname{Var}(n\widehat{\mathcal{L}}(B)) \leq n + n^{2-c} \leq 2n^{2-c}$, so by Chebyshev's inequality, with probability at least $1 - n^{-c/2}$, we have

$$(4.1) \qquad \left|\widehat{\mathcal{L}}(B) - \mathbb{E}[\widehat{\mathcal{L}}(B)]\right| = \left|\widehat{\mathcal{L}}(B) - \mathcal{L}(B)\right| \leq 2n^{-c/4}.$$

Now, consider a family $\mathfrak{B}$ of $O(n^{c/8}q^d)$ half-open boxes with side lengths at most $D = n^{-c/(8d)}$ that partition the (big) box Q. By the union bound, with probability $1 - O(q^d n^{-c/8})$, the bound (4.1) holds for all $B \in \mathfrak{B}$. Also, since $\mathbb{E}[\widehat{\mathcal{L}}(Q^c)] = \mathcal{L}(Q^c) \leq n^{-c}$, by Markov's inequality we have $\widehat{\mathcal{L}}(Q^c) \leq n^{-c/2}$ with probability at least $1 - n^{-c/2}$. Now, it is a routine matter to deduce the desired conclusion from these two facts. The details are as follows.

For any semi-infinite box $A = (-\infty, a_1] \times \cdots \times (-\infty, a_d]$, we can find subcollections $\mathfrak{B}_-, \mathfrak{B}_+ \subseteq \mathfrak{B}$ such that

$$\bigcup_{B \in \mathfrak{B}_-} B \subseteq A \cap Q \subseteq \bigcup_{B \in \mathfrak{B}_+} B,$$

and $|\mathfrak{B}_+ \setminus \mathfrak{B}_-| = O((q/D)^{d-1})$. Then

$$\sum_{B \in \mathfrak{B}_-} \widehat{\mathcal{L}}(B) \leq \widehat{\mathcal{L}}(A \cap Q) \leq \sum_{B \in \mathfrak{B}_+} \widehat{\mathcal{L}}(B).$$

Furthermore, using (3) and (4.1) for all $B \in \mathfrak{B}$, we see that both the sum $\sum_{B \in \mathfrak{B}_-} \widehat{\mathcal{L}}(B)$ and the sum $\sum_{B \in \mathfrak{B}_+} \widehat{\mathcal{L}}(B)$ differ from $\mathcal{L}(A \cap Q)$ by at most

$$O\left(|\mathfrak{B}_+ \setminus \mathfrak{B}_-|(qD^d) + |\mathfrak{B}|(2n^{-c/4})\right) = O\left(q^d n^{-c/(8d)}\right).$$

So, we have, with probability $1 - O(q^d n^{-c/8})$,

$$\left|\mathcal{L}(A) - \widehat{\mathcal{L}}(A)\right| = O\left(\mathcal{L}(Q^c) + \widehat{\mathcal{L}}(Q^c) + q^d n^{-c/(8d)}\right) = O\left(q^d n^{-c/(8d)}\right),$$

proving the lemma. □

Now we use Lemma 4.5 to prove the base case of Proposition 4.3.

Proof of the $k = 0$ case of Proposition 4.3. First, we have $|V_0| = |A_0| = |V_1| = |B_0| = n/2$, proving **(A1)**. Furthermore, for a sufficiently large $C_{\alpha,k} > 0$, given a vertex v, we have $|\deg_{V_i}(v) - n/4| < C_{\alpha,k} n^{1/2}\sqrt{\log n}$ with probability at least $1 - 1/n^2$, say, just by the Chernoff bound, whence a union bound demonstrates **(A3)**.

It remains to prove **(A2)** and **(A4)**. It is enough to prove them for $x = (0)$, by symmetry. We will take $\mathcal{L}_0$ to be the distribution of the random vector

$$\vec{d}(v) = \left(|\deg_{V_0}(v) - n/4|/\sqrt{n}, \ |\deg_{V_1}(v) - n/4|/\sqrt{n}\right),$$

where $v \in V_0$ is arbitrary; clearly, this distribution does not actually depend on the specific choice of $v \in V_0$. Then, $\mathcal{L}_0$ has a simple description in terms of independent binomial distributions. Although it will not be necessary for the proof, we remark that $\mathcal{L}_0$ is well-approximated by the bivariate normal distribution $N(0, 1/2)^2$, and it is possible to take $\mathcal{L}_0$ to be this distribution as well.

Before proceeding further, we note that the aforementioned Chernoff bound shows that with $Q = (-C_{\alpha,k}\sqrt{\log n}, C_{\alpha,k}\sqrt{\log n}]^2$, we have $\mathcal{L}_0(Q^c) \leq 2/n^2$. Now, for every individual point $\vec{d} \in \mathbb{R}^2$, we have $\mathcal{L}_0(\{\vec{d}\}) = O((1/\sqrt{n})^2) = O(1/n)$ (by the Erdős–Littlewood–Offord theorem applied to each coordinate, say). Since $\mathcal{L}_0$ is supported on the lattice $((\mathbb{Z} - n/4)/\sqrt{n})^2$, for a box B with side lengths at least $1/\sqrt{n}$, we have

$$(4.2) \qquad\qquad\qquad \mathcal{L}_0(B) = O\left(\text{vol}(B)\right),$$

establishing **(A4)**. Now, we claim that for every pair of vertices u, v and every pair of points $\vec{s}, \vec{t} \in Q$, we have

$$\mathbb{P}(\vec{d}(u) = \vec{t} \text{ and } \vec{d}(v) = \vec{s}) = (1 \pm O(\sqrt{\log n/n}))\mathbb{P}(\vec{d}(u) = \vec{t})\mathbb{P}(\vec{d}(v) = \vec{s}).$$

Indeed, we will then be able to apply Lemma 4.5 to establish that **(A2)** holds with high probability. The claim follows from the following explicit calculation. The only dependence between $\vec{d}(u)$ and $\vec{d}(v)$ comes from the potential edge between u and v, but we can check that if we condition on this edge being present (or not), the probabilities $\mathbb{P}(\vec{d}(u) = \vec{t})$ and $\mathbb{P}(\vec{d}(v) = \vec{s})$ vary only by a factor of $(1 \pm O(\sqrt{\log n/n}))$, which in itself boils down to the observation that $\binom{n/2-1}{t}/\binom{n/2-1}{t-1} = (n/2 - t)/t = 1 + O(|1/4 - t/n|)$. (Note that we are only considering $\vec{t} = (t_1, t_2)$ for which $|t - n/4| \leq O(\sqrt{n \log n})$ for $i \in \{1, 2\}$). □

4.2. **Preliminaries for the inductive step.** We start with some preparations before proceeding to the details of the inductive step. First, we provide a proof of Lemma 4.4; actually we prove the following more general lemma.

Lemma 4.6. *For fixed $c > 0, d \in \mathbb{N}$ and any $q \geq 1$, let $\mathcal{L}, \mathcal{L}'$ be probability distributions on $\mathbb{R}^d$ satisfying $d_K(\mathcal{L}, \mathcal{L}') \leq n^{-c}$, $\mathcal{L}'\left((-q, q]^d\right) = 1$, and $\mathcal{L}(B) \leq q\,\text{vol}(B)$ for all boxes B with side lengths at least n^{-c}. Then the following conclusions hold.*

(1) *For any region $H \subseteq \mathbb{R}^d$ defined as the intersection of $O(1)$ (closed or open) affine half-spaces, we have $\mathcal{L}'(H) = \mathcal{L}(H) \pm O(q^d n^{-c/(2d)})$.*

(2) *For any $R \subseteq \mathbb{R}^d$ obtained as the region between two parallel (closed or open) affine hyperplanes separated by a distance of at most n^{-c}, we have $\mathcal{L}(R) = O(q^d n^{-c/(2d)})$.*

Here, the constants suppressed by the asymptotic notation in Lemma 4.6 are allowed to depend on the fixed parameters c and d.

Proof of Lemma 4.6. Let $Q = (-q, q]^d$, and note that $\mathcal{L}(Q) = \mathcal{L}'(Q) \pm 2^d n^{-c} \geq 1 - 2^d n^{-c}$. As in the proof of the base case of Proposition 4.3 (in Section 4.1), we consider a family $\mathfrak{B}$ of $O(q^d n^{c/2})$ half-open boxes with side lengths at most $D = n^{-c/(2d)}$ that partition Q.

For the first point, let $\mathfrak{B}_+ \subseteq \mathfrak{B}$ be the subcollection of boxes which intersect H, and let $\mathfrak{B}_- \subseteq \mathfrak{B}$ be the subcollection of boxes fully included in H, so that $|\mathfrak{B}_+ \setminus \mathfrak{B}_-| = O((q/D)^{d-1})$. We then observe that $|\mathcal{L}'(H) - \mathcal{L}(H)|$ is bounded by

$$O\left(|\mathfrak{B}_+ \setminus \mathfrak{B}_-|qD^d + |\mathfrak{B}|n^{-c} + \mathcal{L}(Q^c)\right) = O\left(q^d n^{-c/(2d)}\right).$$

For the second part, let $\mathfrak{B}_+$ be the subcollection of boxes that intersect R. The distance between the bounding hyperplanes of R, which is at most n^{-c}, is less than the width D of each box in our partition, so each box intersecting R must in fact intersect one of its bounding hyperplanes, whence $|\mathfrak{B}_+| = O((q/D)^{d-1})$ as earlier. We then observe as before that

$$\mathcal{L}(R) = O\left(|\mathfrak{B}_+|qD^d + \mathcal{L}(Q^c)\right) = O\left(q^d n^{-c/(2d)}\right). \qquad \square$$

Second, we isolate the part of the proof of Lemma 4.5 in which we approximated Kolmogorov distance via small boxes.

Lemma 4.7. *For fixed $c > 0$ and $d \in \mathbb{N}$, there exists a $c' = c'(c, d) > 0$ for which the following holds. Let $\mathcal{L}, \mathcal{L}'$ be probability distributions on $\mathbb{R}^d$, where $\mathcal{L}'$ is (possibly) a random object. Let $Q = (-q, q]^d \subseteq \mathbb{R}^d$ be a box for $q \geq 1$, and let $\mathfrak{B}$ be a partition of it into at most $(3q)^d n^{c/2}$ boxes with side lengths at most $n^{-c/(2d)}$. Suppose the following conditions are satisfied.*

(1) *For each $B \in \mathfrak{B}$, we have $|\mathcal{L}'(B) - \mathcal{L}(B)| \leq n^{-c}$ with probability at least $1 - n^{-c}$.*

(2) *$\mathcal{L}(Q^c) \leq n^{-c}$, and $\mathcal{L}'(Q^c) \leq n^{-c}$ with probability at least $1 - n^{-c}$.*

(3) *For each box $B \subseteq Q$ with side lengths at least n^{-c}, we have $\mathcal{L}(B) \leq q\,\mathrm{vol}(B)$.*

Then, with probability at least $1 - O(q^d n^{-c'})$, we have $\mathrm{d}_K(\mathcal{L}, \mathcal{L}') \leq O(q^d n^{-c'})$.

We will also need some lemmas for working with random graphs with constrained degree sequences. These lemmas will be deduced from powerful enumeration theorems due to McKay and Wormald [27] and Canfield, Greenhill, and McKay [12]. Before stating these lemmas, we define a notion of 'closeness' between two degree sequences. This definition is chosen to be convenient for the proof of Proposition 4.3; it has two cases which will both arise in different parts of the proof.

Definition 4.8. Consider a pair of sequences $(a(v))_{v \in V}$ and $(b(w))_{w \in W}$. Let $\widehat{\mathcal{A}}, \widehat{\mathcal{B}}$ be the uniform measures on these sequences (obtained by choosing a random element of each of these sequences). We say that $(a(v))_{v \in V}$ and $(b(w))_{w \in W}$ are *C-proximate* if at least one of the following two conditions holds.

(1) There is a bijection $\psi : V \to W$ such that $\sum_{v \in V} |a(v) - b(\psi(v))| \leq C|V|$.

(2) $||V| - |W|| \leq (|V| + |W|)^{1-1/C}$ and $d_K(\widehat{\mathcal{A}}, \widehat{\mathcal{B}}) \leq (|V| + |W|)^{-1/C}$.

In applications, we simply say that $(a(v))_{v \in V}$ and $(b(w))_{w \in W}$ are *proximate* if they are C-proximate for some $C = O(1)$ (recalling that implicit constants in asymptotic notation are allowed to depend on k and α).

We are now ready to state the promised pair of lemmas. We defer the details of their proofs to Appendix A. The first of these lemmas is for the non-bipartite setting. Recall that $\simeq$ means equality up to a multiplicative factor $(1 \pm n^{-\Omega(1)})$.

Lemma 4.9. *Let $(d_w)_{w \in W}$ be a sequence with even sum on a set W of n vertices such that*

- $d_w = n/2 \pm O(\sqrt{n \log n})$ *for each $w \in W$,*
- $\sum_{w \in T} d_w = n|T|/2 \pm O(n^{3/2})$ *for all $T \subseteq W$, and*
- $\sum_{w \in W}(d_w - n/2)^2 = O(n^2)$.

Such a sequence is a graphic sequence for all sufficiently large n. Let G be a uniformly random graph on W with this degree sequence. Then, for any fixed $v \in W$ and $S \subseteq W$ satisfying $|S|, n - |S| = \Omega(n)$, the following hold.

(1) *For any integer $0 \leq t \leq |S|$, parameterising $t = |S|/2 + \tau\sqrt{n}$, if $|\tau| > n^{1/10}$, then we have*

$$\mathbb{P}(\deg_S(v) = t) \leq \exp(-\Omega(\tau^2)),$$

and if $|\tau| \leq n^{1/10}$, then we have

$$\mathbb{P}(\deg_S(v) = t) \leq \exp\left(O\left(|\tau| + \sqrt{\log n}\right)\right)\mathbb{P}(Z = t),$$

where $Z = |R \cap S|$ for a random subset $R \subseteq W$ of size d_v, i.e.,

$$Z \sim \text{Hypergeometric}(n, |S|, d_v).$$

(2) *Let us write*

$$\mathbb{P}(\deg_S(v) = t) = p(v, (d_w)_{w \in S}, (d_w)_{w \notin S}, t)$$

as a function of v, the relevant degree sequences, and t. Then, for $t = |S|/2 \pm O(\sqrt{n \log n})$ and the other parameters as constrained above, this function $p(\cdot)$ depends continuously on its parameters, in the following sense: if

- $|t - t'|, |d_v - d'_{v'}| \leq n^{1/2-\Omega(1)}$,
- $(d_w)_{w \in S}$ *and $(d'_w)_{w \in S'}$ are proximate, and*
- $(d_w)_{w \in W \setminus S}$ *and $(d'_w)_{w \in W' \setminus S'}$ are proximate,*

then

$$p(v, (d_w)_{w \in S}, (d_w)_{w \in W \setminus S}, t) \simeq p(v', (d'_w)_{w \in S'}, (d'_w)_{w \in W' \setminus S'}, t').$$

Next, the second of the promised pair of lemmas is for the bipartite setting.

Lemma 4.10. *Let $((d_v)_{v \in V}, (d_w)_{w \in W})$ be a pair of sequences with identical sums on a bipartition $V \cup W$ with $|V|, |W| = \Theta(n)$ such that*

- $d_v = |W|/2 \pm O(\sqrt{n \log n})$ *for all $v \in V$ and $d_w = |V|/2 \pm O(\sqrt{n \log n})$ for all $w \in W$,*
- $\sum_{v \in T} d_v = |W||T|/2 \pm O(n^{3/2})$ *for all $T \subseteq V$ and $\sum_{w \in T} d_w = |V||T|/2 \pm O(n^{3/2})$ for all $T \subseteq W$, and*
- $\sum_{v \in V}(d_v - |W|/2)^2 = O(n^2)$ *and $\sum_{w \in W}(d_w - |V|/2)^2 = O(n^2)$.*

Such a pair of sequences form a bipartite-graphic sequence for all sufficiently large n. Let G be a uniformly random bipartite graph between V and W with this degree sequence. Then, for any fixed $u \in V$ and $S \subseteq W$ satisfying $|S|, n - |S| = \Omega(n)$, the following hold.

(1) *For any integer $0 \leq t \leq |S|$, parameterising $t = |S|/2 + \tau\sqrt{n}$, if $|\tau| > n^{1/10}$, then we have*

$$\mathbb{P}(\deg_S(u) = t) \leq \exp(-\Omega(\tau^2)),$$

and if $|\tau| \leq n^{1/10}$, then we have

$$\mathbb{P}(\deg_S(u) = t) \leq \exp\left(O\left(|\tau| + \sqrt{\log n}\right)\right)\mathbb{P}(Z = t),$$

where $Z = |R \cap S|$ for a random subset $R \subseteq W$ of size d_v, i.e.,

$$Z \sim \text{Hypergeometric}(|W|, |S|, d_v).$$

(2) *Let us write*

$$\mathbb{P}(\deg_S(u) = t) = p(u, (d_v)_{v \in V}, (d_w)_{w \in S}, (d_w)_{w \in W \setminus S}, t)$$

as a function of u, the relevant degree sequences, and t. Then, for $t = |S|/2 \pm O(\sqrt{n \log n})$ and the other parameters as constrained above, this function $p(\cdot)$ depends continuously on its parameters, in the following sense: if
  - *$|t - t'|, |d_u - d'_{u'}| \leq n^{1/2 - \Omega(1)}$,*
  - *$(d_v)_{v \in V}$ and $(d'_v)_{v \in V'}$ are proximate,*
  - *$(d_w)_{w \in S}$ and $(d'_w)_{w \in S'}$ are proximate, and*
  - *$(d_w)_{w \in W \setminus S}$ and $(d'_w)_{w \in W' \setminus S'}$ are proximate,*

then

$$p(u, (d_v)_{v \in V}, (d_w)_{w \in S}, (d_w)_{w \in W \setminus S}, t) \simeq p(u', (d'_v)_{v \in V'}, (d'_w)_{w \in S'}, (d'_w)_{w \in W' \setminus S'}, t').$$

Finally, we require the following concentration properties of the edge-counts in a random graph.

Lemma 4.11. *There are absolute constants $C, c > 0$ such that if $G \sim \mathbb{G}(n, 1/2)$ is a random graph, then with probability at least $1 - \exp(-cn)$ we have for all disjoint S, T that*

(1) $\sum_{v \in T}(\deg_S(v) - |S|/2)^2 \leq Cn^2$,
(2) $\sum_{v \in T}(\deg_T(v) - (|T| - 1)/2)^2 \leq Cn^2$,
(3) $|\sum_{v \in T}(\deg_S(v) - |S|/2)| \leq Cn^{3/2}$, and
(4) $|\sum_{v \in T}(\deg_T(v) - (|T| - 1)/2)| \leq Cn^{3/2}$.

The proof of Lemma 4.11 is an immediate application of a Chernoff bound and the union bound, similar to the proof of Lemma 2.1, so we omit the details.

Now we are ready to finish the proof of Proposition 4.3 by establishing its inductive step.

4.3. **Proof of the inductive step.** Consider $k - 1$ iterations of the α-swap process, giving rise to a partition of the vertices into sets V_x, for $x \in \{0, 1\}^k$, as defined in Definition 4.1. An additional iteration of the α-swap process will refine this to a partition into sets V_x, for $x \in \{0, 1\}^{k+1}$; to emphasise the difference between these two partitions we write W_x instead of V_x when $x \in \{0, 1\}^k$.

By the inductive hypothesis, there are real numbers $\pi_x \geq \alpha^{4(k-1)}/2$ and distributions $\mathcal{L}_x$ for $x \in \{0, 1\}^k$ such that the following properties are satisfied with high probability.

(B1) For each $x \in \{0,1\}^k$, we have

$$\big||W_x| - \pi_x n\big| \le n^{1-c_{\alpha,k-1}}.$$

(B2) For each $x \in \{0,1\}^k$, we have

$$d_K(\widehat{\mathcal{L}}_x, \mathcal{L}_x) \le n^{-c_{\alpha,k-1}}.$$

(B3) For each vertex $v \in V(G)$ and each $x \in \{0,1\}^k$, we have

$$\big|\deg_{W_x}(v) - |W_x|/2\big| \le C_{\alpha,k-1} n^{1/2}\sqrt{\log n}.$$

(B4) For each $x \in \{0,1\}^k$, and each box B with side lengths $n^{-c_{\alpha,k-1}}$ we have

$$\mathcal{L}_x(B) \le \mathrm{vol}(B)\exp(C_{\alpha,k-1}\sqrt{\log n}).$$

Here, we remind the reader that $\widehat{\mathcal{L}}_x$ is an empirical distribution measuring the degrees of vertices in W_x into the various sets W_y. Also, we remark that although **(B4)** as written only concerns boxes with side lengths *exactly* $n^{-c_{\alpha,k}}$, a simple covering argument shows that the same conclusion holds when B is a box with side lengths *at least* $n^{-c_{\alpha,k}}$ (up to a constant factor).

Next, let

$$\mathcal{R} = \Big((W_x)_{x \in \{0,1\}^k}, (\deg_{W_x}(v))_{v \in V(G), x \in \{0,1\}^k}\Big)$$

record the part and degree information after $k-1$ iterations of the α-swap process, so **(B1)** to **(B4)** are all really properties of $\mathcal{R}$. Let $\mathcal{E}$ be the event that all the conclusions of Lemma 4.11 hold for all disjoint subsets of vertices S and T. By Lemma 4.11, we have

$$\mathbb{E}[\mathbb{P}(\mathcal{E}^c \mid \mathcal{R})] = \mathbb{P}(\mathcal{E}^c) \le e^{-cn}$$

for some universal $c > 0$, so by Markov's inequality, with high probability, $\mathcal{R}$ has the property that

(B5) $\mathbb{P}(\mathcal{E} \mid \mathcal{R}) \ge 1 - e^{-(c/2)n}$.

Now, let us condition on an outcome of $\mathcal{R}$ satisfying **(B1)** to **(B5)**; we say that such an outcome is *well-behaved*. It suffices to prove that, in the resulting conditional probability space, **(A1)** to **(A4)** hold with high probability. Note that, conditionally, G is now a random graph with certain degree constraints. To be precise, for each $x \in \{0,1\}^k$, the induced subgraph $G[W_x]$ is uniform over all graphs in which each $v \in W_x$ has degree $\deg_{W_x}(v)$, and for each pair of distinct $x, y \in \{0,1\}^k$, the subgraph $G[W_x, W_y]$ (consisting of the edges of G between W_x and W_y) is uniform over all bipartite graphs in which each $v \in W_x$ has degree $\deg_{W_y}(v)$ and each $v \in W_y$ has degree $\deg_{W_x}(v)$. Furthermore, all these random subgraphs of the form $G[W_x], G[W_x, W_y]$ are independent, and **(B1)**, **(B3)** and **(B5)** in particular ensure that either Lemma 4.9 or Lemma 4.10 applies to all these subgraphs.

Recalling that we have performed $k-1$ iterations of the α-swap procedure so far, we now consider the effect of a kth α-swap. Recall that this α-swap has two steps. First, the $\lfloor \alpha n \rfloor$ unfriendliest vertices on each side are swapped. The information recorded in $\mathcal{R}$ is enough to determine the outcome of this first step. Second, a random pair of sets of $\lfloor \alpha^4 n \rfloor$ vertices on each side are swapped; let $\mathcal{S}$ be the random pair of sets that are swapped in this second step, and note that $\mathcal{S}$ is independent from G conditional on the partition at that time.

For the remainder of this proof, asymptotic notation should be understood to be treating k, α as fixed constants, so, for example, the inequality in **(B2)** can be described as saying $d_K(\widehat{\mathcal{L}}_x, \mathcal{L}_x) \le n^{-\Omega(1)}$.

4.3.1. *Concentration of the part sizes.* First we prove that **(A1)** holds with high probability. Let $S_i = \{z \in \{0,1\}^k : z_k = i\}$, and recall that the bisection resulting from the first $k-1$ iterations of the α-swap process has parts $A_{k-1} = \bigcup_{z \in S_0} W_z$ and $B_{k-1} = \bigcup_{z \in S_1} W_z$. (Recall that z_k records whether a vertex is in A_{k-1} or B_{k-1}.)

Consider any $z \in \{0,1\}^k$, and let W_z' be the portion of W_z that is swapped during the first step of the kth α-swap (i.e., these vertices are among the $\lfloor \alpha n \rfloor$ unfriendliest vertices in their part of the bisection $A_{k-1} \cup B_{k-1}$; this is determined by the outcome of $\mathcal{R}$ we have conditioned on). It suffices to prove that $|W_z'| = \pi_z' n \pm n^{1-\Omega(1)}$, for some π_z' that does not depend on the specific choice of $\mathcal{R}$ that we are conditioning on (but demanding no lower bound on π_z'). Indeed, for any $b \in \{0,1\}$, the second part of the α-swap process (in which we randomly swap sets A', B' of $\lfloor \alpha^4 n \rfloor$ vertices on both sides) will then, with high probability, yield $|V_{(z,b)}| = \pi_{(z,b)} n \pm n^{1-\Omega(1)}$, where

$$\pi_{(z,b)} = \begin{cases} \alpha^4 \pi_z' + (1-\alpha^4)(\pi_z - \pi_z') & \text{if } z_k = b, \\ \alpha^4(\pi_z - \pi_z') + (1-\alpha^4)\pi_z' & \text{if } z_k \ne b \end{cases}$$
$$\ge \alpha^4 \pi_z \ge \alpha^4 \cdot \alpha^{4(k-1)}/2 = \alpha^{4k}/2.$$

Here we have used **(B1)** and a Chernoff bound for the hypergeometric distribution; see for example [22, Theorem 2.10].

To this end, we study the sets W_z'. Assume without loss of generality that $z_k = 0$ (i.e., $W_z' \subseteq A_{k-1}$). Let A' be the set of the $\lfloor \alpha n \rfloor$ unfriendliest vertices in A_{k-1} (so $W_z' = W_z \cap A'$), and let $A^{(\zeta)}$ be the set of vertices in A_{k-1} with friendliness at most $\zeta \sqrt{n}$. We will approximate A' with $A^{(\zeta)}$, for an appropriate choice of ζ.

For $\zeta \in \mathbb{R}$, define the affine half-space

$$H_\zeta = \left\{ d \in \mathbb{R}^{\{0,1\}^k} : \sum_{y \in S_0} d_y - \sum_{y \in S_1} d_y \le \zeta \right\}.$$

Then, $|A^{(\zeta)}| = \sum_{y \in S_0} |W_y| \widehat{\mathcal{L}}_y(H_\zeta)$. Let us set

$$f(\zeta) = \sum_{y \in S_0} \pi_y \mathcal{L}_y(H_\zeta).$$

By the second point in Lemma 4.6, the function f satisfies a Lipschitz-like property: if $|\zeta - \zeta'| \le n^{-\Omega(1)}$ then $|f(\zeta) - f(\zeta')| \le n^{-\Omega(1)}$. Since $\lim_{\zeta \to -\infty} f(\zeta) = 0$ and $\lim_{\zeta \to \infty} f(\zeta) = \sum_{y \in S_0} \pi_y = 1/2 + o(1)$, there is some ζ_α such that $|f(\zeta_\alpha) - \alpha| \le n^{-\Omega(1)}$.

By the first point in Lemma 4.6, we then have $||A'| - |A^{(\zeta_\alpha)}|| \le n^{1-\Omega(1)}$. That is to say, the set A' differs from the set $A^{(\zeta_\alpha)}$ by only $n^{1-\Omega(1)}$ elements (noting that either $A' \subseteq A^{(\zeta)}$ or $A^{(\zeta)} \subseteq A'$ always). Again using the first point in Lemma 4.6, it follows that

$$|W_z'| = |W_z \cap A'| = |W_z \cap A^{(\zeta_\alpha)}| \pm n^{1-\Omega(1)} = |W_z| \widehat{\mathcal{L}}(H_{\zeta_\alpha}) \pm n^{1-\Omega(1)} = \pi_z' n \pm n^{1-\Omega(1)},$$

as desired, where $\pi_z' = \pi_z \mathcal{L}(H_{\zeta_\alpha})$.

4.3.2. *Some intermediate empirical degree distributions.* For a vertex v, define the degree vector

(4.3) $$\vec{g}(v) = \left((\deg_{W_y}(v) - |W_y|/2)/\sqrt{n}\right)_{y \in \{0,1\}^k}$$

(which is determined by $\mathcal{R}$), and recall that for $z \in \{0,1\}^k$, $\widehat{\mathcal{L}}_z$ is the uniform measure on the sequence $(\vec{g}(v))_{v \in W_z}$. For $b \in \{0,1\}$, let $\widehat{\mathcal{D}}_{(z,b)}$ be the uniform measure on $(\vec{g}(v))_{v \in V_{(z,b)}}$ (which depends on $\mathcal{R}, \mathcal{S}$, but not the remaining randomness of G). This can be thought of as an 'intermediate' empirical degree distribution between $\widehat{\mathcal{L}}_z$ and $\widehat{\mathcal{L}}_{(z,b)}$, where we consider the degrees from vertices in $V_{(z,b)}$ into the sets W_y.

The considerations in the previous section give us quite strong control over the $\widehat{\mathcal{D}}_{(z,b)}$. Indeed, for any box $B \subseteq \mathbb{R}^{\{0,1\}^k}$ let $W_z(B)$ be the set of all $v \in W_z$ with $\vec{g}(v) \in B$, and as in the last section, assume without loss of generality that $z_k = 0$. Let $\rho'_z(B) = \pi_z \mathcal{L}_z(B \cap H_{\zeta_\alpha})$, so that $|W_z(B) \cap W'_z| = \rho'_z(B)n + n^{1-\Omega(1)}$, and a concentration inequality for the hypergeometric distribution shows that with probability $1 - O(1/n)$ over the randomness of $\mathcal{S}$, we have $|W_z(B) \cap V_{(z,b)}| = \rho_z(B)n \pm n^{1-\Omega(1)}$, where

$$\rho_{(z,b)}(B) = \begin{cases} \alpha^4 \rho'_z(B) + (1 - \alpha^4)(\pi_z \mathcal{L}_z(B) - \rho'_z(B)) & \text{if } z_k = b, \\ \alpha^4 (\pi_z \mathcal{L}_z(B) - \rho'_z(B)) + (1 - \alpha^4)\rho'_z(B) & \text{if } z_k \neq b. \end{cases}$$

Since $\widehat{\mathcal{D}}_{(z,b)}(B) = |W_z(B) \cap V_{(z,b)}|/|V_{(z,b)}|$, **(A1)** implies that $\widehat{\mathcal{D}}_{(z,b)}(B) = \mathcal{D}_{(z,b)}(B) \pm n^{-\Omega(1)}$, where $\mathcal{D}_{(z,b)}$ is the probability distribution for which $\mathcal{D}_{(z,b)}(S)$ is proportional to $\rho_{(z,b)}(S)$ for all boxes $S \subseteq \mathbb{R}^{\{0,1\}^k}$. Recalling **(B3)** and **(B4)**, and partitioning the big box

$$Q = \left(-C_{\alpha,k-1}\sqrt{\log n}, C_{\alpha,k-1}\sqrt{\log n}\right]^{2^k}$$

into $n^{c/2+o(1)}$ boxes with side lengths $n^{-c/(2 \cdot 2^k)}$ for a sufficiently small $c > 0$, it follows from Lemma 4.7 that $d_K(\widehat{\mathcal{D}}_{(z,b)}, \mathcal{D}_{(z,b)}) \leq n^{-\Omega(1)}$ with high probability over the randomness of $\mathcal{S}$.

4.3.3. *Controlling the outlier degrees.* We next prove that **(A3)** holds with high probability. In addition to our conditioning on $\mathcal{R}$, in this subsection we also condition on an outcome of $\mathcal{S}$ such that each $|V_x| = \Omega(n)$. Note we proved in Section 4.3.1 that $|V_x| = \pi_x n + o(n)$ with high probability and $\pi_x \geq \alpha^{4(k-1)}/2$, so in particular $|V_x| \geq \Omega(n)$ with high probability.

Fix an arbitrary $x \in \{0,1\}^{k+1}$ and $y \in \{0,1\}^k$. We wish to show that with high probability, for every $v \in W_y$ we have $\left|\deg_{V_x}(v) - |V_x|/2\right| \leq C_{\alpha,k}\sqrt{n \log n}$, for some $C_{\alpha,k} > 0$. This suffices, since we will then be able to take the union bound over all $O(1)$ choices of x, y. The desired bound follows from part (1) of Lemma 4.9 and part (1) of Lemma 4.10 along with a Chernoff bound for the hypergeometric distribution and a union bound over $v \in W_y$: if $z = (x_1, \ldots, x_k)$ satisfies $z = y$, then we consider the degree-constrained random graph $G[W_y]$, and if we instead have $z \neq y$, then we consider the degree-constrained bipartite graph $G[W_y, W_z]$.

4.3.4. *Defining the ideal distributions.* We shall address **(A4)** first before turning to **(A2)** (which is by far the most involved of the four properties). Therefore, at this juncture, we take a moment to say something about how we will define the distributions $\mathcal{L}_x$ for $x \in \{0,1\}^{k+1}$. First, for specific outcomes of $\mathcal{R}, \mathcal{S}$ (which determine the sets V_x

for $x \in \{0,1\}^{k+1}$), we let $\mathcal{L}_x^{\mathcal{R},\mathcal{S}}$ be the distribution obtained by choosing a random $v \in V_x$ and sampling its degree vector

$$\vec{d}(v) = \left(\left(\deg_{V_y}(v) - |V_y|/2\right)/\sqrt{n}\right)_{y \in \{0,1\}^{k+1}}$$

according to the remaining randomness in G. We will later show that if $\mathcal{R}$ is well-behaved, and $\mathcal{S}$ also satisfies certain properties that hold with high probability, then $\mathcal{L}_x^{\mathcal{R},\mathcal{S}}$ is actually not very sensitive to the specific choice of $\mathcal{R}$ and $\mathcal{S}$, whence we will be able to prove that **(A2)** holds with high probability when we take $\mathcal{L}_x$ to be any such $\mathcal{L}_x^{\mathcal{R},\mathcal{S}}$.

4.3.5. *Anti-concentration.*

Here, we show that **(A4)** holds. As in Section 4.3.3, we condition on a well-behaved outcome of $\mathcal{R}$ as well as on an outcome of $\mathcal{S}$ such that each $|V_x| = \Omega(n)$. By the above discussion, it suffices to show that $\mathcal{L}_x^{\mathcal{R},\mathcal{S}}$ satisfies the anti-concentration property in **(A4)**. The rough idea for establishing this involves combining Lemmas 4.9 and 4.10 (which provide anti-concentration subject to the remaining randomness in G) with the anti-concentration property in **(B4)** coming from the outcome of the process so far.

Fix a vertex $v \in W_z$ for some $z \in \{0,1\}^k$. By part (1) of Lemma 4.9 and part (1) of Lemma 4.10, for $y \in \{0,1\}^k$ and $t \in \mathbb{N}$, parameterising $t = |V_{(y,0)}|/2 + \tau\sqrt{n}$ and writing $d_v = \deg_{W_y}(v)$, we have

$$\mathbb{P}\left(\deg_{V_{(y,0)}}(v) = t\right) \leq \exp\left(O\left(\sqrt{\log n}\right)\right) n^{-1/2}$$

uniformly in t. Indeed, when applying Lemma 4.9, this holds with room to spare when $|\tau| > |V_{(y,0)}|^{1/10} = \Omega(n^{1/10})$, and when $|\tau| \leq |V_{(y,0)}|^{1/10}$, we may see that we uniformly have

$$\mathbb{P}\left(\deg_{V_{(y,0)}}(v) = t\right) \leq \exp\left(O\left(|\tau| + \sqrt{\log n}\right)\right) \frac{\binom{|V_{(y,0)}|}{t}\binom{|W_y|-|V_{(y,0)}|}{d_v-t}}{\binom{m-1}{d_v}}$$

$$\leq \exp\left(O\left(\sqrt{\log n}\right)\right) n^{-1/2}$$

by a standard anti-concentration inequality for the hypergeometric distribution (see for example [14, Lemma 3.2]).

Since we are conditioning on $\mathcal{R}, \mathcal{S}$, the degree-constrained random graph $G[W_z]$ and the degree-constrained bipartite graphs $G[W_z, W_y]$ are all independent, so the 2^k different degrees $\deg_{V_{(y,0)}}(v)$, for $y \in \{0,1\}^k$, are all independent as well. Thus, we obtain the uniform joint anti-concentration bound

$$\mathbb{P}\left(\deg_{V_{(y,0)}}(v) = t_y \text{ for all } y \in \{0,1\}^k\right) \leq \exp\left(O\left(\sqrt{\log n}\right)\right)\left(n^{-1/2}\right)^{2^k}.$$

Note that for each $y \in \{0,1\}^k$, the degrees $\deg_{V_{(y,0)}}(v)$ and $\deg_{V_{(y,1)}}(v)$ are certainly not independent, since $\deg_{V_{(y,0)}}(v) + \deg_{V_{(y,1)}}(v) = \deg_{W_y}(v)$ is determined by $\mathcal{R}$. Nonetheless, our joint anti-concentration bound does imply that for any box $B \subseteq \mathbb{R}^{\{0,1\}^{k+1}}$ with side lengths $D \geq 1/\sqrt{n}$, we have

$$(4.4) \qquad \mathbb{P}\left(\vec{d}(v) \in B\right) \leq \exp\left(O\left(\sqrt{\log n}\right)\right) D^{2^k}.$$

Note that $\mathrm{vol}(B) = D^{2^{k+1}}$, so (4.4) only provides 'half as much anti-concentration' as we desire for **(A4)**. So far, we have only considered anti-concentration of $\vec{d}(v)$ when v is a fixed vertex; we will next establish the remainder of our anti-concentration and **(A4)** proper by allowing v to vary and appealing to **(B2)** and **(B4)**.

Recall the definition of the degree vectors $\vec{g}(v)$ and the intermediate empirical distributions $\widehat{\mathcal{D}}_{(z,b)}$ from Section 4.3.2: the empirical distribution $\widehat{\mathcal{L}}_z$ is defined in terms of a uniformly random vertex $v \in W_z$, and $\widehat{\mathcal{D}}_{(z,b)}$ is then obtained by conditioning on the event that $v \in V_{(z,b)}$. Recall that $|V_{(z,b)}| = \Omega(n) = \Omega(1)|W_z|$, meaning that we are conditioning on an event that holds with probability at least $\Omega(1)$. So, **(B4)** implies the same anti-concentration property in the $\widehat{\mathcal{D}}_{(z,b)}$, i.e.,

$$(4.5) \qquad \widehat{\mathcal{D}}_x(B) \leq \exp\left(O\left(\sqrt{\log n}\right)\right) \mathrm{vol}(B),$$

for all boxes $B \subseteq \mathbb{R}^{\{0,1\}^k}$ with side lengths at least n^{-c}, where $c = c_{\alpha,k-1}$, and all $x \in \{0,1\}^{k+1}$.

Now, let $\pi : \mathbb{R}^{\{0,1\}^{k+1}} \to \mathbb{R}^{\{0,1\}^k}$ be the linear map $(d_x)_{x \in \{0,1\}^{k+1}} \mapsto (d_{(y,0)} + d_{(y,1)})_{y \in \{0,1\}^k}$. Note that $\vec{g}(v) = \pi(\vec{d}(v))$ for all v, and note that if $B \subseteq \mathbb{R}^{\{0,1\}^{k+1}}$ is a box with side lengths n^{-c}, then $\pi(B)$ is contained in a box with side lengths $2n^{-c}$. So, by (4.4) and (4.5), we have

$$\begin{aligned}
\mathcal{L}_x^{\mathcal{R},\mathcal{S}}(B) &= \sum_{v \in V_x \,:\, \vec{g}(v) \in \pi(B)} \frac{1}{|V_x|} \cdot \mathbb{P}\left(\vec{d}(v) \in B\right) \\
&\leq \widehat{\mathcal{D}}_x(\pi(B)) \sup_{v \in V_x} \mathbb{P}\left(\vec{d}(v) \in B\right) \\
&\leq \exp\left(O\left(\sqrt{\log n}\right)\right) (2n^{-c})^{2^k} (n^{-c})^{2^k} \\
&\leq \exp\left(O\left(\sqrt{\log n}\right)\right) \mathrm{vol}(B)
\end{aligned}$$

for all $x \in \{0,1\}^{k+1}$, as desired.

4.3.6. *Concentration of the empirical degree distributions.* In this subsection we use a second moment calculation as in Section 4.1 to show that if we condition on appropriate outcomes of $\mathcal{R}$ and $\mathcal{S}$, then with high probability, for any $x \in \{0,1\}^{k+1}$, we have

$$\mathrm{d_K}\left(\widehat{\mathcal{L}}_x, \mathcal{L}_x^{\mathcal{R},\mathcal{S}}\right) \leq n^{-\Omega(1)}.$$

We shall later prove that the distributions $\mathcal{L}_x^{\mathcal{R},\mathcal{S}}$, for appropriate $\mathcal{R}, \mathcal{S}$, are all Kolmogorov-close to each other; it will then follow that **(A2)** holds with high probability.

As in the previous two subsections, we condition on a well-behaved outcome of $\mathcal{R}$ and an outcome of $\mathcal{S}$ for which $|V_x| = \Omega(n)$ for all $x \in \{0,1\}^{k+1}$. Fix an $x \in \{0,1\}^{k+1}$, and as before, let $Q = (-C_{\alpha,k}\sqrt{\log n}, C_{\alpha,k}\sqrt{\log n}]^{2^{k+1}}$, where $C_{\alpha,k}$ is as chosen in Section 4.3.3 (so, we have say $\mathcal{L}_x(Q^c) \leq n^{-2}$).

We wish to apply Lemma 4.5. To this end, we shall, for an arbitrary pair of vertices u and v, study conditional probabilities of the form

$$\mathbb{P}\left(\deg_{V_{(z,b)}}(v) = t \,\Big|\, N_{W_z}(u) = T\right),$$

where $z \in \{0,1\}^k$, $b \in \{0,1\}$, and T is a set of $\deg_{W_z}(u)$ elements of $W_z \setminus \{u\}$. Let $R_{(z,b)} = \left\{ t : |t - |V_{(z,b)}|/2| \leq C_{\alpha,k} \sqrt{n \log n} \right\}$. We will show that for such data u, v, z, b, and each $t \in R_{(z,b)}$, the value of the above conditional probability is not very sensitive to the choice of T.

Let $y \in \{0,1\}^k$ be such that $v \in W_y$. As usual, we need to consider separately the case where $y = z$ and where $y \neq z$; in the former case, we study the degree-constrained random graph $G[W_y]$, and in the latter case we study the degree-constrained random bipartite graph $G[W_y, W_z]$.

If $y = z$, then having conditioned on the event $N_{W_y}(u) = T$, now $G[W_y \setminus \{u\}]$ is a random graph with a particular degree sequence (namely, the degree sequence where we delete u if it is in W_y, and if so we also decrement the degree of every vertex in T by one). Considering how this degree sequence varies for different choices of T, T', it follows from part (2) of Lemma 4.9 (and the first part of Definition 4.8) that for each u, v, z, b as above, each $t \in R_{(z,b)}$, and each such pair T, T', we have

$$\mathbb{P}\left(\deg_{V_{(z,b)}}(v) = t \,\middle|\, N_{W_z}(u) = T \right) \simeq \mathbb{P}\left(\deg_{V_{(z,b)}}(v) = t \,\middle|\, N_{W_z}(u) = T' \right).$$

We obtain the same conclusion if $z \neq y$ by considering the bipartite graph $G[W_y, W_z]$, except now relying on Lemma 4.10.

The above argument implies that for all u, v, z, b, t as above, we in fact have

$$\mathbb{P}\left(\deg_{V_{(z,b)}}(v) = t \,\middle|\, N_{W_z}(u) = T \right) \simeq \mathbb{P}\left(\deg_{V_{(z,b)}}(v) = t \right).$$

Observing that all the random subgraphs of the form $G[W_y], G[W_y, W_z]$ are independent, we deduce that for any $\vec{\tau}, \vec{\sigma} \in Q$, we have

$$\mathbb{P}(\vec{d}(v) = \vec{\tau} \text{ and } \vec{d}(u) = \vec{\sigma}) \simeq \mathbb{P}(\vec{d}(v) = \vec{\tau})\mathbb{P}(\vec{d}(u) = \vec{\sigma}).$$

Therefore we can apply Lemma 4.5, using **(A4)** (which we have already proved) and the fact that $\mathcal{L}_x(Q^c) \leq 1/n^2$ for all $x \in \{0,1\}^{k+1}$, to conclude that **(A2)** holds with high probability.

4.3.7. *Sensitivity to the conditioned information.* To finish, we wish to show that for all $x \in \{0,1\}^{k+1}$, well-behaved $\mathcal{R}$ and $\mathcal{R}'$, and almost all outcomes $\mathcal{S}$ and $\mathcal{S}'$, we have

$$d_K\left(\mathcal{L}_x^{\mathcal{R},\mathcal{S}}, \mathcal{L}_x^{\mathcal{R}',\mathcal{S}'} \right) \leq n^{-\Omega(1)}.$$

This will complete the proof of the inductive step of Proposition 4.3.

Recall the definitions of the degree vectors $\vec{g}(v)$ and the intermediate degree distributions $\widehat{\mathcal{D}}_x, \mathcal{D}_x$ from Section 4.3.2. In that subsection, we showed for all well-behaved $\mathcal{R}$ that, with high probability over $\mathcal{S}$, we have $d_K(\widehat{\mathcal{D}}_x, \mathcal{D}_x) \leq n^{-\Omega(1)}$. Let c (depending on α, k) be sufficiently small such that $d_K(\widehat{\mathcal{D}}_x, \mathcal{D}_x) \leq n^{-c}$ with high probability, and let us now call an outcome of $\mathcal{S}$ *well-behaved* if this is the case for all $x \in \{0,1\}^{k+1}$.

Let $\pi : \mathbb{R}^{\{0,1\}^{k+1}} \to \mathbb{R}^{\{0,1\}^k}$ be the linear map $(d_x)_{x \in \{0,1\}^{k+1}} \mapsto (d_{(y,0)} + d_{(y,1)})_{y \in \{0,1\}^k}$, as was the case in Section 4.3.5. If we condition on any $\mathcal{R}, \mathcal{S}$, then for any $v \in V_x$ and any $\vec{\tau} \in \mathbb{R}^{\{0,1\}^{k+1}}$ with $\vec{g}(v) = \pi(\vec{\tau})$, we have

$$\mathbb{P}(\vec{d}(v) = \vec{\tau}) = \prod_{y \in \{0,1\}^k} \mathbb{P}(\deg_{V_{(y,0)}}(v) = t_y),$$

where $(t_y - |V_{(y,0)}|/2)/\sqrt{n} = \tau_y$. Now, probabilities of the form $\mathbb{P}(\deg_{V_x}(v) = t)$ are actually not very sensitive to the specific choice of $v, t, \mathcal{R}, \mathcal{S}$, in the following sense. Suppose $\mathcal{R}, \mathcal{S}, \mathcal{R}', \mathcal{S}'$ are all well-behaved, and for some $y \in \{0, 1\}^k$, let $v \in W_y^{\mathcal{R}}$ and $v' \in W_y^{\mathcal{R}'}$ be vertices in the 'same part' with respect to $\mathcal{R}$ and $\mathcal{R}'$. Moreover, suppose that

$$\left\| \vec{g}^{\mathcal{R}}(v) - \vec{g}^{\mathcal{R}'}(v') \right\|_\infty \leq n^{1/2 - \Omega(1)}.$$

Then for any $x \in \{0, 1\}^{k+1}$ and $t, t' = \pi_x n/2 \pm n^{1/2 - \Omega(1)}$, by part (2) of Lemma 4.9 and part (2) of Lemma 4.10 (and using the second part of Definition 4.8), we have

$$(4.6) \qquad \mathbb{P}\left(\deg_{V_x^{\mathcal{R},\mathcal{S}}}(v) = t \,\middle|\, \mathcal{R}, \mathcal{S}\right) \simeq \mathbb{P}\left(\deg_{V_x^{\mathcal{R}',\mathcal{S}'}}(v') = t' \,\middle|\, \mathcal{R}', \mathcal{S}'\right).$$

Now, consider well-behaved data $\mathcal{R}, \mathcal{S}, \mathcal{R}', \mathcal{S}'$, and fix some $x \in \{0, 1\}^{k+1}$. Our next objective is to construct an injective mapping ϕ between subsets of $V_x^{\mathcal{R},\mathcal{S}}$ and $V_x^{\mathcal{R}',\mathcal{S}'}$ that maps a vertex $v \in V_x^{\mathcal{R},\mathcal{S}}$ to a vertex $\phi(v) \in V_x^{\mathcal{R}',\mathcal{S}'}$ with 'roughly the same statistics' as v. This will allow us to compare probabilities conditional on the outcomes $(\mathcal{R}, \mathcal{S})$ with probabilities conditional on the outcomes $(\mathcal{R}', \mathcal{S}')$.

First, let $Q = (-C_{\alpha,k}\sqrt{\log n}, C_{\alpha,k}\sqrt{\log n}]^{2^{k+1}}$, so by the same considerations as in Section 4.3.3, we know that

$$\mathcal{L}_x^{\mathcal{R},\mathcal{S}}(Q^c) \leq 1/n \text{ and } \mathcal{L}_x^{\mathcal{R}',\mathcal{S}'}(Q^c) \leq 1/n.$$

Now, partition Q into a collection $\mathfrak{B}$ of $n^{c/2 + o(1)}$ boxes with side lengths $n^{-c/(2 \cdot 2^{k+1})}$. Since $\mathcal{R}, \mathcal{S}, \mathcal{R}', \mathcal{S}'$ are all well-behaved, we have

$$\mathrm{d}_{\mathrm{K}}\left(\widehat{\mathcal{D}}_x^{\mathcal{R},\mathcal{S}}, \widehat{\mathcal{D}}_x^{\mathcal{R}',\mathcal{S}'}\right) \leq n^{-c}.$$

Also, we may assume with no loss of generality that c is sufficiently small, and in particular, that $c < c_{\alpha,k}$, so by **(A1)**, we have $|V_x^{\mathcal{R},\mathcal{S}}| = |V_x^{\mathcal{R}',\mathcal{S}'}| \pm n^{1-c}$. It follows that, for each $B \in \mathfrak{B}$, if we consider the sets

$$V_x^{\mathcal{R},\mathcal{S}}(B) = \{v \in V_x^{\mathcal{R},\mathcal{S}} : \vec{g}(v) \in \pi(B)\} \text{ and } V_x^{\mathcal{R}',\mathcal{S}'}(B) = \{v \in V_x^{\mathcal{R}',\mathcal{S}'} : \vec{g}(v) \in \pi(B)\},$$

then we have

$$|V_x^{\mathcal{R},\mathcal{S}}(B)| = |V_x^{\mathcal{R}',\mathcal{S}'}(B)| \pm O\left(n^{1-c}\right).$$

Now, let

$$m(B) = \min\left\{|V_x^{\mathcal{R},\mathcal{S}}(B)|, |V_x^{\mathcal{R}',\mathcal{S}'}(B)|\right\},$$

and let $U \subseteq V_x^{\mathcal{R},\mathcal{S}}$ be obtained by choosing $m(B)$ elements from each $V_x^{\mathcal{R},\mathcal{S}}(B)$ for $B \in \mathfrak{B}$, so that

$$|U| \geq |V_x^{\mathcal{R},\mathcal{S}}| - O\left(n^{1-c/2+o(1)}\right).$$

Let $\phi : U \to V_x^{\mathcal{R}',\mathcal{S}'}$ be an injection such that $\phi(v) \in V_x^{\mathcal{R}',\mathcal{S}'}(B)$ for each $v \in U \cap V_x^{\mathcal{R},\mathcal{S}}(B)$ (we think of ϕ as providing a coupling between a uniformly random vertex in $V_x^{\mathcal{R},\mathcal{S}}$ and a uniformly random vertex in $V_x^{\mathcal{R}',\mathcal{S}'}$, up to a change of $O(n^{-c})$ in total variation distance). Each $B \in \mathfrak{B}$ has ℓ^∞-diameter $O(n^{-c/(2 \cdot 2^{k+1})})$, so applying (4.6) and summing over points in B, we see for all $v \in U$ that

$$\mathbb{P}\left(\vec{d}(v) \in B \,|\, \mathcal{R}, \mathcal{S}\right) = (1 \pm n^{-c'})\mathbb{P}\left(\vec{d}(\phi(v)) \in B \,|\, \mathcal{R}', \mathcal{S}'\right),$$

for some $c' > 0$ depending on c and k. Now, if we coarsen $\mathfrak{B}$ into a partition $\mathfrak{B}'$ of $n^{c'/2+o(1)}$ boxes with side lengths at most $n^{-c'/(2\cdot 2^{k+1})}$, then we easily see that the conditions of Lemma 4.7 are satisfied, and we deduce that $d_K(\mathcal{L}_x^{\mathcal{R},\mathcal{S}}, \mathcal{L}_x^{\mathcal{R}',\mathcal{S}'}) \leq n^{-\Omega(1)}$ as desired. This finishes the inductive proof of Proposition 4.3.

APPENDIX A. PROBABILITIES IN DEGREE-CONSTRAINED GRAPH MODELS

We start by showing how Lemma 4.9 follows from a series of results of increasing precision about random graphs with specified degree sequences.

Proposition A.1. *Let* $(d_w)_{w\in W}$ *be a sequence with even sum on a set* W *of* n *vertices such that, defining* β_w *by* $d_w = (n-1)/2 + \beta_w\sqrt{(n-1)}/2$, *we have*

- $|\beta_w| \leq \log n$ *for each* $w \in W$, *and*
- $\sum_{w\in W} \beta_w^2 \leq n(\log n)^{1/9}$.

Such a sequence is a graphic sequence for all sufficiently large n. *Let* G *be a uniformly random graph with this degree sequence on the vertex set* W. *Consider any fixed* $v \in W$, *any fixed subset* $S \subseteq W$ *of size* h *satisfying* $\min(h, n-h) \geq n/(\log n)^{1/8}$, *and an integer* $t \in [0, d_v]$. *If* $|t - h/2| > n^{3/5}$, *then we have*

$$(A.1) \qquad \mathbb{P}(\deg_S(v) = t) \leq \exp(-\Omega((t-h/2)^2/n)).$$

If $|t - h/2| \leq n^{3/5}$ *on the other hand, then we have*

$$(A.2) \qquad \mathbb{P}(\deg_S(v) = t) = (1 \pm O(n^{-1/10}))\frac{\binom{h}{t}\binom{n-h-1}{d_v-t}}{\binom{n-1}{d_v}} \exp(\Lambda_1 - \Lambda_2 - \Lambda_3 + \Lambda_4),$$

where $\Lambda_1, \Lambda_2, \Lambda_3$ *and* Λ_4 *are given by*

$$\Lambda_1 = \frac{1}{2n^2}\left(\sum_{i\in W}\beta_i\right)\left(\sum_{i\in W}\beta_i - 2n\beta_v\right),$$

$$\Lambda_2 = \sum_{i\in S\setminus v}\left(1 - \frac{2t}{h}\right)\frac{\beta_i}{\sqrt{n-1}} + \sum_{i\in S^c\setminus v}\left(1 - \frac{2(d_v-t)}{(n-h)}\right)\frac{\beta_i}{\sqrt{n-1}},$$

$$\Lambda_3 = \frac{1}{2}\sum_{i\in W\setminus v}\frac{\beta_i^2}{n-1}, \ and$$

$$\Lambda_4 = \frac{1}{2nh}\sum_{i,j\in S\setminus v}(\beta_i - \beta_j)^2 + \frac{1}{2n(n-h)}\sum_{i,j\in S^c\setminus v}(\beta_i - \beta_j)^2,$$

the sums in the definition of Λ_4 *being over all (unordered) two-element subsets.*

First, we deduce Lemma 4.9 from Proposition A.1. To this end, we need Lemma A.2 comparing the moments of distributions that are bounded and Kolmogorov-close.

Lemma A.2. *Fix* $c > 0$ *and* $k \in \mathbb{N}$. *If* $(a_v)_{v\in V}$ *and* $(b_u)_{u\in U}$ *are two sequences of* $\Omega(n)$ *real numbers with* $\big||V| - |U|\big| \leq n^{1-c}$ *satisfying* $|a_v|, |b_u| < q$, *and such that the uniform measures* $\widehat{\mathcal{A}}, \widehat{\mathcal{B}}$ *on the two lists satisfy* $d_K(\widehat{\mathcal{A}}, \widehat{\mathcal{B}}) \leq n^{-c}$, *then we have*

$$\left|\sum_{v\in V}a_v^k - \sum_{u\in U}b_u^k\right| = O(q^k n^{1-c}).$$

Proof. First, note that

$$\frac{1}{|V|} \sum_{v \in V} a_v^k = \int_0^q kt^{k-1}(1 - \widehat{\mathcal{A}}((-\infty, t])) \, dt - \int_{-q}^0 kt^{k-1}\widehat{\mathcal{A}}((-\infty, t]) \, dt$$

$$= \int_0^q kt^{k-1}(1 - \widehat{\mathcal{B}}((-\infty, t])) \, dt - \int_{-q}^0 kt^{k-1}\widehat{\mathcal{B}}((-\infty, t]) \, dt \pm O(q^k n^{-c})$$

$$= \frac{1}{|U|} \sum_{u \in U} b_u^k \pm O(q^k n^{-c}).$$

The desired result now follows from the fact that $|V| = (1 \pm O(n^{-c}))|U|$.   □

We are now ready for the proof of Lemma 4.9.

Proof of Lemma 4.9. We shall estimate the probabilities in question using Proposition A.1. Indeed, the hypothesis in the statement of Lemma 4.9, in the language of Proposition A.1, may be stated as

- $|\beta_w| = O(\sqrt{\log n})$ (and hence $|\beta_w| \le \log n$) for each $w \in W$,
- $|\sum_{w \in T} \beta_w| = O(n)$ for all $T \subseteq W$, and
- $\sum_{w \in W} \beta_w^2 = O(n)$,

whence it is clear that Proposition A.1 applies.

For part (1) of Lemma 4.9, we may argue as follows. If $|t - h/2| > n^{3/5}$, then (A.1) gives us what we need. If $|t - h/2| \le n^{3/5}$, we claim that (A.2) implies the bound in part (1) of Lemma 4.9. To see this, it suffices to verify in this regime that each of $|\Lambda_1|$, $|\Lambda_2|$, $|\Lambda_3|$ and $|\Lambda_4|$ is $O(|\tau| + \sqrt{\log n})$, where τ is defined by $t = h/2 + \tau\sqrt{n}$.

Using the facts that $|\sum_{i \in W} \beta_i| = O(n)$ and $\sum_{i \in W} \beta_i^2 = O(n)$, we may bound $|\Lambda_1|$ by

$$|\Lambda_1| = \left| \frac{1}{2n^2} \left(\sum_{i \in W} \beta_i \right) \left(\sum_{i \in W} \beta_i - 2n\beta_v \right) \right|$$

$$\le \frac{1}{2n^2} \left(\sum_{i \in W} \beta_i \right)^2 + \frac{|\beta_v|}{n} \left| \sum_{i \in W} \beta_i \right|$$

$$= O(1) + O(|\beta_v|) = O\left(\sqrt{\log n}\right).$$

Next, we bound $|\Lambda_2|$ using the facts that $h, n - h = \Omega(n)$, $|\sum_{i \in S \setminus v} \beta_i| = O(n)$ and $|\sum_{i \in S^c \setminus v} \beta_i| = O(n)$ by

$$|\Lambda_2| \le \left| \sum_{i \in S \setminus v} \left(1 - \frac{2t}{h}\right) \frac{\beta_i}{\sqrt{n-1}} \right| + \left| \sum_{i \in S^c \setminus v} \left(1 - \frac{2(d_v - t)}{(n - h)}\right) \frac{\beta_i}{\sqrt{n-1}} \right|$$

$$\le O\left(|\tau|/n\right) \left| \sum_{i \in S \setminus v} \beta_i \right| + O\left(|\tau|/n + \sqrt{\log n}/n\right) \left| \sum_{i \in S^c \setminus v} \beta_i \right|$$

$$= O(|\tau| + \sqrt{\log n}).$$

Finally, since $\sum_{i \in W} \beta_i^2 = O(n)$, it is immediate that $|\Lambda_3| = O(1)$, and it follows from the facts that $\sum_{i \in W} \beta_i^2 = O(n)$, $|\sum_{i \in S \setminus v} \beta_i| = O(n)$ and $|\sum_{i \in S^c \setminus v} \beta_i| = O(n)$ that $|\Lambda_4| = O(1)$ as well.

For part (2) of Lemma 4.9, it is sufficient to verify that the expression in (A.2) is polynomially-stable when the parameters in question vary by the amounts specified in the statement of Lemma 4.9; here, we say that an expression is polynomially-stable if it varies by at most a multiplicative factor of $1 \pm n^{-\Omega(1)}$. This may be done term by term, as we outline below.

Suppose $(d'_w)_{w \in W'}$, $|W'| = n'$, $v' \in W'$, $S' \subseteq W'$, $|S'| = h'$ and t' satisfy the hypothesis in the statement of the lemma, and additionally, are such that

- $|t - t'|, |d_v - d'_{v'}| \leq n^{1/2 - \Omega(1)}$,
- $(d_w)_{w \in S}$ and $(d'_w)_{w \in S'}$ are proximate, and
- $(d_w)_{w \in W \setminus S}$ and $(d'_w)_{w \in W' \setminus S'}$ are proximate.
- $|n - n'|, |h - h'| \leq n^{1 - \Omega(1)}$, this being a consequence of the previous two points.

In the regime where $h, n - h = \Omega(n)$, $d = n/2 \pm O(\sqrt{n \log n})$, $t = h/2 \pm O(\sqrt{n \log n})$, the expression

$$\binom{h}{t}\binom{n - h - 1}{d - t}\binom{n - 1}{d}^{-1}$$

is polynomially-stable when n and h vary by $n^{1 - \Omega(1)}$, and d and t vary by $n^{1/2 - \Omega(1)}$, which in particular tells us that

$$\binom{h}{t}\binom{n - h - 1}{d_v - t}\binom{n - 1}{d_v}^{-1} \simeq \binom{h'}{t'}\binom{n' - h' - 1}{d'_{v'} - t'}\binom{n' - 1}{d'_{v'}}^{-1}.$$

This can be seen via a careful (and rather tedious) application of Stirling's approximation, or alternately, by using a sufficiently precise form of the de Moivre–Laplace normal approximation, as in [40] for example.

Next, we need to verify that each of $\exp(\Lambda_1)$, $\exp(-\Lambda_2)$, $\exp(-\Lambda_3)$ and $\exp(\Lambda_4)$ is similarly polynomially-stable, and this may be accomplished in a straightforward manner using Lemma A.2. To illustrate, we spell out the details for $\exp(-\Lambda_3)$ below.

Recall that

$$\Lambda_3 = \frac{1}{2} \sum_{i \in W \setminus v} \frac{\beta_i^2}{n - 1} = \frac{1}{2} \sum_{i \in W} \frac{\beta_i^2}{n - 1} \pm O(\log n / n).$$

Our goal is to show, with β'_i defined by $d'_i = (n' - 1)/2 + \beta'_i \sqrt{(n' - 1)}/2$ for $i \in W'$, that

$$\Lambda'_3 = \frac{1}{2} \sum_{i \in W' \setminus v'} \frac{(\beta'_i)^2}{n' - 1} = \frac{1}{2} \sum_{i \in W'} \frac{(\beta'_i)^2}{n' - 1} \pm O(\log n / n)$$

is close enough to Λ_3 to ensure $\exp(-\Lambda_3) \simeq \exp(-\Lambda'_3)$.

Since $(d_w)_{w \in S}$ and $(d'_w)_{w \in S'}$ are proximate, we claim that

$$\left| \sum_{i \in S} \beta_i^2 - \sum_{i \in S'} (\beta'_i)^2 \right| \leq n^{1 - \Omega(1)}.$$

This is true with room to spare if the two sequences are proximate on account of the first part of Definition 4.8, since in this case, we know that

$$\sum_{i \in S} \left| \beta_i - \beta'_{\psi(i)} \right| = O(\sqrt{n})$$

for some bijection $\psi : S \to S'$, from which it follows that

$$\left| \sum_{i \in S} \beta_i^2 - \sum_{i \in S'} (\beta_i')^2 \right| \leq \left(\max_{i \in S} |\beta_i + \beta_{\psi(i)}'| \right) \left(\sum_{i \in S} |\beta_i - \beta_{\psi(i)}'| \right) = O(\sqrt{n \log n}).$$

If the two sequences are proximate on account of the second part of Definition 4.8, then since $|n - n'| \leq n^{1-\Omega(1)}$, it is easily checked that the Kolmogorov distance between the uniform measures on $(\beta_i)_{i \in S}$ and $(\beta_i')_{i \in S'}$ is at most $n^{-\Omega(1)}$, so by Lemma A.2 (with $k = 2$ and $q = \log n$), we have

$$\left| \sum_{i \in S} \beta_i^2 - \sum_{i \in S'} (\beta_i')^2 \right| \leq n^{1-\Omega(1)}$$

as claimed. Reasoning similarly about the proximate pair $(d_w)_{w \in W \setminus S}$ and $(d_w')_{w \in W' \setminus S'}$, we deduce that

$$\left| \sum_{i \in W \setminus S} \beta_i^2 - \sum_{i \in W' \setminus S'} (\beta_i')^2 \right| \leq n^{1-\Omega(1)}$$

as well. Putting these pairs of estimates together shows that $|\Lambda_3 - \Lambda_3'| \leq n^{-\Omega(1)}$, whence it is clear that $\exp(-\Lambda_3) \simeq \exp(-\Lambda_3')$.

The details in the other three cases (i.e., for Λ_1, Λ_2 and Λ_4) are similar, and we leave them to the reader. $\qquad\square$

Proposition A.1 is a consequence of the following more general statement, the proof of which will be given in Appendix C once we have collected the requisite machinery in Appendix B.

Proposition A.3. *Let $(d_w)_{w \in W}$ be a sequence with even sum on a set W of n vertices such that, defining β_w by $d_w = (n - 1)/2 + \beta_w \sqrt{(n - 1)}/2$, we have $|\beta_w| \leq \log n$ for each $w \in W$. Such a sequence is a graphic sequence for all sufficiently large n. Let G be a uniformly random graph with this degree sequence on the vertex set W. For any fixed $v \in W, S \subseteq W$ of size h satisfying $\min(h, n-h) \geq n/(\log n)^{1/8}$, and an integer $t \in [0, d_v]$, we have*

$$\mathbb{P}(\deg_S(v) = t) = (1 \pm O(n^{-1/6})) \frac{\binom{h-\mathbb{1}_S(v)}{t}\binom{n-h-\mathbb{1}_{S^c}(v)}{d_v-t}}{\binom{n-1}{d_v}} \exp(\Lambda_1 - \Lambda_3)\mathbb{E}_T\left[\exp(-\Lambda_T)\right],$$

where $T = T_1 \cup T_2$ is a random set chosen by picking T_1 uniformly from $\binom{S \setminus v}{t}$ and T_2 uniformly from $\binom{S^c \setminus v}{d_v - t}$, and where Λ_1, Λ_3 and Λ_T are given by

$$\Lambda_1 = \frac{1}{2n^2}\left(\sum_{i \in W} \beta_i\right)\left(\sum_{i \in W} \beta_i - 2n\beta_v\right),$$

$$\Lambda_3 = \frac{1}{2}\sum_{i \in W \setminus v} \frac{\beta_i^2}{n - 1}, \text{ and}$$

$$\Lambda_T = \sum_{i \in W \setminus v} (-1)^{\mathbb{1}_T(i)} \frac{\beta_i}{\sqrt{n - 1}}.$$

To proceed, we will need to understand expressions as appearing on the right side of Proposition A.3. To this end, we state two general results about sums of random variables constrained to live on a 'slice' of the Boolean hypercube.

Lemma A.4. *Let $a_1, \ldots, a_n \in \mathbb{R}$, $s \in [0, n]$ be an integer and let $X = \sum_{i=1}^n a_i \xi_i$, where $\xi = (\xi_1, \ldots, \xi_n)$ is uniform on the subset of binary vectors in $\{0, 1\}^n$ which have sum s. Writing $\eta^2 = \sum_{i=1}^n a_i^2 - (\sum_{i=1}^n a_i)^2/n$, we have*

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq 2 \exp(-t^2/(4\eta^2))$$

and

$$\mathbb{E}\left[e^{X - \mathbb{E}[X]}\right] \leq 2 e^{O(\eta^2)}.$$

Proof. The first part follows from the Azuma–Hoeffding inequality, as outlined in [23, Lemma 2.2], for example. The second part follows from integrating the first; see [38, Proposition 2.5.2]. $\square$

Lemma A.5. *Let $a_1, \ldots, a_n \in \mathbb{R}$, s be an integer with $\min(s, n-s) \geq n(\log n)^{-2}$, and let $X = \sum_{i=1}^n a_i \xi_i$, where $\xi = (\xi_1, \ldots, \xi_n)$ is uniform on the subset of $\{0, 1\}^n$ with sum s. Suppose that $|a_i| \leq n^{-1/2}(\log n)^2$ and $\eta^2 = \sum_{i=1}^n a_i^2 - (\sum_{i=1}^n a_i)^2/n \leq \sqrt{\log n}$. Then we have*

$$\mathbb{E}\left[e^X\right] = \exp\left(\mathbb{E}[X] + \frac{1}{2}\operatorname{Var}[X] \pm O(n^{-1/9})\right).$$

Proof. Writing $\sigma^2 = \operatorname{Var}[X]$, we clearly have

$$\sigma^2 = \sum_{i \neq j} a_i a_j (\mathbb{E}[\xi_i \xi_j] - \mathbb{E}[\xi_i]\mathbb{E}[\xi_j]) + \sum_i a_i^2 (\mathbb{E}[\xi_i^2] - \mathbb{E}[\xi_i]^2)$$

$$= \sum_{i \neq j} a_i a_j \left(\frac{s(s-1)}{n(n-1)} - \frac{s^2}{n^2}\right) + \sum_i a_i^2 \left(\frac{s}{n} - \frac{s^2}{n^2}\right) = \frac{s(n-s)}{n(n-1)}\eta^2.$$

First, by Lemma A.4, we have

$$\mathbb{P}[|X - \mathbb{E}[X]| \geq t] \leq 2 \exp(-t^2/(4\eta^2))$$

for all $t \geq 0$. Now

$$\mathbb{E}\left[e^{X - \mathbb{E}[X]}\right] = \int_{-\infty}^{\infty} e^t \mathbb{P}(X - \mathbb{E}[X] \geq t)\, dt$$

$$= \int_{-\infty}^{8\eta\sqrt{\log n}} e^t \mathbb{P}(X - \mathbb{E}[X] \geq t)\, dt + O\left(\int_{8\eta\sqrt{\log n}}^{\infty} e^{t - t^2/(4\eta^2)}\, dt\right)$$

$$= \int_{-\infty}^{8\eta\sqrt{\log n}} e^t \mathbb{P}(X - \mathbb{E}[X] \geq t)\, dt + O\left(\int_{8\eta\sqrt{\log n}}^{\infty} e^{-t^2/(8\eta^2)}\, dt\right)$$

$$= \int_{-\infty}^{8\eta\sqrt{\log n}} e^t \mathbb{P}(X - \mathbb{E}[X] \geq t)\, dt + O(n^{-4}).$$

If $\sigma \leq n^{-1/8}$, then $\eta = (n(n-1)/(s(n-s)))^{1/2}\sigma \leq n^{-2/17}$ and therefore we obtain the upper bound $\mathbb{E}[e^{X - \mathbb{E}[X]}] \leq 1 + O(n^{-1/9})$. Combining with $\mathbb{E}e^X \geq e^{\mathbb{E}X}$, the result follows. If $\sigma > n^{-1/8}$, then a combinatorial central limit theorem of Bolthausen [11] (applied to the $n \times n$ matrix in which the first s rows are copies of the vector $(a_1, \ldots, a_n)$, and the last $n - s$ rows are zero) shows that

$$d_K(X - \mathbb{E}[X], \mathcal{N}(0, \sigma^2)) = O\left(\sum_{i=1}^n |a_i|^3/\sigma^3\right) = O(n^{-2/17}).$$

This allows us the replace the integrand above with the cumulative distribution function of a Gaussian, and we easily derive

$$\mathbb{E}\left[e^{X-\mathbb{E}[X]}\right] = e^{\frac{\sigma^2}{2}} \pm O\left(n^{-2/17}e^{8\eta\sqrt{\log n}}\right) = \exp(\sigma^2/2 \pm O(n^{-1/9})). \qquad \square$$

Proposition A.1 is now easily deduced from Proposition A.3.

Proof of Proposition A.1. With $T = T_1 \cup T_2$ a random set chosen by picking T_1 uniformly from $\binom{S \setminus v}{t}$ and T_2 uniformly from $\binom{S^c \setminus v}{d_v - t}$, we have

$$\mathbb{E}_T[\Lambda_T] = \mathbb{E}_T\left[\sum_{i \in W \setminus v} (-1)^{\mathbb{1}_T(i)} \frac{\beta_i}{\sqrt{n-1}}\right]$$

$$= \sum_{i \in S \setminus v}\left(1 - \frac{2t}{h}\right)\frac{\beta_i}{\sqrt{n-1}} + \sum_{i \in S^c \setminus v}\left(1 - \frac{2(d_v - t)}{(n-h)}\right)\frac{\beta_i}{\sqrt{n-1}} \pm O(n^{-1/3})$$

$$= \Lambda_2 \pm O(n^{-1/3}),$$

where the small additive error term comes from the fact that whether $v \in S$ or $v \in S^c$ slightly change the fractions listed above, but not by much.

At this point, if $|t - h/2| > n^{3/5}$, we have

$$\frac{\binom{h - \mathbb{1}_S(v)}{t}\binom{n - h - \mathbb{1}_{S^c}(v)}{d_v - t}}{\binom{n-1}{d_v}} \le \exp(-\Omega((t - h/2)^2/n))$$

by a standard tail bound for the hypergeometric distribution (see [22, Theorem 2.10], for example). Since $|\beta_w| \le \log n$ for each $w \in W$, clearly both $|\Lambda_1|$ and $|\Lambda_3|$ are $O((\log n)^2)$, whence $\exp(\Lambda_1 - \Lambda_3) \le \exp(O((\log n)^4))$, and we are left with estimating $\mathbb{E}[\exp(-\Lambda_T)]$. Now Lemma A.4 demonstrates

$$\mathbb{E}[\exp(-\Lambda_T)] \le \exp(\mathbb{E}[-\Lambda_T] + O((\log n)^2)),$$

since the coefficient variance in $-\Lambda_T$ is $O((\log n)^2/n)$ by the given conditions. The above explicit expression for $\mathbb{E}[\Lambda_T]$ demonstrates that

$$|\mathbb{E}[-\Lambda_T]| = O\left(\frac{|t - h/2|(\log n)^2}{\sqrt{n}}\right)$$

when $|t - h/2| > n^{3/5}$. These estimates together immediately yield a bound of the claimed quality.

From now on we assume $|t - h/2| \le n^{3/5}$. We next compute the variance of Λ_T. Following the computation in the proof of Lemma A.5, we see

$$\mathrm{Var}[\Lambda_T] = \frac{4}{(n-1)}\left(\frac{t(h-t)}{h(h-1)}\frac{\sum_{i,j \in S \setminus v}(\beta_i - \beta_j)^2}{h}\right.$$

$$\left. + \frac{(d_v - t)((n-h) - (d_v - t))}{(n-h)(n-h-1)}\frac{\sum_{i,j \in S^c \setminus v}(\beta_i - \beta_j)^2}{n-h}\right) + O(n^{-1/4}),$$

these sums being over all (unordered) two-element subsets; here, we again use the fact that the fraction $t/|S \setminus v|$ is close to t/h regardless of if $v \in S$ or $v \in S^c$. Now using

$t = h/2 \pm n^{3/5}$ and $d_v = n/2 + O(\sqrt{n}(\log n))$, we find

$$\mathrm{Var}[\Lambda_T] = \frac{1}{nh} \sum_{i,j \in S \setminus v} (\beta_i - \beta_j)^2 + \frac{1}{n(n-h)} \sum_{i,j \in S^c \setminus v} (\beta_i - \beta_j)^2 \pm O(n^{-1/4})$$

$$= 2\Lambda_4 \pm O(n^{-1/4}).$$

Note that $\mathrm{Var}[\Lambda_T] \leq \sum_{i=1}^{n} \beta_i^2 / \min(h, n-h) = O(n(\log n)^{1/9}/\min(h, n-h))$, and apply Lemma A.5 to the two slices defining Λ_T. Note that the condition $\eta^2 \leq \sqrt{\log n}$ follows from the inequalities $(n/h)(\log n)^{1/9} < \sqrt{\log n}$ and the relation between σ^2 and η^2 in the proof of Lemma A.5. Therefore

$$\mathbb{E}[\exp(-\Lambda_T)] = \exp\left(\mathbb{E}[-\Lambda_T] + \frac{1}{2}\mathrm{Var}[\Lambda_T] \pm O(n^{-1/9})\right)$$

$$= \exp\left(-\Lambda_2 + \Lambda_4 \pm O(n^{-1/9})\right).$$

Plugging this last estimate into Proposition A.3, we obtain

$$\mathbb{P}(\deg_S(v) = t) = (1 \pm O(n^{-1/10}))\frac{\binom{h}{t}\binom{n-h-1}{d_v-t}}{\binom{n-1}{d_v}} \exp(\Lambda_1 - \Lambda_2 - \Lambda_3 + \Lambda_4),$$

as desired, using the fact that the product of binomials in question changes by a small factor depending on whether $v \in S$ or $v \in S^c$, a factor which is nonetheless subsumed by the error term with room to spare. $\qquad\square$

The proof of Lemma 4.10 is analogous to the argument above, so in this case, we only record the appropriate intermediate results needed, and omit the details.

Lemma 4.10 may be deduced from the following result in the same fashion as Lemma 4.9 was from Proposition A.1.

Proposition A.6. *Let $((d_v)_{v \in V}, (d_w)_{w \in W})$ be a pair of sequences with identical sums on a bipartition $V \cup W$ with $|V| = m$, $|W| = n$ such that, defining α_v by $d_v = (n-1)/2 + \alpha_v\sqrt{(n-1)}/2$ for $v \in V$ and β_w by $d_w = (n-1)/2 + \beta_w\sqrt{(n-1)}/2$ for $w \in W$, we have*

- *$(\log n)^{-1/4} \leq m/n \leq (\log n)^{1/4}$,*
- *$|\alpha_v| \leq \log n$ for each $v \in V$ and $|\beta_w| \leq \log n$ for each $w \in W$, and*
- *$(n/m)\sum_{w \in W} \beta_w^2 \leq n(\log n)^{1/9}$.*

Then $(d_v)_{v \in V}, (d_w)_{w \in W}$ are the degree sequences of the two parts of some bipartite graph (assuming n is sufficiently large). Let G be a uniformly random bipartite graph with these degree sequences on the vertex set $V \cup W$. Consider any fixed $u \in V$, any fixed subset $S \subseteq W$ of size h satisfying $\min(h, n-h) \geq n/(\log n)^{1/8}$, and an integer $t \in [0, d_u]$. If $|t - h/2| > n^{3/5}$, then we have

$$\text{(A.3)} \qquad\qquad \mathbb{P}(\deg_S(u) = t) \leq \exp(-\Omega((t - h/2)^2/n)).$$

If $|t - h/2| \leq n^{3/5}$ on the other hand, then we have

$$\text{(A.4)} \qquad \mathbb{P}(\deg_S(u) = t) = (1 \pm O(n^{-1/10}))\frac{\binom{h}{t}\binom{n-h}{d_u-t}}{\binom{n}{d_u}} \exp(\Lambda_1 - \Lambda_2 - \Lambda_3 + \Lambda_4),$$

where Λ_1, Λ_2, Λ_3 and Λ_4 are given by

$$\Lambda_1 = \frac{1}{2mn}\left(\sum_{i \in W}\beta_i\right)\left(\sum_{i \in W}\beta_i - 2\sqrt{mn}\alpha_u\right),$$

$$\Lambda_2 = \sum_{i \in S}\left(1 - \frac{2t}{h}\right)\frac{\beta_i}{\sqrt{m}} + \sum_{i \in W \setminus S}\left(1 - \frac{2(d_v - t)}{(n - h)}\right)\frac{\beta_i}{\sqrt{m}},$$

$$\Lambda_3 = \frac{1}{2}\sum_{i \in W}\frac{\beta_i^2}{m}, \ and$$

$$\Lambda_4 = \frac{1}{2mh}\sum_{i,j \in S}(\beta_i - \beta_j)^2 + \frac{1}{2m(n - h)}\sum_{i,j \in W \setminus S}(\beta_i - \beta_j)^2,$$

the sums in the definition of Λ_4 being over all (unordered) two-element subsets.

As before, Proposition A.6 itself is a consequence of the following result, whose proof will be sketched in Appendix C once we have collected the requisite machinery in Appendix B.

Proposition A.7. *Let $((d_v)_{v \in V}, (d_w)_{w \in W})$ be a pair of sequences of identical sums on a bipartition $V \cup W$ with $|V| = m$, $|W| = n$ such that, defining α_v by $d_v = (n - 1)/2 + \alpha_v\sqrt{(n - 1)}/2$ for $v \in V$ and β_w by $d_w = (n - 1)/2 + \beta_w\sqrt{(n - 1)}/2$ for $w \in W$, we have*

- *$(\log n)^{-1/4} \le m/n \le (\log n)^{1/4}$,*
- *$|\alpha_v| \le \log n$ for each $v \in V$ and $|\beta_w| \le \log n$ for each $w \in W$.*

Then $(d_v)_{v \in V}, (d_w)_{w \in W}$ are the degree sequences of the two parts of some bipartite graph (assuming n is sufficiently large). Let G be a uniformly random bipartite graph with these degree sequences on the vertex set $V \cup W$. For any fixed $u \in V$, $S \subseteq W$ of size h satisfying $\min(h, n - h) \ge n/(\log n)^{1/8}$, and an integer $t \in [0, d_u]$, we have

$$\mathbb{P}(\deg_S(u) = t) = (1 \pm O(n^{-1/8}))\frac{\binom{h}{t}\binom{n-h}{d_u-t}}{\binom{n}{d_u}}\exp(\Lambda_1 - \Lambda_3)\mathbb{E}_T\left[\exp(-\Lambda_T)\right],$$

where $T = T_1 \cup T_2$ is a random set chosen by picking T_1 uniformly from $\binom{S}{t}$ and T_2 uniformly from $\binom{W \setminus S}{d_u - t}$, and where Λ_1, Λ_3 and Λ_T are given by

$$\Lambda_1 = \frac{1}{2mn}\left(\sum_{i \in W}\beta_i\right)\left(\sum_{i \in W}\beta_i - 2\sqrt{mn}\alpha_u\right),$$

$$\Lambda_3 = \frac{1}{2}\sum_{i \in W}\frac{\beta_i^2}{m}, \ and$$

$$\Lambda_T = \sum_{i \in W}(-1)^{\mathbb{1}_T(i)}\frac{\beta_i}{\sqrt{m}}.$$

APPENDIX B. GRAPH ENUMERATION RESULTS AND RELATED ESTIMATES

The main tools needed to prove Propositions A.3 and A.7 are the following enumeration theorems of McKay and Wormald [27] and of Canfield, Greenhill, and McKay [12].

Theorem B.1. *There exists a fixed constant $\varepsilon > 0$ such that the following holds. Consider a sequence $\mathbf{d} = (d_1, \ldots, d_n)$ with even sum such that, writing $\overline{d} = (1/n)\sum_{i=1}^{n}d_i$, we have*

- $|d_i - \overline{d}| \leq n^{1/2+\varepsilon}$ for $1 \leq i \leq n$, and
- $\overline{d} \geq n/\log n$.

Writing $m = \overline{d}n/2 \in \mathbb{Z}$, $\mu = \overline{d}/(n-1)$, and $\gamma_2^2 = (1/(n-1)^2)\sum_{i=1}^n (d_i - \overline{d})^2$, the number of labelled graphs with degree sequence $\mathbf{d}$ is

$$(1 \pm O(n^{-1/4})) \exp\left(\frac{1}{4} - \frac{\gamma_2^2}{4\mu^2(1-\mu)^2}\right)\binom{n(n-1)/2}{m}\binom{n(n-1)}{2m}^{-1}\prod_{i=1}^n\binom{n-1}{d_i}.$$

Theorem B.2. *There exists a fixed constant $\varepsilon > 0$ such that the following holds. Consider a pair of sequences $(\mathbf{s} = (s_1,\ldots,s_n), \mathbf{t} = (t_1,\ldots,t_m))$ with identical sums such that, writing $\overline{s} = (1/n)\sum_{i=1}^n s_i$ and $\overline{t} = (1/m)\sum_{i=1}^m t_i$, we have*

- $n/(\log n)^{1/2} \leq m \leq n(\log n)^{1/2}$,
- $|s_i - \overline{s}| \leq n^{1/2+\varepsilon}$ for $1 \leq i \leq n$ and $|t_i - \overline{t}| \leq m^{1/2+\varepsilon}$ for $1 \leq i \leq m$, and
- $\overline{s} \geq n/(\log n)^{1/2}$ and $\overline{t} \geq m/(\log m)^{1/2}$.

Writing $\mu = \sum_{i=1}^n s_i/(mn) = \sum_{i=1}^m t_i/(mn)$, $\gamma_2(\mathbf{s})^2 = (1/(mn))\sum_{i=1}^n(s_i - \overline{s})^2$ and $\gamma_2(\mathbf{t})^2 = (1/(mn))\sum_{i=1}^m(t_i - \overline{t})^2$, the number of labelled bipartite graphs whose partition classes have degree sequences $\mathbf{s}$ and $\mathbf{t}$ is

$$(1 \pm O(n^{-1/8})) \exp\left(-\frac{1}{2}\left(1 - \frac{\gamma_2(\mathbf{s})^2}{\mu(1-\mu)}\right)\left(1 - \frac{\gamma_2(\mathbf{t})^2}{\mu(1-\mu)}\right)\right)\binom{mn}{mn\mu}^{-1}\prod_{i=1}^n\binom{m}{s_i}\prod_{i=1}^m\binom{n}{t_i}.$$

We remark that these enumeration results are now known to hold under even broader conditions on the degree sequences (i.e., $\mathbf{d}$, $\mathbf{s}$ and $\mathbf{t}$) due to works of Barvinok and Hartigan [6], and for essentially all sparsities by recent work of Liebenau and Wormald [24, 25]. We refer the reader to [39] for an excellent survey of these results.

In order to estimate the expressions in Theorems B.1 and B.2, we shall also require the following estimates for binomial coefficients. These follow from sufficiently precise versions of Stirling's approximation for the factorial. These estimates are nonetheless somewhat non-standard, and so we include proofs, following the exceptionally clean approach in [33].

Lemma B.3. *We have the following pair of estimates.*

(1) *For integers $e, m, d \in \mathbb{N}$, let $\Delta_1 = e - m(m-1)/4$ and $\Delta_2 = (m-1)/2 - d$. If $|\Delta_1| = O(m^{3/2})$ and $|\Delta_2| = O(\sqrt{m\log m})$, then*

$$\frac{\binom{m(m-1)/2}{e}\binom{m(m-1)}{2e}^{-1}}{\binom{(m-1)(m-2)/2}{e-d}\binom{(m-1)(m-2)}{2e-2d}^{-1}} = (1 \pm O(m^{-2/5}))2^{-(m-1)}\exp(-8(\Delta_1^2 + \Delta_1\Delta_2 m)/m^3).$$

(2) *For integers $e, m, d, n \in \mathbb{N}$, let $\Delta_1 = e - mn/2$ and $\Delta_2 = n/2 - d$. If $|\Delta_1| \leq O(m^{3/2})$, $|\Delta_2| \leq O(\sqrt{m\log m})$, and $m = \Theta(n)$, then*

$$\binom{mn}{e}^{-1}\binom{(m-1)n}{e-d} = (1 \pm O(m^{-2/5}))2^{-n}\exp(-2(2m\Delta_1\Delta_2 + \Delta_1^2)/(m^2 n)).$$

Proof. We start with an auxiliary estimate. Consider any N, i satisfying $|i| \leq N^{4/5}$; we claim that

$$(\text{B.1}) \qquad \binom{N}{(N+i)/2} = \binom{N}{N/2}(1 - i^2/(2N) - i^4/(12N^3) \pm O(N^{-1/5})).$$

Indeed, note that

$$\binom{N}{(N+i)/2}\binom{N}{N/2}^{-1} = \prod_{j=1}^{i/2} \frac{N/2-j+1}{N/2+j} = \prod_{j=1}^{i/2} \frac{N/2-j}{N/2+j} \prod_{j=1}^{i/2} \frac{N/2-j+1}{N/2-j}.$$

The final product on the right hand side is $(1 \pm O(N^{-1/5}))$. For the first of the two products on the right hand side, note that

$$\sum_{j=1}^{i/2} \log((N/2-j)/(N/2+j)) = \sum_{j=1}^{i/2} -4j/N - 2(2j/N)^3/3 \pm O(N^{-1/5})$$

$$= -i^2/(2N) - i^4/(12N^3) \pm O(N^{-1/5}),$$

proving (B.1).

Now, for (1), we have $\Delta_1 = e - m(m-1)/4$ and $\Delta_2 = (m-1)/2 - d$. Applying (B.1) to each term, we find that

$$\frac{\binom{m(m-1)/2}{e}\binom{m(m-1)}{2e}^{-1}}{\binom{(m-1)(m-2)/2}{e-d}\binom{(m-1)(m-2)}{2e-2d}^{-1}} = (1 \pm O(m^{-2/5}))2^{-(m-1)} \exp(-8(\Delta_1^2 + \Delta_1\Delta_2 m)/m^3),$$

proving (1). (Note that in each application of (B.1) we have $N = \Theta(m^2)$).

For (2), we have $\Delta_1 = e - mn/2$ and $\Delta_2 = n/2 - d$. Then note that

$$\binom{mn}{e}^{-1}\binom{(m-1)n}{e-d} = (1 \pm O(m^{-2/5}))2^{-n} \exp(-2(2m\Delta_1\Delta_2 + \Delta_1^2)/(m^2 n)),$$

which proves (2).   □

APPENDIX C. PROOFS OF THE MAIN TECHNICAL ESTIMATES

With the results in Appendix B in hand, we are now ready to prove Propositions A.3 and A.7. We start with Proposition A.3.

Proof of Proposition A.3. Given $\mathbf{d} = (d_w)_{w \in W}$, $v \in W$ and $T \subseteq W \setminus v$ of size d_v, we shall estimate the probability of the neighbourhood of v in G being exactly T.

To this end, let $\mathbf{d}_T = (d_w - \mathbb{1}_T(w))_{w \in W}$. As in Theorem B.1, let

$$\bar{d} = \frac{1}{n} \sum_{i \in W} d_i, \qquad \bar{d}_T = \frac{1}{n-1} \sum_{i \in W \setminus v} (d_i - \mathbb{1}_T(i)) = \frac{n\bar{d} - 2d_v}{n-1},$$

$$r = \frac{\bar{d}n}{2}, \qquad r_T = \frac{\bar{d}_T(n-1)}{2} = r - d_v,$$

$$\mu = \frac{\bar{d}}{n-1}, \qquad \mu_T = \frac{\bar{d}_T}{n-2} = \frac{n}{n-2}\mu - \frac{2d_v}{(n-1)(n-2)},$$

$$\gamma_2^2 = \frac{1}{(n-1)^2} \sum_{i \in W} (d_i - \bar{d})^2, \qquad \gamma_2^2(T) = \frac{1}{(n-2)^2} \sum_{i \in W \setminus v} (d_{T,i} - \bar{d}_T)^2.$$

Note that $\mathbf{d}$ and $\mathbf{d}_T$ both clearly satisfy the conditions of Theorem B.1 due to our hypotheses, and that

$$\gamma_2^2(T) = \gamma_2^2 \pm O(n^{-1/4}) \text{ and } \mu_T = \mu \pm O(1/n),$$

again, from the given hypotheses. Now define

$$\Phi = \frac{\binom{(n-1)(n-2)/2}{r-d_v}\binom{(n-1)(n-2)}{2r-2d_v}^{-1}}{\binom{n(n-1)/2}{r}\binom{n(n-1)}{2r}^{-1}} 2^{-(n-1)}$$

and recall $d_i = (n-1)/2 + \beta_i\sqrt{(n-1)}/2$. We have

$$r - \frac{1}{2}\binom{n}{2} = \frac{1}{2}\sum_{i\in W}(d_i - (n-1)/2) = \frac{\sqrt{(n-1)}}{4}\sum_{i\in W}\beta_i.$$

From our hypotheses and the first estimate in Lemma B.3, we then deduce that

$$\Phi = \exp\left(\frac{(\sum_{i\in W}\beta_i)(\sum_{i\in W}\beta_i - 2n\beta_n)}{2n^2} \pm O(n^{-1/6})\right)$$

$$= \exp(\Lambda_1 \pm O(n^{-1/6})).$$

The above estimates for $\gamma_2^2(T)$ and μ_T imply that

$$\frac{\exp\left(\frac{1}{4} - \frac{\gamma_2^2(T)}{4\mu_T^2(1-\mu_T)^2}\right)}{\exp\left(\frac{1}{4} - \frac{\gamma_2^2}{4\mu^2(1-\mu)^2}\right)} = 1 \pm O(n^{-1/4}),$$

and this fact in conjunction with Theorem B.1 yields

$$\mathbb{P}[N(v) = T] = (1 \pm O(n^{-1/4}))\frac{\binom{(n-1)(n-2)/2}{r_T}\binom{(n-1)(n-2)}{2r_T}^{-1}\prod_{i\in W\setminus v}\binom{n-2}{d_i-\mathbb{1}_T(i)}}{\binom{n(n-1)/2}{r}\binom{n(n-1)}{2r}^{-1}\prod_{i\in W}\binom{n-1}{d_i}}$$

$$= (1 \pm O(n^{-1/4}))\frac{\Phi 2^{n-1}}{\binom{n-1}{d_v}}\prod_{i\in T}\frac{d_i}{n-1}\prod_{i\notin T}\frac{n-1-d_i}{n-1}$$

$$= (1 \pm O(n^{-1/4}))\frac{\Phi}{\binom{n-1}{d_v}}\prod_{i\in T}\left(1 + \frac{\beta_i}{\sqrt{n-1}}\right)\prod_{i\notin T}\left(1 - \frac{\beta_i}{\sqrt{n-1}}\right)$$

$$= \frac{\Phi}{\binom{n-1}{d_v}}\exp\left(-\sum_{i\in W\setminus v}(-1)^{\mathbb{1}_T(i)}\frac{\beta_i}{\sqrt{n-1}} - \frac{1}{2}\sum_{i\in W\setminus v}\frac{\beta_i^2}{n-1} \pm O(n^{-1/4})\right)$$

$$= \frac{\Phi}{\binom{n-1}{d_v}}\exp\left(-\Lambda_T - \Lambda_3 \pm O(n^{-1/4})\right).$$

Since the above estimate holds for every choice of $T \subseteq W \setminus v$, we may finish by noting that

$$(1 \pm O(n^{-1/4}))\frac{\binom{n-1}{d_v}}{\Phi\binom{h-\mathbb{1}_S(v)}{t}\binom{n-h-\mathbb{1}_{S^c}(v)}{d_v-t}}\mathbb{P}[\deg_S(v) = t] = \exp(-\Lambda_3)\mathbb{E}_T[\exp(-\Lambda_T)],$$

where $T = T_1 \cup T_2$ is a random set chosen by picking T_1 uniformly from $\binom{S}{t}$ and T_2 uniformly from $\binom{W\setminus S}{d_v-t}$. Rearranging this, and recalling that $\Phi = \exp(\Lambda_1 \pm O(n^{-1/6}))$, gives us the desired result. $\square$

To finish, we outline the proof of Proposition A.7.

Proof of Proposition A.7. The proof of this proposition mirrors that of Proposition A.3, except now using Theorem B.2 instead of Theorem B.1, and the second estimate in Lemma B.3 instead of the first. Since the requisite calculations are routine (and are analogous to those spelled out in the proof of Proposition A.3), we leave the details of these calculations to the reader. □

Acknowledgment

References

[1] Louigi Addario-Berry, Luc Devroye, Gábor Lugosi, and Roberto I. Oliveira, *Local optima of the Sherrington-Kirkpatrick Hamiltonian*, J. Math. Phys. **60** (2019), no. 4, 043301, 13, DOI 10.1063/1.5020662. MR3940914 ↑380

[2] R. Aharoni, E. C. Milner, and K. Prikry, *Unfriendly partitions of a graph*, J. Combin. Theory Ser. B **50** (1990), no. 1, 1–10, DOI 10.1016/0095-8956(90)90092-E. MR1070461 ↑380

[3] Noga Alon and Joel H. Spencer, *The probabilistic method*, 4th ed., Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2016. MR3524748 ↑383

[4] Omer Angel, Sébastien Bubeck, Yuval Peres, and Fan Wei, *Local max-cut in smoothed polynomial time*, STOC'17—Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, ACM, New York, 2017, pp. 429–437, DOI 10.1145/3055399.3055402. MR3678200 ↑380

[5] Amir Ban and Nati Linial, *Internal partitions of regular graphs*, J. Graph Theory **83** (2016), no. 1, 5–18, DOI 10.1002/jgt.21909. MR3529944 ↑380, 381

[6] Alexander Barvinok and J. A. Hartigan, *The number of graphs and a random graph with a given degree sequence*, Random Structures Algorithms **42** (2013), no. 3, 301–348, DOI 10.1002/rsa.20409. MR3039682 ↑411

[7] Tom Bohman, Alan Frieze, and Eyal Lubetzky, *Random triangle removal*, Adv. Math. **280** (2015), 379–438, DOI 10.1016/j.aim.2015.04.015. MR3350225 ↑382

[8] Tom Bohman and Peter Keevash, *The early evolution of the H-free process*, Invent. Math. **181** (2010), no. 2, 291–336, DOI 10.1007/s00222-010-0247-x. MR2657427 ↑382

[9] Béla Bollobás, *Modern graph theory*, Graduate Texts in Mathematics, vol. 184, Springer-Verlag, New York, 1998, DOI 10.1007/978-1-4612-0619-4. MR1633290 ↑383

[10] B. Bollobás and A. D. Scott, *Problems and results on judicious partitions*, Random Structures Algorithms **21** (2002), no. 3-4, 414–430, DOI 10.1002/rsa.10062. Random structures and algorithms (Poznan, 2001). MR1945378 ↑380, 381

[11] E. Bolthausen, *An estimate of the remainder in a combinatorial central limit theorem*, Z. Wahrsch. Verw. Gebiete **66** (1984), no. 3, 379–386, DOI 10.1007/BF00533704. MR751577 ↑407

[12] E. Rodney Canfield, Catherine Greenhill, and Brendan D. McKay, *Asymptotic enumeration of dense 0-1 matrices with specified line sums*, J. Combin. Theory Ser. A **115** (2008), no. 1, 32–66, DOI 10.1016/j.jcta.2007.03.009. MR2378856 ↑383, 390, 393, 410

[13] Xi Chen, Chenghao Guo, Emmanouil V. Vlatakis-Gkaragkounis, Mihalis Yannakakis, and Xinzhi Zhang, *Smoothed complexity of local Max-Cut and binary Max-CSP*, STOC '20—Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, ACM, New York, [2020] ©2020, pp. 1052–1065, DOI 10.1145/3357713.3384325. MR4141821 ↑380

[14] Asaf Ferber, Matthew Kwan, and Lisa Sauermann, *Singularity of sparse random matrices: simple proofs*, Combin. Probab. Comput. **31** (2022), no. 1, 21–28, DOI 10.1017/s0963548321000146. MR4356454 ↑399

[15] Gonzalo Fiz Pontiveros, Simon Griffiths, and Robert Morris, *The triangle-free process and the Ramsey number R*(3, *k*), Mem. Amer. Math. Soc. **263** (2020), no. 1274, v+125, DOI 10.1090/memo/1274. MR4073152 ↑382

[16] Z. Füredi, Personal communication. 380

[17] David Gamarnik, *The overlap gap property: a topological barrier to optimizing over random structures*, Proc. Natl. Acad. Sci. USA **118** (2021), no. 41. 383

[18] David Gamarnik and Quan Li, *On the max-cut of sparse random graphs*, Random Structures Algorithms **52** (2018), no. 2, 219–262, DOI 10.1002/rsa.20738. MR3758958 ↑380

[19] Michael U. Gerber and Daniel Kobler, *Algorithmic approach to the satisfactory graph partitioning problem*, European J. Oper. Res. **125** (2000), no. 2, 283–291, DOI 10.1016/S0377-2217(99)00459-2. Combinatorial optimization (Copenhagen, 1998). MR1785664 ↑380

[20] Reza Gheissari, Charles M. Newman, and Daniel L. Stein, *Zero-temperature dynamics in the dilute Curie-Weiss model*, J. Stat. Phys. **172** (2018), no. 4, 1009–1028, DOI 10.1007/s10955-018-2087-9. MR3830296 ↑380

[21] B. Green, *100 open problems*, Manuscript, Available on request. 380

[22] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000, DOI 10.1002/9781118032718. MR1782847 ↑386, 397, 408

[23] Matthew Kwan, Benny Sudakov, and Tuan Tran, *Anticoncentration for subgraph statistics*, J. Lond. Math. Soc. (2) **99** (2019), no. 3, 757–777, DOI 10.1112/jlms.12192. MR3977889 ↑407

[24] A. Liebenau and N. Wormald, *Asymptotic enumeration of digraphs and bipartite graphs by degree sequence*, arXiv:2006.15797, 2020. 411

[25] A. Liebenau and N. Wormald, *Asymptotic enumeration of graphs by degree sequence, and the degree sequence of a random graph*, arXiv:1702.08373, 2017. 411

[26] Nathan Linial and Sria Louis, *Asymptotically almost every* 2*r-regular graph has an internal partition*, Graphs Combin. **36** (2020), no. 1, 41–50, DOI 10.1007/s00373-019-02116-0. MR4054839 ↑380

[27] Brendan D. McKay and Nicholas C. Wormald, *Asymptotic enumeration by degree sequence of graphs of high degree*, European J. Combin. **11** (1990), no. 6, 565–580, DOI 10.1016/S0195-6698(13)80042-X. MR1078713 ↑382, 386, 390, 393, 410

[28] Vojtěch Rödl, *On a packing and covering problem*, European J. Combin. **6** (1985), no. 1, 69–78, DOI 10.1016/S0195-6698(85)80023-8. MR793489 ↑382

[29] Ashwin Sah and Mehtaab Sawhney, *Majority dynamics: the power of one*, arXiv:2105.13301, 2021. 381

[30] Khurram H. Shafique and Ronald D. Dutton, *On satisfactory partitioning of graphs*, Proceedings of the Thirty-third Southeastern International Conference on Combinatorics, Graph Theory and Computing (Boca Raton, FL, 2002), Congr. Numer. **154** (2002), 183–194. MR1980038 ↑380

[31] Saharon Shelah and E. C. Milner, *Graphs with no unfriendly partitions*, A tribute to Paul Erdős, Cambridge Univ. Press, Cambridge, 1990, pp. 373–384, DOI 10.1177/016502549001300309. MR1117030 ↑380

[32] Eric Yilun Song, Reza Gheissari, Charles M. Newman, and Daniel L. Stein, *Local minima in disordered mean-field ferromagnets*, J. Stat. Phys. **180** (2020), no. 1-6, 576–596, DOI 10.1007/s10955-019-02480-4. MR4131002 ↑380

[33] Joel Spencer, *Asymptopia*, Student Mathematical Library, vol. 71, American Mathematical Society, Providence, RI, 2014. With Laura Florescu, DOI 10.1090/stml/071. MR3185739 ↑411

[34] Michael Stiebitz, *Decomposing graphs under degree constraints*, J. Graph Theory **23** (1996), no. 3, 321–324, DOI 10.1002/(SICI)1097-0118(199611)23:3<321::AID-JGT12>3.3.CO;2-B. MR1413661 ↑380

[35] Terence Tao and Van H. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2010. Paperback edition [of MR2289012]. MR2573797 ↑386

[36] Carsten Thomassen, *Graph decomposition with constraints on the connectivity and minimum degree*, J. Graph Theory **7** (1983), no. 2, 165–167, DOI 10.1002/jgt.3190070204. MR698697 ↑380

[37] L. Tran and V. Vu, *Reaching a consensus on random networks: the power of few*, Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2020) (Dagstuhl, Germany), vol. 176, Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020, pp. 20:1–20:15. 381

[38] Roman Vershynin, *High-dimensional probability*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 47, Cambridge University Press, Cambridge, 2018. An introduction with applications in data science; With a foreword by Sara van de Geer, DOI 10.1017/9781108231596. MR3837109 ↑407

[39] Nicholas Wormald, *Asymptotic enumeration of graphs with given degree sequence*, Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures, World Sci. Publ., Hackensack, NJ, 2018, pp. 3245–3264. MR3966531 ↑411

[40] A. M. Zubkov and A. A. Serov, *A complete proof of universal inequalities for the distribution function of the binomial law*, Theory Probab. Appl. **57** (2013), no. 3, 539–544, DOI 10.1137/S0040585X97986138. MR3196787 ↑405

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, IRVINE, CALIFORNIA 92697
Email address: `asaff@uci.edu`

INSTITUTE OF SCIENCE AND TECHNOLOGY AUSTRIA (ISTA), 3400 KLOSTERNEUBURG, AUSTRIA
Email address: `matthew.kwan@ist.ac.at`

DEPARTMENT OF MATHEMATICS, RUTGERS UNIVERSITY, PISCATAWAY, NEW JERSEY 08854
Email address: `narayanan@math.rutgers.edu`

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MAS-
SACHUSETTS 02139
Email address: `asah@mit.edu`

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MAS-
SACHUSETTS 02139
Email address: `msawhney@mit.edu`