ON PERFECTLY FRIENDLY BISECTIONS OF RANDOM GRAPHS

DOR MINZER, ASHWIN SAH, AND MEHTAAB SAWHNEY

ABSTRACT. We prove that there exists a constant $\gamma_{\text{crit}} \approx .17566$ such that if $G \sim \mathbb{G}(n, 1/2)$ then for any $\varepsilon > 0$ with high probability G has a equipartition such that each vertex has $(\gamma_{\text{crit}} - \varepsilon)\sqrt{n}$ more neighbors in its own part than in the other part and with high probability no such partition exists for a separation of $(\gamma_{\text{crit}} + \varepsilon)\sqrt{n}$. The proof involves a number of tools ranging from isoperimetric results on vertex-transitive sets of graphs coming from Boolean functions, switchings, degree enumeration formulas, and the second moment method. Our results substantially strengthen recent work of Ferber, Kwan, Narayanan, and the last two authors on a conjecture of Füredi from 1988 and in particular prove the existence of fully-friendly bisections in $\mathbb{G}(n, 1/2)$.

1. INTRODUCTION

In this paper we consider the circumstances under which the random graph $\mathbb{G}(n, 1/2)$ can be partitioned into two roughly equal size sets $A \cup B$ such that every vertex in A has more substantially more neighbors in A and analogously for B. We formalize this with the notion of an H-friendly equipartition.

Definition 1.1. Given a graph G with n vertices an H-friendly equipartition is a partition $A_1 \cup A_2 = V(G)$ such that $||A_1| - |A_2|| \le 1$ and $\deg(v, A_i) \ge H + \deg(v, A_{i+1})$ for all choices of $i \in \{1, 2\}$ and $v \in A_i$. (Here we take indices (mod 2), which we do without further notice.)

We now define a constant $\gamma_{\rm crit}$ which will be used throughout the paper.

Definition 1.2. Define $\gamma_{\rm crit} > 0$ to be the constant such that

$$\log 2 + \sup_{\alpha \in \mathbb{R}} (-\alpha^2 + \log(\mathbb{P}_{Z \sim \mathcal{N}(0,1)}[Z \ge (\gamma + \alpha)\sqrt{2}])) = 0.$$

Remark. Noting that $\gamma \mapsto \log 2 + \sup_{\alpha \in \mathbb{R}} (-\alpha^2 + \log(\mathbb{P}_{Z \sim \mathcal{N}(0,1)}[Z \geq (\gamma + \alpha)\sqrt{2}]))$ is a strictly decreasing function therefore γ_{crit} is unique. Furthermore one can prove via numerical computation that .2484195 $\leq \gamma_{\text{crit}} \leq .2484196$ (Claim B.11).

Our main theorem proves, conditional on a numerical hypothesis Assumption 1.5, that an *H*-friendly partition in $\mathbb{G}(n, 1/2)$ exists with *H* being order $\Omega(\sqrt{n})$ and pins down the precise leading order behavior for the best-possible *H*.

Theorem 1.3. Fix $\varepsilon > 0$ and let $G \sim \mathbb{G}(n, 1/2)$. Given Assumption 1.5, with high probability, G has a $(\gamma_{\text{crit}}/\sqrt{2} - \varepsilon)\sqrt{n}$ -friendly equipartition. Furthermore with high probability G does not have a $(\gamma_{\text{crit}}/\sqrt{2} + \varepsilon)\sqrt{n}$ -friendly equipartition.

We state Assumption 1.5 precisely in Section 1.2.3 and discuss a rather careful floating-point verification of Assumption 1.5 which is carried out in Appendix B which gives a procedure to reduce the verification of Assumption 1.5 to a series of non-exact integration computations. We then verify the most delicate of these numerical computations using interval arithmetic; the necessary numerical procedure and implementation details are discussed in Section 1.2.3 and in Appendix B.

Minzer was supported by a Sloan Research Fellowship and NSF CCF award 2227876. Sah was supported by the PD Soros Fellowship. Sawhney was supported by the Churchill Foundation. Sah and Sawhney were supported by NSF Graduate Research Fellowship Program DGE-2141064.

1.1. Background on friendly partitions. Questions about friendly bisections and the analogous concept of friendly partitions (where one drops the condition that the partition is equitable) have been studied under a host of names as recognized in recent work of Behrens, Arpino, Kivva, and Zdeborová [5]; these include satisfactory graph partition [4,23], generalized matching cuts [24], local minimum cut [2], and assortative partitions [5]. By complementing the graph, one is also naturally led to considering the partitioning problem where one aims to have more neighbors across the cut instead of fewer; this variant has similarly gone under a host of different names.

Our primary motivation comes from seeking to understand the behavior dictating an old conjecture due to Füredi [18] from 1988 which was popularized by its inclusion in Green's list of 100 open problems [25, Problem 91]. Füredi conjectured the existence of an bisection where only n - o(n)vertices on both sides are required to be 0-friendly; note in particular that this conjecture is weaker even than the existence of a 0-friendly bisection. This conjecture of Füredi was recently resolved in work of Ferber, Kwan, Narayanan, and the last two authors [15]. However as mentioned in [15], based on computer simulations, and even to Füredi¹, the stronger conjecture that there exists a partition where all nodes are friendly appeared plausible.

This conjecture is also closely related to the existence of gapped states in the Sherrington– Kirkpatrick (SK) model [43] predicted in work of Treves and Amit [44]. To state the model precisely, considered the a random symmetric matrix J where entries are equally likely to be ± 1 . Treves and Amit consider the gap of a signing $\vec{x} \in \{\pm 1\}^n$

$$\frac{1}{\sqrt{n}}\min_{i\in[n]}\sum_{j=1}^n J_{ij}x_ix_j$$

and conjectured that such a signing \vec{x} exists for gap at least $2\gamma_{\text{crit}}$. A straightforward translation via assigning edges of $\mathbb{G}(n, 1/2)$ to the state +1 and assigning the partition according to the vector \vec{x} in the obvious manner, Theorem 1.3 immediately gives a rigorous proof of the existence of such states (in fact with $|\sum_i x_i| \leq 1$).

Finally, recent work of Behrens, Arpino, Kivva, and Zdeborová [5] conjectures, based on the replica method in statistical physics, that the analogue of Theorem 1.3 holds when the underlying graph G is a sparse random regular graph. When translating the predictions of these authors into the dense regime corresponding to $\mathbb{G}(n, 1/2)$ (see [5, Section 3, Gapped states in spin glasses are computationally hard to find]), the results of Behrens, Arpino, Kivva, and Zdeborová [5] suggest that gapped states such that

$$\frac{1}{\sqrt{n}} \min_{i \in [n]} \sum_{j=1}^{n} J_{ij} x_i x_j \ge \varepsilon \sqrt{n}$$

are likely computationally difficult to find for any $\varepsilon > 0$. Furthermore the proof of Theorem 1.3 implicitly proves that states which are $(\gamma_{\text{crit}}/\sqrt{2}-\varepsilon)\sqrt{n}$ -friendly exhibit the overlap gap property (OGP) whp; in particular there exists $[\beta_1, \beta_2] \subseteq (0, 1/2)$ such that whp any two *H*-friendly partitions do not differ on a fraction of nodes between β_1 and β_2 . The overlap gap property (and variants) have been used to provide evidence for the computation hardness of numerous constraint satisfaction problems (see e.g. [9, 11, 21, 22, 28, 38]); we refer the reader to [19] for a recent survey on the key role of the overlap gap property in optimization over random structures. In particular this evidence in part explains why we divert from the constructive approach adopted in [15] (which gives an $O_{\varepsilon}(n^2)$ -time algorithm to compute a partition where $(1 - \varepsilon)n$ nodes on both sides are friendly) and instead opt for a nonconstructive approach based on the second moment method.

We note here briefly that in fact [5] suggest a frozen 1-RSB structure for the space of H-friendly partitions and providing further mathematical proof of these predictions remains of interest. In particular we find the question of whether their exists a polynomial time algorithm to compute a

¹We thank Benny Sudakov for this remark.

 $\gamma\sqrt{n}$ -friendly partition in $\mathbb{G}(n, 1/2)$ for $\gamma > 0$ a fascinating open question; in particular the work of [5] (based on evidence in the sparse graph analog) suggests that the above is likely computationally difficult and giving evidence based on the overlap gap property or more recent variants would be enticing.

1.2. Discussion of techniques. The proof of Theorem 1.3 broadly proceeds via the second moment method on the number of *H*-friendly partitions. The first and second moment of the number of *H*-friendly partitions is computed via specially adapting machinery developed in the context of enumeration of graphs with a given degree sequence. However it appears likely that the underlying moments do not match up to a (1 + o(1)) factor but only a constant factor.

Thus, we can only show that the number of H-friendly partitions will be nonzero with some positive constant probability directly, and therefore we require separate tools to boost the constant probability result to high probability. For this we rely on a specially tailored isoperimetric result which is ultimately derived from recent work on Talagrand's conjecture in Boolean function analysis. As such the proof of Theorem 1.3 breaks into a series of essentially separate steps which we now discuss in detail.

1.2.1. Vertex-adapted isoperimetry in graphs. As mentioned, we prove that the first and second moment of the number of bisections which are $(\gamma_{\rm crit}/\sqrt{2}-\varepsilon)\sqrt{n}$ -friendly match up to a constant factor. Therefore an application of Paley–Zygmund inequality implies the existence of a partition in Theorem 1.3 with at least constant probability. We wish to boost this to 1 - o(1) probability. Before delving into precise statement proved, let \mathcal{F} denote the family of graphs on $\{0,1\}^{\binom{[n]}{2}}$ with a $(\gamma_{\rm crit}/\sqrt{2}-\varepsilon)\sqrt{n}$ -friendly partition and let us consider what naive considerations provide for. Simply viewing a graph as an element in $\{0,1\}^{\binom{[n]}{2}}$, an application of the edge isoperimetric inequality on the hypercube shows that at least $1-\varepsilon$ fraction of graphs are within an edit distance of $O_{\varepsilon}(n)$ edges from a graph in \mathcal{F} . This immediately gives a partition with at most $O_{\varepsilon}(\sqrt{n})$ many $(\gamma_{\rm crit}/\sqrt{2}-\varepsilon)^{1/2}\sqrt{n}$ -unfriendly vertices. While this is a promising first step, this proof ignores the underlying \mathcal{S}_n -invariant nature of \mathcal{F} (since the property of having a $(\gamma_{\rm crit}/\sqrt{2}-\varepsilon)\sqrt{n}$ -friendly partition is a graph property).

We are thus led to the idea that the optimal vertex isoperimetric statement might be that given a family of graphs on $\{0,1\}^{\binom{[n]}{2}}$ of constant density, if we are allowed to modify at most $O_{\varepsilon}(1)$ edges at each vertex we may reach nearly all graphs. However, one can consider Δ so that the probability $\mathbb{G}(n, 1/2)$ has maximum degree at most Δ is ε . By a result of Bollobás [8], one can see that we must be allowed to modify roughly order $\log(1/\varepsilon)\sqrt{n/\log n}$ edges at some vertex to have a chance at reaching a $1 - \varepsilon$ fraction of all graphs.

Our main isoperimetric result (Theorem 2.2) proves that the above example is the worst possible up to an absolute constant. This immediately allows one to derive the necessary sharp threshold result, since modifying $o(\sqrt{n})$ edges at every vertex leaves a graph essentially unchanged with respect to having an $(\gamma_{\text{crit}}/\sqrt{2} - \varepsilon)\sqrt{n}$ -friendly partition. Furthermore we note that while this is certainly not the first use of Boolean functions to establish a sharp threshold theorem (see e.g. [7, 16, 17]), the underlying property considered here is not monotone on the Boolean cube and statement is also novel in proving isoperimetry with respect to a metric (maximum degree distance) that is tailored to the problem at hand.

The proof of this isoperimetric result proceeds via an iterative application of Talagrand's inequality [13,14]. By using (a variant of) Talagrand's inequality applied to symmetric functions one finds that given a family a graphs with density in $[\varepsilon, 1 - \varepsilon]$, it must have many points of reasonably large (positive) sensitivity, i.e., graphs such that modification of many edges will yield many new graphs. Specifically, we obtain a trichotomy of possible good properties (Lemma 2.6). Using this one can consider a sequence of "expansion moves" of the underlying family \mathcal{F} of graphs depending on which case we are in at each step. We can use this to show that it is possible to expand the family to double its density without traversing very far in the relevant metric (maximum degree distance), which is the key Proposition 2.3. Proving closeness in the metric requires one to reverse the expansion process and "trace backwards" how a graph may have been added to the family under consideration. This isoperimetry argument is carried out in Section 2 and is independent from the remainder of the paper.

Finally, we note that the main isoperimetry result can also be used to re-derive the threshold of the symmetric binary perceptron (originally proven by Abbe, Li, and Sly [1] and independently Perkins and Xu [36]); we comment on this relation further in Section 2.1.

1.2.2. Second-moment computation. For the remainder of the paper the focus is on proving that if X_{γ} is is the number of $\gamma\sqrt{n}$ -friendly partitions then $\mathbb{E}X_{\gamma}^2 \leq (\mathbb{E}X_{\gamma})^2$ for $\gamma < \gamma_{\rm crit}$ (see Lemmas 3.2 and 3.3). We will actually only prove such a statement in the neighborhood of the critical threshold; this is an artifact of our verification of Assumption 1.5 and extending the range of γ in this computation naturally extends the range of validity of Lemmas 3.2 and 3.3. Before proceeding further, it is worth noting that the matching of the first and second moment of the number of solutions all the way to the critical threshold is by no means a priori obvious; in particular this relies on the the frozen 1-RSB nature of the model which suggest that near criticality the associated solution space consists of subexponential size clusters.

To get a sense for $\mathbb{E}X_{\gamma}$, note that it suffices via linearity of expectation to compute the probability that a given partition is $\gamma\sqrt{n}$ -friendly. At a heuristic level, one can replace each edge with a Gaussian random variable of mean 1/2 and variance 1/4; such a model has matching mean and variance to $\mathbb{G}(n, 1/2)$. Note however in this Gaussian model that the degree sequence of the vertices in $\mathbb{G}(n, 1/2)$ form a multidimensional Gaussian vector and a straightforward computation implies that the degrees can be expressed in terms of the sum of this vector as well as a collection of independent random variables, one for each vertex. This heuristic, which explicitly appears in work of Riordan and Selby [39, Section 3], was used to compute the log-probability that a given graph in $\mathbb{G}(n, 1/2)$ has maximum degree bounded by n/2. As such this approach is not directly useful since the $\exp(o(n))$ order multiplicative error term is far to large to be directly useful (we remark that however if one only asks for a friendly partition with up to o(n) exceptions such a moment based computation is sufficient due to the edge isoperimetry argument sketched in the Section 1.2.1). However, the work of Riordan and Selby was substantially refined in work of McKay, Wanless, and Wormald [32] to give the asymptotic probability that a graph in $\mathbb{G}(n, 1/2)$ has maximum degree bounded by n/2. These results in turn relied on powerful tools of McKay and Wormald [33] which enumerate dense graphs with sufficiently regular degree sequences.

The key difficulty given the results of [33] within the work of McKay, Wanless, and Wormald [32] is summing over all degree sequence with a given maximum degree; in the work of McKay, Wanless. and Wormald [32] this is accomplished by modeling the degree sequence with a certain tilted binomial distribution and reducing the question to certain local central limit theorem considerations for these distributions. Such an approach, while in theory possible for our problem, appears rather unsuitable since one would require (for the second moment) tracking 10 different tilting patterns and the computations quickly appear to become completely infeasible. We instead adopt a framework given in work of McKay and Wormald [34] (extended to bipartite graphs by McKay and Skerman [31]) which models the degree sequence of $\mathbb{G}(n, 1/2)$ by first sampling some p near 1/2 (which controls the total edge density) and then treats each degree as independent binomial distribution conditional on the total sum being even. One may note that this result in fact immediately makes precise the heuristic degree distribution given in the previous paragraph. This is used in work of McKay and Wormald [34], to (say) compute the asymptotic distribution of the median degree in $\mathbb{G}(n, 1/2)$. However, directly as stated the work of McKay and Wormald [34] is not suitable for exponentially rare events. Instead, a sufficiently close examination reveals that a modification of their methods allows one to, losing constant factors in the probability, handle such exponentially rare events. This is modification is handled in Section 4.

Using the results in Section 4, a direct but lengthy computation reduces computing the first and second moments to certain explicit 3 and 10 variable Gaussian integrals. While in theory one could apply a grid search to compute the maximum of these integrals, this appears computationally intractable. Instead we use an observation of Gamarnik and Li [20] which considered friendly partitions in regular constant degree random graphs, and noted that a certain Gaussian optimization problem could be simplified via using the log-concavity of certain tail probabilities of Gaussian random variables. Via a delicate application of these techniques the first and second moment (involving careful uses of symmetry present in combination with the log-concavity) can be reduced to the problem to checking the optima of a certain univariate and 3-dimensional functions (which appear in Assumption 1.5) match in an appropriate manner.

Finally, we remark that implicit in the above discussion we have assumed that the degree enumeration results of McKay and Wormald [33] (and the extension to bipartite graphs by McKay and Skerman [31]) apply uniformly to all degree sequences; unfortunately they only naturally apply to degree sequences where the maximum and minimum degrees are within $O(n^{1/2+\delta})$ of the median degree. As such we use switchings to a priori prove that the contribution of atypical degree sequences, even conditional on the exponentially small probability of a given pair of overlapping partitions being $\gamma\sqrt{n}$ -friendly, is small. A similar strategy was used in the work of McKay, Wanless, and Wormald [32], but that work relies on the fact that the event of being below a given maximum degree is monotone in all the edges and this is used crucially in their argument. When considering a pair of overlapping partitions one encounters more complex constraints and thus a more involved argument is necessary.

The argument instead proceeds by using the subset of edges which are naturally decreasing or increasing with respect to the constraints in order to prove that, with super-polynomially high probability, at most a $O(n^{-1/2+\eta})$ fraction of vertices will exactly meet the threshold constraints. Then under the event that there are very few such critical vertices, a relatively straightforward switching argument can then be used prove the necessary maximum and minimum degree events; the fact that there are few critical vertices is used here in order to prove that one has essentially the expected number of switches.

1.2.3. Verification of Assumption 1.5. We now finally state the crucial numerical computation. We first formally define a series of special functions which will appear throughout the paper.

Definition 1.4. Fix $\gamma \in \mathbb{R}$. Then

$$F_1(\alpha) := \log 2 - \alpha^2 + \log(\mathbb{P}_{Z \sim \mathcal{N}(0,1)}[Z \ge (\gamma + \alpha)\sqrt{2}]),$$

$$f(\beta, \alpha) := \mathbb{P}\left[\sqrt{\frac{\beta}{2}}Z_1 + \sqrt{\frac{1-\beta}{2}}Z_2 \ge \gamma + \alpha \wedge \sqrt{\frac{\beta}{2}}Z_1 - \sqrt{\frac{1-\beta}{2}}Z_2 \ge \gamma + \alpha\right],$$

$$F_2(\beta, \alpha_1, \alpha_2) := 2\log 2 - 2\beta \log \beta - 2(1-\beta)\log(1-\beta) - 2\alpha_1^2 - 2\alpha_2^2 + 2\beta \log f(\beta, \alpha_1) + 2(1-\beta)\log f(1-\beta, \alpha_2).$$

The crucial computational component which is verified with a computer assisted computation is the following claim.

Assumption 1.5. Fix $\varepsilon_{1.5} = 10^{-25}$ and $\gamma \in [\gamma_{\text{crit}} - \varepsilon_{1.5}, \gamma_{\text{crit}} + \varepsilon_{1.5}]$. We have that

$$\sup_{\substack{\beta \in [0,.001]\\\alpha_1,\alpha_2 \in \mathbb{R}}} F_2(\beta,\alpha_1,\alpha_2) = 2 \sup_{\alpha \in \mathbb{R}} F_1(\alpha)$$

and

$$\sup_{\substack{\beta \in [.001,.999]\\\alpha_1,\alpha_2 \in \mathbb{R}}} F_2(\beta,\alpha_1,\alpha_2) = 4 \sup_{\alpha \in \mathbb{R}} F_1(\alpha).$$

Furthermore let $\alpha(\gamma) := \arg \max_{\alpha \in \mathbb{R}} F_1(\alpha)$. Then the Hessian of F_2 evaluated at $(1/2, \alpha(\gamma), \alpha(\gamma))$ is strictly negative definite and the unique optimizer of $F_2(\beta, \alpha_1, \alpha_2)$ for $\beta \in [.001, .999]$ occurs at $(1/2, \alpha(\gamma), \alpha(\gamma))$.

As stated Assumption 1.5 asserts an equality between two presumably transcendental quantities and a bit of care is required in verifying such an assumption. An equivalent manner of phrasing Assumption 1.5 is that the maximum value of $F_2(\beta, \alpha_1, \alpha_2)$ occurs either at $(1/2, \alpha(\gamma), \alpha(\gamma))$, or at $(0, 0, \alpha(\gamma))$, or at $(1, \alpha(\gamma), 0)$. Naively, this suggests verifying that F_2 is concave in the neighborhood of these points and then performing a grid search in the remainder of the parameter space. However such a procedure appears to be computationally prohibitive and thus we opt for a more delicate approach. The crucial feature of $F_2(\beta, \alpha_1, \alpha_2)$ is that it is strictly convex for fixed β ; therefore at least for fixed β one can simply produce a pair of values (α_1, α_2) and certify that $F_2(\beta, \alpha_1, \alpha_2)$ is sufficiently small, with sufficiently small derivative in α_1, α_2 that it cannot reach the desired supremum.

To handle a range of β simultaneously, we produce a majorant for $F(\alpha_1, \alpha_2)$ which handles a range of β uniformly. Furthermore this majorant has a particular structure that allows one to produce a near-optimum for $\tilde{F}(\alpha_1, \alpha_2)$ via a certain Newton iteration-type procedure thus allowing for efficient optimization. This allows one to essentially handle $\beta \notin [0, .001] \cup [.495, .505] \cup [.999, 1]$. Due to symmetry it suffices to handle $\beta \in [0, .001] \cup [.495, .50]$.

For the first interval, one can verify that for all relevant (α_1, α_2) , we have $F(0, \alpha_1, \alpha_2) \geq F(\beta, \alpha_1, \alpha_2)$ via a direct mathematical argument; such an argument is plausible as in the neighborhood of 0 the derivative in β of $F(\beta, \alpha_1, \alpha_2)$ is $-\infty$ (this is a manifestation of the model exhibiting frozen 1-RSB behavior). Finally for $\beta \in [.495, .5]$, one can use a similar numerical procedure to the initial parts handling β not near 0, 1/2, 1 to localize (α_1, α_2) to a certain small ball around the desired optimizer and one can computationally verify via grid-search that the Hessian is negative definite.

The various numerical claims in Appendix B are carried out in Python 3 using numpy and scipy, and then separately in mpmath which use floating point arithmetic and python-flint which is a rigorous library for interval arithmetic. (We note here that python-flint is wrapper for the interval arithmetic package Arb package in C++ [29].) We give formal proofs for all claims in the numerical appendix except Claims B.19 to B.21 which are used to verify Claim B.4; these numerical computations have only been verified in numpy and scipy, and separately in mpmath. This is due primarily to the authors not being aware of any package which has rigorous multidimensional numerical integration; however, the necessary bounds only require estimates to 2 or 3 decimal places and are not nearly as sensitive as the others. We have attached Python code in the arXiv listing; the computer assisted portions of Claims B.1 to B.3 are carried out in InitialLocalizationOfHessianConditions.py and IntegralSupplement.py.

1.3. Notation. We use standard asymptotic notation throughout, as follows. For functions f = f(n) and g = g(n), we write f = O(g) or $f \leq g$ to mean that there is a constant C such that $|f(n)| \leq C|g(n)|$ for sufficiently large n. Similarly, we write $f = \Omega(g)$ or $f \geq g$ to mean that there is a constant c > 0 such that $f(n) \geq c|g(n)|$ for sufficiently large n. Finally, we write $f \approx g$ or $f = \Theta(g)$ to mean that $f \leq g$ and $g \leq f$, and we write f = o(g) or $g = \omega(f)$ to mean that $f(n)/g(n) \to 0$ as $n \to \infty$. Subscripts on asymptotic notation indicate quantities that should be treated as constants.

Furthermore logarithms are natural unless a base is specified. For an integer $k \ge 0$ we write $\log^{(k)} x$ for the k times iterated logarithm $\log \cdots \log x$. Then $\log^* x$ for $x \ge 1$ is defined as the unique nonnegative integer k so that $\log^{(k)} x \in [1, e]$. Finally, we write $C_{2,2}$ or $c_{2,6}$ to mean that there is such a positive absolute constant which makes the relevant proposition hold as stated, but we do not care to specify an explicit value.

1.4. Organization of the paper. In Section 2 we prove the main vertex adapted isoperimetry result Theorem 2.2. In Section 3 we reduce Theorem 1.3 to a pair of moment computations (Lemmas 3.2 and 3.3). In Section 4 we provide the necessary comparison of degree sequence to a degree sequence models. In Section 5 we use switchings to eliminate the contributions for degree sequences which are atypical. In Section 6, we state and prove the necessary log-convexity claims which are used in the proof of Theorem 1.3. Finally in Section 7 we prove Lemmas 3.2 and 3.3.

In Appendix A we compute precisely various tails of binomial coefficients which are used in the main moments claims and in Appendix B we carry out the mathematical part of computer-assisted verification of Assumption 1.5.

1.5. Concurrent work. In concurrent work, Dandi, Gamarnik and Zdeborová prove an analogue of Theorem 1.3, allowing for o(n) exceptional vertices. The proof proceeds via the Lindeberg exchange method to reduce the problem to a question on sparse graphs. Here they then extend work of Gamarnik and Li [20] to prove the desired result for sparse graphs; here sparsity is used in a crucial manner as it allows one to compute various probabilities in the configuration model.

Acknowledgments. The second and third authors thank David Gamarnik for bringing [20] to our attention as this served as the starting point for our work. The third author thanks George Barbulescu and Naveen Rameen for help installing the python-flint library and for helping write multiprocessing code to improve the runtime.

2. Concentration of Friendliness via Boolean Functions

We now prove the desired expansion result regarding subset of graphs under the metric of maxdegree differences at a vertex.

Definition 2.1. For any two graphs on a fixed (labeled) vertex set V, let

$$d(G,H) = \Delta(G \triangle H),$$

where \triangle denotes taking the graph with edge set equal to the symmetric difference of the given edge sets, namely, $(E(G) \setminus E(H)) \cup (E(H) \setminus E(G))$, and where \triangle denotes maximum degree.

Note that d(G, H) trivially defines a metric on the set of graphs. The main result of this section is the following expansion result.

Theorem 2.2. Let \mathcal{G} be a family of graphs on a labeled vertex set V that is \mathfrak{S}_n -invariant ² and let $\mu(\cdot)$ be the uniform measure on labeled graphs on n vertices³. Let $\varepsilon \in (0,1)$, possibly depending on n. Suppose that $\mu(\mathcal{G}) \geq \varepsilon$ and let

$$\mathcal{G}_T = \{H \colon \min_{G \in \mathcal{G}} d(G, H) \le T\}.$$

Then

$$\mu(\mathcal{G}_{C_{2,2}\log(1/\varepsilon)\sqrt{n/\log n}}) \ge 1-\varepsilon.$$

Remark. This theorem is optimal up to a constant factor (both in n and ε dependence) due to considering graphs in $\mathbb{G}(n, 1/2)$ with maximum degree below a fixed threshold, using the result of Bollobás [8] on this distribution.

Remark. An analogue of this result holds for properties of bipartite graphs with sides n, m satisfying $n = \Theta(m)$ which are $\mathfrak{S}_n \times \mathfrak{S}_m$ -invariant.

We first make a preliminary reduction.

 $^{{}^{2}\}mathcal{G}$ can be equivalently be viewed as a family of undirected graphs.

³Equivalently, $\mu(\mathcal{S}) = \mathbb{P}_{G \sim \mathbb{G}(n, 1/2)}[G \in \mathcal{S}]$

Proposition 2.3. Let $\varepsilon \in (0, 1/4]$. Let \mathcal{G} be a family of graphs on a labeled vertex set V that is \mathfrak{S}_n -invariant and let $\mu(\cdot)$ be the uniform measure on labeled graphs on n vertices. Suppose that $\mu(\mathcal{G}) \geq \varepsilon$ and let

$$\mathcal{G}_T = \{H \colon \min_{G \in \mathcal{G}} d(G, H) \le T\}.$$

Then

$$\mu(\mathcal{G}_{C_{2,3}\sqrt{n/\log n}}) \geq 2\varepsilon.$$

We briefly sketch the reduction.

Proof of Theorem 2.2 given Proposition 2.3. Iterating Proposition 2.3 allows one to immediately prove $\mu(\mathcal{G}_{C\log(1/\varepsilon)\sqrt{n/\log n}}) \geq 1/2$ for an appropriately chosen absolute constant C. To boost from probability 1/2 to $1 - \varepsilon$, let $\mathcal{G}_0 = \mathcal{G}_{C\log(1/\varepsilon)\sqrt{n/\log n}}$. Suppose that $\mu(\mathcal{G}_0) = 1 - \delta$ for $\delta \in (0, 1/2]$. Consider the graphs which are distance at least $C\sqrt{n/\log n}$ in distance away from \mathcal{G}_0 , call it \mathcal{G}_1 . Applying Proposition 2.3 implies that $\mu(\mathcal{G}_1) \leq \delta/2$ (since the $C\sqrt{n/\log n}$ -neighborhood is disjoint from \mathcal{G}_0). Iterating this procedure $O(\log(1/\varepsilon))$ times gives the claimed result.

To prove Proposition 2.3 we will require a number of standard definitions from Boolean function analysis.

Definition 2.4. Given a function $f: \{0,1\}^n \to \{0,1\}$, define the *influence* of the *i*th variable to be $I_i[f] = \mathbb{P}[f(x) \neq f(x \oplus e_i)]$, where e_i is the *i*th basis vector and \oplus is \mathbb{F}_2 addition. Furthermore define the *total influence* of a function f to be $I[f] = \sum_{i \in [n]} I_i[f]$. Finally define the *positive sensitivity* of a function f at a point x to be

$$s_{f}^{+}(x) = \begin{cases} 0 & \text{if } f(x) = 1\\ \sum_{i \in [n]} f(x \oplus e_{i}) & \text{if } f(x) = 0. \end{cases}$$

We will need the following variant of a strengthening of Talagrand's inequality, which follows from the proof given by Eldan, Kindler, Lifshitz, and the first author [14, Theorem 1.2]. The only change required is noting that [14, Lemma 3.1] can be replaced with $\mathbb{E}\left[\sqrt{s_f^+(x)}\right] \ge \sqrt{\sum_{i \in [n]} \hat{f}(\{i\})^2}$.

Theorem 2.5. For any function $f: \{0,1\}^n \to \{0,1\}$,

$$\mathbb{E}\left[\sqrt{s_f^+(x)}\right] \gtrsim \operatorname{var}(f) \cdot \sqrt{\log\left(1 + \frac{1}{\sum_{i \in [n]} I_i[f]^2}\right)}.$$

In particular, if $\sum_{i \in [n]} I_i[f]^2 \le n^{-1/4}$ then

$$\mathbb{E}\left[\sqrt{s_f^+(x)}\right] \gtrsim \operatorname{var}(f)\sqrt{\log n}.$$

We in fact will only apply Theorem 2.5 for functions f on the set of graphs $\{0, 1\}^{\binom{[n]}{2}}$ which are symmetric under relabeling the vertices. Therefore we have that $I_{(i_1,i_2)}[f]$ is independent of the pair (i_1, i_2) . From this we now derive the following trichotomy for functions f which are symmetric.

Lemma 2.6. Let $\varepsilon \in (0, 1/4]$ and $f: \{0, 1\}^{\binom{[n]}{2}} \to \{0, 1\}$ such that $\mathbb{E}f \in [\varepsilon, 1 - \varepsilon]$. Then if f is symmetric under the natural \mathfrak{S}_n -action, then one of the following holds, where we sample x from the uniform measure on labeled graphs μ :

Infl $\mathbb{P}[s_f^+(x) > 0] \ge c_{2.6} \varepsilon \sqrt{\log n/n}$ Infl There is an integer $j \in [\lfloor (\log_2 n)/2 \rfloor, \lfloor \log_2 \binom{n}{2} \rfloor]$ such that $2^j \mathbb{P}[s_f^+(x) \in [2^j, 2^{j+1}]] \ge c_{2.6} \varepsilon \sqrt{n} (\log n)^5.$ Inf3 There are integers $\ell \in [2, \log^* n - 1], j \in [\lfloor \log_2 n + 30 \log^{(\ell+1)} n \rfloor, \lfloor \log_2 n + 30 \log^{(\ell)} n \rfloor]$ such that $2^{j/2 \operatorname{Tr}[\ell+(\ell)]} \in [2^j, 2^{j+1}] > \sqrt{1 - (\ell)} > 2^{j/2}$

$$2^{j/2}\mathbb{P}[s_f^+(x) \in [2^j, 2^{j+1}]] \ge c_{2.6}\varepsilon \sqrt{\log n} / (\log^{(\ell)} n)^2.$$

Proof. Note that we may assume n is larger than an absolute constant (not dependent on ε). Suppose that Infl fails, i.e., $\mathbb{P}[s_f^+(x) > 0] < c_{2.6}\varepsilon \sqrt{\log n/n}$ where the constant will be chosen later. Furthermore note that

$$I[f] \asymp \sum_{j=0}^{\lfloor \log_2 \binom{n}{2} \rfloor} 2^j \mathbb{P}[s_f^+(x) \in [2^j, 2^{j+1}]] \lesssim \sum_{j=\lfloor (\log_2 n)/2 \rfloor}^{\lfloor \log_2 \binom{n}{2} \rfloor} 2^j \mathbb{P}[s_f^+(x) \in [2^j, 2^{j+1}]] + n^{1/2}.$$

If $I[f] \ge n^{1/2} (\log n)^6$ then there exists $j \in [\lfloor (\log_2 n)/2 \rfloor, \lfloor \log_2 \binom{n}{2} \rfloor]$ such that

$$2^j \mathbb{P}[s_f^+(x) \in [2^j, 2^{j+1}]] \gtrsim \sqrt{n} (\log n)^5$$

which will satisfy Inf2 for appropriately chosen $c_{2,6} > 0$.

Thus, if we further assume that Inf2 fails, then in the remaining case we may assume $I[f] \leq n^{1/2}(\log n)^6$, $\mathbb{P}[s_f^+(x) > 0] \leq c_{2.6} \varepsilon \sqrt{\log n/n}$, and $w_j := 2^j \mathbb{P}[s_f^+(x) \in [2^j, 2^{j+1}]] \leq c_{2.6} \varepsilon \sqrt{n}(\log n)^5$ for all $j \in [\lfloor (\log_2 n)/2 \rfloor, \lfloor \log_2 {n \choose 2} \rfloor]$. Using that f is symmetric we have

$$\sum_{(i_1,i_2)\in \binom{[n]}{2}} I_{(i_1,i_2)}[f]^2 = \binom{n}{2}^{-1} I[f]^2 \le n^{-1/2}.$$

Thus we have by Theorem 2.5 and the condition on $\mathbb{E}f$ that $\mathbb{E}[\sqrt{s_f^+(x)}] \gtrsim \varepsilon \sqrt{\log n}$. Note that

$$\mathbb{E}[\sqrt{s_f^+(x)}] \asymp \sum_{j=0}^{\lfloor \log_2 \binom{n}{2} \rfloor} 2^{-j/2} w_j$$

and that

$$\sum_{j=0}^{\lfloor \log_2 n + 100 \rfloor} 2^{-j/2} w_j \le 2^{(\log_2 n + 100)/2} \mathbb{P}[s_f^+(x) > 0] \lesssim c_{2.6} \varepsilon \sqrt{\log n}$$

Furthermore

1

$$\sum_{i=\lfloor \log_2 n+30\log\log n \rfloor}^{\lfloor \log_2 \binom{n}{2} \rfloor} 2^{-j/2} w_j \le 2^{-(\log_2 n+30\log\log n)/2} c_{2.6} \varepsilon \sqrt{n} (\log n)^5 \lesssim c_{2.6} \varepsilon (\log n)^{-1} N_{10} \log \log n$$

Combining these three inequalities we have (choosing $c_{2.6}$ appropriately small) that

$$\sum_{\substack{j=\lfloor \log_2 n+100\rfloor}}^{\log_2 n+30\log\log n\rfloor} 2^{-j/2} w_j \gtrsim \varepsilon \sqrt{\log n}.$$
(2.1)

Now $\ell \in [2, \log^* n - 1]$, $j \in [\lfloor \log n + 30 \log^{(\ell+1)} n \rfloor, \lfloor \log n + 30 \log^{(\ell)} n \rfloor]$ cover the interval of values of j between $\lfloor \log_2 n + 100 \rfloor$ and $\lfloor \log_2 n + 30 \log \log n \rfloor$. Thus, if none of the desired indices satisfy **Inf3** then the total contribution on the left of (2.1) would be bounded via

$$\sum_{j=\lfloor \log_2 n+100 \rfloor}^{\lfloor \log_2 n+30\log\log n \rfloor} 2^{-j/2} w_j \le \sum_{\ell \in [2,\log^* n-1]} \frac{c_{2.6}\varepsilon\sqrt{\log n}}{\log^{(\ell)} n} \lesssim c_{2.6}\varepsilon\sqrt{\log n}$$

which provides a contradiction to (2.1) if $c_{2.6}$ is sufficiently small.

Given these facts we are now in position to prove Proposition 2.3 and hence Theorem 2.2.

Proof of Proposition 2.3. We define a series of expansions of $\mathcal{F}_0 = \mathcal{G}$ as follows, each of being an \mathfrak{S}_n -symmetric Boolean function. For each $i \geq 1$, if $\mu(\mathcal{F}_{i-1}) \geq 2\varepsilon$ we stop. Otherwise suppose that $\mu(\mathcal{F}_{i-1}) \in [\varepsilon, 2\varepsilon)$. Let $f_{i-1} = \mathbb{1}_{\mathcal{F}_{i-1}}$ and note that f_{i-1} is a symmetric Boolean function.

Alg1 If $\mathbb{P}[s_{f_{i-1}}^+(x) > 0] \ge c_{2.6} \varepsilon \sqrt{\log n/n}$, define $\mathcal{F}_i = \mathcal{F}_{i-1} \cup \{x \colon s_{f_{i-1}}^+(x) > 0\}$. **Alg2** Else if there is $j \in [\lfloor (\log_2 n)/2 \rfloor, \lfloor \log_2 \binom{n}{2} \rfloor]$ satisfying

$$2^{j}\mathbb{P}[s_{f_{i-1}}^{+}(x) \in [2^{j}, 2^{j+1}]] \ge c_{2.6}\varepsilon\sqrt{n}(\log n)^{5}$$

then define $\mathcal{F}_i = \mathcal{F}_{i-1} \cup \{x : s_{f_{i-1}}^+(x) \in [2^j, 2^{j+1}]\}.$

Alg3 Else if there are $\ell \in [2, \log^* n - 1], j \in [\lfloor \log_2 n + 30 \log^{(\ell+1)} n \rfloor, \lfloor \log_2 n + 30 \log^{(\ell)} n \rfloor]$ such that

$$2^{j/2} \mathbb{P}[s_{f_{i-1}}^+(x) \in [2^j, 2^{j+1}]] \ge c_{2.6} \varepsilon \sqrt{\log n} / (\log^{(\ell)} n)^2$$

then define $\mathcal{F}_i = \mathcal{F}_{i-1} \cup \{x : s_{f_{i-1}}^+(x) \in [2^j, 2^{j+1}]\}.$

By Lemma 2.6, at least one of these situations occurs so this is well-defined. Note that at each stage of this expansion, symmetry of f_{i-1} implies that f_i is also symmetric. Combined with the fact $2\varepsilon \leq 1 - \varepsilon$, we see the application of Lemma 2.6 is always valid. Additionally, note that every $H \in \mathcal{F}_i$ is either in \mathcal{F}_{i-1} or obtained from a graph in \mathcal{F}_{i-1} by flipping an edge in or out.

Note that $s_{f_{i-1}}^+(x) > 0$ is only possible for $x \notin \mathcal{F}_{i-1}$. Thus $\mu(\mathcal{F}_i \setminus \mathcal{F}_{i-1}) \gtrsim \varepsilon \sqrt{\log n/n}$ if Alg1 occurs, $\mu(\mathcal{F}_i \setminus \mathcal{F}_{i-1}) \gtrsim \varepsilon \sqrt{n}(\log n)^5/2^j$ if Alg2 with parameter j occurs, and $\mu(\mathcal{F}_i \setminus \mathcal{F}_{i-1}) \gtrsim \varepsilon \sqrt{\log n}/(2^{j/2}(\log^{(\ell)} n)^2)$ if Alg3 with parameters ℓ and j occurs.

Since we stop the process once the density is at least 2ε , there must be at most $O(\sqrt{n/\log n})$ steps of Alg1, at most $O(2^{j}/(\sqrt{n}(\log n)^5) + 1)$ steps of Alg2 for each possible parameter j, and at most $O(2^{j/2}(\log^{(\ell)} n)^2/\sqrt{\log n})$ steps of Alg3 for each possible choice of parameters ℓ, j , where the implicit constants are independent of ε . Thus the process described above must eventually stop at some time t with $\mu(\mathcal{F}_t) \geq 2\varepsilon$. We now prove any graph in \mathcal{F}_t is close to \mathcal{F}_0 in graph distance in the sense of Definition 2.1, which will finish.

Let us consider some $H \in \mathcal{F}_t$. Let $H_t = H$. Additionally, let R_j for $j \in [\lfloor (\log_2 n)/2 \rfloor, \lfloor \log_2 \binom{n}{2} \rfloor]$ and $R_{\ell,j}$ for $\ell \in [2, \log^* n - 1]$ and $j \in [\lfloor \log_2 n + 30 \log^{(\ell+1)} n \rfloor, \lfloor \log_2 n + 30 \log^{(\ell)} n \rfloor]$ be "record graphs" which start as the empty graph on V = V(H). For $i \in [t]$ in decreasing order, we perform the following:

Alg0' If $H_i \in \mathcal{F}_{i-1}$ let $H_{i-1} = H_i$. Decrement *i* and restart the loop.

- Alg1' Otherwise, if further $\mathcal{F}_i \setminus \mathcal{F}_{i-1}$ was generated using Alg1, let $H_{i-1} \in \mathcal{F}_{i-1}$ be chosen to be one of the graphs distance 1 (a single edge flip) from H_i arbitrarily.
- **Alg2'** Else if $\mathcal{F}_i \setminus \mathcal{F}_{i-1}$ was generated using **Alg2** with parameter j, let $H_{i-1} \in \mathcal{F}_{i-1}$ be chosen to be one of the graphs at distance 1 from H_i . If possible, we choose it with the additional property that if e_i is the differing edge, adding e_i to record graph R_j does not increase the maximum degree of R_j . (If e_i is already present in R_j , adding it in leaves R_j the same.)
- **Alg3'** Else if $\mathcal{F}_i \setminus \mathcal{F}_{i-1}$ was generated using **Alg3** with parameters ℓ, j , let $H_{i-1} \in \mathcal{F}_{i-1}$ be chosen to be one of the graphs at distance 1 from H_i . If possible, we choose it with the additional property that if e_i is the differing edge, adding e_i to record graph R_j does not increase the maximum degree of R_j .

This is clearly well-defined since every $\mathcal{F}_i \setminus \mathcal{F}_{i-1}$ is generated via one of Alg1 to Alg3 and also any graph in $\mathcal{F}_i \setminus \mathcal{F}_{i-1}$ is generated by flipping an edge of some graph in \mathcal{F}_{i-1} . Furthermore, we have $H_0 \in \mathcal{F}_0 = \mathcal{G}$.

For all $i \in [t]$, let e_i be the differing edge between H_{i-1}, H_i . Clearly $H_0 \triangle H_t$ is a subgraph of the union of all these edges, call it R ($H_0 \triangle H_t$ may be a strict subgraph of R if say at one step we flip an edge, and at another step flip it again). It suffices to bound the maximum degree of R. Let R^*

be the union of edges from Alg1, and let $R_j^*, R_{\ell,j}^*$ be the final values of the record graphs $R_j, R_{\ell,j}$ for the various parameter choices.

The edges e_i corresponding to Alg1 are at most $O(\sqrt{n/\log n})$ in number by the earlier analysis and hence $\Delta(R^*) = O(\sqrt{n/\log n})$.

For Alg2 with parameter j, let $k = 2^j$ and recall that there are at most $O(k/(\sqrt{n}(\log n)^5) + 1)$ corresponding relevant steps i. For j such that $k \leq n(\log n)^2$, this amounts to $O(\sqrt{n}/(\log n)^3)$ total edges, so certainly $\Delta(R_j^*) = O(\sqrt{n}/(\log n)^3)$.

For j such that $n(\log n)^2 < k \leq n^2$, consider any relevant step i. We have a graph $H_i \in \mathcal{F}_i$ and a current record graph R_j , and we choose $H_{i-1} \in \mathcal{F}_{i-1}$ at distance 1 from H_i . Let $\Delta = \lfloor \sqrt{n}/(\log n)^2 \rfloor$. If $\Delta(R_j) < \Delta$ before step i then clearly $\Delta(R_j) \leq \Delta$ after step i. Otherwise $\Delta(R_j) \geq \Delta$ before step i. By the definition of Alg2, there are at least k choices of $H_{i-1} \in \mathcal{F}_{i-1}$ at distance 1 to H_i . Equivalently, there are at least k choices of e_i . If all of these choices have at least one endpoint with degree at least Δ in R_j , then R_j has at least k/n vertices of degree at least Δ , and hence at least $(k\Delta)/(2n)$ total edges. This contradicts our bound of $O(k/(\sqrt{n}(\log n)^5) + 1)$ on the number of steps of Alg2 with parameter j. That is, there will always be a choice which maintains the maximum degree once if it ever hits Δ . Thus we see that $\Delta(R_j^*) \leq \Delta = O(\sqrt{n}/(\log n)^2)$.

Finally for Alg3 with parameters ℓ, j , let $k = 2^j$ and recall that earlier analysis showed that there are at most $O(\sqrt{k}(\log^{(\ell)} n)^2/\sqrt{\log n})$ total relevant steps *i*. We can perform a similar argument as the previous case. Let $\Delta_{\ell,j} = Kn(\log^{(\ell)} n)^4/\sqrt{2^j \log n}$ where *K* is a sufficiently large absolute constant. By the definition of Alg3, there are at least *k* choices of H_{i-1} or equivalently edges e_i . If all of these choices have at least one endpoint with degree at least $\Delta_{\ell,j}$ in $R_{\ell,j}$, then $R_{\ell,j}$ has at least $(k\Delta_{\ell,j})/(2n)$ total edges. This violates our bound on the number of relevant steps *i* as long as *K* was chosen sufficiently large. Hence a similar argument as in the analysis of Alg2 shows that $\Delta(R^*_{\ell,j}) \leq \Delta_{\ell,j}$.

Finally,

$$\begin{split} \Delta(R) &\leq \Delta(R^*) + \sum_{j=\lfloor (\log_2 n)/2 \rfloor}^{\lfloor \log_2 \binom{n}{2} \rfloor} \Delta(R_j^*) + \sum_{\ell=2}^{\log^* n-1} \sum_{j=\lfloor \log_2 n+30 \log^{(\ell)} n \rfloor}^{\lfloor \log_2 n+30 \log^{(\ell)} n \rfloor} \Delta(R_{\ell,j}^*) \\ &\lesssim \sqrt{\frac{n}{\log n}} + (\log n) \cdot \frac{\sqrt{n}}{(\log n)^2} + \sum_{\ell=2}^{\log^* n-1} \frac{n(\log^{(\ell)} n)^4}{\sqrt{2^{\lfloor \log_2 n+30 \log^{(\ell+1)} n \rfloor} \log n}} \\ &\lesssim \sqrt{\frac{n}{\log n}} + \sqrt{\frac{n}{\log n}} \sum_{\ell=2}^{\log^* n-1} \frac{(\log^{(\ell)} n)^4}{2^{15 \log^{(\ell+1)} n}} \lesssim \sqrt{\frac{n}{\log n}} \left(1 + \sum_{\ell=2}^{\log^* n-1} \frac{1}{\log^{(\ell)} n}\right) \lesssim \sqrt{\frac{n}{\log n}}. \end{split}$$

We are done, since this establishes that every $H \in \mathcal{F}_t$ satisfies $\Delta(H \triangle H_0) = O(\sqrt{n/\log n})$ for some $H \in \mathcal{G}$, and since $\mu(\mathcal{F}_t) \ge 2\varepsilon$.

2.1. Applications beyond the present work. We finally elaborate on a remark made in Section 1.2.1 stating that Theorem 2.2 may be used to derive the threshold for the capacity of the symmetric binary perceptron. Recall that the threshold for the symmetric binary perceptron with parameter κ (with Bernoulli disorder) is defined via taking a series of random vectors $x_1, \ldots \in \{\pm 1\}^n$ and asking for the first time τ at which there does not exist $y \in \{\pm 1\}^n$ such that $|\langle x_i, y \rangle| \leq \kappa \sqrt{n}$ for all $i \in [\tau]$. The work of Aubins, Perkins, and Zdeborová [3] establish (based on the second moment method) the existence of a continuous function $\alpha(\kappa)$ such that with high probability the threshold is less than $(\alpha(\kappa) + o(1))n$ and with $\Omega(1)$ probability the threshold is at least $(\alpha(\kappa) - o(1))n$, if the disorder is Gaussian, and their technique applies in the case of Bernoulli disorder as well. Treating the vectors x_1, \ldots, x_m as an $m \times n$ adjacency matrix of a bipartite graph, the remark following Theorem 2.2 regarding bipartite graphs immediately implies that at $(\alpha(\kappa) - o(1))n$ whp there exists

a vector y such that $|\langle x_i, y \rangle| \leq (\kappa + o(1))\sqrt{n}$. Via continuity of $\alpha(\kappa)$, this establishes the sharp threshold for the symmetric perceptron model. We also note that an essentially identical argument proves that the number of solutions for a given m concentrates on the exponential scale. (We note sharpness and counting solutions was shown without Theorem 2.2 by Abbe, Li, and Sly [1] for Bernoulli disorder and Perkins and Xu [36] for Gaussian disorder.)

We note that in general arguments in the flavor of Theorem 2.2 in the context of perceptron model can prove sharp thresholds given "monotonicity of the activation functions with respect to deformation" in an appropriate sense. If one can prove a priori that the (not necessarily sharp) threshold is continuous with respect to an appropriate "deformation parameter", then Theorem 2.2 or similar arguments may apply to achieve sharpness. General results of this nature were established in work of Xu [46], with a substantially simplified proof given in work of Nakajima and Sun [35]. We note here that the techniques of Perkins and Xu [46] and Nakajima and Sun [35] rely on the "row by row" independence of the perceptron model and hence appear fundementally unsuited to give sharpness for our situation. The work of Xu [46] more similarly relies on Boolean functions and symmetry; however it requires as input deep work of Hatami [27]. Finally the work of Abbe, Li, and Sly [1] is based on an analogue of small subgraph conditioning for dense graphs and at the very least requires one to be able to compute the moments underlying Theorem 1.3 to a multiplicative 1+o(1) accuracy. This appears to be nigh computationally infeasible. Furthermore, one would need to show asymptotic normality for subgraph counts in various associated degree-constrained models (which has only recently seen progress [42]).

3. Setup for Second Moment Calculation

We now define the necessary definitions for the remainder of the paper.

Definition 3.1. Let $G \sim \mathbb{G}(2n, 1/2)$ and V denote the vertex set of G. Furthermore, let X_{γ} denote the number of equipartitions $V = Y_1 \cup Y_2$ of G such that for every vertex $y \in Y_1$ we have $\deg_{Y_1}(y) \geq \deg_{Y_2}(y) + \gamma \sqrt{n}$ and for every $y \in Y_2$ we have $\deg_{Y_2}(y) \geq \deg_{Y_1}(y) + \gamma \sqrt{n}$. We will always consider $\gamma \in [-1, 1]$ with the property that $\gamma \sqrt{n} \in \mathbb{Z}$.

The two crucial computations give a precise (up to constant) asymptotic for $\mathbb{E}[X_{\gamma}]$ and $\mathbb{E}[X_{\gamma}^2]$.

Lemma 3.2. Suppose that $|\gamma| \leq 1$ and $\gamma \sqrt{n} \in \mathbb{Z}$. Then with notation as in Definition 3.1,

$$\mathbb{E}X_{\gamma} \asymp \binom{2n}{n} \left(\sup_{\alpha \in \mathbb{R}} \exp(-\alpha^2) \mathbb{P}_{Z \sim \mathcal{N}(0,1)} [Z \ge (\gamma + \alpha)\sqrt{2}] \right)^{2n} = 4^{-n} \binom{2n}{n} \exp(2n \sup_{\alpha \in \mathbb{R}} F_1(\alpha)).$$

Lemma 3.3. Suppose that Assumption 1.5 holds. Fix $\eta > 0$ and suppose $\gamma \in [\gamma_{\text{crit}} - \varepsilon_{1.5}, \gamma_{\text{crit}} + \varepsilon_{1.5}]$ and $\gamma \sqrt{n} \in \mathbb{Z}$. Then with notation as in Definition 3.1,

$$\mathbb{E}X_{\gamma}^{2} \lesssim_{\eta} (1+\eta)^{n} \mathbb{E}X_{\gamma} + 16^{-n} \binom{2n}{n}^{2} \exp\left(n \sup_{\substack{\beta \in [0,1]\\\alpha_{1},\alpha_{2} \in \mathbb{R}}} F_{2}(\beta,\alpha_{1},\alpha_{2})\right).$$

Remark. We certainly believe that Lemma 3.3 holds for all γ ; the restriction on the range of γ in consideration comes precisely from Assumption 1.5.

We now prove Theorem 1.3 given Assumption 1.5 and Lemmas 3.2 and 3.3. In the following argument we will not be too sensitive about $\gamma\sqrt{n}$ being an integer as it is inconsequential here, but we note that in Section 7 where Lemmas 3.2 and 3.3 are proven one must be careful about integrality concerns.

Proof. Choose $\varepsilon > 0$ sufficiently small and note direct numerical computation verifies $0 \le \gamma_{\text{crit}} \le 1/2$. We first prove that with high probability, $G \sim \mathbb{G}(2n, 1/2)$ has no equipartitions where each vertex has at least $(\gamma_{\text{crit}}/\sqrt{2} + \varepsilon)\sqrt{2n}$ more vertices in its own part than in the other side. Note

that $\gamma \mapsto \sup_{\alpha \in \mathbb{R}} (-\alpha^2 + \mathbb{P}_{Z \sim \mathcal{N}(0,1)}[Z \geq (\gamma + \alpha)\sqrt{2}])$ is trivially a strictly decreasing function. Therefore by Lemma 3.2 and Definition 1.2, the number of $(\gamma_{\rm crit}/\sqrt{2}+\varepsilon)\sqrt{2n}$ -friendly equipartitions in expectation is exponentially small. By Markov's inequality, there are no such equipartitions with probability $1 - \exp(-\Omega_{\varepsilon}(n))$. The analogous result for odd numbers can be deduced as follows: if there exists a sufficiently friendly partition in $\mathbb{G}(2n-1,1/2)$ then adding a vertex with uniformly random neighbors to the smaller part, making the total number of vertices even, gives a partition which is $(\gamma_{\rm crit}/\sqrt{2} + \varepsilon/2)\sqrt{2n}$ -friendly with probability $\Omega(1)$ (namely, approximately the probability that the extra vertex is sufficiently friendly), a contradiction to the even case.

We next prove that with high probability, $G \sim \mathbb{G}(2n, 1/2)$ has a $(\gamma_{\text{crit}}/\sqrt{2} - \varepsilon)\sqrt{2n}$ -friendly equipartition whp. Let $\gamma = \gamma_{\text{crit}} - \varepsilon/2$. By $\sup_{\alpha \in \mathbb{R}} F_1(\alpha)$ decreasing as γ increases and the definition of γ_{crit} , along with Assumption 1.5, we know that $\sup_{\alpha \in \mathbb{R}} F_1(\alpha) > 0$ hence $\sup_{\beta \in [0,1]} F_2(\beta, \alpha_1, \alpha_2) = 0$

 $4F_1(\alpha)$. Then Lemmas 3.2 and 3.3 with η sufficiently small in terms of ε demonstrates

$$\mathbb{E}X_{\gamma}^2 \lesssim_{\varepsilon} (\mathbb{E}X_{\gamma})^2.$$

Therefore, since X_{γ} takes on nonnegative integer values, we have $\mathbb{P}[X_{\gamma} \neq 0] = \Omega_{\varepsilon}(1)$. That is, we have at least a constant lower bound on the desired probability. The next step is to use the techniques of Section 2 to boost this probability to 1 - o(1).

Let \mathcal{P} be the property of graphs on 2n vertices defined by having a $(\gamma_{\rm crit} - \varepsilon/2)$ -friendly bisection. We have $\mu(\mathcal{P}) = \Omega_{\varepsilon}(1)$ by the above analysis, borrowing the notation from Section 2. Then Theorem 2.2 implies $\mu(\mathcal{P}_{\sqrt{n}(\log n)^{-1/4}}) = 1 - o(1)$. But notice that under the metric defined in Definition 2.1, at every vertex the set of neighbors differs by at most $\sqrt{n}(\log n)^{-1/4}$ additions and deletions when moving from a graph in \mathcal{P} to $\mathcal{P}_{\sqrt{n}(\log n)^{-1/4}}$. Therefore a $(\gamma_{\text{crit}} - \varepsilon/2)$ -friendly bisection guaranteed in \mathcal{P} will certainly remain at least $(\gamma_{\text{crit}} - \varepsilon)$ -friendly in $\mathcal{P}_{\sqrt{n}(\log n)^{-1/4}}$. The result follows for graphs with 2n vertices.

Finally, to prove this part of Theorem 1.3 for graphs with 2n-1 vertices, simply add a vertex with uniformly random neighborhood to form $\mathbb{G}(2n, 1/2)$, find a $(\gamma_{\text{crit}} - \varepsilon/2)$ -friendly bisection, and then remove the extra vertex and check that the result is still $(\gamma_{\rm crit} - \varepsilon)$ -friendly. We are done.

In the remainder of the body of the paper we focus on proving Lemmas 3.2 and 3.3.

4. Degree Models

4.1. **Degree enumeration.** We will first require the following results regarding degree enumeration due to [33] and [10], respectively.

Theorem 4.1 ([33]). There exists a fixed constant $\varepsilon = \varepsilon_{4,1} > 0$ 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

• $|d_i - \overline{d}| \le n^{1/2+\varepsilon}$ for $1 \le i \le n$, and

•
$$\overline{d} \ge n/\log n$$
.

Writing $m = \overline{dn}/2 \in \mathbb{Z}$, $\mu = \overline{d}/(n-1)$, and $\gamma_2^2 = (1/(n-1)^2) \sum_{i=1}^n (d_i - \overline{d})^2$, the number of labeled 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 4.2 ([10]). There exists a fixed constant $\varepsilon = \varepsilon_{4,2} > 0$ 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 $\overline{t} = (1/n) \sum_{i=1}^m t_i$, we have

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

• $\overline{s} \ge n/(\log n)^{1/2}$ and $\overline{t} \ge m/(\log m)^{1/2}$.

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 labeled 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 now define a plethora of degree sequence models for random graphs that will be needed for the computations. At a high level, the work of McKay and Wormald [34] and McKay and Skerman [31] demonstrate that degrees of random graphs look independent conditional on, for example, total edge count. These models provide a way to encapsulate these facts quantitatively; however the precise result in [31, 34] are not sufficient for our work and we will require a number of suitable modifications.

Definition 4.3 (Degree sequence domains). Let $I_n = \{0, \ldots, n-1\}^n$, E_n be the even sum sequences in this set, and I_n^{ℓ} be the sum ℓ sequences. We will typically denote elements of these sets by **d**. Let $I_{m,n} = \{0, \ldots, n\}^m \times \{0, \ldots, m\}^n$, $E_{m,n}$ be the sequences with equal sums on both sides, and $E_{m,n}^{\ell}$ be the sequences with equal sums ℓ . We will typically denote elements of these sets by **s** of length m and **t** of length n. We will denote random variable versions of these by capital boldface instead.

Definition 4.4 (True degree models). \mathcal{D}_p^n is the degree sequence distribution of $\mathbb{G}(n,p)$, which is a random variable supported on $E_n \subseteq I_n$. $\mathcal{D}_p^{m,n}$ is the degree sequence distribution of a bipartite graph with m vertices on one side and n on the other, each edge included independently with probability p, which is a random variable supported on $E_{m,n} \subseteq I_{m,n}$.

Definition 4.5 (Independent degree models). \mathcal{B}_p^n is the distribution of n independent $\operatorname{Bin}(n-1,p)$ random variables, supported on I_n . $\mathcal{B}_p^{m,n}$ is the distribution of m independent $\operatorname{Bin}(n,p)$ and n independent $\operatorname{Bin}(m,p)$ variables, supported on $I_{m,n}$.

Definition 4.6 (Conditioned degree models). \mathcal{E}_p^n is the distribution of \mathcal{B}_p^n conditioned on having even sum, supported on E_n . $\mathcal{E}_p^{m,n}$ is the distribution of $\mathcal{B}_p^{m,n}$ conditioned on having equal sums on both sides, supported on $E_{m,n}$.

Definition 4.7 (Integrated degree models). \mathcal{I}_p^n is the distribution sampled as follows. Sample $p' \sim \mathcal{N}(p, p(1-p)/(n^2-n))$, conditional on being in (0,1). Then sample from $\mathcal{E}_{p'}^n$. $\mathcal{I}_p^{m,n}$ is the distribution sampled as follows. Sample $p' \sim \mathcal{N}(p, p(1-p)/(2mn))$, conditional on being in (0,1). Then sample from $\mathcal{E}_{p'}^{m,n}$.

The key reason these degree models will prove crucial in our analysis is that they allow, losing a constant, an estimate for the probability that a random graph or bipartite graph has a particular degree sequence:

Lemma 4.8. Fix $\varepsilon \in (0, \varepsilon_{4,8}]$ and C > 0. Let a sequence $\vec{d} \in \mathbb{Z}^n$ be (C, ε) -regular if:

 $\begin{array}{l} \boldsymbol{A1} \sup_{i \in [n]} |d_i - (n-1)/2| \leq n^{1/2+\varepsilon}, \\ \boldsymbol{A2} \sum_{i \in [n]} |d_i - (n-1)/2|^2 \leq Cn^2, \\ \boldsymbol{A3} \ 2|\sum_{i \in [n]} d_i. \end{array}$

Then there exists $C' = C'_{4,8}(C,\varepsilon)$ such that for all (C,ε) -regular \vec{d} we have

$$\frac{1}{C'} \le \frac{\mathbb{P}_{\mathcal{D}_{1/2}^n}[\mathbf{D} = \vec{d}]}{\mathbb{P}_{\mathcal{I}_{1/2}^n}[\mathbf{D} = \vec{d}]} \le C'.$$

Lemma 4.9. Fix $\varepsilon \in (0, \varepsilon_{4,9}]$, $\theta > 0$, and C > 0. Let a pair of degree sequences $(\vec{s}, \vec{t}) \in \mathbb{Z}^{n_1} \times \mathbb{Z}^{n_2}$ be (C, θ, ε) -regular if:

B1 $\theta \le n_1/n_2 \le \theta^{-1}$, $\begin{array}{l} \textbf{B1} \ b \leq n_1/n_2 \leq b \ , \\ \textbf{B2} \ \sum_{i \in [n_1]} s_i = \sum_{i \in [n_2]} t_i, \\ \textbf{B3} \ |s_i - n_2/2| \leq n_2^{1/2 + \varepsilon} \ for \ i \in [n_1], \\ \textbf{B4} \ |t_i - n_1/2| \leq n_1^{1/2 + \varepsilon} \ for \ i \in [n_2], \\ \textbf{B5} \ \sum_{i \in [n_1]} |s_i - n_2/2|^2 + \sum_{i \in [n_2]} |t_i - n_1/2|^2 \leq Cn_1^2. \end{array}$

Then there exists $C' = C'_{4,9}(C,\theta,\varepsilon)$ such that for all (C,θ,ε) -regular (\vec{s},\vec{t}) we have

$$\frac{1}{C'} \leq \frac{\mathbb{P}_{\mathcal{D}_{1/2}^{n_1,n_2}}[\mathbf{S} = \vec{s} \wedge \mathbf{T} = \vec{t}]}{\mathbb{P}_{\mathcal{I}_{1/2}^{n_1,n_2}}[\mathbf{S} = \vec{s} \wedge \mathbf{T} = \vec{t}]} \leq C'.$$

These lemmas follow from Theorems 4.1 and 4.2 and some basic computations, which we now provide. We will need the following technical binomial coefficient estimate.

Lemma 4.10. For $|t| \le n^{4/5}$,

$$\binom{n}{n/2+t}\binom{n}{n/2}^{-1} = \left(1 + O\left(\frac{1}{n} + \frac{t^2}{n^2} + \frac{t^6}{n^5}\right)\right) \exp\left(-\frac{2t^2}{n} - \frac{4t^4}{3n^3}\right)$$

Proof. Note that

$$\binom{n}{n/2+t}\binom{n}{n/2}^{-1} = \frac{(n/2)!^2}{(n/2+t)!(n/2-t)!}$$
$$= \left(1+O\left(\frac{1}{n}\right)\right)\sqrt{\frac{n^2}{n^2-4t^2}}\left(1+\frac{2t}{n}\right)^{-(n/2+t)}\left(1-\frac{2t}{n}\right)^{-(n/2-t)}$$
$$= \left(1+O\left(\frac{1}{n}+\frac{t^2}{n^2}\right)\right)\exp\left(-\frac{n}{2}\left(\frac{4t^2}{n^2}+\frac{8t^4}{3n^4}+O\left(\frac{t^6}{n^6}\right)\right)$$
$$= \left(1+O\left(\frac{1}{n}+\frac{t^2}{n^2}+\frac{t^6}{n^5}\right)\right)\exp\left(-\frac{2t^2}{n}-\frac{4t^4}{3n^3}\right)$$

using Stirling's formula.

Now we prove Lemmas 4.8 and 4.9.

Proof of Lemma 4.8. Let N = n(n-1)/2 and $2m = \sum_{i \in [n]} d_i$. Applying Theorem 4.1 and noting that the associated statistics satisfy $\gamma_2^2, \mu = \Theta_C(1)$ under the given conditions A1 and A2, we have

$$\mathbb{P}_{\mathcal{D}_{1/2}^{n}}[\mathbf{D} = \vec{d}] \asymp_{C,\varepsilon} 2^{-N} \binom{2N}{2m}^{-1} \binom{N}{m} \prod_{i=1}^{n} \binom{n-1}{d_{i}} \asymp_{C,\varepsilon} 2^{-2N} \exp\left(\frac{(2m-N)^{2}}{2N}\right) \prod_{i=1}^{n} \binom{n-1}{d_{i}}.$$
(4.1)

The second part comes from

$$\binom{2N}{2m}^{-1} \binom{N}{m} (1/2)^N \asymp \frac{\binom{2N}{N}}{\binom{2N}{2m}} \cdot \frac{\binom{N}{m}}{\binom{N}{N/2}} \cdot (1/2)^{2N} \asymp_C (1/2)^{2N} \exp\left(\frac{(2m-N)^2}{2N}\right)$$

by Lemma 4.10, since $m = N/2 + O_C(N^{3/4})$ follows from A2 and Cauchy–Schwarz. On the other hand, the definition of $\mathcal{I}_{1/2}^n$ and [34, Lemma 2.2] demonstrate

$$\mathbb{P}_{\mathcal{I}_{1/2}^{n}}[\mathbf{D}=\vec{d}] \asymp \prod_{i=1}^{n} \binom{n-1}{d_{i}} \int_{0}^{1} \frac{2}{1+(2p'-1)^{2N}} (p')^{2m} (1-p')^{2N-2m} \cdot \sqrt{\frac{4N}{\pi}} \exp(-4N(p'-1/2)^{2}) dp'$$

$$\approx \prod_{i=1}^{n} \binom{n-1}{d_i} \int_0^1 2\sqrt{\frac{4N}{\pi}} \exp(-4N(p'-1/2)^2)(p')^{2m}(1-p')^{2N-2m}dp'.$$
(4.2)

It suffices to prove that the expressions (4.1) and (4.2) are within multiplicative constants.

To compute the desired integral we set $2m = N + y\sqrt{N}$ (recall that $|y| \leq_C N^{1/4}$) and let $p' = 1/2 + z/(2\sqrt{N}), z_0 = -\sqrt{N}, z_1 = \sqrt{N}$. Then the integral under question can be reparametrized as

$$T(m) = \frac{2}{\sqrt{\pi}} \int_{z_0}^{z_1} t(z) dz$$

where

$$t(z) = \exp\left(-z^2 + (N + y\sqrt{N})\log(1/2 + z/(2\sqrt{N})) + (N - y\sqrt{N})\log(1/2 - z/(2\sqrt{N}))\right).$$

We break into a number of cases in order to show this integral is, up to constants, of size $2^{-2N} \exp((2m-N)^2/(2N))$. We first bound the contribution to the integral from z such that $|z| \ge 2|y| + N^{1/4}$. Note that

$$\begin{aligned} (N+y\sqrt{N})\log\left(\frac{1}{2}+\frac{z}{2\sqrt{N}}\right) + (N-y\sqrt{N})\log\left(\frac{1}{2}-\frac{z}{2\sqrt{N}}\right) \\ &\leq (N+y\sqrt{N})\log\left(\frac{1}{2}+\frac{y}{2\sqrt{N}}\right) + (N-y\sqrt{N})\log\left(\frac{1}{2}-\frac{y}{2\sqrt{N}}\right) \leq -2N\log 2 + 2y^2 \end{aligned}$$

hence we have

$$t(z) \le 2^{-2N} \exp(-z^2 + 2y^2)$$

and thus the contribution from $|z| \ge 2|y| + N^{1/4}$ is trivially seen to be negligible compared to the target value. Now when $|z| \le 2|y| + N^{1/4}$ we find from Taylor series that

$$t(z) = 2^{-2N} \exp\left(\frac{y^2 - 4(z - y/2)^2}{2} + O\left(\frac{z^4}{N} + \frac{z^2y}{N}\right)\right).$$

Noting that $z = O_C(N^{1/4})$ in this range, up to a constant multiplicative factor depending on C we see this trivially integrates to give the desired result noting that $\exp((2m-N)^2/(2N)) = \exp(y^2/2)$.

Proof of Lemma 4.9. Let $N = n_1 n_2$ and $m = \sum_{i \in [n_1]} s_i$. Applying Theorem 4.2 and Lemma 4.10, similar to the proof of Lemma 4.8 we have $m = N/2 + O_C(N^{3/4})$ and find

On the other hand, the definition of $\mathcal{I}_{1/2}^{n_1,n_2}$ demonstrates

$$\frac{\mathbb{P}_{\mathcal{I}_{1/2}^{n_1,n_2}}[\mathbf{S}=\vec{s}\wedge\mathbf{T}=\vec{t}]}{\prod_{i=1}^{n_1}\binom{n_2}{s_i}\prod_{i=1}^{n_2}\binom{n_1}{t_i}} \approx \int_0^1 \left(\frac{N}{\pi(1/2)^2}\right)^{1/2} \exp\left(-\frac{N}{(1/2)^2}(p'-1/2)^2\right) \frac{(p')^{2m}(1-p')^{2(N-m)}}{\mathbb{P}[\operatorname{Bin}(N,p')=\operatorname{Bin}(N,p')]} dp \\ \approx N \int_0^1 \exp\left(-\frac{N}{(1/2)^2}(p'-1/2)^2\right) (p')^{2m}(1-p')^{2(N-m)} dp', \tag{4.4}$$

where the second line uses that $\mathbb{P}[\operatorname{Bin}(N, p') = \operatorname{Bin}(N, p')] = \Theta(1/\sqrt{N})$ for $p' \in [1/3, 2/3]$ and that the integrand is so small when $p \notin [1/3, 2/3]$ that it contributes only lower order terms. Therefore it suffices to prove that $4^{-N} \exp((2m - N)^2/(2N))$ and

$$T(m) = \int_0^1 \left(\frac{N}{\pi(1/2)^2}\right)^{1/2} \exp\left(-\frac{N}{(1/2)^2}(p'-1/2)^2\right) (p')^{2m}(1-p')^{2(N-m)}dp'$$

are within constants. We now proceed in an identical manner to the proof of Lemma 4.8. Let $2m = N + y\sqrt{N}$, $p' = 1/2 + z/(2\sqrt{N})$ and note that

$$T(m) = \frac{1}{\sqrt{\pi}} \int_{-\sqrt{N}}^{\sqrt{N}} t(z) dz$$

where

$$t(z) = \exp\left(-z^2 + (N + y\sqrt{N})\log(1/2 + z/(2\sqrt{N})) + (N - y\sqrt{N})\log(1/2 - z/(2\sqrt{N}))\right).$$

This is the same situation as the end of the proof of Lemma 4.8, and we finish in the same way. \Box

5. Eliminating Atypical Degree Sequences

In order to apply Lemmas 4.8 and 4.9 we need to eliminate degree sequences which are too atypical. This involves eliminating graphs where the largest degrees are $n^{1/2+\eta}$ bigger than the smallest and graphs for which there exists a subset S such that $\sum_{v \in V(G)} |\deg(v, S) - |S|/2|^2 \ge Cn^2$ for C sufficiently large.

To eliminate high and low degrees we use switchings in the style of [32, Section 4]; however the work of [32, Section 4] uses that the class of graphs considered is downwards-closed. In our model, however, there are constraints which are "two-sided" for various vertices and therefore our analysis is more delicate. The crucial observation is that conditional on the graph not containing a very dense subgraph, only an $n^{-1/2+\eta}$ fraction of vertices can be critical with respect to the constraints (see Lemma 5.9). This itself is proved via switchings, but under this assumption a straightforward argument proves the necessary degree bounds.

The second condition is substantially easier to guarantee and only requires noting that for sufficiently large C the above occurs with (very) exponentially small probability. The necessary statement essentially appears in [15, Lemma 4.11] but with an extra disjointness condition that we easily remove.

5.1. Eliminating highly irregular subgraphs. We first show that it is typical for the degree sequence of $\mathbb{G}(n, 1/2)$ and all its dense subgraphs in an appropriate sense to have variance of order n. This will hold with sufficiently high probability that it will be robust to conditioning even on very unlikely events.

We will require a version of the classical Bernstein inequality.

Theorem 5.1 ([45, Theorem 2.8.1]). For a random variable X define the ψ_1 -norm

$$||X||_{\psi_1} = \inf\{t > 0 \colon \mathbb{E}[\exp(|X|/t)] \le 2\}.$$

There is an absolute constant $c = c_{5,1} > 0$ such that the following holds. If X_1, \ldots, X_N are independent random variables then

$$\mathbb{P}\left[\left|\sum_{i=1}^{N} X_{i}\right| \ge t\right] \le 2\exp\left(-c\min\left(\frac{t^{2}}{\sum_{i=1}^{N} \|X_{i}\|_{\psi_{1}}^{2}}, \frac{t}{\max_{i} \|X_{i}\|_{\psi_{1}}}\right)\right)$$

for all $t \geq 0$.

Lemma 5.2. Given c, K > 0, there is $C = C_{5,2}(K) > 0$ so the following holds. Let $G \sim \mathbb{G}(2n, 1/2)$ and let \mathcal{E}_{big} be the event that for some $S \subseteq [2n]$, we have

$$\sum_{v \in [2n]} (\deg(v, S) - |S|/2)^2 \ge Cn^2.$$

Then $\mathbb{P}[\mathcal{E}_{\text{big}}] \leq \exp(-Kn)$ for n large.

Proof. Suppose there is some S with

$$\sum_{v \in [2n]} (\deg(v, S) - |S|/2)^2 \ge Cn^2,$$

with C large to be chosen later. Let T be a uniformly random subset of S and $U = S \setminus T$. Let us denote by \mathbb{E}_T the result of averaging only over the randomness of T. Then given some $v \in [2n]$,

$$\mathbb{E}_{T}[\mathbb{1}_{v \notin T}(\deg(v, T) - |T|/2)^{2} + \mathbb{1}_{v \notin U}(\deg(v, U) - |U|/2)^{2}]$$

$$\geq \frac{1}{2}\mathbb{E}_{T}[(\deg(v, T) - |T|/2)^{2} + (\deg(v, U) - |U|/2)^{2}]$$

$$\geq \frac{1}{4}\mathbb{E}_{T}[(\deg(v, S) - |S|/2)^{2}]$$

where the first line comes from the fact that $\deg(v, U)$, $\deg(v, T)$ are determined purely by the information of $T \cap (S \setminus \{v\})$, and that $v \notin T$ and $v \notin U$ then conditionally occur with probability at least 1/2, and the second line follows from Cauchy–Schwarz and $S = T \cup U$ with T, U disjoint. Summing over v, we find that

$$\mathbb{E}_T \left[\sum_{v \in [2n]} \left(\mathbb{1}_{v \notin T} (\deg(v, T) - |T|/2)^2 + \mathbb{1}_{v \notin U} (\deg(v, U) - |U|/2)^2 \right) \right] \ge Cn^2/4$$

hence there is some $T \subseteq S$ such that $\sum_{v \in [2n] \setminus T} (\deg(v, T) - |T|/2)^2 \ge Cn^2/8$. That is, if \mathcal{E}_{big} holds then one of 2^{2n} possible T satisfy this inequality. Therefore, by the union bound and adjusting the value of K appropriately, it suffices to understand for a fixed T the chances of $\sum_{v \in [2n] \setminus T} (2 \deg(v, T) - |T|)^2 \ge Cn^2/2$ occurring.

Now note that over the randomness of $\mathbb{G}(n, 1/2)$, each $(2 \deg(v, T) - |T|)^2$ is distributed as X^2 where X is the sum of |T| independent Rademacher variables (uniform on $\{\pm 1\}$). Furthermore, as we vary $v \in [2n] \setminus T$, these variables are independent. We easily check that $||X^2||_{\psi_1} = O(|T|) = O(n)$ and hence by Theorem 5.1 we have

$$\mathbb{P}\bigg[\sum_{v\in[2n]\backslash T} (2\deg(v,T)-|T|)^2 \ge t\bigg] \le 2\exp\big(-\Omega(\min\{t^2/n^3,t/n\})\big).$$

Plugging in $t = Cn^2/2$ and choosing C sufficiently large compared to K, the result follows.

Next we show friendly partitions are not super-exponentially unlikely (in particular, although conditioning on having two fixed partitions be friendly will tilt the probability space, the event \mathcal{E}_{var} from Lemma 5.2 with appropriately chosen parameters will still hold).

Lemma 5.3. Given c > 0 there is $C = C_{5,3}(c) > 0$ such that the following holds for all $|\gamma| \leq 1$. Fix a pair of equipartitions of $[2n] = A_1 \cup A_2 = B_1 \cup B_2$ such that $cn \leq |A_1 \cap B_1| \leq (1-c)n$. Sample $G \sim \mathbb{G}(2n, 1/2)$ and let \mathcal{E}_2 be the event that $\deg(v, A_i) \geq \deg(v, A_{i+1}) + \gamma\sqrt{n}$ for all $v \in A_i$ and $\deg(v, B_i) \geq \deg(v, B_{i+1}) + \gamma\sqrt{n}$ for all $v \in B_i$. Then for n large,

$$\mathbb{P}[\mathcal{E}_2] \ge \exp(-Cn).$$

Proof. Let us consider the event \mathcal{E} that for every $i, j \in \{1, 2\}$ and $v \in A_i \cap B_j$, we have $\deg(v, A_i \cap B_j) - |A_i \cap B_j|/2 \in [4\sqrt{n}, 8\sqrt{n}]$ and further for every $(i', j') \neq (i, j)$ we have $\deg(v, A_{i'} \cap B_{j'}) - |A_{i'} \cap B_{j'}|/2 \in [-\sqrt{n}, \sqrt{n}]$. It is easy to check that \mathcal{E}_2 is satisfied in such a circumstance. We now compute a lower bound for $\mathbb{P}[\mathcal{E}]$ using Theorems 4.1 and 4.2.

Notice that there are 4 degree sequences for each $G[A_i \cap B_j]$, call them $(d_v^{(i,j)})_{v \in A_i \cap B_j}$ and 12 degree sequences for each $G[A_i \cap B_j, A_{i'} \cap B_{j'}]$, call them $(d_v^{(i,j,i',j')})_{v \in A_i \cap B_j}$ (switching the roles of (i,j), (i',j') gives the degrees on both sides of a given pair of parts). The event \mathcal{E} is merely a system

of constraints on the elements of these degree sequences. Let us consider any sequence of values satisfying the constraints given by \mathcal{E} and such that the two following additional properties hold:

•
$$2|\sum_{v \in A_i \cap B_i} d_v^{(i,j)}$$
 for all $i, j \in \{1, 2\};$

•
$$\sum_{v \in A_i \cap B_j} d_v^{(i,j,i',j')} = \sum_{v \in A_{i'} \cap B_{j'}} d_v^{(i',j',i,j)}$$
 for all $i, j, i', j' \in \{1,2\}$ with $(i,j) \neq (i',j')$.

We easily see that there are at least $(\Omega(\sqrt{n}))^{8n}$ many choices of degree sequences with this property: each individual value deg $(v, A_i \cap B_j)$ for $v \in [2n]$ and $i, j \in \{1, 2\}$ has at least \sqrt{n} choices, and it is easy to see that at least a $\exp(-n)$ fraction of all these choices satisfy the above constraints. To solve the parity constraints, it is easy to see around half of the sequences $d_v^{(i,j)}$ work for each i, j. To solve the equality constraints one can check that most pairs of sequences $(d_v^{(i,j,i',j')})_{v \in A_i \cap B_j}$ and $(d_v^{(i',j',i,j)})_{v \in A_{i'} \cap B_{j'}}$ have sums differing by say $O(n \log n)$ due to variance considerations, and then modifying $O(\sqrt{n}(\log n)^2)$ values in both sequences by approximately $\sqrt{n}/\log n$ each will allow for the appropriate balancing. The fraction of possible sequences attained from this procedure is at least $(1/\sqrt{n})^{O(\sqrt{n}(\log n)^2)} \ge \exp(-n)$.

Furthermore, Theorems 4.1 and 4.2 apply to these degree sequences. If we let $n_{ij} = |A_i \cap B_j|$ and $2m_{ij} = \sum_{v \in A_i \cap B_j} d_v^{(i,j)}$, then Theorem 4.1 applied to $(d_v^{(i,j)})_{v \in A_i \cap B_j}$ leads to $\mu = 1/2 + O(1/\sqrt{n})$, $\gamma_2^2 = \Theta(1)$, and therefore on the order of at least

$$\gtrsim \binom{n_{ij}(n_{ij}-1)/2}{m_{ij}} \binom{n_{ij}(n_{ij}-1)}{2m_{ij}}^{-1} \prod_{v \in A_i \cap B_j} \binom{n_{ij}-1}{d_v^{(i,j)}}.$$

One can check that each binomial coefficient in the product is $\Omega(2^{n_{ij}-1}/\sqrt{n})$, while the two initial binomial coefficients when divided contribute at least $\exp(-Kn)2^{-n_{ij}(n_{ij}-1)/2}$ for some constant K depending only on c. Multiplying we have a contribution of at least say

$$\exp(-2Kn) \cdot 2^{\binom{n_{ij}}{2}} (1/\sqrt{n})^{n_{ij}}$$

many possibilities for the graph $G[A_i \cap B_j]$ given this degree sequence.

Similarly, the number of choices for $G[A_i \cap B_j, A_{i'} \cap B_{j'}]$ can be seen to be at least say

$$\exp(-2Kn) \cdot 2^{n_{ij}n_{i'j'}} (1/\sqrt{n})^{n_{ij}+n_{i'j'}}$$

Multiplying over 4 graphs $G[A_i \cap B_j]$ and 6 graphs $G[A_i \cap B_j, A_{i'} \cap B_{j'}]$, this becomes at least $\exp(-20Kn)2^{\binom{2n}{2}}(1/\sqrt{n})^{8n}$ total. Multiplying by the number of choices of degree sequences from earlier, we have at least say $\exp(-40Kn)2^{\binom{2n}{2}}$ total choices of realizations $G \sim \mathbb{G}(2n, 1/2)$ satisfying the desired property \mathcal{E} (and hence \mathcal{E}_2), if K was chosen sufficiently large. The desired result immediately follows taking C = 40K, which depends only on c.

As an immediate consequence we have the following proposition.

Proposition 5.4. Given the setup and notation of Lemma 5.3, there is $K = K_{5.4}(c) > 0$ such that the following holds. Let $G \sim \mathbb{G}(2n, 1/2)$ and let \mathcal{E}_{var} be the event that

$$\sum_{v \in [2n]} (\deg(v, S) - |S|/2)^2 \le Kn^2$$

for all $S \subseteq [2n]$. Then $\mathbb{P}[\mathcal{E}_{var}|\mathcal{E}_2] \ge 1 - \exp(-n)$ for n large.

Proof. Let $C = C_{5,3}(c)$ and apply Lemma 5.2 with K replaced by $1+2C_{5,3}(c)$. We have $\mathbb{P}[\mathcal{E}_{\text{big}}|\mathcal{E}_2] \leq \mathbb{P}[\mathcal{E}_{\text{big}}]/\mathbb{P}[\mathcal{E}_2] \leq \exp(-(2C+1)n)/\exp(-Cn) \leq \exp(-n)$, and the result follows. \Box

5.2. Eliminating high degree vertices. We require a tail bound of the binomial random model based on the first value in the tail, and a generalization of this estimate to arbitrary upwards-closed families of sets.

Lemma 5.5. For $0 \le \ell \le n$ we have that

$$\frac{\binom{n}{\ell}}{\sum_{k=\ell}^{n}\binom{n}{k}} \lesssim \frac{|\ell - n/2|}{n} + \frac{1}{\sqrt{n}}.$$

Proof. For $\ell \leq n/2 + \sqrt{n}$, we have that $\binom{n}{\ell} \leq 2^n/\sqrt{n}$ and $\sum_{k=\ell}^n \binom{n}{k} \geq 2^n$ and hence the result follows. For $|\ell - n/2| \geq n/5$ the result is trivial. Finally when $n/2 + \sqrt{n} \leq \ell \leq n/2 + n/5$ we have that

$$\frac{\sum_{k=\ell}^{n} \binom{n}{k}}{\binom{n}{\ell}} \geq \sum_{t=0}^{\sqrt{n}} \binom{n}{\ell+t} \binom{n}{\ell}^{-1} \\
\geq \sum_{t=0}^{\sqrt{n}} ((\ell+t)/(n-\ell-t+1))^{-t} \\
\gtrsim \sum_{t=0}^{\sqrt{n}} (1+16|\ell-n/2|/n)^{-t} \gtrsim n/|\ell-n/2|.$$

Lemma 5.6. Let \mathcal{F} be an upwards-closed family of subsets of [n], i.e., if $F_1 \in \mathcal{F}$ and $F_2 \supseteq F_1$ then $F_2 \in \mathcal{F}$. Let \mathcal{F}_{ℓ} be the elements of \mathcal{F} of size ℓ . Then

$$\frac{|\mathcal{F}_{\ell}|}{|\mathcal{F}_{\ell}| + |\mathcal{F}_{\ell+1}| + \dots + |\mathcal{F}_{n}|} \leq \frac{\binom{n}{\ell}}{\sum_{i=\ell}^{n} \binom{n}{i}} \lesssim \frac{|\ell - n/2|}{n} + \frac{1}{\sqrt{n}}.$$

Proof. Every size ℓ set within \mathcal{F} has exactly $\binom{n-\ell}{s-\ell}$ size s sets in \mathcal{F} containing it by the upwardsclosed condition. Furthermore, every size s set within \mathcal{F} has at most $\binom{s}{\ell}$ size ℓ sets in \mathcal{F} contained within it. Therefore, $\binom{n-\ell}{s-\ell}|\mathcal{F}_{\ell}| \leq \binom{s}{\ell}|\mathcal{F}_{s}|$ for all $\ell \leq s \leq n$. The result easily follows noting that $\binom{n-\ell}{s-\ell}/\binom{s}{\ell} = \binom{n}{s}/\binom{n}{\ell}$ and then using Lemma 5.5.

Now we prove various structural properties hold whp for friendly partitions in $\mathbb{G}(2n, 1/2)$.

Proposition 5.7. Fix $\eta > 0$ and consider $|\gamma| \leq 1$. Fix an equipartitions $[2n] = A_1 \cup A_2$. Sample $G \sim \mathbb{G}(2n, 1/2)$ and let \mathcal{E}_1 be the event that $\deg(v, A_i) \geq \deg(v, A_{i+1}) + \gamma \sqrt{n}$ for all $v \in A_i$. Let $\mathcal{E}_{irreg,1}$ denote the event that there exists a vertex $v \in [2n]$ and $i \in \{1, 2\}$ such that $|\deg(v, A_i) - n| \geq n^{1/2+\eta}$. Then for n large,

$$\mathbb{P}[\mathcal{E}_{\text{irreg},1}|\mathcal{E}_1] \le n^{-\omega(1)}.$$

Proposition 5.8. Fix $c, \eta > 0$ and consider $|\gamma| \leq 1$. Fix a pair of equipartitions of $[2n] = A_1 \cup A_2 = B_1 \cup B_2$ such that $cn \leq |A_1 \cap B_1| \leq (1-c)n$. Sample $G \sim \mathbb{G}(2n, 1/2)$ and let \mathcal{E}_2 be the event that $\deg(v, A_i) \geq \deg(v, A_{i+1}) + \gamma\sqrt{n}$ for all $v \in A_i$ and $\deg(v, B_i) \geq \deg(v, B_{i+1}) + \gamma\sqrt{n}$ for all $v \in B_i$. Let $\mathcal{E}_{irreg,2}$ denote the event that there exists a vertex $v \in [2n]$ and $i, j \in \{1, 2\}$ such that $|\deg(v, A_i \cap B_j) - |A_i \cap B_j|/2| \geq n^{1/2+\eta}$. Then for n large,

$$\mathbb{P}[\mathcal{E}_{\mathrm{irreg},2}|\mathcal{E}_2] \le n^{-\omega(1)}$$

To prove these propositions, we need the following estimate on "critical" vertices in friendly partitions.

Lemma 5.9. Fix c > 0 and a positive integer k, and consider $|\gamma| \leq 1$ with $\gamma\sqrt{n} \in \mathbb{Z}$. Fix a pair of equipartitions of $[2n] = A_1 \cup A_2 = B_1 \cup B_2$ such that $cn \leq |A_1 \cap B_1| \leq (1-c)n$. Sample $G \sim \mathbb{G}(2n, 1/2)$ and let \mathcal{E}_2 be the event that $\deg(v, A_i) \geq \deg(v, A_{i+1}) + \gamma\sqrt{n}$ for all $v \in A_i$ and

 $\deg(v, B_i) \geq \deg(v, B_{i+1}) + \gamma \sqrt{n}$ for all $v \in B_i$. Let \mathcal{E}_1 be the event that this holds just for A. Define a vertex to be critical for A if $v \in A_i$ and $\deg(v, A_i) = \deg(v, A_{i+1}) + \gamma \sqrt{n}$ and let the number of critical for A vertices be X. Then

$$\mathbb{E}[X^k | \mathcal{E}_b] \lesssim_{c,k} n^{k/2} \quad \text{for } b \in \{1,2\}.$$

Proof. Let Y be the number of critical for A vertices inside $A_1 \cap B_1$. We will show $\mathbb{E}[Y^k | \mathcal{E}_2] \leq_{c,k} n^{k/2}$. By symmetry we will find the same holds for each $A_i \cap B_j$, which will imply the result for b = 2. The case for b = 1 is similar so we forgo the details.

Let us further reveal the values $\deg(v, A_i \cap B_j)$ for all $i, j \in \{1, 2\}$ and $v \in V(G) \setminus (A_i \cap B_j)$. Thus the remaining randomness is over $\mathbb{G}(A_i \cap B_j, 1/2)$ for each $i, j \in \{1, 2\}$, and we condition these graphs on certain degree inequalities. Notice that the four parts are independent since we have revealed the information of the degrees between the parts. Whether a vertex is critical thus also only depends on its degree within a part. Let us focus on the part $A_1 \cap B_1$, which will allow us to control Y. The graph $G[A_i \cap B_j]$ is a uniform random graph satisfying various lower bounds on the degrees of its vertices. For each vertex, given the revealed information of the degrees to the outside, the conditions $\deg(v, A_1) \geq \deg(v, A_2) + \gamma \sqrt{n}$ and $\deg(v, B_1) \geq \deg(v, B_2) + \gamma \sqrt{n}$ lead to precisely these conditions on $\deg(v, A_i \cap B_j)$. If the latter inequality provides a sharper inequality than the former, then v will not be critical for A. If the former inequality is sharper, then it can be critical for A precisely when it equals the minimum allowed value.

Let \mathcal{G} be the information of all edges in G outside of $A_1 \cap B_1$. We have by linearity of expectation that

$$\mathbb{E}[Y(Y-1)\cdots(Y-k+1)|\mathcal{E}_2,\mathcal{G}] \leq \sum_{\substack{v_1,\dots,v_k \in A_1 \cap B_1 \\ \text{distinct}}} \mathbb{P}[v_1,\dots,v_k \text{ are critical for } A|\mathcal{E}_2,\mathcal{G}].$$
(5.1)

Now given a choice of v_1, \ldots, v_k , let us further reveal everything in the graph $G[A_1 \cap B_1]$ except the edges between $\{v_1, \ldots, v_k\}$ and $A_1 \cap B_1$ and furthermore reveal the edges internal to $\{v_1, \ldots, v_k\}$. The remaining random portion is a uniformly random bipartite graph between $\{v_1, \ldots, v_k\}$ and $(A_1 \cap B_1) \setminus \{v_1, \ldots, v_k\}$ subject to certain lower bounds for degrees on both sides (coming from our original inequalities). We now consider the maximum probability that v_1 is critical given the neighborhoods of v_2, \ldots, v_k . Let us think of the possible neighborhoods of v_1 (disregarding the revealed part among $\{v_2, \ldots, v_k\}$ as subsets of $(A_1 \cap B_1) \setminus \{v_1, \ldots, v_k\}$. Since the conditioned information only gives lower bounds on certain degrees (which may or may not force certain vertices to connect to v_1 to meet their "quota"), we see that the family \mathcal{F} of possibilities is upwards-closed. By Lemma 5.6 we find that the resulting probability that v is critical is at most

$$O\left(\frac{|\ell - (|A_1 \cap B_1| - k)/2|}{cn} + \frac{1}{\sqrt{cn}}\right)$$

where ℓ is the lower bound on the size of $N(v) \cap ((A_1 \cap B_1) \setminus \{v_1, \ldots, v_k\})$ coming from the condition $\deg(v, A_1) \geq \deg(v, A_2) + \lceil \gamma \sqrt{n} \rceil$. That is,

$$\ell = \deg(v_1, A_2) - \deg(v_1, A_1 \cap B_2) - \deg(v_1, \{v_2, \dots, v_k\}) + \lceil \gamma \sqrt{n} \rceil.$$

Since k is constant, the bound becomes

$$O\left(\frac{|\deg(v_1, A_2) - \deg(v_1, A_1 \cap B_2) - |A_1 \cap B_1|/2|}{n} + \frac{1}{\sqrt{n}}\right).$$

Multiply over v_1, \ldots, v_k to obtain a bound on the probability that all of v_1, \ldots, v_k are critical (note that the above is a bound on v_1 even conditioned on any outcomes for the other vertices, so this is valid). Plugging into (5.1), we find

$$\mathbb{E}[Y(Y-1)\cdots(Y-k+1)|\mathcal{E}_2,\mathcal{G}] \lesssim \left(\sum_{v \in A_1 \cap B_1} \frac{|\deg(v,A_2) - \deg(v,A_1 \cap B_2) - |A_1 \cap B_1|/2|}{n} + \frac{1}{\sqrt{n}}\right)^k$$

$$\lesssim n^{k/2} + n^{-k} \left(\sum_{v \in A_1 \cap B_1} |\deg(v, A_2) - \deg(v, A_1 \cap B_2) - |A_1 \cap B_1|/2| \right)^k$$

$$\lesssim n^{k/2} + n^{-k/2} \left(\sum_{v \in A_1 \cap B_1} |\deg(v, A_2) - \deg(v, A_1 \cap B_2) - |A_1 \cap B_1|/2|^2 \right)^{k/2}.$$

Now by Proposition 5.4, there is some appropriately large constant A so that conditional on \mathcal{E}_2 , with probability at least $1 - \exp(-n)$ we have $\sum_{v \in [2n]} (\deg(v, S) - |S|/2)^2 \leq An^2$ for all $S \subseteq [2n]$ (event \mathcal{E}_{var}). In particular, note that

$$\sum_{v \in A_1 \cap B_1} |\deg(v, A_2) - \deg(v, A_1 \cap B_2) - |A_1 \cap B_1|/2|^2 \le 4An^2$$
(5.2)

holds under the event \mathcal{E}_{var} . Furthermore, the truth of this particular inequality is a function only of the information of \mathcal{G} . Therefore, we easily see

$$\mathbb{E}[Y(Y-1)\cdots(Y-k+1)|\mathcal{E}_2] \lesssim (n^{k/2} + n^{-k/2}(4An^2)^{k/2}) + \exp(-n) \cdot n^k \lesssim n^{k/2}$$

by considering whether \mathcal{G} satisfies (5.2) or not. The desired estimate follows immediately.

Now we use switchings to prove Propositions 5.7 and 5.8.

Proof of Propositions 5.7 and 5.8. We focus on Proposition 5.8; the case for Proposition 5.7 is analogous but simpler, so we truncate the details. Furthermore, it suffices to prove the estimate for the probability for a single choice of $i^*, j^* \in \{1, 2\}$ and $v \in V(G)$. Let us assume $v \in A_1 \cap B_1$ without loss of generality. We wish to understand $\mathbb{P}[|\deg(v, A_{i^*} \cap B_{j^*}) - |A_{i^*} \cap B_{j^*}|/2| \ge n^{1/2+\eta}|\mathcal{E}_2].$

Let us reveal the identities of all edges in G except the potential edges between v and $V(G) \setminus \{v\}$. Call this information \mathcal{G} . Conditional on $\mathcal{E}_2, \mathcal{G}$, our distribution on the neighborhood of v is uniform over choices such that v satisfies its two degree inequalities and such that every other vertex satisfies its inequalities as well. But note that for the typical other vertex, whether v connects to it will not influence whether the inequalities are satisfied. In fact, this will only be the case for vertices that are ultimately critical (for A or B) in the graph G. By Lemma 5.9 with k growing sufficiently slowly as a function of n, we see that there are at most $n^{1/2+\eta/4}$ critical vertices for A with probability $1 - n^{-\omega(1)}$, and similar for critical vertices for B by symmetry. Thus, we may assume \mathcal{G} is such that the number of vertices with "forced" edges and non-edges to v is at most $n^{1/2+\eta/2}$ in total.

Now the neighborhood of v has the following distribution: there are at most $n^{1/2+\eta/2}$ total designated forced edges and non-edges to the four parts $A_i \cap B_j$, and then the remaining edges are uniform subject to the constraints $\deg(v, A_1) - \deg(v, A_2) \geq \gamma \sqrt{n}$ and $\deg(v, B_1) - \deg(v, B_2) \geq \gamma \sqrt{n}$. We wish to understand the chance that $|\deg(v, A_{i^*} \cap B_{j^*}) - |A_{i^*} \cap B_{j^*}|/2| \geq n^{1/2+\eta}$, call this event \mathcal{E}_{bad} , in this model. Let C be the set of vertices which are forced to either connect or be disconnected from v, with C_1 connected and C_0 disconnected. Heuristically, we are close to a binomial random situation with some "planted" deviations coming from C on the order of $n^{1/2+\eta/2}$, so the chance that some degree deviates by $n^{1/2+\eta}$ ought to be extremely unlikely.

We make this heuristic precise in order to complete the proof. Having revealed the above information, let us consider the distribution Δ of sampling the neighborhood of v within $V(G) \setminus (C \cup \{v\})$ uniformly at random. The above analysis shows it is enough to give an upper bound on $\mathbb{P}[\mathcal{E}_{\text{bad}} \cap \mathcal{E}_2]/\mathbb{P}[\mathcal{E}_2]$ in this model (given the revealed information including C_0, C_1).

First, we claim $\mathbb{P}[\mathcal{E}_2] \geq \exp(-n^{3\eta/2})$. Indeed, choose positive integers n_{ij} for $i, j \in \{1, 2\}$ such that $n_{11} + n_{12} - n_{21} - n_{22} \geq \gamma \sqrt{n} + n^{1/2+\eta/2}$ and $n_{11} + n_{21} - n_{12} - n_{22} \geq \gamma \sqrt{n} + n^{1/2+\eta/2}$ and $|n_{ij} - |A_i \cap B_j|/2| \leq 4n^{1/2+\eta/2}$ for $i, j \in \{1, 2\}$. This is easily seen to be possible since $|A_i \cap B_j| \in [cn, (1-c)n]$ and $|A_i| = |B_j| = n$. Then the probability that v has n_{ij} neighbors to $(A_i \cap B_j) \setminus C$ for each $i, j \in \{1, 2\}$ in the distribution Δ is easily seen to be at least $\exp(-O((n^{1/2+\eta/2})^2/n)) \geq 1$

 $\exp(-n^{3\eta/2})$ for n large. Here we are using the easy estimate $\binom{x}{x/2+t}/2^x = \exp(-\Theta(t^2/x))$ for $t \in [-x/2, x/2]$ satisfying $|t| \ge x^{1/2+\eta/4}$.

Similarly, we see that $\mathbb{P}[\mathcal{E}_{bad}] \leq \exp(-\Omega(n^{2\eta}))$ since if \mathcal{E}_{bad} holds then $\deg(v, (A_i \cap B_j) \setminus C) \geq$ $|(A_i \cap B_i) \setminus C|/2 + n^{1/2+\eta}/2$ is easily verified, and under the independent distribution Δ this can be bounded as above. Dividing, we obtain

$$\mathbb{P}[\mathcal{E}_{\text{bad}} \cap \mathcal{E}_2] / \mathbb{P}[\mathcal{E}_2] \le n^{-\omega(1)}$$

and combining with the earlier analysis we are done.

6. Convexity via the Prékopa-Leindler Inequality

In this section we collect a pair of log-convexity claims which are proved via observing that the marginals of a log-concave function are log-concave. This fact is a well known consequence of the Prékopa-Leindler inequality [30, 37]. We note here that the use of log-concavity of such Gaussian functionals also appears in the work of Gamarnik and Li [20, Lemma 4.6] (and for the similar purpose of simplifying an associated maximization problem arising in the second moment computation); we provide a short proof of the necessary results in this section.

Theorem 6.1. If f(x,y) is a log-concave function over $(x,y) \in \mathbb{R}^m \times \mathbb{R}^n$, then $g(x) = \int_{\mathbb{R}^n} f(x,y) dy$ is a log-concave function on \mathbb{R}^m .

For the first moment we will require the following log-concavity.

Lemma 6.2. The following function is log-concave:

$$\alpha \mapsto \mathbb{P}_{Z \sim \mathcal{N}(0,1)}[Z \ge (\gamma + \alpha)\sqrt{2}].$$

Proof. Via shifting and rescaling it suffices to prove that

$$\alpha \mapsto \mathbb{P}_{Z \sim \mathcal{N}(0,1)}[Z \ge \alpha]$$

is log-concave. Note that

$$\mathbb{P}_{Z \sim \mathcal{N}(0,1)}[Z \ge \alpha] = \int_0^\infty \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(x+\alpha)^2}{2}\right) dx$$

and that

$$\mathbb{1}_{x \ge 0} \exp\left(\frac{-(x+\alpha)^2}{2}\right)$$

is a log-concave function of x and α . Theorem 6.1 finishes the proof, taking the marginal in x.

For the second moment we will repeatedly used the following generalization of Lemma 6.2 (which can be seen as the special case $\beta = 1/2$).

Lemma 6.3. Fix $\beta \in [0, 1]$. The function

$$(\alpha_1, \alpha_2) \mapsto \mathbb{P}_{Z_1, Z_2 \sim \mathcal{N}(0, 1)} \left[\sqrt{\frac{\beta}{2}} Z_1 + \sqrt{\frac{1 - \beta}{2}} Z_2 \ge (\gamma + \alpha_1) \sqrt{2} \wedge \sqrt{\frac{\beta}{2}} Z_1 - \sqrt{\frac{1 - \beta}{2}} Z_2 \ge (\gamma + \alpha_2) \sqrt{2} \right]$$

is log-concave.

Proof. Via shifting and applying a linear transformation it suffices to prove that

$$(\alpha_1, \alpha_2) \mapsto \mathbb{P}_{Z_1, Z_2 \sim \mathcal{N}(0, 1)} \left[\sqrt{\frac{\beta}{2}} (Z_1 - \alpha_1) + \sqrt{\frac{1 - \beta}{2}} (Z_2 - \alpha_2) \ge 0 \land \sqrt{\frac{\beta}{2}} (Z_1 - \alpha_1) - \sqrt{\frac{1 - \beta}{2}} (Z_2 - \alpha_2) \ge 0 \right]$$

is log-concave. Note that

$$\mathbb{P}_{Z_1, Z_2 \sim \mathcal{N}(0, 1)} \left[\sqrt{\frac{\beta}{2}} (Z_1 - \alpha_1) + \sqrt{\frac{1 - \beta}{2}} (Z_2 - \alpha_2) \ge 0 \land \sqrt{\frac{\beta}{2}} (Z_1 - \alpha_1) - \sqrt{\frac{1 - \beta}{2}} (Z_2 - \alpha_2) \ge 0 \right]_{23}$$

$$= \int_{\sqrt{\frac{\beta}{2}}z_1 + \sqrt{\frac{1-\beta}{2}}z_2 \ge 0} \frac{1}{2\pi} \exp\left(\frac{-(z_1 + \alpha_1)^2 - (z_2 + \alpha_2)^2}{2}\right) dz_1 dz_2.$$
$$\sqrt{\frac{\beta}{2}}z_1 - \sqrt{\frac{1-\beta}{2}}z_2 \ge 0$$

Applying Theorem 6.1 to the log-concave function of $z_1, z_2, \alpha_1, \alpha_2$ given by

$$\mathbb{1}_{\sqrt{\frac{\beta}{2}}z_{1}+\sqrt{\frac{1-\beta}{2}}z_{2}\geq0}\mathbb{1}_{\sqrt{\frac{\beta}{2}}z_{1}-\sqrt{\frac{1-\beta}{2}}z_{2}\geq0}\exp\left(\frac{-(z_{1}+\alpha_{1})^{2}-(z_{2}+\alpha_{2})^{2}}{2}\right)$$

and taking the marginal in z_1, z_2 , the result follows.

7. Moment Computations

Now we are ready to attack the moment computation. Recall $|\gamma| \leq 1$ is such that $\gamma \sqrt{n} \in \mathbb{Z}$.

7.1. The second moment of X_{γ} . We first perform the second moment in detail, proving Lemma 3.3. We defer the much simpler first moment to Section 7.2, in which we truncate various repeated or similar details. Let $c \in (0, 1/2)$ be a small constant to be chosen later.

7.1.1. Setup and initial truncation. Consider $n'_1 \in \{0, \ldots, n\}$ and $n'_2 = n - n'_1$, and let $A_1 = [n]$ and $A_2 = \{n + 1, \ldots, 2n\}$ and $B_1 = [n'_1] \cup \{n + n'_1 + 1, \ldots, 2n\}$ and $B_2 = \{n'_1 + 1, n'_1 + 2, \ldots, n + n'_1\}$. Let $n_{ij} = |A_i \cap B_j|$, so that $n_{ij} = n'_{i+j-1}$ where we take indices (mod 2). Also, define $\beta = n_{11}/n$. Given this setup let $Y_A(n'_1) = 1$ if $G \sim \mathbb{G}(2n, 1/2)$ is such that $A_1 \cup A_2$ is a $\gamma \sqrt{n}$ -friendly partition, and 0 otherwise; let $Y_B(n'_1)$ be defined similarly. By linearity of expectation we have

$$\mathbb{E}X_{\gamma} = \binom{2n}{n} \mathbb{P}[Y_A(n_1') = 1], \tag{7.1}$$

$$\mathbb{E}X_{\gamma}^{2} = \binom{2n}{n} \sum_{n_{1}^{\prime}=0}^{n} \binom{n}{n_{1}^{\prime}}^{2} \mathbb{P}[Y_{A}(n_{1}^{\prime}) = 1 \land Y_{B}(n_{1}^{\prime}) = 1].$$
(7.2)

This means that as long as c is sufficiently smaller than η , we have

$$\binom{2n}{n} \sum_{n_1' < cn} \binom{n}{n_1'}^2 \mathbb{P}[Y_A(n_1') = 1 \land Y_B(n_1') = 1] + \binom{2n}{n} \sum_{n_1' > (1-c)n} \binom{n}{n_1'}^2 \mathbb{P}[Y_A(n_1') = 1 \land Y_B(n_1') = 1]$$

$$\leq 2 \sum_{n_1' < cn} \binom{n}{n_1'}^2 \binom{2n}{n} \mathbb{P}[Y_A(n_1') = 1]$$

$$\leq (1+\eta)^n \mathbb{E}X_{\gamma}$$
(7.3)

for n large. Thus it suffices to consider the terms for which $n'_1 \in [cn, (1-c)n]$ (equivalently, $\beta \in [c, 1-c]$), and we wish to compute $\mathbb{P}[Y_A(n'_1) = Y_B(n'_1) = 1]$. Note that whether the desired event holds or not is purely a function of the degree sequences of the 4 graphs $G[A_i \cap B_j]$ for $i, j \in \{1, 2\}$ and the 6 bipartite graphs $G[A_i \cap B_j, A_{i'} \cap B_{j'}]$ for $i, j, i', j' \in \{1, 2\}$ such that either i < i' or i = i' and j < j'. Thus, we wish to apply Lemmas 4.8 and 4.9 to transfer to a simple model for such degree sequences in which it is easier to compute the desired probabilities.

7.1.2. Transference and defining the integrated degree model. To do this, we use the results of Section 5. Let \mathcal{E}_2 be the event that $Y_A(n'_1) = Y_B(n'_1) = 1$ (which is in line with usage in Section 5). Let K be some large constant in terms of c to be determined. First apply Proposition 5.4. Let \mathcal{E}_{var} be the event that $\sum_{v \in [2n]} (\deg(v, S) - |S|/2)^2 \ge Kn^2$ for all $S \subseteq [2n]$, where K is some appropriate constant coming from the lemma. We have $\mathbb{P}[\mathcal{E}_{\text{var}}|\mathcal{E}_2] \ge 1 - \exp(-n)$. By Proposition 5.8 with η replaced by $\min(\varepsilon_{4.8}, \varepsilon_{4.9})/2$, we have $\mathbb{P}[\mathcal{E}_{\text{irreg},2}|\mathcal{E}_2] \le n^{-\omega(1)}$. Thus

$$\mathbb{P}[\mathcal{E}_2] = (1 \pm n^{-\omega(1)}) \mathbb{P}[\mathcal{E}_2 \wedge \mathcal{E}_{\mathrm{irreg},2}^c \wedge \mathcal{E}_{\mathrm{var}}].$$
(7.4)

Note that $\mathcal{E}_{irreg,2}^{c}$, \mathcal{E}_{var} ensure that for the 4 + 6 graphs defined above, they have degree sequences which satisfy A1 to A3 of Lemma 4.8 with appropriate parameters (namely, the degree sequences of $G[A_i \cap B_j]$ are $(O(1), \varepsilon_{4,8})$ -regular) or satisfy **B1** to **B5** of Lemma 4.9 with appropriate parameters (namely, the degree sequences of the 6 bipartite portions of G are $(O(1), \Omega(1), \varepsilon_{4,9})$ -regular). Here we are also using that $cn \leq n'_1, n'_2 \leq (1-c)n$, e.g., to establish **B1**.

We deduce

$$\mathbb{P}[\mathcal{E}_2 \wedge \mathcal{E}_{\mathrm{irreg},2}^c \wedge \mathcal{E}_{\mathrm{var}}] \lesssim \rho, \tag{7.5}$$

where ρ is the following probability:

- For $i, j \in \{1, 2\}$ sample the "degrees" $d_v^{(i,j)}$ for $v \in A_i \cap B_j$ (which should be thought of as corresponding to the internal degrees $\deg(v, A_i \cap B_j)$ for $v \in A_i \cap B_j$ from $\mathcal{I}_{1/2}^{|A_i \cap B_j|}$;
- For $i, j, i', j' \in \{1, 2\}$ with i < i' or i = i' and j < j', sample the "degrees" $d_v^{(i,j,i',j')}$ for $v \in A_i \cap B_j$ and $d_v^{(i',j',i,j)}$ for $v \in A_{i'} \cap B_{j'}$ (which correspond to the bipartite degrees $\deg(v, A_{i'} \cap B_{j'}) \text{ for } v \in A_i \cap B_j \text{ and vice versa}) \text{ jointly from } \mathcal{I}_{1/2}^{|A_i \cap B_j|, |A_{i'} \cap B_{j'}|};$
- Compute the probability that for each $i, j \in \{1, 2\}$ and $v \in A_i \cap B_j$ we have

$$d_v^{(i,j)} - d_v^{(i,j,i+1,j+1)} + d_v^{(i,j,i,j+1)} - d_v^{(i,j,i+1,j)} \ge \gamma \sqrt{n},$$
(7.6)

$$d_v^{(i,j)} - d_v^{(i,j,i+1,j+1)} - d_v^{(i,j,i,j+1)} + d_v^{(i,j,i+1,j)} \ge \gamma \sqrt{n},$$
(7.7)

where indices are taken $(\mod 2)$.

This can be interpreted as the probability that if the degrees of our 4+6 graphs are sampled as in Lemmas 4.8 and 4.9, we have that $A_1 \cup A_2$ and $B_1 \cup B_2$ are both $\gamma \sqrt{n}$ -friendly.

Understanding the "degree" distribution is relatively simple:

- Sample 4 + 6 parameters $p_{ij} \sim \mathcal{N}(1/2, 1/(4n_{ij}^2 4n_{ij}))$ for $i, j \in \{1, 2\}$ and $p_{ij,i'j'} \sim \mathcal{N}(1/2, 1/(8n_{ij}n_{i'j'}))$ for $i, j, i', j' \in \{1, 2\}$ with i < i' or i = i' and j < j' (all conditional on being in (0,1)), and let $p_{i'j',ij} = p_{ij,i'j'}$; • Sample $d_v^{(i,j)} \sim \text{Bin}(n_{ij} - 1, p_{ij})$ and $d_v^{(i,j,i',j')} \sim \text{Bin}(n_{i'j'}, p_{ij,i'j'})$ completely independently;
- Condition on the 4 divisibility conditions and 6 equality conditions

$$2 \mid \sum_{v \in A_i \cap B_i} d_v^{(i,j)},\tag{7.8}$$

$$\sum_{v \in A_i \cap B_j} d_v^{(i,j,i',j')} = \sum_{v \in A_{i'} \cap B_{j'}} d_v^{(i',j',i,j)}.$$
(7.9)

The first step can be handled by taking a Gaussian integral over the 4+6 parameters, and the second step is independent. Thus the major difficulty lies in the conditioning introduced by the third step. We will handle this by use of Bayes' theorem, local central limit theorems, and anticoncentration inequalities (see e.g. [41, Sections 2.3.2, 2.3.3]), but one should think of this more as a technical detail: the probability ρ ought to be mostly unchanged if this conditioning were not present, although one cannot actually directly drop the equal sum conditions.

7.1.3. Bayes' theorem in the transferred model. For simplicity, let us assume that any event involving Bin(n,p) for $p \notin [0,1]$ does not hold. Let $p_{ij} = 1/2 + \alpha_{ij}/\sqrt{4n_{ij}^2 - 4n_{ij}}$ and $p_{ij,i'j'} = 1/2 + \alpha_{ij}/\sqrt{4n_{ij}^2 - 4n_{ij}}$ $\alpha_{ij,i'j'}/\sqrt{8n_{ij}n_{i'j'}}$. Additionally, we will use the subscript \vec{d} to denote the randomness over the joint "degree" distribution, and the subscript $\vec{d^*}$ for this distribution prior to conditioning on (7.8) and (7.9) (which is therefore a collection of independent binomials if we reveal \vec{p}). Now we have

$$\rho = \mathbb{P}_{\vec{p},\vec{d}}[(7.6) \text{ and } (7.7)] = (1+o(1)) \int_{\alpha_{ij},\alpha_{ij,i'j'}} \mathbb{P}_{\vec{d}}[(7.6) \text{ and } (7.7)] \cdot \frac{e^{-\frac{1}{2}\sum \alpha_{ij}^2 - \frac{1}{2}\sum \alpha_{ij,i'j'}^2}}{(2\pi)^5} d\alpha$$

$$\lesssim e^{-32n} + \int_{\|\alpha\|_{\infty} \le 8\sqrt{n}} \mathbb{P}_{\vec{d}}[(7.6) \text{ and } (7.7)] e^{-\frac{1}{2}\sum \alpha_{ij}^{2} - \frac{1}{2}\sum \alpha_{ij,i'j'}^{2} d\alpha} \\ \lesssim e^{-32n} + \int_{\|\alpha\|_{\infty} \le 8\sqrt{n}} \mathbb{P}_{\vec{d}^{*}}[(7.6) \text{ and } (7.7)|(7.8) \text{ and } (7.9)] e^{-\frac{1}{2}\sum \alpha_{ij}^{2} - \frac{1}{2}\sum \alpha_{ij,i'j'}^{2} d\alpha} \\ \lesssim e^{-32n} + \int_{\|\alpha\|_{\infty} \le 8\sqrt{n}} \mathbb{P}_{\vec{d}^{*}}[(7.6) \text{ and } (7.7)] \frac{\mathbb{P}_{\vec{d}^{*}}[(7.8) \text{ and } (7.9)|(7.6) \text{ and } (7.7)]}{\mathbb{P}_{\vec{d}^{*}}[(7.8) \text{ and } (7.9)]} e^{-\frac{1}{2}\sum \alpha_{ij}^{2} - \frac{1}{2}\sum \alpha_{ij}^{2} - \frac{1}{2}\sum \alpha_{ij,i'j'}^{2} d\alpha}$$

The 1 + o(1) comes from the probability that the α are such that each $p_{ij} \in (0, 1)$, etc., and the e^{-32n} comes from the fact that there is a very low contribution to the integral when any coordinate of α is bigger than $8\sqrt{n}$ in magnitude. We remark that the sum $\sum \alpha_{ij,i'j'}^2$ is over pairs with i < i' or i = i' and j < j', and that furthermore the integration over $\alpha_{ij,i'j'}$ is only done over such pairs; recall we are letting $\alpha_{ij,i'j'} = \alpha_{i'j',ij}$.

7.1.4. Estimating the Bayes ratio. Now we estimate the ratio of probabilities in the final integrand using local central limit theorems. As we only care about ρ up to constant factors, we will be somewhat loose with details (see e.g. [41, Section 2] for a similar computation carried out with precision). The denominator is the chance that in the purely independent distribution (which we labeled by $\vec{d^*}$), we have the given divisibilities and equalities. For the denominator, note that each condition in (7.8) and (7.9) is independent of each other. Furthermore, each equality in (7.9) is tantamount to a sum of independent binomials being equal to another sum of independent binomials (and each divisibility is making sure a sum of independent binomials is even). We easily deduce that the denominator has order of magnitude $\Theta((1/2)^4(1/n)^6) = \Theta(n^{-6})$ by e.g. using a local central limit theorem (or explicit binomial calculations). Here we are implicitly using that $||\alpha||_{\infty}$ is not too large.

Now we focus on the numerator. It is similar to the denominator, but we first condition on the inequalities (7.6) and (7.7). At this stage it suffices to give an upper bound, so we focus just on bounding $\mathbb{P}[(7.9)|(7.6)$ and (7.7)].

Note that even after conditioning, the "degrees" $(d_v^{(i,j)}, d_v^{(i,j,i,j+1)}, d_v^{(i,j,i+1,j+1)}, d_v^{(i,j,i+1,j)})$ are independent over all $v \in A_i \cap B_j$ (and choices of $i, j \in \{1, 2\}$). Therefore, the resulting upper bound will follow from an Erdős–Littlewood–Offord-style argument. Explicitly, first reveal the "degrees" for all $v \in A_1 \cap B_1$. Then, reveal them for $v \in A_1 \cap B_2$ and consider the probability that (7.9) holds for i = 1, j = 1, i' = 1, j' = 2. Then, reveal them for $v \in A_2 \cap B_1$ and consider the probability that (7.9) holds for i' = 2, j' = 1 and $(i, j) \in \{(1, 1), (1, 2)\}$. Finally, reveal them for $v \in A_2 \cap B_2$ and consider the probability that (7.9) holds for i' = 2, j' = 1 and $(i, j) \in \{(1, 1), (1, 2)\}$. Finally, reveal them for $v \in A_2 \cap B_2$ and consider the probability that (7.9) holds for i' = 2, j' = 2 and $(i, j) \in \{(1, 1), (1, 2), (2, 1)\}$. We claim that the probabilities are, conditional on the prior revelations, bounded by $O(n^{-1}), O(n^{-2})$, and $O(n^{-3})$, respectively. We focus on the last claim: note that for each $v \in A_2 \cap B_2$ the distribution of $(d_v^{(2,2)}, d_v^{(2,2,1)}, d_v^{(2,2,1,1)}, d_v^{(2,2,1,2)})$ is that of

$$(\operatorname{Bin}(n_{22}-1, p_{22}), \operatorname{Bin}(n_{21}, p_{22,21}), \operatorname{Bin}(n_{11}, p_{22,11}), \operatorname{Bin}(n_{12}, p_{22,12}))$$

conditional on inequalities coming from (7.6) and (7.7). It is not hard to see that there is a 4dimensional axis-aligned lattice cube of dimensions $\Omega(\sqrt{n})$ such that every lattice point in the cube is obtained by this conditional distribution with probability $\Omega(n^{-4/2})$ (this follows almost immediately from the same property for the unconditioned binomial distribution). Projecting, this means that there is a 3-dimensional axis-aligned cube of dimensions $\Omega(\sqrt{n})$ such that every lattice point in the cube is obtained by the distribution $(d_v^{(2,2,2,1)}, d_v^{(2,2,1,2)})$ with probability $\Omega(n^{-3/2})$.

Now it suffices to show that the sum of $n_{22} = \Omega(n)$ independent copies of this distribution has the property that every point in \mathbb{R}^3 is obtained with probability $O(n^{-3})$. If we knew the distribution was in fact uniform on a 3-dimensional cube of dimensions $\Omega(\sqrt{n})$, the result would be immediate by considering the dimensions independently and some explicit computations (or the Lévy–Kolmogorov–Rogozin inequality [40]). We can reduce to this situation by the following process: after sampling $X = x \in \mathbb{R}^3$ to add to our sum, we sample a weighted coin which turns heads with probability $p_x \in [0, 1]$, choosing the weights p_x to ensure that conditional on heads appearing, the outcome of X is uniform on the aforementioned lattice cube. It is easy to check that we can further do this so that with probability $1 - \exp(-\Omega(n))$, there will be $\Omega(n)$ many heads appearing. Reveal which indices are heads and tails, and then notice that the sum corresponding to indices where heads appears is an independent sum of random variables uniform on lattice cubes.

With the claims in hand, the ultimate upshot is that

$$\frac{\mathbb{P}_{\vec{d^*}}[(7.8) \text{ and } (7.9)|(7.6) \text{ and } (7.7)]}{\mathbb{P}_{\vec{d^*}}[(7.8) \text{ and } (7.9)]} \le O(1).$$

7.1.5. The independent model. We now deduce

$$\rho \lesssim e^{-32n} + \int_{\|\alpha\|_{\infty} \leq 8\sqrt{n}} \mathbb{P}_{\vec{d^*}}[(7.6) \text{ and } (7.7)] e^{-\frac{1}{2}\sum \alpha_{ij}^2 - \frac{1}{2}\sum \alpha_{ij,i'j'}^2} d\alpha,$$

and now we are integrating over probabilities coming from a purely independent model (which depends on the parameters α). Now at this stage let us define $p_{ij} = 1/2 + \alpha_{ij}^* / \sqrt{4n_{ij}^2}$ and $p_{ij,i'j'} = 1/2 + \alpha_{ij,i'j'}^* / \sqrt{8n_{ij}n_{i'j'}}$. Let

$$\rho^* := \int_{\|\alpha^*\|_{\infty} \le 8\sqrt{n}} \prod_{v \in V(G)} \mathbb{P}_{\vec{d^*}}[(7.6) \text{ and } (7.7) \text{ for } v] e^{-\frac{1}{2} \sum \alpha_{ij}^{*2} - \frac{1}{2} \sum \alpha_{ij,i'j'}^{*2} d\alpha^*}$$

and note that $\alpha^* = (1 + O(1/n))\alpha$ and $\|\alpha\|_{\infty} \le 8\sqrt{n}$ in the integrand easily imply that

$$\rho \lesssim e^{-32n} + \rho^*. \tag{7.10}$$

(Note that the shift from α to α^* slightly changes the region of integration, but this error can be folded into the e^{-32n} error term already present.)

Now we focus on ρ^* . Recall that $n_{11} = n_{22} = \beta n$ and $n_{12} = n_{21} = (1 - \beta)n$, and note that each condition (7.6) and (7.7) for v corresponds to some binomial sum and difference inequalities. These probabilities can be estimated to within O(1/n) error by the results in Appendix A. For the second moment we use Lemma A.4 to find

$$\rho^* = \int_{\|\alpha\|_{\infty} \le 8\sqrt{n}} \prod_{v \in V(G)} \mathbb{P}_{\vec{d^*}}[(7.6) \text{ and } (7.7) \text{ for } v] e^{-\frac{1}{2}\sum \alpha_{ij}^{*2} - \frac{1}{2}\sum \alpha_{ij,i'j'}^{*2} d\alpha^*}$$
$$= \int_{\|\alpha^*\|_{\infty} \le 8\sqrt{n}} (q_{11}(\alpha)q_{22}(\alpha))^{\beta n} (q_{12}(\alpha)q_{21}(\alpha))^{(1-\beta)n} e^{-\frac{1}{2}\sum \alpha_{ij}^{*2} - \frac{1}{2}\sum \alpha_{ij,i'j'}^{*2} d\alpha^*}$$

where

$$q_{ij}(\alpha) := \mathbb{P} \begin{bmatrix} X^{(i,j)} - X^{(i,j,i+1,j+1)} + X^{(i,j,i,j+1)} - X^{(i,j,i+1,j)} \ge \gamma \sqrt{n} \\ X^{(i,j)} - X^{(i,j,i+1,j+1)} - X^{(i,j,i,j+1)} + X^{(i,j,i+1,j)} \ge \gamma \sqrt{n} \end{bmatrix}$$

with $X^{(i,j)} \sim \text{Bin}(n_{ij} - 1, p_{ij})$ and $X^{(i,j,i',j')} \sim \text{Bin}(n_{i'j'}, p_{ij,i'j'})$.

We respectively apply Lemma A.4 with the following parameters:

$$k = \beta n, \qquad (a_1, a_2, a_3, a_4) = \left(\frac{\alpha_{11}^*}{\sqrt{\beta^2 n}}, \frac{\alpha_{11,12}^*}{\sqrt{2\beta(1-\beta)n}}, \frac{\alpha_{11,22}^*}{\sqrt{2\beta^2 n}}, \frac{\alpha_{11,21}^*}{\sqrt{2\beta(1-\beta)n}}\right);$$

$$k = \beta n, \qquad (a_1, a_2, a_3, a_4) = \left(\frac{\alpha_{22}}{\sqrt{\beta^2 n}}, \frac{\alpha_{22,12}}{\sqrt{2\beta(1-\beta)n}}, \frac{\alpha_{22,11}}{\sqrt{2\beta^2 n}}, \frac{\alpha_{22,21}}{\sqrt{2\beta(1-\beta)n}}\right);$$

$$k = (1 - \beta)n, \quad (a_1, a_2, a_3, a_4) = \left(\frac{\alpha_{12}^*}{\sqrt{(1 - \beta)^2 n}}, \frac{\alpha_{12,11}}{\sqrt{2\beta(1 - \beta)n}}, \frac{\alpha_{12,21}}{\sqrt{2(1 - \beta)^2 n}}, \frac{\alpha_{12,22}}{\sqrt{2\beta(1 - \beta)n}}\right);$$

$$k = (1 - \beta)n, \quad (a_1, a_2, a_3, a_4) = \left(\frac{\alpha_{21}^*}{\sqrt{(1 - \beta)^2 n}}, \frac{\alpha_{21, 11}^*}{\sqrt{2\beta(1 - \beta)n}}, \frac{\alpha_{21, 12}^*}{\sqrt{2(1 - \beta)^2 n}}, \frac{\alpha_{21, 22}^*}{\sqrt{2\beta(1 - \beta)n}}\right).$$

These provide expressions of the form $g(\tau, a_1, a_2, a_3, a_4)$ (Definition A.3) where $\tau \in \{\beta, 1 - \beta\}$ for the functions $q_{11}(\alpha^*), q_{22}(\alpha^*), q_{12}(\alpha^*), q_{21}(\alpha^*)$, respectively, within an additive error of O(1/n); call these expressions $q_{ij}^{\dagger}(\alpha^*)$, which are continuous functions (suppressing the dependence on γ, β). In fact, each $q_{ij}^{\dagger}(\alpha^*)$ is log-concave in \mathbb{R}^{10} (recall $\alpha_{ij,i'j'}^* = \alpha_{i'j',ij}^*$) by Lemma 6.3. This is because $g(\tau, a_1, a_2, a_3, a_4)$ is seen to be log-concave in (a_1, a_2, a_3, a_4) . Additionally, they are lower-bounded by an absolute constant in the range $\|\alpha^*\|_{\infty} \leq 8\sqrt{n}$, which means

$$\rho^* \lesssim \int_{\|\alpha^*\|_{\infty} \le 8\sqrt{n}} \left[(q_{11}^{\dagger}(\alpha^*) q_{22}^{\dagger}(\alpha^*))^{\beta n} (q_{12}^{\dagger}(\alpha^*) q_{21}^{\dagger}(\alpha^*))^{(1-\beta)n} \right] e^{-\frac{1}{2} \sum \alpha_{ij}^{*2} - \frac{1}{2} \sum \alpha_{ij,i'j'}^{*2} d\alpha^*}.$$
(7.11)

Now notice that the initial portion of the integrand in square brackets has the property that it is symmetric under simultaneously switching $\alpha_{12}^* \leftrightarrow \alpha_{21}^*, \alpha_{12,11}^* \leftrightarrow \alpha_{21,11}^*, \alpha_{12,22}^* \leftrightarrow \alpha_{21,22}^*$ (doing so in a way that continues to enforce the symmetry $\alpha_{ij,i'j'}^* = \alpha_{i'j',ij}^*$). This can be seen as "switching the roles of partitions A and B", and one can see that it amounts to switching the indices "12" and "21" wherever they appear. Additionally, the initial portion of the integrand is symmetric under simultaneously switching $\alpha_{11}^* \leftrightarrow \alpha_{22}^*, \alpha_{11,12}^* \leftrightarrow \alpha_{22,12}^*, \alpha_{11,21}^* \leftrightarrow \alpha_{22,21}^*$. This can be seen as "switching the roles of A_1 and B_2 , as well as A_2 and B_1 ", and it amounts to switching "11" and "22" wherever they appear.

We can apply these symmetries and log-concavity to obtain an upper bound on the integrand in (7.11). (As a basic example, if g(x, y) is log-concave and symmetric then $g(x, y) = \sqrt{g(x, y)g(y, x)} \le g((x+y)/2, (x+y)/2)$.) We obtain

$$(q_{11}^{\dagger}(\alpha^{*})q_{22}^{\dagger}(\alpha^{*}))^{\beta n}(q_{12}^{\dagger}(\alpha^{*})q_{21}^{\dagger}(\alpha^{*}))^{(1-\beta)n} \le (q_{11}^{\dagger}(\alpha')q_{22}^{\dagger}(\alpha'))^{\beta n}(q_{12}^{\dagger}(\alpha')q_{21}^{\dagger}(\alpha'))^{(1-\beta)n}$$
(7.12)

where

$$\begin{aligned} \alpha'_{11} &= \alpha'_{22} = \frac{\alpha^*_{11} + \alpha^*_{22}}{2}, \quad \alpha'_{12} = \alpha'_{21} = \frac{\alpha^*_{12} + \alpha^*_{21}}{2}, \\ \alpha'_{11,12} &= \alpha'_{11,21} = \alpha'_{22,12} = \alpha'_{22,21} = \frac{\alpha^*_{11,12} + \alpha^*_{11,21} + \alpha^*_{22,12} + \alpha^*_{22,21}}{4} \\ \alpha'_{12,21} &= \alpha^*_{12,21}, \quad \alpha'_{11,22} = \alpha^*_{11,22}. \end{aligned}$$

We have

$$q_{11}^{\dagger}(\alpha') = g\Big(\beta, \frac{\alpha'_{11}}{\sqrt{\beta^2 n}}, \frac{\alpha'_{11,12}}{\sqrt{2\beta(1-\beta)n}}, \frac{\alpha'_{11,22}}{\sqrt{2\beta^2 n}}, \frac{\alpha'_{11,21}}{\sqrt{2\beta(1-\beta)n}}\Big)$$

and similar for the other values. But from Definition A.3 we see that $g(\tau, a_1, a_2, a_3, a_4)$ depends only on $\beta, a_1 - a_3, a_2 - a_4$. By $\alpha'_{11,12} = \alpha'_{11,21}$ and similar we may deduce that

$$\begin{aligned} q_{11}^{\dagger}(\alpha') &= g\Big(\beta, \frac{\alpha'_{11}\sqrt{2} - \alpha'_{11,22}}{\sqrt{2\beta^2 n}}, 0, 0, 0\Big), \qquad \qquad q_{22}^{\dagger}(\alpha') = g\Big(\beta, \frac{\alpha'_{22}\sqrt{2} - \alpha'_{22,11}}{\sqrt{2\beta^2 n}}, 0, 0, 0\Big), \\ q_{12}^{\dagger}(\alpha') &= g\Big(1 - \beta, \frac{\alpha'_{12}\sqrt{2} - \alpha'_{12,21}}{\sqrt{2(1 - \beta)^2 n}}, 0, 0, 0\Big), \qquad \qquad q_{21}^{\dagger}(\alpha') = g\Big(1 - \beta, \frac{\alpha'_{21}\sqrt{2} - \alpha'_{21,12}}{\sqrt{2(1 - \beta)^2 n}}, 0, 0, 0\Big). \end{aligned}$$

Now since $g(\tau, a_1, a_2, a_3, a_4)$ is log-concave in (a_1, a_2, a_3, a_4) , we see

$$q_{11}^{\dagger}(\alpha')q_{22}^{\dagger}(\alpha') \le g\left(\beta, \frac{\alpha'_{11}\sqrt{2} - \alpha'_{11,22}}{\sqrt{2\beta^2 n}}, 0, 0, 0\right)^2 = f\left(\beta, \frac{\alpha'_{11,22} - \alpha'_{11}\sqrt{2}}{2\sqrt{2n}}\right)^2, \tag{7.13}$$

$$q_{12}^{\dagger}(\alpha')q_{21}^{\dagger}(\alpha') \le g\left(1-\beta, \frac{\alpha'_{12}\sqrt{2}-\alpha'_{12,21}}{\sqrt{2(1-\beta)^2n}}, 0, 0, 0\right)^2 = f\left(1-\beta, \frac{\alpha'_{12,21}-\alpha'_{12}\sqrt{2}}{2\sqrt{2n}}\right)^2, \tag{7.14}$$

using $\alpha'_{11} = \alpha'_{22}$, $\alpha'_{11,22} = \alpha'_{22,11}$, and similar, as well as the relation between g and f described in Definition A.3 (recall f from Definition 1.4). Now (7.11) to (7.14) yield

$$\rho^* \lesssim \int_{\|\alpha^*\|_{\infty} \le 8\sqrt{n}} f\left(\beta, \frac{\alpha'_{11,22} - \alpha'_{11}\sqrt{2}}{2\sqrt{2n}}\right)^{2\beta n} f\left(1 - \beta, \frac{\alpha'_{12,21} - \alpha'_{12}\sqrt{2}}{2\sqrt{2n}}\right)^{2(1-\beta)n} e^{-\frac{1}{2}\sum \alpha^{*2}_{ij} - \frac{1}{2}\sum \alpha^{*2}_{ij,i'j'} d\alpha^*}.$$
(7.15)

Now,

$$\frac{\alpha_{11,22}' - \alpha_{11}'\sqrt{2}}{\sqrt{2}}, \frac{\alpha_{11,22}' + \alpha_{11}'\sqrt{2}}{\sqrt{2}}, \frac{\alpha_{12,21}' - \alpha_{12}'\sqrt{2}}{\sqrt{2}}, \frac{\alpha_{12,21}' + \alpha_{12}'\sqrt{2}}{\sqrt{2}}$$

are linear combinations of α^* , i.e., $v_1 \cdot \alpha^*$, $v_2 \cdot \alpha^*$, $v_3 \cdot \alpha^*$, $v_4 \cdot \alpha^*$ for vectors $v_1, v_2, v_3, v_4 \in \mathbb{R}^{10}$ (here we are assuming α^* only contains one of each symmetric pair $\alpha^*_{ij,i'j'} = \alpha^*_{i'j',ij}$). One can easily check that v_1, v_2, v_3, v_4 are orthonormal vectors in \mathbb{R}^{10} . Thus, there exists a completion to an orthonormal basis $\{v_1, \ldots, v_{10}\}$ so that

$$\sum \alpha_{ij}^{*2} + \sum \alpha_{ij,i'j'}^{*2} = \frac{(\alpha_{11,22}' - \alpha_{11}'\sqrt{2})^2 + (\alpha_{11,22}' + \alpha_{11}'\sqrt{2})^2 + (\alpha_{12,21}' - \alpha_{12}'\sqrt{2})^2 + (\alpha_{12,21}' + \alpha_{12}'\sqrt{2})^2}{2} + \sum_{j=5}^{10} (v_j \cdot \alpha^*)^2.$$

Reparametrizing the integration coordinates of (7.15) via this basis change, we find

$$\rho^* \lesssim \int_{\|x\|_{\infty} \leq 32\sqrt{n}} f\left(\beta, \frac{x_1}{2\sqrt{n}}\right)^{2\beta n} f\left(1-\beta, \frac{x_2}{2\sqrt{n}}\right)^{2(1-\beta)n} e^{-\frac{1}{2}\sum_{j=1}^{10} x_j^2} dx
\lesssim \int_{|x_1|, |x_2| \leq 32\sqrt{n}} f\left(\beta, \frac{x_1}{2\sqrt{n}}\right)^{2\beta n} f\left(1-\beta, \frac{x_2}{2\sqrt{n}}\right)^{2(1-\beta)n} e^{-\frac{1}{2}(x_1^2+x_2^2)} dx
\lesssim n \int_{|x_1|, |x_2| \leq 16} f(\beta, x_1)^{2\beta n} f(1-\beta, x_2)^{2(1-\beta)n} \exp\left(-(2x_1^2+2x_2^2)n\right) dx
\lesssim n \int_{|x_1|, |x_2| \leq 16} \exp\left(n\left(-2x_1^2-2x_2^2+2\beta\log f(\beta, x_1)+2(1-\beta)\log f(1-\beta, x_2)\right)\right) dx. \quad (7.16)$$

Let Q be the region $[-16, 16]^2 \subseteq \mathbb{R}^2$. Finally, (7.4), (7.5), (7.10), and (7.16) demonstrate

$$\mathbb{P}[Y_A(n_1') = Y_B(n_1') = 1] = \mathbb{P}[\mathcal{E}_2] \lesssim e^{-32n} + n \int_Q e^{n(-2x_1^2 - 2x_2^2 + 2\beta \log f(\beta, x_1) + 2(1-\beta) \log f(1-\beta, x_2))} dx.$$

$$\lesssim n \int_Q e^{n(-2x_1^2 - 2x_2^2 + 2\beta \log f(\beta, x_1) + 2(1-\beta) \log f(1-\beta, x_2))} dx.$$

for $n'_1/n \in [c, 1-c]$. Here we eliminated e^{-32n} by noting that the region near $x_1 = x_2 = -2$ contributes at least $\exp(-32n)$, using $|\gamma| \leq 1$.

Combining this with (7.2) and (7.3) we find

$$\begin{split} \mathbb{E}X_{\gamma}^{2} &\lesssim (1+\eta)^{n} \mathbb{E}X_{\gamma} + \binom{2n}{n} \sum_{c \leq \beta \leq 1-c} \binom{n}{\beta n}^{2} n \int_{Q} e^{n(-2x_{1}^{2}-2x_{2}^{2}+2\beta \log f(\beta,x_{1})+2(1-\beta)\log f(1-\beta,x_{2}))} dx \\ &\lesssim \frac{1}{\sqrt{n}} \sum_{c \leq \beta \leq 1-c} \int_{Q} e^{n(2\log 2-2\beta \log \beta-2(1-\beta)\log(1-\beta)-2x_{1}^{2}-2x_{2}^{2}+2\beta \log f(\beta,x_{1})+2(1-\beta)\log f(1-\beta,x_{2}))} dx \\ &+ (1+\eta)^{n} \mathbb{E}X_{\gamma} \\ &\lesssim (1+\eta)^{n} \mathbb{E}X_{\gamma} + \frac{1}{\sqrt{n}} \sum_{c \leq \beta \leq 1-c} \int_{Q} e^{nF_{2}(\beta,x_{1},x_{2})} dx. \end{split}$$

Here the sum is over β with $\beta n \in \mathbb{Z}$. We used Stirling's formula.

Now we split the sum over β into two regions. First we consider $\beta \in [c, .001] \cup [.999, 1 - c]$. Note that F_2 is symmetric under $\beta \leftrightarrow 1 - \beta$, so up to a constant factor loss we may assume $\beta \in [c, .001]$. By the first part of Assumption 1.5, we see $F_2(\beta, x_1, x_2) \leq 2 \sup_{y \in \mathbb{R}} F_1(y)$, so

$$\frac{1}{\sqrt{n}} \sum_{\beta \in [c,.001] \cup [.999,1-c]} \int_Q e^{nF_2(\beta,x_1,x_2)} dx \lesssim (1+\eta/2)^n \exp(2\sup_{y \in \mathbb{R}} F_1(y))$$

Combining with the lower bound in Lemma 3.2 (established below), this is bounded by $(1+\eta)^n \mathbb{E} X_{\gamma}$ so can be absorbed.

Finally, we consider $\beta \in [.001, .999]$. By the second part of Assumption 1.5, the unique optimizer of $F_2(\beta, x_1, x_2)$ for $\beta \in [.001, .999]$ and $x_1, x_2 \in \mathbb{R}$ occurs at $(1/2, x^*, x^*)$ where $x^* = \arg \max_{x \in \mathbb{R}} F_1(x)$. Additionally, the Hessian of F_2 evaluated at $(1/2, x^*, x^*)$ is strictly negative definite. Therefore there is some small absolute constants $\theta, \theta' > 0$ and a box $Q^* = (1/2, x^*, x^*) + [-\theta, \theta]^3$ such that for all $(\beta, x_1, x_2) \in Q^*$ we have

$$F_2(\beta, \alpha_1, \alpha_2) \le F_2(1/2, x^*, x^*) - \theta((\beta - 1/2)^2 + (x_1 - x^*)^2 + (x_2 - x^*)^2)$$

and also $F_2(\beta, x_1, x_2) \leq F_2(1/2, x^*, x^*) - \theta'$ for $(\beta, x_1, x_2) \notin Q^*$ satisfying $\beta \in [.001, .999]$. Letting $Q' = (x^*, x^*) + [-\theta, \theta]^2$, we therefore find

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{.001 \le \beta \le .999} \int_{Q} e^{nF_{2}(\beta,x_{1},x_{2})} dx &\lesssim \frac{1}{\sqrt{n}} \sum_{\beta \in [1/2 \pm \theta]} \int_{Q'} e^{nF_{2}(1/2,x^{*},x^{*}) - \theta n ((\beta - 1/2)^{2} + (x_{1} - x^{*})^{2} + (x_{2} - x^{*})^{2})} dx_{1} dx_{2} \\ &+ e^{(F_{2}(1/2,x^{*},x^{*}) - \theta')n} \\ &\lesssim \frac{e^{nF_{2}(1/2,x^{*},x^{*})}}{n^{3/2}} \sum_{\beta \in [1/2 \pm \theta]} e^{-\theta n (\beta - 1/2)^{2}} + e^{(F_{2}(1/2,x^{*},x^{*}) - \theta')n} \\ &\lesssim n^{-1} e^{nF_{2}(1/2,x^{*},x^{*})} \lesssim 16^{-n} \binom{2n}{n}^{2} \exp\left(n \sup_{\substack{\beta \in [0,1] \\ \alpha_{1},\alpha_{2} \in \mathbb{R}}} F_{2}(\beta,\alpha_{1},\alpha_{2})\right). \end{aligned}$$

We deduce

$$\mathbb{E}X_{\gamma}^{2} \lesssim (1+\eta)^{n} \mathbb{E}X_{\gamma} + 16^{-n} \binom{2n}{n}^{2} \exp\left(n \sup_{\substack{\beta \in [0,1]\\\alpha_{1},\alpha_{2} \in \mathbb{R}}} F_{2}(\beta,\alpha_{1},\alpha_{2})\right),$$

as desired, proving Lemma 3.3.

7.2. The first moment of X_{γ} . Now we prove Lemma 3.2. As discussed earlier, it is significantly easier and we truncate some details (although take note that Lemma 3.2 provides a matching lower bound unlike Lemma 3.3, which we used in the proofs of Theorem 1.3 and Lemma 3.3).

7.2.1. Setup and transference. Let $A_1 = [n]$ and $A_2 = \{n + 1, ..., 2n\}$. Let \mathcal{E}_1 be the event that $G \sim \mathbb{G}(2n, 1/2)$ is such that $A_1 \cup A_2$ is a $\gamma \sqrt{n}$ -friendly partition. By linearity of expectation we have

$$\mathbb{E}X_{\gamma} = \binom{2n}{n} \mathbb{P}[\mathcal{E}_1]. \tag{7.17}$$

By Propositions 5.4 and 5.7 with η replaced by $\min(\varepsilon_{4.8}, \varepsilon_{4.9})/2$, we have

$$\mathbb{P}[\mathcal{E}_1] = (1 \pm n^{-\omega(1)}) \mathbb{P}[\mathcal{E}_1 \wedge \mathcal{E}_{\mathrm{irreg},1}^c \wedge \mathcal{E}_{\mathrm{var}}^*].$$
(7.18)

Here we define $\mathcal{E}_{\text{var}}^*$ just to be the appropriate set of conditions (in a sense a subset of those from \mathcal{E}_{var}) guaranteeing that the degree sequences of $G[A_1], G[A_2], G[A_1, A_2]$ satisfy **A2** of Lemma 4.8 or **B5** of Lemma 4.9. So $\mathcal{E}_{\text{irreg},1}^c, \mathcal{E}_{\text{var}}^*$ ensure that for the 2 + 1 graphs $G[A_1]$ and $G[A_2]$ and bipartite

 $G[A_1, A_2]$, the degree sequences satisfy A1 to A3 of Lemma 4.8 and B1 to B5 of Lemma 4.9, respectively. We deduce

$$\mathbb{P}[\mathcal{E}_1 \wedge \mathcal{E}_{\mathrm{irreg},1}^c \wedge \mathcal{E}_{\mathrm{var}}^*] \asymp \rho', \tag{7.19}$$

where ρ' is the following probability:

- For $i \in \{1, 2\}$ sample the "degrees" d_v^i for $v \in A_i$ (which should be thought of as corresponding to the internal degrees $\deg(v, A_i)$ for $v \in A_i$) from $\mathcal{I}_{1/2}^n$;
- Sample the "degrees" $d_v^{(1,2)}$ for $v \in A_1$ and $d_v^{(2,1)}$ for $v \in A_2$ (which correspond to the bipartite degrees deg (v, A_2) for $v \in A_1$ and vice versa) jointly from $\mathcal{I}_{1/2}^{n,n}$;
- Compute the probability that (a) for each $i \in \{1, 2\}$ and $v \in A_i$ we have

$$d_v^i - d_v^{(i,i+1)} \ge \gamma \sqrt{n},\tag{7.20}$$

where indices are taken (mod 2), and (b) the resulting degree sequences satisfy $\mathcal{E}_{irreg,1}^c \wedge \mathcal{E}_{var}^*$. This can be interpreted as the probability that if the degrees of our 1+2 graphs are sampled as in Lemmas 4.8 and 4.9, we have that $A_1 \cup A_2$ is $\gamma \sqrt{n}$ -friendly (and the degree sequences are not too irregular). Let ρ be the same probability without condition (b).

Understanding the "degree" distribution is relatively simple:

- Sample 2 + 1 parameters $p_1, p_2 \sim \mathcal{N}(1/2, 1/(4n^2 4n))$ and $p_{1,2} \sim \mathcal{N}(1/2, 1/(8n^2))$ (all conditional on being in (0, 1)), and let $p_{2,1} = p_{1,2}$;
- Sample $d_v^i \sim \operatorname{Bin}(n-1, p_i)$ and $d_v^{(i,i+1)} \sim \operatorname{Bin}(n, p_{i,i+1})$ completely independently;
- Condition on the 2 divisibility conditions and 1 equality condition

$$2 \mid \sum_{v \in A_i} d_v^i, \tag{7.21}$$

$$\sum_{v \in A_1} d_v^{(1,2)} = \sum_{v \in A_2} d_v^{(2,1)}.$$
(7.22)

The first step can be handled by taking a Gaussian integral over the 2 + 1 parameters, and the second step is independent. Again we handle the conditioning from the third step by use of Bayes' theorem and local central limit theorems similar to Section 7.1.4, but now the local central limit results used will be purely 1-dimensional. We will also require a two-sided bound unlike before.

7.2.2. Upper bound. We clearly have $\rho' \leq \rho$. Now let $p_i = 1/2 + \alpha_i/\sqrt{4n^2 - 4n}$ and $p_{1,2} = 1/2 + \alpha_{1,2}/\sqrt{8n^2}$. Let \vec{d} denote the distribution over the joint "degree" distribution, and $\vec{d^*}$ denote the distribution prior to conditioning on (7.21) and (7.22). We have, similar to in Section 7.1.3,

$$\begin{split} \rho &\lesssim \int_{\alpha} \mathbb{P}_{\vec{d}}[(7.20)] e^{-\frac{1}{2}(\alpha_{1}^{2} + \alpha_{2}^{2} + \alpha_{1,2}^{2})} d\alpha \\ &\lesssim e^{-8n} + \int_{\|\alpha\|_{\infty} \le 4\sqrt{n}} \mathbb{P}_{\vec{d^{*}}}[(7.20)] \frac{\mathbb{P}_{\vec{d^{*}}}[(7.21) \text{ and } (7.22)|(7.20)]}{\mathbb{P}_{\vec{d^{*}}}[(7.21) \text{ and } (7.22)]} e^{-\frac{1}{2}(\alpha_{1}^{2} + \alpha_{2}^{2} + \alpha_{1,2}^{2})} d\alpha \end{split}$$

Now, the denominator is easily seen to be $\Theta(1/n)$. Indeed, for the denominator we require two independent sums of random variables to be even, and a signed sum of random variables to be 0. It follows e.g. from a local central limit theorem for log-concave random variables ([6]) and the fact that the mean of this signed sum can be seen to be within standard-deviation-range (up to a constant factor) of 0. (Here we are using that the binomial random variable is log-concave, hence a pair of independent binomials is log-concave, hence conditioning on a single linear inequality keeps the distribution log-concave, and finally that the marginal of a log-concave discrete distribution is log-concave [26].) To estimate the numerator, first note that upon conditioning on (7.20) and all values $d_v^{(1,2)}, d_v^{(2,1)}$, we are left with independent randomness for the d_v^1 and d_v^2 values. The probability both sums are even is easily seen to be approximately 1/4. Thus, we find

$$\mathbb{P}_{\vec{d^*}}[(7.21) \text{ and } (7.22)|(7.20)] = (1/4 + o(1))\mathbb{P}_{\vec{d^*}}[(7.22)|(7.20)].$$

Now note that the $d_v^{(1,2)}, d_v^{(2,1)}$ are jointly independent and thus it is easy to see (e.g. via the Lévy–Kolmogorov–Rogozin inequality [40]) that $\mathbb{P}_{\vec{d^*}}[(7.22)|(7.20)] \leq 1/n$. Overall, we deduce

$$\rho' \le \rho \lesssim e^{-8n} + \int_{\|\alpha\|_{\infty} \le 4\sqrt{n}} \mathbb{P}_{\vec{d^*}}[(7.20)] d\alpha.$$

Now at this stage let us define $p_i = 1/2 + \alpha_i^*/(2n)$ and $p_{i,i+1} = 1/2 + \alpha_{i,i+1}^*/(2n\sqrt{2})$. We see $\alpha^* = (1 + O(1/n))\alpha$ so it is not hard to deduce

$$\rho' \lesssim e^{-8n} + \int_{\|\alpha^*\|_{\infty} \le 4\sqrt{n}} \mathbb{P}_{\vec{d^*}}[(7.20)] e^{-\frac{1}{2}(\alpha_1^{*2} + \alpha_2^{*2} + \alpha_{1,2}^{*2})} d\alpha^*$$
(7.23)

similar to the beginning of Section 7.1.5. At this stage, note that

$$\mathbb{P}_{\vec{d^*}}[(7.20)] = \mathbb{P}[\operatorname{Bin}(n-1, 1/2 + \alpha_1^*/(2n)) - \operatorname{Bin}(n, 1/2 + \alpha_{1,2}^*/(2n\sqrt{2})) \ge \gamma\sqrt{n}]^n \\ \cdot \mathbb{P}[\operatorname{Bin}(n-1, 1/2 + \alpha_2^*/(2n)) - \operatorname{Bin}(n, 1/2 + \alpha_{1,2}^*/(2n\sqrt{2})) \ge \gamma\sqrt{n}]^n.$$

By Lemma A.2, we deduce

$$\mathbb{P}_{\vec{d^*}}[(7.20)] \asymp \left(\mathbb{P}_{Z \sim \mathcal{N}(0,1)} \left[Z \ge \gamma \sqrt{2} + \frac{\alpha_{1,2}^* - \alpha_1^* \sqrt{2}}{2\sqrt{n}} \right] \mathbb{P}_{Z \sim \mathcal{N}(0,1)} \left[Z \ge \gamma \sqrt{2} + \frac{\alpha_{1,2}^* - \alpha_2^* \sqrt{2}}{2\sqrt{n}} \right] \right)^n \\ \lesssim \mathbb{P}_{Z \sim \mathcal{N}(0,1)} \left[Z \ge \gamma \sqrt{2} + \frac{\alpha_{1,2}^* - (\alpha_1^* + \alpha_2^*) \sqrt{2}/2}{2\sqrt{n}} \right]^{2n}.$$
(7.24)

by Lemma 6.2.

Now similar to Section 7.1.5, we may combine (7.23) and (7.24) and integrate the Gaussian in the direction orthogonal to $\alpha_{1,2}^* - (\alpha_1^* + \alpha_2^*)\sqrt{2}/2$. Using an appropriate identity

$$\alpha_1^{*2} + \alpha_2^{*2} + \alpha_{1,2}^{*2} = \frac{(\alpha_{1,2}^* \sqrt{2} - \alpha_1^* - \alpha_2^*)^2 + (v_2 \cdot \alpha^*) + (v_3 \cdot \alpha^*)}{4}$$

and changing variables (with $x = (\alpha_{1,2}^*\sqrt{2} - \alpha_1^* - \alpha_2^*)/2)$, we deduce

$$\begin{split} \rho' &\lesssim e^{-8n} + \int_{|x| \le 8\sqrt{n}} \mathbb{P}_{Z \sim \mathcal{N}(0,1)} \left[Z \ge \gamma \sqrt{2} + \frac{x}{\sqrt{2n}} \right]^{2n} e^{-\frac{1}{2}x^2} dx \\ &\lesssim e^{-8n} + \int_{|x| \le 4} \mathbb{P}_{Z \sim \mathcal{N}(0,1)} [Z \ge (\gamma + x)\sqrt{2}]^{2n} e^{-2x^2 n} dx. \end{split}$$

Finally, note that $h(x) := \log(\mathbb{P}_{Z \sim \mathcal{N}(0,1)}[Z \geq (\gamma + x)\sqrt{2}])$ is concave hence adding $-x^2$ makes it strictly concave. Furthermore, let x^* be the unique maximizer of $-x^2 + h(x)$ on \mathbb{R} . We easily find

$$\rho' \lesssim e^{-8n} + e^{2n \sup_{\alpha \in \mathbb{R}} (F_1(\alpha) - \log 2)},$$

similar to Section 7.1.5. Furthermore, we easily check that for $|\gamma| \leq 1$, e^{-8n} is lower order than the supremum here.

Hence, $\rho' \leq 4^{-n} \exp(2n \sup_{\alpha \in \mathbb{R}} F_1(\alpha))$. Combining with (7.17) to (7.19) we deduce

$$\mathbb{E}X_{\gamma} \lesssim 4^{-n} \binom{2n}{n} \exp(2n \sup_{\alpha \in \mathbb{R}} F_{1}(\alpha))$$

as desired for the upper bound.

7.2.3. Lower bound. For the lower bound, we must first find a way of removing the condition (b) above. To this end, let us further consider the event defined by ρ' with the additional conditions $|p_1 - p_2| \leq 2/n$ and $||p - 1/2||_{\infty} \leq 1/\sqrt{n}$, and say it has probability ρ'_0 . Similarly, the event defined by ρ with the additional condition $|p_1 - p_2| \leq 2/n$ is ρ_0 . Clearly $\rho' \geq \rho'_0$.

We claim $\rho'_0 = (1 \pm n^{-\omega(1)})\rho_0$. To see this, note that ρ'_0/ρ_0 is the following probability: sample the "degree" distribution as above with the additional conditions $|p_1-p_2| \leq 2/n$ and $||p-1/2||_{\infty} \leq 1/\sqrt{n}$, then further condition on (7.20). Then consider the probability both $\mathcal{E}^c_{\text{irreg},1}$ and $\mathcal{E}^*_{\text{var}}$ hold. It is easy to see that this new distribution can be thought of as follows: sample p appropriately, then for each $v \in A_i$ sample $(d^i_v, d^{(i,i+1)}_v) \sim (\text{Bin}(n-1, p_i), \text{Bin}(n, p_{i,i+1}))$ conditional on $d^i_v - d^{(i,i+1)}_v \geq \gamma\sqrt{n}$, then further condition on $2|\sum_{v \in A_i} d^i_v$ and $\sum_{v \in A_1} d^{(1,2)}_v = \sum_{v \in A_2} d^{(2,1)}_v$. We claim that these latter conditions then occur with probability $\Omega((1/2)^2(1/n))$. The evenness

We claim that these latter conditions then occur with probability $\Omega((1/2)^2(1/n))$. The evenness conditions can be handled as in Section 7.2.2; the only condition of interest is (7.22). For this, if $p_1 = p_2$ then it is the probability an alternating sum of i.i.d. binomials equals 0, and a log-concave local limit theorem [6] will show the result. For $|\alpha_1 - \alpha_2| = O(1/n)$, we again use [6] and simply note that the mean of the signed sum of binomials is within standard-deviation-range (i.e., O(n)) of 0.

Letting the random model formed merely by sampling $p, d_v^i, d_v^{(i,i+1)}$ and no further conditioning be labeled $\vec{d^*}$, we therefore see it suffices to show

$$\mathbb{P}_{\vec{d^{\dagger}}}[\mathcal{E}_{\text{irreg},1}^c \wedge \mathcal{E}_{\text{var}}^*] \le n^{-\omega(1)}.$$
(7.25)

This is easy to show using concentration of measure. Note that, e.g., the $d_v^{(1,2)}$ and $d_v^{(2,1)}$ are jointly independent in $\vec{d^{\dagger}}$. It is thus not too hard to use concentration of measure to show (7.25), though we truncate the (relatively standard) details.

Thus, $\rho' \gtrsim \rho_0$. Now, ρ_0 is very similar to ρ and can be analyzed as in Section 7.2.2. Let R denote the region in \mathbb{R}^3 defined by $\|\alpha\|_{\infty} \leq 2\sqrt{n}$ and $|\alpha_1 - \alpha_2| \leq 4$. We have

$$\begin{split} \rho' \gtrsim \int_{R} \mathbb{P}_{\vec{d}}[(7.20)] e^{-\frac{1}{2}(\alpha_{1}^{2} + \alpha_{2}^{2} + \alpha_{1,2}^{2})} d\alpha \\ \geq \int_{R} \mathbb{P}_{\vec{d}^{*}}[(7.20)] \frac{\mathbb{P}_{\vec{d}^{*}}[(7.21) \text{ and } (7.22)|(7.20)]}{\mathbb{P}_{\vec{d}^{*}}[(7.21) \text{ and } (7.22)]} e^{-\frac{1}{2}(\alpha_{1}^{2} + \alpha_{2}^{2} + \alpha_{1,2}^{2})} d\alpha. \end{split}$$

Now, the denominator is easily seen to be $\Theta(1/n)$ similar to Section 7.2.2 and above we deduced that the numerator is $\Omega(1/n)$. Further reparametrizing to $p_i = 1/2 + \alpha_1^*/(2n)$ and $p_{i,i+1} = 1/2 + \alpha_{i,i+1}^*/(2n\sqrt{2})$, overall we find

$$\rho' \gtrsim \int_{R^*} \mathbb{P}_{\vec{d^*}}[(7.20)] e^{-\frac{1}{2}(\alpha_1^2 + \alpha_2^2 + \alpha_{1,2}^2)} d\alpha \tag{7.26}$$

where R^* is the region $\|\alpha^*\|_{\infty} \leq \sqrt{n}$ and $|\alpha_1 - \alpha_2| \leq 2$.

Now, using the first line of (7.24), $|\alpha_1^* - \alpha_2^*| = O(1)$, and $||\alpha^*||_{\infty} = O(\sqrt{n})$, we easily deduce that

$$\mathbb{P}_{\vec{d^*}}[(7.20)] \gtrsim \mathbb{P}_{Z \sim \mathcal{N}(0,1)} \left[Z \ge \gamma \sqrt{2} + \frac{\alpha_{1,2}^* - (\alpha_1^* + \alpha_2^*) \sqrt{2}/2}{2\sqrt{n}} \right]^{2n}$$

Let us focus on the region Q^* where additionally

$$|\alpha_{1,2}^* - (\alpha_1^* + \alpha_2^*)\sqrt{2}/2 - (2\sqrt{2n})x^*| \le \sqrt{2},$$

where $x^* = \arg \max_{x \in \mathbb{R}} F_1(x)$.

Change coordinates in (7.26) similar to Section 7.2.2, with $x = (\alpha_{1,2}^* \sqrt{2} - \alpha_1^* - \alpha_2^*)/2$. Integrating with respect to the other coordinates, we find

$$\rho' \gtrsim \int_{|x-(2\sqrt{n})x^*| \le 1} \mathbb{P}_{Z \sim \mathcal{N}(0,1)}[Z \ge (\gamma + x/(2\sqrt{n}))\sqrt{2}]^{2n} e^{-\frac{1}{2}x^2} dx$$

$$\gtrsim \sqrt{n} \int_{|x-x^*| \le 1/(2\sqrt{n})} \mathbb{P}_{Z \sim \mathcal{N}(0,1)} [Z \ge (\gamma+x)\sqrt{2}]^{2n} e^{-2x^2 n} dx \gtrsim \sqrt{n} \int_{|x-x^*| \le 1/(2\sqrt{n})} \exp(2n(F_1(x) - \log 2)) dx.$$

Since $x = x^*$ maximizes the expression in the exponent of the integrand, we see that $2n(F_1(x) - \log 2)$ stays within a range of size O(1) when $|x - x^*| \le 1/(2\sqrt{n})$, hence

$$\rho' \gtrsim \sqrt{n} \cdot \frac{1}{\sqrt{n}} \cdot \exp(2n \sup_{x \in \mathbb{R}} (F_1(x) - \log 2)).$$

Combining this with (7.17) to (7.19), we obtain

$$\mathbb{E}X_{\gamma} \gtrsim 4^{-n} \binom{2n}{n} \exp(2n \sup_{\alpha \in \mathbb{R}} F_1(\alpha)),$$

completing the lower bound of Lemma 3.2 and hence the proof.

References

- Emmanuel Abbe, Shuangping Li, and Allan Sly, Proof of the Contiguity Conjecture and Lognormal Limit for the Symmetric Perceptron, 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science—FOCS 2021, IEEE Computer Soc., Los Alamitos, CA, [2022] (©)2022, pp. 327–338. 4, 12
- [2] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke, Local algorithms for maximum cut and minimum bisection on locally treelike regular graphs of large degree. 2
- Benjamin Aubin, Will Perkins, and Lenka Zdeborová, Storage capacity in symmetric binary perceptrons, J. Phys. A 52 (2019), 294003, 32. 11
- [4] Cristina Bazgan, Zsolt Tuza, and Daniel Vanderpooten, Satisfactory graph partition, variants, and generalizations, European J. Oper. Res. 206 (2010), 271–280.
- [5] Freya Behrens, Gabriel Arpino, Yaroslav Kivva, and Lenka Zdeborová, (Dis)assortative partitions on random regular graphs, Journal of Physics A: Mathematical and Theoretical 55 (2022), 395004. 2, 3
- [6] Edward A. Bender, Central and local limit theorems applied to asymptotic enumeration, J. Combinatorial Theory Ser. A 15 (1973), 91–111. 31, 33
- [7] Itai Benjamini, Gil Kalai, and Oded Schramm, Noise sensitivity of Boolean functions and applications to percolation, Inst. Hautes Études Sci. Publ. Math. (1999), 5–43 (2001). 3
- [8] Béla Bollobás, The distribution of the maximum degree of a random graph, Discrete Math. 32 (1980), 201–203.
 3, 7
- [9] Guy Bresler and Brice Huang, The algorithmic phase transition of random k-SAT for low degree polynomials, 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science—FOCS 2021, IEEE Computer Soc., Los Alamitos, CA, [2022] ©2022, pp. 298–309. 2
- [10] E. R. Canfield, C. Greenhill, and B. D. McKay, Asymptotic enumeration of dense 0-1 matrices with specified line sums, J. Combin. Theory Ser. A 115 (2008), 32–66. 13
- [11] Wei-Kuo Chen, David Gamarnik, Dmitry Panchenko, and Mustazee Rahman, Suboptimality of local algorithms for a class of max-cut problems, Ann. Probab. 47 (2019), 1587–1618. 2
- [12] Zvi Drezner and G. O. Wesolowsky, On the computation of the bivariate normal integral, J. Statist. Comput. Simulation 35 (1990), 101–107. 40
- [13] Ronen Eldan and Renan Gross, Concentration on the Boolean hypercube via pathwise stochastic analysis, Invent. Math. 230 (2022), 935–994. 3
- [14] Ronen Eldan, Guy Kindler, Noam Lifshitz, and Dor Minzer, Isoperimetric Inequalities Made Simpler, arXiv:2204.06686. 3, 8
- [15] Asaf Ferber, Matthew Kwan, Bhargav Narayanan, Ashwin Sah, and Mehtaab Sawhney, Friendly bisections of random graphs, Comm. Amer. Math. Soc. 2 (2022), 380–416. 2, 17
- [16] Ehud Friedgut, Sharp thresholds of graph properties, and the k-sat problem, J. Amer. Math. Soc. 12 (1999), 1017–1054, With an appendix by Jean Bourgain. 3
- [17] Ehud Friedgut and Gil Kalai, Every monotone graph property has a sharp threshold, Proc. Amer. Math. Soc. 124 (1996), 2993–3002. 3
- [18] Z. Füredi, personal communication. 2
- [19] David Gamarnik, The overlap gap property: A topological barrier to optimizing over random structures, Proc. Natl. Acad. Sci. USA 118 (2021). 2

- [20] David Gamarnik and Quan Li, On the max-cut of sparse random graphs, Random Structures Algorithms 52 (2018), 219–262. 5, 7, 23
- [21] David Gamarnik and Madhu Sudan, Limits of local algorithms over sparse random graphs, Ann. Probab. 45 (2017), 2353–2376. 2
- [22] David Gamarnik and Madhu Sudan, Performance of sequential local algorithms for the random NAE-K-SAT problem, SIAM J. Comput. 46 (2017), 590–619. 2
- [23] Michael U. Gerber and Daniel Kobler, Algorithmic approach to the satisfactory graph partitioning problem, vol. 125, 2000, Combinatorial optimization (Copenhagen, 1998), pp. 283–291. 2
- [24] Guilherme C. M. Gomes and Ignasi Sau, Finding cuts of bounded degree: complexity, FPT and exact algorithms, and kernelization, Algorithmica 83 (2021), 1677–1706. 2
- [25] B. Green, 100 open problems, Manuscript, available on request. 2
- [26] Diana Halikias, Bo'az Klartag, and Boaz A Slomka, Discrete variants of Brunn-Minkowski type inequalities, arXiv:1911.04392. 31
- [27] Hamed Hatami, A structure theorem for Boolean functions with small total influences, Ann. of Math. (2) 176 (2012), 509–533. 12
- [28] Brice Huang and Mark Sellke, Tight Lipschitz hardness for optimizing mean field spin glasses, 2022 IEEE 63rd Annual Symposium on Foundations of Computer Science—FOCS 2022, IEEE Computer Soc., Los Alamitos, CA, [2022] ©2022, pp. 312–322. 2
- [29] Fredrik Johansson, Arb: efficient arbitrary-precision midpoint-radius interval arithmetic, IEEE Transactions on Computers 66 (2017), 1281–1292. 6
- [30] L. Leindler, On a certain converse of Hölder's inequality. II, Acta Sci. Math. (Szeged) 33 (1972), 217–223. 23
- [31] Brendan D. McKay and Fiona Skerman, Degree sequences of random digraphs and bipartite graphs, J. Comb. 7 (2016), 21–49. 4, 5, 14
- [32] Brendan D. McKay, Ian M. Wanless, and Nicholas C. Wormald, Asymptotic enumeration of graphs with a given upper bound on the maximum degree, Combin. Probab. Comput. 11 (2002), 373–392. 4, 5, 17
- [33] Brendan D. McKay and Nicholas C. Wormald, Asymptotic enumeration by degree sequence of graphs of high degree, European J. Combin. 11 (1990), 565–580. 4, 5, 13
- [34] Brendan D. McKay and Nicholas C. Wormald, The degree sequence of a random graph. I. The models, Random Structures Algorithms 11 (1997), 97–117. 4, 14, 15
- [35] Shuta Nakajima and Nike Sun, Sharp threshold sequence and universality for Ising perceptron models, arXiv:2204.03469. 12
- [36] Will Perkins and Changji Xu, Frozen 1-RSB structure of the symmetric Ising perceptron, STOC '21—Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing, ACM, New York, [2021] ©2021, pp. 1579–1588. 4, 12
- [37] András Prékopa, Logarithmic concave measures with application to stochastic programming, Acta Sci. Math. (Szeged) 32 (1971), 301–316. 23
- [38] Mustazee Rahman and Bálint Virág, Local algorithms for independent sets are half-optimal, Ann. Probab. 45 (2017), 1543–1577. 2
- [39] Oliver Riordan and Alex Selby, The maximum degree of a random graph, Combin. Probab. Comput. 9 (2000), 549–572. 4
- [40] B. A. Rogozin, On the increase of dispersion of sums of independent random variables, Teor. Verojatnost. i Primenen 6 (1961), 106–108. 27, 32
- [41] Ashwin Sah and Mehtaab Sawhney, Majority Dynamics: The Power of One, arXiv:2105.13301. 25, 26
- [42] Ashwin Sah and Mehtaab Sawhney, Subgraph distributions in dense random regular graphs, arXiv:2209.00734.
 12
- [43] David Sherrington and Scott Kirkpatrick, Solvable model of a spin-glass, Physical review letters 35 (1975), 1792.
 2
- [44] Alessandro Treves and Daniel J Amit, Metastable states in asymmetrically diluted Hopfield networks, Journal of Physics A: Mathematical and General 21 (1988), 3155. 2
- [45] R. 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. 17
- [46] Changji Xu, Sharp threshold for the Ising perceptron model, Ann. Probab. 49 (2021), 2399–2415. 12

Appendix A. Binomial Distribution Computations

We will require a number of precise asymptotics associated to the binomial random variables.

Lemma A.1. Fix $C \ge 1$ such that $\max\{|a_1|, |a_2|\} \le C$. Then for n large,

$$\mathbb{P}[\operatorname{Bin}(n, 1/2 + a_2/(2\sqrt{n})) - \operatorname{Bin}(n-\ell, 1/2 + a_1/(2\sqrt{n})) = t] = \frac{1}{\sqrt{\pi n}} \exp\left(-\left(\frac{a_2 - a_1}{2} + \frac{-t+\ell/2}{\sqrt{n}}\right)^2 + O_{C,|\ell|}\left(\frac{1}{n} + \frac{t^4}{n^3}\right)\right) \pm \exp(-\Omega((\log n)^2)).$$

Proof. If $|t| \ge \sqrt{n} \log n/4$ then the result easily follows by the Azuma–Hoeffding inequality so assume $|t| < \sqrt{n} \log n/4$. Consider the range of values $k \in [0, n]$ and let $\tau = k - n/2$. Note that $\mathbb{P}[\operatorname{Bin}(n, 1/2 + a_2/(2\sqrt{n})) - \operatorname{Bin}(n - \ell, 1/2 + a_1/(2\sqrt{n})) = t]$

$$\begin{split} &= \sum_{|k-n/2| \le \sqrt{n} \log n} \binom{n}{k} \binom{n-\ell}{k-1} \left(\frac{1}{2} + \frac{a_2}{2\sqrt{n}} \right)^k \left(\frac{1}{2} - \frac{a_2}{2\sqrt{n}} \right)^{n-k} \left(\frac{1}{2} + \frac{a_1}{2\sqrt{n}} \right)^{k-\ell} \left(\frac{1}{2} - \frac{a_1}{2\sqrt{n}} \right)^{n-\ell-k+\ell} \\ &\pm \exp(-\Omega((\log n)^2)) \\ &= \sum_{|k-n/2| \le \sqrt{n} \log n} \binom{n}{n/2 + \tau} \binom{n-\ell}{n/2 - t + \tau} \left(\frac{1}{4} - \frac{a_2^2}{4n} \right)^{n/2} \left(\frac{1/2 + a_2/(2\sqrt{n})}{1/2 - a_2/(2\sqrt{n})} \right)^{\tau} \\ &\left(\frac{1}{4} - \frac{a_1^2}{4n} \right)^{(n-\ell)/2} \left(\frac{1/2 + a_1/(2\sqrt{n})}{1/2 - a_1/(2\sqrt{n})} \right)^{\tau+\ell/2-\ell} \pm \exp(-\Omega((\log n)^2)) \\ &= \frac{2}{\pi n} \sum_{|k-n/2| \le \sqrt{n} \log n} \exp\left(-\frac{a_1^2}{2} - \frac{a_2^2}{2} + \frac{2a_2\tau}{\sqrt{n}} + \frac{2a_1(\tau + \ell/2 - t)}{\sqrt{n}} - \frac{2\tau^2}{n} - \frac{2(\tau - t + \ell/2)^2}{n} \right) \\ &\cdot \left(1 + O_{C,|\ell|} \left(\frac{1}{n} + \frac{\tau^4 + t^4}{n^3} \right) \right) \pm \exp(-\Omega((\log n)^2)) \\ &= \frac{e^{-\frac{1}{4n}(\ell - a_1\sqrt{n} + a_2\sqrt{n} - 2t)^2}}{\sqrt{\pi n}} \sum_{|k-n/2| \le \sqrt{n} \log n} \frac{1}{\sqrt{2\pi(n/8)}} \exp\left(-\frac{4}{n} \left(\tau + \frac{\ell - a_1\sqrt{n} - a_2\sqrt{n} - 2t}{4} \right)^2 \right) \\ &\cdot \left(1 + O_{C,|\ell|} \left(\frac{1}{n} + \frac{\tau^4 + t^4}{n^3} \right) \right) \pm \exp(-\Omega((\log n)^2)) \\ &= \frac{1}{\sqrt{\pi n}} \exp\left(- \left(\frac{a_2 - a_1}{2} + \frac{-t + \ell/2}{\sqrt{n}} \right)^2 + O_{C,|\ell|} \left(\frac{1}{n} + \frac{t^4}{n^3} \right) \right) \pm \exp(-\Omega((\log n)^2)) \end{split}$$

where we have used Stirling's formula and that $\exp(2t) - (1+t)/(1-t) = O(t^3)$ for $|t| \le 1/2$ and the approximation of sum via a Riemann integral.

We will now derive a number of further estimates which ultimately are based on Lemma A.1.

Lemma A.2. Fix $C \ge 1$ such that $\max\{|a_1|, |a_2|\} \le C$. Then

$$\mathbb{P}\left[\operatorname{Bin}\left(n-1,\frac{1}{2}+\frac{a_1}{2\sqrt{n}}\right)-\operatorname{Bin}\left(n,\frac{1}{2}+\frac{a_2}{2\sqrt{n}}\right) \ge k\right] = \mathbb{P}_{Z\sim\mathcal{N}(0,1)}\left[Z \ge \frac{k\sqrt{2}}{\sqrt{n}}+\frac{a_2-a_1}{\sqrt{2}}\right] + O_C\left(\frac{1}{n}\right).$$
Proof. Applying Lemma A.1 we have that

Proof. Applying Lemma A.1 we have that

$$\mathbb{P}\left[\operatorname{Bin}\left(n-1,\frac{1}{2}+\frac{a_{1}}{2\sqrt{n}}\right)-\operatorname{Bin}\left(n,\frac{1}{2}+\frac{a_{2}}{2\sqrt{n}}\right)\geq k\right] \\
=\mathbb{P}\left[\operatorname{Bin}\left(n,\frac{1}{2}+\frac{a_{2}}{2\sqrt{n}}\right)-\operatorname{Bin}\left(n-1,\frac{1}{2}+\frac{a_{1}}{2\sqrt{n}}\right)\leq -k\right] \\
=\sum_{-\sqrt{n}\log n\leq t\leq -k}\frac{1}{\sqrt{\pi n}}\exp\left(-\left(\frac{a_{2}-a_{1}}{2}+\frac{-t+1/2}{\sqrt{n}}\right)^{2}+O_{C}\left(\frac{1}{n}+\frac{|t|^{4}}{n^{3}}\right)\right)\pm\exp(-\Omega((\log n)^{2}))$$

$$= \sum_{-\sqrt{n}\log n \le t \le -k} \frac{1}{\sqrt{\pi n}} \exp\left(-\left(\frac{a_2 - a_1}{2} + \frac{-t + 1/2}{\sqrt{n}}\right)^2\right) + O_C\left(\frac{1}{n}\right)$$
$$= \int_{-\infty}^{-k+1/2} \frac{1}{\sqrt{\pi n}} \exp\left(-\left(\frac{a_2 - a_1}{2} + \frac{-t + 1/2}{\sqrt{n}}\right)^2\right) dt + O_C\left(\frac{1}{n}\right)$$
$$= \int_{-\infty}^{-k} \frac{1}{\sqrt{\pi n}} \exp\left(-\left(\frac{a_1 - a_2}{2} + \frac{t}{\sqrt{n}}\right)^2\right) dt + O_C\left(\frac{1}{n}\right)$$
$$= \int_{-\infty}^{\frac{-k\sqrt{2}}{\sqrt{n}}} \frac{1}{\sqrt{2\pi}} \exp\left(-\left(\frac{a_1 - a_2}{2} + \frac{z}{\sqrt{2}}\right)^2\right) dz + O_C\left(\frac{1}{n}\right)$$
$$= \mathbb{P}_{Z \sim \mathcal{N}(0,1)} \left[Z \ge \frac{k\sqrt{2}}{\sqrt{n}} + \frac{a_2 - a_1}{\sqrt{2}}\right] + O_C\left(\frac{1}{n}\right)$$

where to replace the sum by the corresponding integral we have used midpoint rule for approximating integrals (which in turn easily follows from the Euler–Maclaurin formula). \Box

For the second estimate, we define notation for ease of use.

Definition A.3. Given $\gamma \in [-1, 1]$ and $\beta \in (0, 1)$ and $a_1, a_2, a_3, a_4 \in \mathbb{R}$, we define

$$g_{\gamma}(\beta, a_1, a_2, a_3, a_4) = \mathbb{P}_{Z_i \sim \mathcal{N}(0, 1)} \left[\sqrt{\frac{\beta}{2}} Z_1 + \sqrt{\frac{1 - \beta}{2}} Z_2 \ge \gamma + \frac{(a_3 - a_1)\beta + (a_4 - a_2)(1 - \beta)}{2} \right]$$
$$\wedge \sqrt{\frac{\beta}{2}} Z_1 - \sqrt{\frac{1 - \beta}{2}} Z_2 \ge \gamma + \frac{(a_3 - a_1)\beta - (a_4 - a_2)(1 - \beta)}{2} \right]$$

We will often suppress the γ in the notation. We see $g(\beta, a_1, 0, 0, 0) = f(\beta, -a_1\beta/2)$ with notation as in Definition 1.4.

Lemma A.4. Fix $C \ge 1$. Let $|a_i| \le C$ for $i \in [4]$ and $k = \beta n$ where $\min(\beta, 1 - \beta) \ge 1/C$. Furthermore let $X_1 \sim \operatorname{Bin}\left(k - 1, \frac{1}{2} + \frac{a_1}{2\sqrt{n}}\right)$, $X_2 \sim \operatorname{Bin}\left(n - k, \frac{1}{2} + \frac{a_2}{2\sqrt{n}}\right)$, $X_3 \sim \operatorname{Bin}\left(k, \frac{1}{2} + \frac{a_3}{2\sqrt{n}}\right)$, and $X_4 \sim \operatorname{Bin}\left(n - k, \frac{1}{2} + \frac{a_4}{2\sqrt{n}}\right)$. Finally let Γ be an integer and $\gamma = \Gamma/\sqrt{n}$. Then

$$\mathbb{P}[X_1 - X_3 + X_2 - X_4 \ge \Gamma \land X_1 - X_3 - X_2 + X_4 \ge \Gamma] \\= g_{\gamma}(\beta, a_1, a_2, a_3, a_4) + O_C\left(\frac{1}{n}\right).$$

Proof. Let $T_1 = X_1 - X_3$ and $T_2 = X_2 - X_4$. By Lemma A.1 we have

$$\begin{split} \mathbb{P}[X_1 - X_3 + X_2 - X_4 \ge \Gamma \land X_1 - X_3 - X_2 + X_4 \ge \Gamma] &= \sum_{t_1 \pm t_2 \ge \Gamma} \mathbb{P}[T_1 = t_1] \mathbb{P}[T_2 = t_2] \\ &= \sum_{\Gamma \le t_1 \pm t_2 \le \sqrt{n} \log n} \frac{1}{\pi \sqrt{k(n-k)}} \exp\left(-\left(\frac{(a_3 - a_1)\sqrt{k}}{2\sqrt{n}} + \frac{t_1 + 1/2}{\sqrt{k}}\right)^2 - \left(\frac{(a_4 - a_2)\sqrt{n-k}}{2\sqrt{n}} + \frac{t_2}{\sqrt{n-k}}\right)^2 + O_C\left(\frac{1}{n} + \frac{t_1^4 + t_2^4}{n}\right)\right) + \exp(-\Omega((\log n)^2)) \\ &= \sum_{\Gamma \le t_1 \pm t_2 \le \sqrt{n} \log n} \frac{1}{\pi \sqrt{k(n-k)}} \exp\left(-\left(\frac{(a_3 - a_1)\sqrt{k}}{2\sqrt{n}} + \frac{t_1 + 1/2}{\sqrt{k}}\right)^2 - \left(\frac{(a_4 - a_2)\sqrt{n-k}}{2\sqrt{n}} + \frac{t_2}{\sqrt{n-k}}\right)^2\right) + O_C\left(\frac{1}{n}\right) \end{split}$$

$$\begin{split} &= \sum_{\Gamma \leq t_1 \pm t_2 \leq \sqrt{n} \log n} \int_{t_1 - 1/2}^{t_1 + 1/2} \int_{t_2 - 1/2}^{t_2 + 1/2} \frac{1}{\pi \sqrt{k(n - k)}} \exp\left(-\left(\frac{(a_3 - a_1)\sqrt{k}}{2\sqrt{n}} + \frac{\ell + 1/2}{\sqrt{k}}\right)^2\right) \\ &\quad - \left(\frac{(a_4 - a_2)\sqrt{n - k}}{2\sqrt{n}} + \frac{m}{\sqrt{n - k}}\right)^2\right) dm d\ell + O_C\left(\frac{1}{n}\right) \\ &= \sum_{\Gamma \leq t_1 \pm t_2 \leq \sqrt{n} \log n} \int_{t_1}^{t_1 + 1} \int_{t_2 - 1/2}^{t_2 + 1/2} \frac{1}{\pi \sqrt{k(n - k)}} \exp\left(-\left(\frac{(a_3 - a_1)\sqrt{k}}{2\sqrt{n}} + \frac{\ell}{\sqrt{k}}\right)^2\right) \\ &\quad - \left(\frac{(a_4 - a_2)\sqrt{n - k}}{2\sqrt{n}} + \frac{m}{\sqrt{n - k}}\right)^2\right) dm d\ell + O_C\left(\frac{1}{n}\right) \\ &= \sum_{\Gamma \leq t_1 \pm t_2 \leq \sqrt{n} \log n} \int_{t_1/\sqrt{k}}^{(t_1 + 1)/\sqrt{k}} \int_{(t_2 - 1/2)/\sqrt{n - k}}^{(t_2 + 1/2)/\sqrt{n - k}} \frac{1}{\pi} \exp\left(-\left(\frac{(a_3 - a_1)\sqrt{\beta}}{2} + \ell\right)^2\right) \\ &\quad - \left(\frac{(a_4 - a_2)\sqrt{1 - \beta}}{2} + m\right)^2\right) dm d\ell + O_C\left(\frac{1}{n}\right) \\ &= \sum_{\Gamma \leq t_1 \pm t_2 \leq \infty} \int_{t_1/\sqrt{k}}^{(t_1 + 1)/\sqrt{k}} \int_{(t_2 - 1/2)/\sqrt{n - k}}^{(t_2 + 1/2)/\sqrt{n - k}} \frac{1}{\pi} \exp\left(-\left(\frac{(a_3 - a_1)\sqrt{\beta}}{2} + \ell\right)^2\right) \\ &\quad - \left(\frac{(a_4 - a_2)\sqrt{1 - \beta}}{2} + m\right)^2\right) dm d\ell + O_C\left(\frac{1}{n}\right) \\ &= \int_{\Gamma \leq \ell\sqrt{k} \pm m\sqrt{n - k}} \frac{1}{\pi} \exp\left(-\left(\frac{(a_3 - a_1)\sqrt{\beta}}{2} + \ell\right)^2 - \left(\frac{(a_4 - a_2)\sqrt{1 - \beta}}{2} + m\right)^2\right) d\ell dm + O_C\left(\frac{1}{n}\right) \\ &= \int_{\Gamma \leq \ell\sqrt{k} \pm m\sqrt{n - k}} \frac{1}{\pi} \exp\left(-\left(\frac{(a_3 - a_1)\sqrt{\beta}}{2} + \ell\right)^2 - \left(\frac{(a_4 - a_2)\sqrt{1 - \beta}}{2} + m\right)^2\right) d\ell dm + O_C\left(\frac{1}{n}\right) \\ &= \mathbb{P}_{Z_i \sim \mathcal{N}(0,1)} \left[\sqrt{\frac{\beta}{2}}Z_1 + \sqrt{\frac{1 - \beta}{2}}Z_2 \geq \frac{\Gamma}{\sqrt{n}} + \frac{(a_3 - a_1)\beta - (a_4 - a_2)(1 - \beta)}{2} + O_C\left(\frac{1}{n}\right). \end{split}$$

In the second-to-last line, one can see the necessary equality by considering the error terms at the line $\Gamma = \ell \sqrt{k} + m \sqrt{n-k}$ versus the given boxes (similar to the Euler-Maclaurin transference between sum and integral in the proof of Lemma A.2).

APPENDIX B. COMPUTER ASSISTED VERIFICATION

We now proceed via a delicate computer assisted verification in order to prove Assumption 1.5. We state a series of numerical claims which will be used to prove Assumption 1.5. Notice by symmetry of the variational problem that we may assume that $\beta \in [0, 1/2]$. The first claim will handle the most numerically unstable part of the claim when β is contained in the initial segment $\beta \in [0, .001]$.

Claim B.1. Let $\gamma \in [\gamma_{crit} - \varepsilon_{1.5}, \gamma_{crit} + \varepsilon_{1.5}]$. Then

$$\sup_{\substack{\beta \in [0,.001]\\\alpha_1,\alpha_2 \in \mathbb{R}}} F_2(\beta,\alpha_1,\alpha_2) = 2 \sup_{\alpha \in \mathbb{R}} F_1(\alpha).$$

We next control the case when $\beta \in [.001, .495]$.

Claim B.2. Let $\gamma \in [\gamma_{\text{crit}} - \varepsilon_{1,5}, \gamma_{\text{crit}} + \varepsilon_{1,5}]$. Then

$$\sup_{\substack{\beta \in [.001,.495]\\\alpha_1,\alpha_2 \in \mathbb{R}}} F_2(\beta, \alpha_1, \alpha_2) < -10^{-5}.$$

We next localize the region of interest to $\beta \in [.495, .5]$ and $\alpha_i \in [-.449, -.441]$.

Claim B.3. Let $\gamma \in [\gamma_{\text{crit}} - \varepsilon_{1,5}, \gamma_{\text{crit}} + \varepsilon_{1,5}]$. Then

$$\sup_{\substack{\beta \in [.495,.5]\\ (\alpha_1, \alpha_2) \notin [-.449, -.441]^2}} F_2(\beta, \alpha_1, \alpha_2) < 10^{-6}$$

Finally in this local region we check that the associated Hessian is strictly negative definite.

Claim B.4. Let $\gamma \in [\gamma_{\text{crit}} - \varepsilon_{1,5}, \gamma_{\text{crit}} + \varepsilon_{1,5}]$. There exists an absolute constant $\delta > 0$ such that $F_2(\beta, \alpha_1, \alpha_2) \colon \mathbb{R}^3 \to \mathbb{R} \text{ satisfies}$

$$\begin{pmatrix} \frac{\partial^2 F_2}{\partial \alpha_1^2} & \frac{\partial^2 F_2}{\partial \alpha_1 \partial \alpha_2} & \frac{\partial^2 F_2}{\partial \alpha_1 \partial \beta} \\ \frac{\partial^2 F_2}{\partial \alpha_1 \partial \alpha_2} & \frac{\partial^2 F_2}{\partial \alpha_2^2} & \frac{\partial^2 F_2}{\partial \alpha_2 \partial \beta} \\ \frac{\partial^2 F_2}{\partial \alpha_1 \partial \beta} & \frac{\partial^2 F_2}{\partial \alpha_2 \partial \beta} & \frac{\partial^2 F_2}{\partial \beta^2} \end{pmatrix} \preceq -\delta I_3$$

for $\beta \in [.495, .505]$ and $\alpha_i \in [-.449, -.441]$, where I_3 is the 3×3 identity matrix and \leq denotes the semidefinite order.

We now deduce the result given the claims outlined above.

Proof of Assumption 1.5. We first notice that

$$\sup_{\substack{\beta \in [0,1]\\ \alpha_1, \alpha_2 \in \mathbb{R}}} F_2(\beta, \alpha_1, \alpha_2) \ge \max(4 \sup_{\alpha \in \mathbb{R}} F_1(\alpha), 2 \sup_{\alpha \in \mathbb{R}} F_1(\alpha)).$$

This follows by evaluating $F_2(\beta, \alpha_1, \alpha_2)$ at $(\beta, \alpha_1, \alpha_2)$ at $(1/2, \alpha(\gamma), \alpha(\gamma))$ and $(0, -1, \alpha(\gamma))$ respectively. Furthermore note that for $\gamma \in [\gamma_{\text{crit}} - \varepsilon_{1.5}, \gamma_{\text{crit}} + \varepsilon_{1.5}]$, we have

$$\log 2 + \sup_{\alpha \in \mathbb{R}} -\alpha^2 + \mathbb{P}_{Z \sim \mathcal{N}(0,1)}[Z \ge (\gamma + \alpha)\sqrt{2}] = \log 2 + \sup_{\alpha \in \mathbb{R}} -\alpha^2 + \mathbb{P}_{Z \sim \mathcal{N}(0,1)}[Z \ge (\gamma_{\text{crit}} + \alpha)\sqrt{2}] \pm \varepsilon_{1.5}$$
$$\in [-\varepsilon_{1.5}, \varepsilon_{1.5}]$$

where we have used that the derivative of $x \to \mathbb{P}[Z \ge x]$ is bounded by $1/\sqrt{2\pi}$ in magnitude. So $\sup_{\alpha \in \mathbb{R}} F_1(\alpha) \in [-\varepsilon_{1.5}, \varepsilon_{1.5}]$ for γ in this range.

Therefore by Claim B.2 and $\varepsilon_{1,5} = 10^{-25}$ it suffices to consider when $\beta \in [0, .001]$ or when $\beta \in [.495, .5]$. For the former case note that the result follows immediately from Claim B.1. For the latter case note that it suffices to check $\beta \in [.495, .5]$ and $\alpha_1, \alpha_2 \in [-.449, -.441]$ by Claim B.3. Note that this also implies for $\gamma \in [\gamma_{\text{crit}} - \varepsilon_{1.5}, \gamma_{\text{crit}} + \varepsilon_{1.5}]$ that $\alpha(\gamma) \in [-.449, -.441]$.

Finally note that F_2 is symmetric under the transformation $F_2(\beta, \alpha_1, \alpha_2) = F_2(1 - \beta, \alpha_2, \alpha_1)$. This implies that the gradient of F_2 vanishes at $(1/2, \alpha(\gamma), \alpha(\gamma))$. By Claim B.4 we know that F_2 is concave on this region, so the desired result follows immediately.

We now proceed with the proof of each individual claim in the remaining subsections. The proof of each of these claims is computer-assisted, but the vast majority of the computational effort occurs in Claim B.2 and Claim B.3.

B.1. Properties of $f(\beta, \alpha)$. We collect a series of general properties of $f(\beta, \alpha)$ which will be used throughout the verification procedure. The first is an explicit formula for $f(\beta, \alpha)$ which serves to make it more amenable to direct computation.

Claim B.5. We have that

$$f(\beta,\alpha) = \int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} \frac{1}{2\pi\sqrt{\beta(1-\beta)}} \exp\left(-\frac{1}{4\beta(1-\beta)}(t_1^2+t_2^2) - \frac{1/2-\beta}{\beta(1-\beta)}t_1t_2\right) dt_1 dt_2$$
39

Proof. This is an immediate consequence of the fact that $T_1 = \sqrt{\beta/2}Z_1 + \sqrt{(1-\beta)/2}Z_2$ and $T_2 = \sqrt{\beta/2}Z_1 - \sqrt{(1-\beta)/2}Z_2$ are jointly Gaussian each having variance 1/2 and with covariance $\beta - 1/2$ between the two. In particular, T_1, T_2 have covariance matrix Σ where

$$\Sigma = \begin{bmatrix} 1/2 & \beta - 1/2 \\ \beta - 1/2 & 1/2 \end{bmatrix}, \quad \Sigma^{-1} = \frac{1}{\beta(1-\beta)} \begin{bmatrix} 1/2 & 1/2 - \beta \\ 1/2 - \beta & 1/2 \end{bmatrix}$$

and thus

$$f(\beta, \alpha) = \int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} \frac{1}{2\pi\sqrt{\beta(1-\beta)}} \exp\left(-\frac{1}{4\beta(1-\beta)}(t_1^2 + t_2^2) - \frac{1/2 - \beta}{\beta(1-\beta)}t_1t_2\right) dt_1 dt_2.$$

We will also require the following substantially more efficient and more numerically stable version of Claim B.5. The gain in efficiency stems from the fact that two-dimensional integrals are substantially more difficult to compute and less accurate than their one-dimensional counterparts.

Claim B.6. We have that

$$f(\beta,\alpha) = \frac{1}{2\pi} \int_0^{\rho} (1-x^2)^{-1/2} \exp\left(\frac{-2(\gamma+\alpha)^2}{1+x}\right) dx + \left(\int_{-\infty}^{-\sqrt{2}(\gamma+\alpha)} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx\right)^2$$

where $\rho = 2\beta - 1$.

Proof. This is precisely [12, (6)] with $\rho = 2\beta - 1$ and $h = k = (\gamma + \alpha)\sqrt{2}$ (note that the Gaussians T_1, T_2 in the proof of Claim B.5 have variance 1/2, not 1, hence the normalization).

We note that when computing $f(\beta, \alpha)$ and the associated derivatives (see Claim B.8) in the rigorous python-flint, we must truncate various integral to a large finite region. We truncate such integrals, treating replacing ∞ by a cutoff amount (10⁸) and straightforward bounds verify that this truncation is easily absorbed in the reported error bounds.

A crucial portion of our analysis will rely on the following upper and lower bounds for $f(\beta, \alpha)$ when β and α are restricted in particular intervals. These will allow us to bound $F_2(\beta, \alpha_1, \alpha_2)$ by a uniform function which handles all β in a specified interval at once. This is crucial as the associated envelope function on the interval will furthermore be convex in α_1, α_2 and therefore finding the corresponding maximum will be amenable to a fixed-point iteration. We return to the precise form of this iteration in the subsequent subsections. In the statement below, the probabilistic event $X \pm Y \ge t$ will serve as shorthand notation for $X + Y \ge t \land X - Y \ge t$.

Lemma B.7. Fix γ and suppose that $\beta \in [\eta_1, \eta_2] \subseteq [0, 1]$. If $-\gamma \leq \alpha \leq 0$ then

 $f(\eta_1, \alpha) \le f(\beta, \alpha) \le f(\eta_2, \alpha).$

Else if $\alpha \leq -\gamma$ then

$$\mathbb{P}\left[\sqrt{\frac{\eta_1}{2}}Z_1 \pm \sqrt{\frac{1-\eta_1}{2}}Z_2 \ge \sqrt{\frac{\eta_1}{\eta_2}}(\gamma+\alpha)\right] \le f(\beta,\alpha) \le \mathbb{P}\left[\sqrt{\frac{\eta_2}{2}}Z_1 \pm \sqrt{\frac{1-\eta_2}{2}}Z_2 \ge \sqrt{\frac{\eta_2}{\eta_1}}(\gamma+\alpha)\right].$$

Proof. We consider the upper bound for the second inequality (the case $\alpha \leq -\gamma$). Note that $f(\beta, \alpha)$ corresponds to integrating the standard bivariate normal density between the lines $\sqrt{\frac{\beta}{2}}z_1 + \sqrt{\frac{1-\beta}{2}}z_2 \geq \gamma + \alpha$ and $\sqrt{\frac{\beta}{2}}z_1 - \sqrt{\frac{1-\beta}{2}}z_2 \geq \gamma + \alpha$. Note that these lines intersect at the point $(\sqrt{\frac{2}{\beta}}(\gamma + \alpha), 0)$ and the lines emanating from this point have slopes $\sqrt{\frac{\beta}{1-\beta}}$ and $-\sqrt{\frac{\beta}{1-\beta}}$. Note that shifting β to η_2 and the intersection point of the lines to $(\sqrt{\frac{2}{\eta_1}}(\gamma + \alpha), 0)$ gives a strictly larger region and corresponds to the upper bound. The remaining three inequalities are obtained in an identical manner.

We will also require the following explicit formula for the partial derivative of $f(\beta, \alpha)$ in α .

Claim B.8. We have that

$$\partial_a f(\beta, a) = -\frac{2}{\sqrt{\pi}} \exp(-(\gamma + a)^2) \mathbb{P}[\sqrt{2\beta(1-\beta)}Z \ge (2-2\beta)(\gamma + a)].$$

Proof. By symmetry we compute that

$$\begin{aligned} \partial_a f(\beta, a) &= -2 \int_{\gamma+a}^{\infty} \frac{1}{2\pi \sqrt{\beta(1-\beta)}} \exp\left(-\frac{1}{4\beta(1-\beta)} (t_1^2 + (\gamma+a)^2) - \frac{1/2-\beta}{\beta(1-\beta)} (\gamma+a)t_1\right) dt_1 \\ &= -\frac{2}{\sqrt{\pi}} \exp(-(\gamma+a)^2) \int_{\gamma+a}^{\infty} \frac{1}{\sqrt{4\pi\beta(1-\beta)}} \exp\left(-\frac{(t+(1-2\beta)(\gamma+a))^2}{4\beta(1-\beta)}\right) dt \\ &= -\frac{2}{\sqrt{\pi}} \exp(-(\gamma+a)^2) \mathbb{P}[\sqrt{2\beta(1-\beta)}Z \ge (2-2\beta)(\gamma+a)]. \end{aligned}$$

Finally, we will repeatedly require the following elementary estimate that a point with small derivative can be used, when combined with an *a priori* second derivative estimate and a gradient bound, to derive an upper bound on the function.

Lemma B.9. Let $f: \mathbb{R} \to \mathbb{R}$ such that $\frac{\partial^2}{\partial x^2} f(x) \leq -M < 0$. Then

$$\sup_{x \in \mathbb{R}} f(x) \le \inf_{z \in \mathbb{R}} f(z) + \frac{f'(z)^2}{2M}.$$

Proof. Given any $z \in \mathbb{R}$, a straightforward application of Taylor's theorem shows that

$$f(x) \le f(z) + f'(z)(x-z) - \frac{M}{2}(x-z)^2.$$

The quadratic on the right is maximized when x = z - f'(z)/M, which implies $f(x) \leq f(z) + f'(z)^2/(2M)$. Now taking a supremum over x and infimum over z finishes.

An identical bivariate version of the above claim follows from the one-variable version, considering a line between the points x, z of interest.

Lemma B.10. Let $f: \mathbb{R}^2 \to \mathbb{R}$ such that $\nabla^2 f(x) \preceq -MI_2 \prec 0$, where I_2 is the 2×2 identity matrix and \preceq denotes the semidefinite order. Then

$$\sup_{\vec{x} \in \mathbb{R}^2} f(x) \le \inf_{\vec{z} \in \mathbb{R}^2} f(\vec{z}) + \frac{\|\nabla f(\vec{z})\|_2^2}{2M}.$$

B.2. Bounds on γ_{crit} . In order to proceed with our analysis we will require sufficiently precise bounds on the value of γ_{crit} .

Claim B.11. We have

$$.24841951 \le \gamma_{\rm crit} \le .24841959.$$

Proof. Notice that $F_1(\alpha)$ satisfies $\frac{\partial^2}{\partial \alpha^2} F_1(\alpha) \leq -2$ due to Lemma 6.2.

Letting $\gamma = .24841951$ and $\alpha = -0.445183267$ we have $F_1(\alpha) \ge 4 \cdot 10^{-8}$ by numerical computation and this provides the necessary lower bound on γ_{crit} .

For the upper bound we set $\gamma = .24841959$ and $\alpha = -0.44518333$. We have $F_1(\alpha) \leq -2 \cdot 10^{-8}$ and $|F'_1(\alpha)| \leq 10^{-5}$ by numerical computation. This implies the desired result via Lemma B.9.

The most important technical point for Claim B.11 is hidden under the hood of producing the candidate values for Claim B.11. In theory these bounds could be produced via a repeated bisection procedure, but a more tailored optimization approach exists which relies on the structure of F_1 and F_2 in a precise manner. The key idea is to rely on a fixed-point iteration where the fixed point corresponds to a critical point of our desired function. In particular note that

$$F_1'(\alpha) = -2\alpha - \frac{\exp(-(\gamma + \alpha)^2)}{\sqrt{\pi}\mathbb{P}[Z \ge (\gamma + \alpha)\sqrt{2}]}$$

Rearranging the equation $F'_1(\alpha) = 0$ we are naturally led to an iterative procedure with

$$\alpha_{i+1} = \frac{-\exp(-(\gamma + \alpha_i)^2)}{2\sqrt{\pi}\mathbb{P}[Z \ge (\gamma + \alpha_i)\sqrt{2}]}$$

Starting with even very crude estimates for α_0 the above procedure converges to a numerical fixed point extraordinarily quickly (being that this is in essence a proxy for Newton iteration tailored to this problem).

Furthermore the separable nature of $F_2(\beta, \alpha_1, \alpha_2)$ with respect to the variables α_1, α_2 allows an essentially identical procedure to produce upper bound on $F_2(\beta, \alpha_1, \alpha_2)$ for fixed β . The crucial issue for Lemma B.7 therefore is that the functional form we use is chosen to bound an interval of β values uniformly, while still allowing for a fixed point iteration of the form specified to succeed.

B.3. **Proof of Claim B.1.** For the proof of Claim B.1, the key idea is that for fixed α_1, α_2 the derivative in β of $F_2(\beta, \alpha_1, \alpha_2)$ is $-\infty$ at $\beta = 0$. This is a manifestation of the fact that the model exhibits a frozen 1-RSB structure and the proof is a sufficiently quantitative version of this fact.

Proof of Claim B.1. Notice that

$$F_2(\beta, \alpha_1, \alpha_2) \le 2\log 2 - 2\beta \log \beta - 2(1-\beta)\log(1-\beta) - 2\alpha_2^2 + 2(1-\beta)\log f(1-\beta, \alpha_2).$$

So, letting

$$G_2(\beta, \alpha_2) = 2\log 2 - 2\beta \log \beta - 2(1-\beta)\log(1-\beta) - 2\alpha_2^2 + 2(1-\beta)\log f(1-\beta, \alpha_2),$$

we will prove the stronger claim that

$$\sup_{\substack{\beta \in [0,.001]\\\alpha \gamma \in \mathbb{R}}} G_2(\beta, \alpha_2) = 2 \sup_{\alpha \in \mathbb{R}} F_1(\alpha).$$

We first prove that it suffices to restrict attention to the case where $-.53 \le \alpha_2 \le -.37$. By numerical computation and $\beta \in [0, .001]$, we find

$$G_2(\beta, \alpha_2) \le 2\log 2 + .01582 + 2(1 - .001)(-\alpha_2^2 + \log \mathbb{P}[Z \ge (\gamma + \alpha_2)\sqrt{2}]).$$

Note that this upper bound, call it $H(\alpha_2)$, is a concave function of α_2 by Lemma 6.2. We can check $H(\alpha_2) \leq -10^{-3}$ for $\alpha_2 \in \{-.53, -.37\}$ and $H(-.45) \geq 10^{-2}$ by numerical computation. Thus H is increasing on $(-\infty, -.53]$ and decreasing on $[-.37, +\infty)$, so $H(\alpha_2) \leq -10^{-3}$ for $\alpha_2 \notin [-.53, -.37]$. From the proof of Assumption 1.5 recall that $\sup_{\alpha \in \mathbb{R}} F_1(\alpha) \in [-\varepsilon_{1.5}, \varepsilon_{1.5}]$. Thus we immediately find that it suffices to consider $\alpha_2 \in [-.53, -.37]$.

A trivial inequality and the second part of Lemma B.7 (recall Claim B.11) yields

$$f(1-\beta,\alpha_2) \ge f(1-\beta,-.37) \ge \mathbb{P}\left[\sqrt{\frac{.999}{2}}Z_1 \pm \sqrt{\frac{.001}{2}}Z_2 \ge \sqrt{.999}(\gamma-.37)\right] \ge .538$$

by numerical computation for the final inequality. We next note that

$$\frac{\partial}{\partial\beta}G_2(\beta,\alpha_2) = 2\log(1-\beta) - 2\log\beta - 2\log f(1-\beta,\alpha_2) + \frac{2(1-\beta)}{f(1-\beta,\alpha_2)}\frac{\partial}{\partial\beta}[f(1-\beta,\alpha_2)]$$
$$\leq 1.239 - 2\log\beta + \frac{2(1-\beta)}{f(1-\beta,\alpha_2)}\frac{\partial}{\partial\beta}[f(1-\beta,\alpha_2)]$$

by the previous inequality, numerical computation, and $\beta \in [0, .001]$.

Notice from Claim B.5 and differentiation under the integral sign that

$$\frac{\partial}{\partial\beta}[f(1-\beta,\alpha_2)] = \int_{[\gamma+\alpha,\infty)^2} \frac{P(t_1,t_2,\beta)}{4(\beta(1-\beta))^2} \left(\frac{\exp\left(-\frac{1}{4\beta(1-\beta)}(t_1^2+t_2^2) - \frac{\beta-1/2}{\beta(1-\beta)}t_1t_2\right)}{2\pi\sqrt{\beta(1-\beta)}}\right) dt_1 dt_2$$

where

$$P(t_1, t_2, \beta) = 2(1 - \beta)(\beta)(-1 + 2\beta) + (1 - \beta)^2(t_1 - t_2)^2 - \beta^2(t_1 + t_2)^2$$

$$\leq 2(1 - \beta)\beta^2 + (1 - \beta)^2((t_1 - t_2)^2 - 2\beta).$$

This implies that

Performing an identical change of variables to that in Claim B.5 (in the reverse direction and with β replaced by $1 - \beta$), we find

$$\begin{aligned} \frac{\partial}{\partial\beta}G_2(\beta,\alpha_2) &\leq 1.239 - 2\log\beta + \frac{2(1-\beta)}{f(1-\beta,\alpha_2)}\frac{\partial}{\partial\beta}[f(1-\beta,\alpha_2)] \\ &\leq 2.239 - 2\log\beta + \frac{(1-\beta)}{\beta}\frac{\mathbb{E}\Big[(Z_2^2 - 1)\mathbb{1}\Big[\sqrt{\frac{1-\beta}{2}}Z_1 \pm \sqrt{\frac{\beta}{2}}Z_2 \geq \gamma + \alpha_2\Big]\Big]}{f(1-\beta,\alpha_2)}. \end{aligned}$$

Notice that a trivial inequality and Lemma B.7 yield

$$f(1-\beta,\alpha_2) \le f(1-\beta,-.53) \le \mathbb{P}\left[Z_1 \ge \frac{1}{\sqrt{.999}}(\gamma-.53)\right] \le 0.611$$

by numerical computation.

Now we find $\mathbb{E}\left[(Z_2^2-1)\mathbb{1}\left[\sqrt{\frac{1-\beta}{2}}Z_1\pm\sqrt{\frac{\beta}{2}}Z_2\geq\gamma+\alpha_2\right]\right]\leq 0$ since it is even true for any fixed value of Z_1 (conditioning a Gaussian to lie in a symmetric interval containing 0 only decreases its variance). In fact, using the transformation $z_1 = \sqrt{2/(1-\beta)}(\gamma+\alpha_2) + \sqrt{\beta/(1-\beta)}t$ we have the inequality

$$\begin{split} \mathbb{E}\Big[(Z_2^2 - 1)\mathbb{1}\Big[\sqrt{\frac{1 - \beta}{2}}Z_1 \pm \sqrt{\frac{\beta}{2}}Z_2 \ge \gamma + \alpha_2\Big]\Big] \\ &= \sqrt{\frac{\beta}{1 - \beta}}\int_{t\ge 0} \frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{1}{2}\Big(\sqrt{\frac{2}{1 - \beta}}(\gamma + \alpha_2) + t\sqrt{\frac{\beta}{1 - \beta}}\Big)^2\Big)\int_{-t}^t \frac{(x^2 - 1)e^{-x^2/2}}{\sqrt{2\pi}}dxdt \\ &\le \frac{\exp(-(\gamma - .53)^2/.999)}{\sqrt{2\pi}} \cdot \sqrt{\beta}\int_0^{10}\int_{-t}^t \frac{(x^2 - 1)e^{-x^2/2}}{\sqrt{2\pi}}dxdt \\ &\le -0.29\sqrt{\beta} \end{split}$$

by numerical computation for the last line. In the first inequality, we used that for $t \in [0, 10]$ we have $\gamma - .53 \leq z_1 \leq -(\gamma - .53)$ by numerical computation. We also used the earlier observation that conditioning a Gaussian to lie in [-t, t] will only decrease its variance, which corresponds to the fact $\int_{-t}^{t} (x^2 - 1)e^{-x^2/2}/\sqrt{2\pi}dx < 0$.

Putting together the previous three centered inequalities we have that

$$\frac{\partial}{\partial\beta}G_2(\beta,\alpha_2) \le 2.239 - 2\log\beta + \frac{1.635}{\beta}\mathbb{E}\Big[(Z_2^2 - 1)\mathbb{1}\Big[\sqrt{\frac{1-\beta}{2}}Z_1 \pm \sqrt{\frac{\beta}{2}}Z_2 \ge \gamma + \alpha_2\Big]\Big] \le 2.239 - \log\beta - \frac{.47}{\sqrt{\beta}}$$

by numerical computation.

The integral of the right side above from $\beta = 0$ to .001 is strictly negative by numerical computation. Therefore, the maximum of $G_2(\beta, \alpha_2)$ is achieved when $\beta = 0$ for $\alpha_2 \in [-.53, -.37]$. The result then follows by definition of F_1 and of G_2 . B.4. **Proof of Claim B.2.** Note that since $f(\beta, \alpha)$ is a decreasing function of γ it suffices to verify Claim B.2 when $\gamma = .2484195 \le \gamma_{\text{crit}} - \varepsilon_{1.5}$ (where we have used Claim B.11).

B.4.1. Initial segment of Claim B.2. We first handle the case where $\beta \in [.001, .005]$ via dropping the variable α_1 in analogy with the proof of Claim B.1; we do this since $f(\beta, \alpha_1)$ is numerically sensitive to compute for sufficiently small values of β .

Claim B.12. Let $\gamma = .2484195$. We have

$$\sup_{\substack{\beta \in [.001,.005]\\\alpha_1,\alpha_2 \in \mathbb{R}\\\alpha_2 \leq -\gamma}} F_2(\beta,\alpha_1,\alpha_2) \leq -10^{-3}.$$

Proof. Suppose that $\beta \in [\eta_1, \eta_2]$ with $\eta_2 \leq 1/2$ and $\alpha_2 \leq -\gamma$. Notice that

$$\begin{aligned} F_2(\beta, \alpha_1, \alpha_2) &= 2\log 2 - 2\beta \log \beta - 2(1-\beta) \log(1-\beta) - 2\alpha_1^2 - 2\alpha_2^2 \\ &+ 2\beta \log f(\beta, \alpha_1) + 2(1-\beta) \log f(1-\beta, \alpha_2) \\ &\leq 2\log 2 - 2\beta \log \beta - 2(1-\beta) \log(1-\beta) - 2\alpha_2^2 + 2(1-\beta) \log f(1-\beta, \alpha_2) \\ &\leq 2\log 2 - 2\eta_2 \log \eta_2 - 2(1-\eta_2) \log(1-\eta_2) - 2\alpha_2^2 + 2(1-\eta_2) \log f(1-\beta, \alpha_2) \\ &\leq 2\log 2 - 2\eta_2 \log \eta_2 - 2(1-\eta_2) \log(1-\eta_2) - 2\alpha_2^2 \\ &+ 2(1-\eta_2) \log \left(\mathbb{P} \left[\sqrt{\frac{1-\eta_1}{2}} Z_1 \pm \sqrt{\frac{\eta_1}{2}} Z_2 \ge \sqrt{\frac{1-\eta_1}{1-\eta_2}} (\gamma + \alpha_2) \right] \right) \end{aligned}$$

where we used Lemma B.7 in the last line.

We now segment β into the intervals $[j \cdot 10^{-4}, (j+1) \cdot 10^{-4}]$ for j = 10 to j = 49. For each of these intervals considered, write it as $[\eta_1, \eta_2]$ and note that the above inequality provides a single univariate function in α_2 , independent of β , which we can use to provide an upper bound to $F_2(\beta, \alpha_1, \alpha_2)$ for β in the specified interval and $\alpha_2 \leq -\gamma$. Furthermore, this function satisfies a second derivative guarantee necessary to apply Lemma B.9, which allows us to reduce the checking to a finite numerical computation.

Claim B.13. Let $\gamma = .2484195$. We have

$$\sup_{\substack{\beta \in [.001,.005]\\\alpha_1,\alpha_2 \in \mathbb{R}\\\alpha_2 > -\gamma}} F_2(\beta,\alpha_1,\alpha_2) \le -10^{-2}.$$

Proof. Suppose that $\beta \in [\eta_1, \eta_2]$ with $\eta_2 \leq 1/2$ and $\alpha_2 \geq -\gamma$. Notice that

$$F_{2}(\beta, \alpha_{1}, \alpha_{2}) = 2 \log 2 - 2\beta \log \beta - 2(1 - \beta) \log(1 - \beta) - 2\alpha_{1}^{2} - 2\alpha_{2}^{2} + 2\beta \log f(\beta, \alpha_{1}) + 2(1 - \beta) \log f(1 - \beta, \alpha_{2}) \leq 2 \log 2 - 2\eta_{2} \log \eta_{2} - 2(1 - \eta_{2}) \log(1 - \eta_{2}) - 2\alpha_{2}^{2} + 2(1 - \eta_{2}) \log f(1 - \eta_{1}, \alpha_{2})$$

by Lemma B.7. We now segment β into the intervals $[j \cdot 10^{-3}, (j+1) \cdot 10^{-3}]$ for j = 1 to j = 4. It is important to note that the $\alpha_2 \in \mathbb{R}$ attaining the maximum of the right will lie outside the range where $\alpha_2 \geq -\gamma$.

Let

$$G(\alpha_2) = 2\log 2 - 2\eta_2 \log \eta_2 - 2(1 - \eta_2) \log(1 - \eta_2) - 2\alpha_2^2 + 2(1 - \eta_2) \log f(1 - \eta_1, \alpha_2)$$

and $\alpha_2^* = \arg \max_{\alpha_2 \in \mathbb{R}} G(\alpha_2)$. By the log-concavity of $f(1 - \eta_2, \alpha_2)$ (Lemma 6.3), we have that $G''(\alpha_2) \leq -2$ and thus

$$G(\alpha_2) = G(\alpha_2) + 2(\alpha_2 - \alpha_2^*)^2$$

is concave. As $\widetilde{G}'(\alpha_2^*) = 0$ and \widetilde{G} is concave, we have that $\widetilde{G}(\alpha_2) \leq \widetilde{G}(\alpha_2^*) = G(\alpha_2^*)$ for all α_2 . This implies that

$$G(\alpha_2) \le G(\alpha_2^*) - 2(\alpha_2 - \alpha_2^*)^2$$

and therefore

$$\sup_{\alpha_2 \ge -\gamma} G(\alpha_2) \le \sup_{\alpha_2 \in R} G(\alpha_2) - 2(\alpha_2^* + \gamma)^2 \mathbb{1}_{\alpha_2^* \le -\gamma}.$$

This allows us to incorporate an additional correction term in certain cases. We verify this by finding a "numerical optimizer" of G and certify it is close to α_2^* by checking it has small derivative. Numerical computation finishes.

B.4.2. Upper segment of Claim B.2 in non-critical regions. For $\gamma \approx \gamma_{\rm crit}$ and $\beta \approx 1/2$ we have that the maximizers satisfy $\alpha_1, \alpha_2 \approx -.445$. Therefore if not both $\gamma + \alpha_1 \leq 0$ and $\gamma + \alpha_2 \leq 0$ then given a sufficiently fine decomposition into intervals one can certify the necessary upper bound, even up to $\beta = 1/2$ instead of just .495. We carry out this procedure in this subsection. As before, since we are concerned only with upper bounds we may assume that $\gamma = .2484195$; additionally, we will be increasingly brief in these proofs since they are quite close in nature to the bounds derived in Claims B.12 and B.13.

Claim B.14. Let $\gamma = .2484195$. We have

$$\sup_{\substack{\beta \in [.005,.5]\\ \alpha_1, \alpha_2 \ge -\gamma}} F_2(\beta, \alpha_1, \alpha_2) \le -10^{-3}.$$

Proof. Suppose that $\beta \in [\eta_1, \eta_2]$ with $\eta_2 \leq 1/2$ and $\alpha_1, \alpha_2 \geq -\gamma$. Notice that by Lemma B.7,

$$F_2(\beta, \alpha_1, \alpha_2) \le 2\log 2 - 2\eta_2 \log \eta_2 - 2(1 - \eta_2) \log(1 - \eta_2) - 2\alpha_1^2 + 2\eta_1 \log f(\eta_2, \alpha_1) - 2\alpha_2^2 + 2(1 - \eta_2) \log f(1 - \eta_1, \alpha_2).$$

We now segment β into $[j \cdot 10^{-3}, (j+1) \cdot 10^{-3}]$ for j = 5 to j = 19 and $[j \cdot 10^{-2}, (j+1) \cdot 10^{-2}]$ for j = 2 to j = 49 and incorporate corrections as carried out in Claim B.13 for both α_1 and α_2 when applicable, i.e., when the optimum over all of \mathbb{R} disagrees with the optimum in the desired interval. (Notice that we can perform the optimizations over the two variables independently and sum the resulting values.) Numerical computation finishes.

Claim B.15. Let $\gamma = .2484195$. We have

$$\sup_{\substack{\beta \in [.005,.5]\\\alpha_1 \leq -\gamma, \alpha_2 \geq -\gamma}} F_2(\beta, \alpha_1, \alpha_2) \leq -10^{-3}.$$

Proof. Suppose that $\beta \in [\eta_1, \eta_2]$ with $\eta_2 \leq 1/2$ and $\alpha_1 \leq -\gamma$ and $\alpha_2 \geq -\gamma$. Lemma B.7 gives

$$F_{2}(\beta, \alpha_{1}, \alpha_{2}) \leq 2 \log 2 - 2\eta_{2} \log \eta_{2} - 2(1 - \eta_{2}) \log(1 - \eta_{2}) - 2\alpha_{2}^{2} + 2(1 - \eta_{2}) \log f(1 - \eta_{1}, \alpha_{2}) - 2\alpha_{1}^{2} + 2\eta_{1} \log \left(\mathbb{P} \left[\sqrt{\frac{\eta_{2}}{2}} Z_{1} \pm \sqrt{\frac{1 - \eta_{2}}{2}} Z_{2} \geq \sqrt{\frac{\eta_{2}}{\eta_{1}}} (\gamma + \alpha_{1}) \right] \right)$$

We now segment β into $[j \cdot 10^{-3}, (j+1) \cdot 10^{-3}]$ for j = 5 to j = 19 and $[j \cdot 10^{-2}, (j+1) \cdot 10^{-2}]$ for j = 2 to j = 49 and incorporate the corrections as carried out in Claim B.13 for both α_1 and α_2 when applicable. Numerical computation finishes.

Claim B.16. Let $\gamma = .2484195$. We have

$$\sup_{\substack{\beta \in [.005,.5]\\ \alpha_1 \ge -\gamma, \alpha_2 \le -\gamma}} F_2(\beta, \alpha_1, \alpha_2) \le -10^{-3}.$$

Proof. Suppose that $\beta \in [\eta_1, \eta_2]$ with $\eta_2 \leq 1/2$ and $\alpha_1 \geq -\gamma$ and $\alpha_2 \leq -\gamma$. Lemma B.7 gives

$$F_{2}(\beta,\alpha_{1},\alpha_{2}) \leq 2\log 2 - 2\eta_{2}\log\eta_{2} - 2(1-\eta_{2})\log(1-\eta_{2}) - 2\alpha_{1}^{2} + 2\eta_{1}\log f(\eta_{2},\alpha_{1}) \\ - 2\alpha_{2}^{2} + 2(1-\eta_{2})\log\left(\mathbb{P}\left[\sqrt{\frac{1-\eta_{1}}{2}}Z_{1} \pm \sqrt{\frac{\eta_{1}}{2}}Z_{2} \geq \sqrt{\frac{1-\eta_{1}}{1-\eta_{2}}}(\gamma+\alpha_{2})\right]\right).$$

We now segment β into $[j \cdot 10^{-3}, (j+1) \cdot 10^{-3}]$ for j = 5 to j = 19 and $[j \cdot 10^{-2}, (j+1) \cdot 10^{-2}]$ for j = 2 to j = 49 and incorporate the corrections as specified in Claim B.13 for both α_1 and α_2 when applicable. Numerical computation finishes.

B.4.3. Upper segment of Claim B.2 in critical region. We now handle the final region of Claim B.2. This procedure is essentially identical to the previous claims, but the set of intervals required is substantially more intricate as the maximum, which is near 0, is attained in this region.

Claim B.17. Let $\gamma = .2484195$. We have that

$$\sup_{\substack{\beta \in [.005,.495]\\\alpha_1,\alpha_2 \le -\gamma}} F_2(\beta, \alpha_1, \alpha_2) \le -10^{-5}.$$

Proof. Suppose that $\beta \in [\eta_1, \eta_2]$ with $\eta_2 \leq 1/2$ and $\alpha_1, \alpha_2 \leq -\gamma$. Notice that

$$F_{2}(\beta, \alpha_{1}, \alpha_{2}) \leq 2\log(2) - 2\eta_{2}\log\eta_{2} - 2(1 - \eta_{2})\log(1 - \eta_{2}) \\ - 2\alpha_{1}^{2} + 2\eta_{1}\log\left(\mathbb{P}\left[\sqrt{\frac{\eta_{2}}{2}}Z_{1} \pm \sqrt{\frac{1 - \eta_{2}}{2}}Z_{2} \geq \sqrt{\frac{\eta_{2}}{\eta_{1}}}(\gamma + \alpha_{1})\right]\right) \\ - 2\alpha_{2}^{2} + 2(1 - \eta_{2})\log\left(\mathbb{P}\left[\sqrt{\frac{1 - \eta_{1}}{2}}Z_{1} \pm \sqrt{\frac{\eta_{1}}{2}}Z_{2} \geq \sqrt{\frac{1 - \eta_{1}}{1 - \eta_{2}}}(\gamma + \alpha_{2})\right]\right).$$

The partition for β in this region is substantially more involved; we take $[j \cdot 10^{-3}, (j+1) \cdot 10^{-3}]$ for $\begin{array}{l} j = 5 \text{ to } j = 19, \ [j \cdot 10^{-2}, (j+1) \cdot 10^{-2}] \text{ for } j = 2 \text{ to } j = 24, \ [j \cdot 10^{-3}, (j+1) \cdot 10^{-3}] \text{ for } j = 250 \text{ to } j = 424, \ [j \cdot 10^{-4}, (j+1) \cdot 10^{-4}] \text{ for } j = 4250 \text{ to } j = 4724, \ [j \cdot (5 \cdot 10^4)^{-1}, (j+1) \cdot (5 \cdot 10^4)^{-1}] \text{ for } j = 23625 \text{ to } j = 24299, \ [j \cdot 10^{-5}, (j+1) \cdot 10^{-5}] \text{ for } j = 48600 \text{ to } j = 48799, \ [j \cdot (2 \cdot 10^5)^{-1}, (j+1) \cdot (2 \cdot 10^5)^{-1}] \text{ for } j = 23625 \text{ to } j = 24299, \ [j \cdot 10^{-5}, (j+1) \cdot 10^{-5}] \text{ for } j = 48600 \text{ to } j = 48799, \ [j \cdot (2 \cdot 10^5)^{-1}, (j+1) \cdot (2 \cdot 10^5)^{-1}] \text{ for } j = 23625 \text{ to } j = 24299, \ [j \cdot 10^{-5}, (j+1) \cdot 10^{-5}] \text{ for } j = 48600 \text{ to } j = 48799, \ [j \cdot (2 \cdot 10^5)^{-1}, (j+1) \cdot (2 \cdot 10^5)^{-1}] \text{ for } j = 23625 \text{ to } j = 24299, \ [j \cdot 10^{-5}, (j+1) \cdot 10^{-5}] \text{ for } j = 48600 \text{ to } j = 48799, \ [j \cdot (2 \cdot 10^5)^{-1}, (j+1) \cdot (2 \cdot 10^5)^{-1}] \text{ for } j = 23625 \text{ to } j = 24299, \ [j \cdot 10^{-5}, (j+1) \cdot 10^{-5}] \text{ for } j = 48600 \text{ to } j = 48799, \ [j \cdot (2 \cdot 10^5)^{-1}, (j+1) \cdot (2 \cdot 10^5)^{-1}] \text{ for } j = 23625 \text{ to } j = 48799, \ [j \cdot (2 \cdot 10^5)^{-1}, (j+1) \cdot (2 \cdot 10^5)^{-1}] \text{ for } j = 48000 \text{ to } j = 48799, \ [j \cdot (2 \cdot 10^5)^{-1}, (j+1) \cdot (2 \cdot 10^5)^{-1}] \text{ for } j = 48000 \text{ to } j = 48799, \ [j \cdot (2 \cdot 10^5)^{-1}, (j+1) \cdot (2 \cdot 10^5)^{-1}] \text{ for } j = 48000 \text{ to } j = 4$ j = 97600 to j = 98749, and finally $[j \cdot (4 \cdot 10^5)^{-1}, (j+1) \cdot (4 \cdot 10^5)^{-1}]$ for j = 197500 to j = 197999. Numerical computation finishes.

We are now in position to prove Claim B.2.

Proof of Claim B.2. This is an immediate consequence of combining Claims B.12 to B.17.

B.5. **Proof of Claim B.3.** The proof of Claim B.3 is closely related to the proofs of various claims given in Claim B.2. Note that Claims B.14 to B.16 handle β up to 1/2, so we only need to refine our estimates in the critical region where $\alpha_1, \alpha_2 \leq -\gamma$. Additionally, we may assume $\beta \in [.495, .5]$ by Claim B.17.

However, in the neighborhood of $(\alpha_1, \alpha_2) = (-.445, -.445)$, our function might be positive. So, some care is required to show that it is negative away from a neighborhood of this point. To do this, we show that there is a point in this neighborhood which has not too large value (potentially positive), and small partial derivatives in α_1, α_2 . However, our function is strongly concave. Therefore, moving far enough from this point will force our function to be negative.

Proof of Claim B.3. As before it suffices to prove the claim conditional on $\gamma = .2484195$. Furthermore, it suffices to consider $\alpha_1, \alpha_2 \leq -\gamma$ by Claims B.14 to B.16 and $\beta \in [.495, .5]$ by Claim B.17. Now, if $\beta \in [\eta_1, \eta_2]$ with $\eta_2 \leq 1/2$ and $\alpha_1, \alpha_2 \leq -\gamma$ then

$$F_{2}(\beta, \alpha_{1}, \alpha_{2}) \leq 2 \log 2 - 2\eta_{2} \log \eta_{2} - 2(1 - \eta_{2}) \log(1 - \eta_{2}) \\ - 2\alpha_{1}^{2} + 2\eta_{1} \log \left(\mathbb{P} \left[\sqrt{\frac{\eta_{2}}{2}} Z_{1} \pm \sqrt{\frac{1 - \eta_{2}}{2}} Z_{2} \geq \sqrt{\frac{\eta_{2}}{\eta_{1}}} (\gamma + \alpha_{1}) \right] \right) \\ 46$$

$$-2\alpha_2^2 + 2(1-\eta_2)\log\left(\mathbb{P}\left[\sqrt{\frac{1-\eta_1}{2}}Z_1 \pm \sqrt{\frac{\eta_1}{2}}Z_2 \ge \sqrt{\frac{1-\eta_1}{1-\eta_2}}(\gamma+\alpha_2)\right]\right)$$

=: $G_2(\alpha_1, \alpha_2).$

We break β into segments $[j \cdot (4 \cdot 10^5)^{-1}, (j+1) \cdot (4 \cdot 10^5)^{-1}]$ for j = 198000 to j = 199999. For each interval of the form $[\eta_1, \eta_2]$ using numerical computation we produce a point (α_1^*, α_2^*) such that

=

- $G_2(\alpha_1^*, \alpha_2^*) \le 10^{-5};$ $|\alpha_1^* + .445| \le 1.5 \cdot 10^{-3} \text{ and } |\alpha_2^* + .445| \le 1.5 \cdot 10^{-3};$ $|\frac{\partial}{\partial \alpha_1} G_2(\alpha_1, \alpha_2)|_{(\alpha_1, \alpha_2) = (\alpha_1^*, \alpha_2^*)}| + |\frac{\partial}{\partial \alpha_2} G_2(\alpha_1, \alpha_2)|_{(\alpha_1, \alpha_2) = (\alpha_1^*, \alpha_2^*)}| \le 10^{-5}.$

These conditions together with the strong concavity of G_2 imply the claimed result along with numerical computation. In particular, we use that the second derivatives in α_1, α_2 are uniformly less than -2, and that G_2 is the sum of a function of α_1 and of α_2 separately.

B.6. **Proof of Claim B.4.** In order to prove the desired Hessian bound we prove various estimates on the coordinate second partial derivatives in the region for Claim B.4. To prove the desired claim we will prove the following set of second derivative estimates.

Lemma B.18. Let $\gamma = .2484195$. At points for which $\beta \in [.495, .505]$ and $\alpha_1, \alpha_2 \in [-.45, -.44]$ we have for $i \in \{1, 2\}$ that

$$\frac{\partial^2 F_2}{\partial \beta^2} \le -2.15, \ \left| \frac{\partial^2 F_2}{\partial \beta \partial \alpha_i} \right| \le 2.05.$$

Given the above pair of estimates we now prove the desired result regarding the Hessian of F_2 .

Proof of Claim B.4. Notice that by the log-concavity of the function $f(\beta, \alpha)$ from Lemma 6.3, by direct computation for all γ and β , α_i such that F_2 is defined we have

$$\frac{\partial^2 F_2}{\partial \alpha_i^2} \le -4.$$

Furthermore the functional form of F_2 immediately implies that

$$\frac{\partial^2 F_2}{\partial \alpha_1 \partial \alpha_2} = 0.$$

Additionally, note that a small shift in γ can be equivalently recast as slightly shifting the values of α_1, α_2 when computing $\frac{\partial^2 F_2}{\partial \beta^2}, \frac{\partial^2 F_2}{\partial \beta \partial \alpha_i}$. Thus, applying Lemma B.18, for γ in the specified range and $\alpha_1, \alpha_2 \in [-.449, -.441]$ the second derivative bounds listed in Lemma B.18 still apply.

The desired result then follows via considering the negation of the Hessian and applying Sylvester's criterion on determinants of minors for a matrix to be positive semidefinite. The matrix is strictly negative semidefinite as the determinant is also seen to be strictly bounded away from zero, and compactness guarantees the existence of an absolute constant $\delta > 0$ for the desired result.

The proof of Lemma B.18 is mostly calculation with the function form given in Claim B.5 and various grid procedures to verify certain elementary two-dimensional inequalities are satisfied. Since the inequalities we which to certify are not particularly delicate, the procedure is mechanical, if not completely pleasant.

We will require the following triplet of claims regarding the values of $f(\beta, \alpha)$ and certain specially chosen combinations of derivatives. The combinations naturally fall out of an analysis based on explicit computation; we present these claims as a series of unmotivated numerical claims for the sake of simplifying the verification.

Claim B.19. Let
$$\gamma = .2484195$$
, $\beta \in [.495, .505]$, and $\alpha \in [-.45, -.44]$. Then we have $f(\beta, \alpha) \in [.36544, .37761]$.

Proof. By Lemma B.7 and numerical computation we have

$$f(\beta,\alpha) \le f(\beta, -.45) \le \mathbb{P}\left[\sqrt{\frac{.505}{2}}Z_1 \pm \sqrt{\frac{.495}{2}}Z_2 \ge (\gamma - .45)\sqrt{\frac{.505}{.495}}\right] \le .37761$$

and

$$f(\beta, \alpha) \ge f(\beta, -.44) \ge \mathbb{P}\left[\sqrt{\frac{.495}{2}}Z_1 \pm \sqrt{\frac{.505}{2}}Z_2 \ge (\gamma - .44)\sqrt{\frac{.495}{.505}}\right] \ge .36544.$$

The next claim bounds the size of the first derivative of $f(\beta, \alpha)$ in β ; the proof involves reducing to a certain tractable two-variable integral whose range can be determined via a direct grid search in combination with a crude bounding procedure. As these computations are substantially less numerically delicate we carried these computations out only in scipy.

Claim B.20. Let $\gamma = .2484195$, $\beta \in [.495, .505]$, and $\alpha \in [-.45, -.44]$. Then we have $\frac{\partial f(\beta, \alpha)}{\partial f(\beta, \alpha)} \in [.2780, .3110]$

$$\frac{\partial f(\beta, \alpha)}{\partial \beta} \in [.2780, .3110]$$

Proof. Let $\eta = 1/2 - \beta$. By differentiation under the integral sign in combination with the form presented in Claim B.5 we have that

$$\frac{\partial f(\beta,\alpha)}{\partial \beta} = \int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} \frac{t_1 t_2 + \eta (2t_1^2 + 2t_2^2 - 1) + 4\eta^2 t_1 t_2 + 4\eta^3}{8\pi (\beta (1-\beta))^{5/2}} \\ \times \exp\left(-\frac{1}{4\beta (1-\beta)} (t_1^2 + t_2^2) - \frac{\eta}{\beta (1-\beta)} t_1 t_2\right) dt_1 dt_2.$$

We break this integral into two pieces. Notice that

$$\begin{aligned} \left| \int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} \frac{\eta(2t_1^2 + 2t_2^2 - 1) + 4\eta^2 t_1 t_2 + 4\eta^3}{8\pi(\beta(1-\beta))^{5/2}} \exp\left(-\frac{1}{4\beta(1-\beta)}(t_1^2 + t_2^2) - \frac{\eta}{\beta(1-\beta)}t_1 t_2\right) dt_1 dt_2 \right| \\ \leq \int_{\gamma-.45}^{\infty} \int_{\gamma-.45}^{\infty} \frac{|\eta(2t_1^2 + 2t_2^2 - 1)| + |4\eta^2 t_1 t_2| + |4\eta^3|}{8\pi(\beta(1-\beta))^{5/2}} \exp\left(-\frac{1}{4\beta(1-\beta)}(t_1^2 + t_2^2) - \frac{\eta}{\beta(1-\beta)}t_1 t_2\right) dt_1 dt_2 \\ \leq \int_{\gamma-.45}^{\infty} \int_{\gamma-.45}^{\infty} \frac{|\eta(2t_1^2 + 2t_2^2 - 1)| + |4\eta^2 t_1 t_2| + |4\eta^3|}{8\pi(\beta(1-\beta))^{5/2}} \exp\left(-.99(t_1^2 + t_2^2)\right) dt_1 dt_2 \leq .00975 \end{aligned}$$

by upper bounding $|\eta| \leq .005$ and $(\beta(1-\beta))^{-5/2} \leq (.495 \cdot .505)^{-5/2}$ and then computing the resulting integral numerically. Furthermore, we have

$$\left| \int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} \left(\frac{t_1 t_2}{8\pi (\beta(1-\beta))^{5/2}} - \frac{4t_1 t_2}{\pi} \right) \exp\left(-\frac{1}{4\beta(1-\beta)} (t_1^2 + t_2^2) - \frac{\eta}{\beta(1-\beta)} t_1 t_2 \right) dt_1 dt_2 \right|$$

$$\leq \int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} |t_1 t_2| \left(\frac{1}{8\pi (.495 \cdot .505)^{5/2}} - \frac{4}{\pi} \right) \exp\left(-.99 (t_1^2 + t_2^2) \right) dt_1 dt_2 \leq .0001$$

by numerical computation.

Thus it suffices to understand the value of

$$\int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} \frac{4t_1 t_2}{\pi} \exp\left(-\frac{1}{4\beta(1-\beta)}(t_1^2+t_2^2) - \frac{1/2-\beta}{\beta(1-\beta)}t_1 t_2\right) dt_1 dt_2$$

for $\beta \in [.495, .505]$ and $\alpha \in [-.45, -.44]$. We first handle shifts in α ; notice that the derivative in absolute value of the above expression in α is at most

$$\frac{8}{\pi} \int_{\gamma+\alpha}^{\infty} |(\gamma+\alpha)t_1| \cdot \exp(-.99(\gamma+\alpha)^2 - .99t_1^2) dt_1 \le \frac{8|\gamma-.45|}{\pi} \int_{\gamma-.45}^{\infty} |t_1| \cdot \exp(-.99t_1^2) dt_1 \le .27$$

and thus shifting α to α' the difference in the integral is bounded by $.27|\alpha - \alpha'|$.

Similarly, by the inequality $|\exp(x) - \exp(y)| \le \exp(\max(x, y))|x - y|$, we have that shifting β to β' induces an error bounded by

$$|\beta - \beta'| \int_{\gamma - .45}^{\infty} \int_{\gamma - .45}^{\infty} \frac{4|t_1 t_2|}{\pi} \exp\left(-..99(t_1^2 + t_2^2)\right) (.041(t_1^2 + t_2^2) + 4.01|t_1 t_2|) dt_1 dt_2 \le 1.1|\beta - \beta'|$$

by numerical computation. We take a net over β and α each of granularity .00025. Applying a grid verification and applying the errors given above we find the specified result by numerical computation.

Finally, we will require an upper bound on the following mixed partial derivative expression. The strategy is essentially identical to the previous proof however the precise formulas are substantially less pleasant and therefore the computation is less clean.

Claim B.21. Let $\gamma = .2484195$, $\beta \in [.495, .505]$, and $\alpha \in [-.45, -.44]$. Then we have

$$\frac{\beta \partial^2 f(\beta, \alpha)}{\partial \beta^2} + \frac{2 \partial f(\beta, \alpha)}{\partial \beta} \le .630$$

Proof. Let $\eta = 1/2 - \beta$ and apply differentiation under the integral sign to Claim B.5. We have

$$\frac{\beta \partial^2 f(\beta, \alpha)}{\partial \beta^2} + \frac{2\partial f(\beta, \alpha)}{\partial \beta} = \int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} \frac{P(t_1, t_2, \beta)}{32\pi \beta^{7/2} (1-\beta)^{9/2}} \exp\left(-\frac{1}{4\beta(1-\beta)}(t_1^2 + t_2^2) - \frac{\eta}{\beta(1-\beta)}t_1 t_2\right) dt_1 dt_2$$

where

$$\begin{split} P(t_1, t_2, \beta) \\ &= 16\eta^5(-1+2t_1t_2) + 4\eta^4(1-4t_1t_2+6t_1^2+6t_2^2+4t_1^2t_2^2) + 8\eta^3(1-(t_1-t_2)^2+2t_1t_2(t_1^2+t_2^2)) \\ &\quad + 2\eta^2(-1-2t_1^2-2t_2^2+8t_1^2t_2^2+2t_1^4+2t_2^4) + \eta(-1+2t_1^2-6t_1t_2+2t_2^2+4t_1t_2(t_1^2+t_2^2)) \\ &\quad + 1/4(1-2(t_1-t_2)^2+4t_1^2t_2^2) \\ &\leq \eta^2(4.5t_1^4+4.5t_2^4+17t_1^2t_2^2) + \eta(-1+2t_1^2-6t_1t_2+2t_2^2+4t_1t_2(t_1^2+t_2^2)) \\ &\quad + 1/4(1-2(t_1-t_2)^2+4t_1^2t_2^2) \end{split}$$

where we have very crudely applied $|\eta| \leq .005$ and the AM-GM inequality to eliminate various higher order terms which appear in the coefficients of powers of η . For the sake of simplicity, let

$$Q_1(t_1, t_2, \beta) = \eta^2 (4.5t_1^4 + 4.5t_2^4 + 17t_1^2t_2^2),$$

$$Q_2(t_1, t_2, \beta) = (-1 + 2t_1^2 - 6t_1t_2 + 2t_2^2 + 4t_1t_2(t_1^2 + t_2^2))\eta + 1/4(1 - 2(t_1 - t_2)^2 + 4t_1^2t_2^2).$$

We have

$$\int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} \frac{Q_1(t_1, t_2, \beta)}{32\pi\beta^{7/2}(1-\beta)^{9/2}} \exp\left(-\frac{1}{4\beta(1-\beta)}(t_1^2+t_2^2) - \frac{\eta}{\beta(1-\beta)}t_1t_2\right) dt_1 dt_2 \\
\leq \int_{\gamma-.45}^{\infty} \int_{\gamma-.45}^{\infty} \frac{Q_1(t_1, t_2, \beta)}{32\pi\beta^{7/2}(1-\beta)^{9/2}} \exp\left(-.99(t_1^2+t_2^2)\right) dt_1 dt_2 \leq .0007$$

by numerical computation and $|\eta| \leq .005$.

We next simplify the remaining integral term further. Notice that

$$\left| \int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} \frac{Q_2(t_1, t_2, \beta)}{32\pi\beta^{7/2}(1-\beta)^{9/2}} - \frac{4Q_2(t_1, t_2, \beta)}{\pi(1-\beta)} \exp\left(-\frac{1}{4\beta(1-\beta)}(t_1^2 + t_2^2) - \frac{\eta}{\beta(1-\beta)}t_1t_2\right) dt_1 dt_2 \right|$$

$$\leq 9 \cdot 10^{-4} \cdot \int_{\gamma-.45}^{\infty} \int_{\gamma-.45}^{\infty} |Q_2(t_1, t_2, \beta)| \exp\left(-..99(t_1^2 + t_2^2)\right) dt_1 dt_2 \leq .00041$$

by numerical computation and $|\eta| \leq .005$ and $\alpha \geq -.45$.

Therefore it suffices to understand

$$\int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} \frac{4Q_2(t_1, t_2, \beta)}{\pi(1-\beta)} \exp\left(-\frac{1}{4\beta(1-\beta)}(t_1^2 + t_2^2) - \frac{\eta}{\beta(1-\beta)}t_1t_2\right) dt_1 dt_2.$$

Note that if the above integral is negative we immediately obtain the desired bound, so it suffices to provide an upper bound to

$$\frac{4}{\pi(.495)} \int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} Q_2(t_1, t_2, \beta) \exp\left(-\frac{1}{4\beta(1-\beta)}(t_1^2 + t_2^2) - \frac{\eta}{\beta(1-\beta)}t_1t_2\right) dt_1 dt_2.$$

We wish to replace this with the following simpler integral to consider:

$$\frac{4}{\pi(.495)} \int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} Q_2(t_1, t_2, \beta) \exp\left(-(1+4\eta^2)(t_1^2+t_2^2)-4\eta t_1 t_2\right) dt_1 dt_2.$$
(B.1)

Note that the difference between the above two integrals can be bounded using that $|\exp(x) - \exp(y)| \le |x-y|\exp(\max(x,y))$ and that $\exp(\cdot)$ is an increasing function. Using this and numerical computation, we can find that those integrals differ by at most

$$\frac{4}{\pi(.495)} \int_{\gamma-.45}^{\infty} \int_{\gamma-.45}^{\infty} |Q_2(t_1, t_2, \beta)| \cdot (2 \cdot 10^{-8} \cdot (t_1^2 + t_2^2) + 2 \cdot 10^{-6} |t_1 t_2|) \exp\left(-.99(t_1^2 + t_2^2)\right) dt_1 dt_2 \le 2 \cdot 10^{-6}.$$

We next remove η out of the exponent of (B.1) by comparing it to the integral

$$\frac{4}{\pi(.495)} \int_{\gamma+\alpha}^{\infty} \int_{\gamma+\alpha}^{\infty} Q_2(t_1, t_2, \beta) (1 - 4\eta^2(t_1^2 + t_2^2) - 4\eta t_1 t_2) \exp(-t_1^2 - t_2^2) dt_1 dt_2$$
(B.2)

at the cost of an additive error bounded by

$$\frac{4}{\pi(.495)} \int_{\gamma-.45}^{\infty} \int_{\gamma-.45}^{\infty} |Q_2(t_1, t_2, \beta)| \exp(-t_1^2 - t_2^2) \\ \times |1 - 4\eta^2(t_1^2 + t_2^2) - 4\eta t_1 t_2) - \exp(-4\eta^2(t_1^2 + t_2^2) - 4\eta t_1 t_2)| dt_1 dt_2 \\ < .0003$$

using numerical computation, relying on the inequality that $|\exp(x) - x - 1| \le \exp(|x|) - |x| - 1$.

We now are in position to perform a grid search on (B.2); notice that the derivative in α is bounded by

$$\frac{8}{\pi(.495)} \int_{\gamma+\alpha}^{\infty} \sup_{\substack{t_2 \in [\gamma-.45, \gamma-.44] \\ |\eta| \le .005}} |Q_2(t_1, t_2, 1/2 - \eta)| \cdot |1 - 4\eta^2(t_1^2 + t_2^2) - 4\eta t_1 t_2| \exp(-t_1^2) dt_1$$

and can be computed to be bounded by 4.2 by numerical computation. Thus shifting α to α' causes a shift bounded by $4.2|\alpha - \alpha'|$. To compute the shift in β , notice that $Q_2(t_1, t_2, \beta)(1 - 4\eta^2(t_1^2 + t_2^2) - 4\eta t_1 t_2)$ is a polynomial in β . So, differentiating under the integral sign it suffices to bound

$$\frac{8}{\pi(.495)} \int_{\gamma-.45}^{\infty} \int_{\gamma-.45}^{\infty} \left| \frac{\partial}{\partial \beta} Q_2(t_1, t_2, \beta) (1 - 4\eta^2(t_1^2 + t_2^2) - 4\eta t_1 t_2) \right| \exp(-t_1^2 - t_2^2) dt_1 dt_2.$$

(Recall $\eta = 1/2 - \beta$.) Via a straightforward numerical computation we find that theisabove is bounded by 8.76 and thus shifting β to β' causes a shift bounded by 8.76 $|\beta - \beta'|$.

Take a grid in β of width .00025 and in α of width .00025 and note that the closest grid points are at most half the width away moving in both coordinate directions. By numerical computation and bounding the error of rounding to grid points, we obtain a bound on (B.2) and can use this along with the prior analysis to find that the maximum of the desired value is bounded by .630 as claimed.

We now proceed with the proof of Lemma B.18. As we will see the estimates in this region are not particularly sharp and in particular we can tolerate a substantial amount of numerical error.

Proof of Lemma B.18. Notice that

$$\begin{aligned} \frac{\partial^2 F_2}{\partial \beta^2} &= \frac{-2}{\beta(1-\beta)} + \frac{\partial^2}{\partial \beta^2} (2\beta \log f(\beta,\alpha_1)) + \frac{\partial^2}{\partial \beta^2} (2(1-\beta) \log f(1-\beta,\alpha_2)) \\ &\leq -8 + 4 \max_{\substack{\beta \in [.495,.505]\\\alpha_1 \in [-.45,-.44]}} \frac{\partial^2}{\partial \beta^2} (\beta \log f(\beta,\alpha_1)) \\ &\leq -8 + 4 \max_{\substack{\beta \in [.495,.505]\\\alpha_1 \in [-.45,-.44]}} \frac{\beta \frac{\partial^2}{\partial \beta^2} f(\beta,\alpha_1) + 2 \frac{\partial}{\partial \beta} f(\beta,\alpha_1)}{f(\beta,\alpha_1)} - \frac{\beta (\frac{\partial}{\partial \beta} f(\beta,\alpha_1))^2}{f(\beta,\alpha_1)^2} \\ &\leq -8 + 4 \left(\frac{.630}{.36544} - .495 \left(\frac{.2775}{.37761} \right)^2 \right) \leq -2.17, \end{aligned}$$

using Claims B.19 to B.21 and numerical computation. This completes the proof of the first item of Lemma B.18. For the second item by symmetry (switching α_1, α_2 and replacing β by $1 - \beta$) we may assume i = 1. Notice that by Claim B.8,

where we have used Claim B.19. We note via crude and direct bounding we have that

$$\frac{(\gamma + \alpha_1) \exp(-(\gamma + \alpha_1)^2 (1 - \beta)/\beta)}{2\sqrt{\pi\beta(1 - \beta)}} \in [-.110, -.103].$$

Therefore it suffices to prove that

$$\mathbb{P}\left[Z \ge (\gamma + \alpha_1)\sqrt{\frac{2(1-\beta)}{\beta}}\right] \left(1 - \frac{\beta \frac{\partial}{\partial \beta} f(\beta, \alpha_1)}{f(\beta, \alpha_1)}\right) \in [-.234, .447].$$

Notice that

$$\mathbb{P}\left[Z \ge (\gamma + \alpha_1)\sqrt{\frac{2(1-\beta)}{\beta}}\right] \le \mathbb{P}\left[Z \ge (\gamma - .45)\sqrt{\frac{2(.505)}{.495}}\right] \le .614$$

Thus it suffices to prove that

$$\left|1 - \frac{\beta \frac{\partial}{\partial \beta} f(\beta, \alpha_1)}{f(\beta, \alpha_1)}\right| \in [-.381, .728], \text{ or equivalently} \frac{\beta \frac{\partial}{\partial \beta} f(\beta, \alpha_1)}{f(\beta, \alpha_1)} \in [-.272, .619].$$

This follows from Claims B.19 and B.20 as well as $\beta \in [.495, .505]$. We are (finally) done.

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139, USA *Email address*: minzer.dor@gmail.com, {asah,msawhney}@mit.edu