The Annals of Probability
2017, Vol. 45, No. 2, 1190–1217
DOI: 10.1214/15-AOP1084

EXTREMAL CUTS OF SPARSE RANDOM GRAPHS

BY AMIR DEMBO[1], ANDREA MONTANARI[2] AND SUBHABRATA SEN[3]

Stanford University

For Erdős–Rényi random graphs with average degree γ, and uniformly random γ-regular graph on n vertices, we prove that with high probability the size of both the Max-Cut and maximum bisection are $n(\frac{\gamma}{4} + \mathsf{P}_*\sqrt{\frac{\gamma}{4}} + o(\sqrt{\gamma})) + o(n)$ while the size of the minimum bisection is $n(\frac{\gamma}{4} - \mathsf{P}_*\sqrt{\frac{\gamma}{4}} + o(\sqrt{\gamma})) + o(n)$. Our derivation relates the free energy of the anti-ferromagnetic Ising model on such graphs to that of the Sherrington–Kirkpatrick model, with $\mathsf{P}_* \approx 0.7632$ standing for the ground state energy of the latter, expressed analytically via Parisi's formula.

1. Introduction. Given a graph $G = (V, E)$, a bisection of G is a partition of its vertex set $V = V_1 \cup V_2$ such that the two parts have the same cardinality (if $|V|$ is even) or differ by one vertex (if $|V|$ is odd). The cut size of any partition is defined as the number of edges $(i, j) \in E$ such that $i \in V_1$, and $j \in V_2$. The minimum (maximum) bisection of G is defined as the bisection with the smallest (largest) size and we will denote this size by $\mathsf{mcut}(G)$ [respectively $\mathsf{MCUT}(G)$]. The related Max-Cut problem seeks to partition the vertices into two parts such that the cut size is maximized. We will denote the size of the Max-Cut by $\mathsf{MaxCut}(G)$. The study of these features is fundamental in combinatorics and theoretical computer science. These properties are also critical for a number of practical applications. For example, minimum bisection is relevant for a number of graph layout and embedding problems [14]. For practical applications of Max-Cut, see [41]. On the other hand, it is hard to even approximate these quantities in polynomial time (see, for instance [16, 25, 26, 31]).

The average case analysis of these features is also of considerable interest. For example, the study of random graph bisections is motivated by the desire to justify and understand various graph partitioning heuristics. Problem instances are usually chosen from the Erdős–Rényi and uniformly random regular graph ensembles. We recall that an Erdős–Rényi random graph $G(n, m)$ on n vertices with m edges is a

Received March 2015; revised August 2015.

[1]Supported in part by NSF Grant DMS-11-06627.

[2]Supported in part by NSF Grants DMS-11-06627 and CCF-1319979 and the AFOSR Grant FA9550-13-1-0036.

[3]Supported in part by the William R. and Sara Hart Kimball Stanford Graduate Fellowship.

MSC2010 subject classifications. 05C80, 68R10, 82B44.

Key words and phrases. Max-cut, bisection, Erdős–Rényi graph, regular graph, Ising model, spin glass, Parisi formula.

graph formed by choosing m edges uniformly at random among all possible edges. A γ-regular random graph on n-vertices $G^{\text{Reg}}(n,\gamma)$ is a graph drawn uniformly from the set of all graphs on n-vertices where every vertex has degree γ (provided γn is even). See [6, 27, 29] for detailed analyses of these graph ensembles.

Both min-bisection and Max-Cut undergo phase transitions on the Erdős–Rényi graph $G(n,[\gamma n])$. For $\gamma < \log 2$, the largest component has less than $n/2$ vertices and minimum bisection is $O(1)$ asymptotically as $n \to \infty$ while above this threshold, the largest component has size greater than $n/2$ and min-bisection is $\Omega(n)$ [33]. Similarly, Max-Cut exhibits a phase transition at $\gamma = 1/2$. The difference between the number of edges and Max-Cut size is $\Omega(1)$ for $\gamma < 1/2$, while it is $\Omega(n)$ when $\gamma > 1/2$ [10]. The distribution of the Max-Cut size in the critical scaling window was determined in [12]. In this paper, we work in the $\gamma \to \infty$ regime, so that both min-bisection and Max-Cut are $\Omega(n)$ asymptotically.

Diverse techniques have been employed in the analysis of minimum and maximum bisection for random graph ensembles. For example, [5] used the Azuma–Hoeffding inequality to establish that

$$\frac{\gamma}{4} - \sqrt{\frac{\gamma \log 2}{4}} \le \frac{1}{n}\mathsf{mcut}\big(G^{\text{Reg}}(n,\gamma)\big) \le \frac{\gamma}{4} + \sqrt{\frac{\gamma \log 2}{4}}.$$

Spectral relaxation based approaches can also be used to bound these quantities. These approaches observe that the minimum and maximum bisection problem can be written as optimization problems over variables $\sigma_i \in \{-1,+1\}$ associated to the vertices of the graph. By relaxing the integrality constraint to an L_2 constraint the resulting problem can be solved through spectral methods. For instance, the minimum bisection is bounded as follows [here $\Omega_n \subseteq \{-1,+1\}^n$ is the set of (± 1)-vectors with $\sum_{i=1}^n \sigma_i = 0$, assuming for simplicity n even]

$$\begin{aligned}\mathsf{mcut}(G) &= \min_{\underline{\sigma}\in\Omega_n}\left\{\frac{1}{4}\sum_{(i,j)\in E}(\sigma_i-\sigma_j)^2\right\} \\ &= \frac{1}{2}\min_{\underline{\sigma}\in\Omega_n}\{\underline{\sigma}\cdot(\mathcal{L}_G\underline{\sigma})\} \ge \frac{1}{2}\lambda_2(\mathcal{L}_G).\end{aligned} \tag{1.1}$$

Here, $\mathcal{L}_G$ is the Laplacian of G, with eigenvalues $0 = \lambda_1(\mathcal{L}_G) \le \lambda_2(\mathcal{L}_G) \le \cdots \le \lambda_n(\mathcal{L}_G)$. For regular graphs, using the result of [19], this implies that $n^{-1}\mathsf{mcut}(G^{\text{Reg}}(n,\gamma)) \ge \frac{\gamma}{4} - \sqrt{\gamma-1}$. However, for Erdős–Rényi graphs $\lambda_2(\mathcal{L}_G) = o(1)$ vanishes with n [30] and this approach fails. A similar spectral relaxation yields, for regular graphs, $\mathsf{MCUT}(G^{\text{Reg}}(n,\gamma))/n \le \frac{\gamma}{4} + \sqrt{\gamma-1}$, but fails for Erdős–Rényi graphs. Nontrivial spectral bounds on Erdős–Rényi graphs can be derived, for instance, from [9, 17].

An alternative approach consists of analyzing algorithms that aim to minimize (maximize) the cut size. This provides upper bounds on $\mathsf{mcut}(G)$ [respectively, lower bounds on $\mathsf{MCUT}(G)$]. For instance, [1] proved that all regular graphs have

$n^{-1}\mathsf{mcut}(G) \le \frac{\gamma}{4} - \sqrt{\frac{9\gamma}{2048}}$ for all n large enough (this method was further developed in [15]).

Similar results have been established for the max-cut problem on Erdős–Rényi random graphs. The recent breakthrough paper [4] shows that there exists $\mathcal{M}(\gamma)$ such that $\mathsf{MaxCut}(G(n,[\gamma n]))/n \xrightarrow{p} \mathcal{M}(\gamma)$ and following upon it, [21] proves that $\mathcal{M}(\gamma) \in [\gamma/2 + 0.47523\sqrt{\gamma}, \gamma/2 + 0.55909\sqrt{\gamma}]$.

To summarize, the general flavor of these results is that if G is an Erdős–Rényi or a random regular graph on n vertices with $[\gamma n/2]$ edges, then $\mathsf{mcut}(G)/n = \gamma/4 - \Theta(\sqrt{\gamma})$ while $\mathsf{MCUT}(G)/n$ and $\mathsf{MaxCut}(G)/n$ behave asymptotically like $\gamma/4 + \Theta(\sqrt{\gamma})$. In other words, the relative spread of cut widths around its average is of order $1/\sqrt{\gamma}$. Despite 30 years of research in combinatorics and random graph theory, even the leading behavior of such a spread remained undetermined.

There are however detailed and intriguing predictions in statistical physics, mainly based on the nonrigorous cavity method [35], which relate the behavior of these features to that of mean field spin glasses. From a statistical physics perspective, determining the minimum (maximum) bisection is equivalent to finding the ground state energy of the ferromagnetic (anti-ferromagnetic) Ising model, constrained to have zero magnetization (see [40] and the references therein). Similarly, the Max-Cut is naturally associated with the ground state energy of an anti-ferromagnetic Ising model on the graph. The cavity method then suggests a surprising conjecture [46] that, with high probability,

$$\begin{aligned}\mathsf{MCUT}\big(G^{\mathrm{Reg}}(n,\gamma)\big) &= \mathsf{MaxCut}\big(G^{\mathrm{Reg}}(n,\gamma)\big) + o(n)\\ &= n\gamma/2 - \mathsf{mcut}\big(G^{\mathrm{Reg}}(n,\gamma)\big) + o(n).\end{aligned}$$

The present paper bridges this gap, by partially confirming some of the physics predictions and by providing estimates of these features which are sharp up to corrections of order $no(\sqrt{\gamma})$. Our estimates are expressed in terms of the celebrated Parisi formula for the free-energy of the Sherrington Kirkpatrick spin glass, and build on its recent proof by Talagrand. In a sense, these results explain the difficulty encountered by classical combinatorics techniques in attacking this problem. In doing so, we develop a new approach based on an interpolation technique from the theory of mean field spin glasses [23, 24, 43]. So far this technique has been used in combinatorics only to prove bounds [18]. We combine and extend these ideas, crucially utilizing properties of both the Poisson and Gaussian distributions to derive an asymptotically sharp estimate.

1.1. *Our contribution.* To state our results precisely, we proceed with a short review of the Sherrington–Kirkpatrick (SK) model of spin glasses. This canonical example of a mean field spin glass has been studied extensively by physicists [36], and seen an explosion of activity in mathematics following Talagrand's proof of the Parisi formula, leading to better understanding of the SK model and its generalizations (cf. the text [39] for an introduction to the subject).

The SK model is a (random) probability distribution on the hyper-cube $\{-1,+1\}^n$ which assigns mass proportional to $\exp(\beta H^{\text{SK}}(\underline{\sigma}))$ to each "spin configuration" $\underline{\sigma} \in \{-1,+1\}^n$. The parameter $\beta > 0$ is interpreted as the inverse temperature, with $H^{\text{SK}}(\cdot)$ called the Hamiltonian of the model. The collection $\{H^{\text{SK}}(\underline{\sigma}) : \underline{\sigma} \in \{-1,+1\}^n\}$ is a Gaussian process on $\{-1,+1\}^n$ with mean $\mathbb{E}[H^{\text{SK}}(\underline{\sigma})] = 0$ and covariance $\mathbb{E}\{H^{\text{SK}}(\underline{\sigma})H^{\text{SK}}(\underline{\sigma}')\} = \frac{1}{2n}(\underline{\sigma} \cdot \underline{\sigma}')^2$. This process is usually constructed by

$$H^{\text{SK}}(\underline{\sigma}) = -\frac{1}{\sqrt{2n}} \sum_{i,j=1}^{n} J_{ij}\sigma_i\sigma_j, \tag{1.2}$$

with $\{J_{ij}\}$ being n^2 independent standard Gaussian variables, and we are mostly interested in the ground state energy of the SK model. That is, the expected (over $\{J_{ij}\}$) minimum (over $\underline{\sigma}$) of the Gaussian process $H^{\text{SK}}(\underline{\sigma})$ introduced above.

DEFINITION 1.1. Let $\mathcal{D}_\beta$ be the space of nondecreasing, right-continuous nonnegative functions $x : [0,1] \to [0,\beta]$. The *Parisi functional* at inverse temperature β is the function $\mathsf{P}_\beta : \mathcal{D}_\beta \to \mathbb{R}$ defined by

$$\mathsf{P}_\beta[x] = f(0,0;x) - \frac{1}{2}\int_0^1 q x(q)\,\mathrm{d}q, \tag{1.3}$$

where $f : [0,1] \times \mathbb{R} \times \mathcal{D}_\beta \to \mathbb{R}$, $(q,y,x) \mapsto f(q,y;x)$ is the unique weak solution of the PDE with boundary condition

$$\begin{aligned} &\frac{\partial f}{\partial q} + \frac{1}{2}\frac{\partial^2 f}{\partial y^2} + \frac{1}{2}x(q)\left(\frac{\partial f}{\partial y}\right)^2 = 0, \\ &f(1,y;x) = (1/\beta)\log\bigl(2\cosh(\beta y)\bigr) \end{aligned} \tag{1.4}$$

among all continuous functions $f(q,y)$ such that $\frac{\partial f}{\partial y} \in L^2([0,1] \times \mathbb{R})$.

The Parisi replica-symmetry-breaking prediction for the SK model is

$$\mathsf{P}_{*,\beta} \equiv \inf\bigl\{\mathsf{P}_\beta[x] : x \in \mathcal{D}_\beta\bigr\}. \tag{1.5}$$

We refer to [28], Proposition 7, for the uniqueness of such a solution of (1.4), and to [3] for the strict convexity of $x \mapsto \mathsf{P}_\beta[x]$, which implies the existence of a unique global minimizer of $\mathsf{P}_{*,\beta}$. We are interested here in the zero-temperature limit

$$\mathsf{P}_* \equiv \lim_{\beta\to\infty} \mathsf{P}_{*,\beta}, \tag{1.6}$$

which exists because the free energy density (and hence $\mathsf{P}_{*,\beta}$, by [44]), is uniformly continuous in $1/\beta$. It follows from the Parisi formula [44], that

$$\lim_{n\to\infty} n^{-1}\mathbb{E}\Bigl[\max_{\underline{\sigma}}\bigl\{H^{\text{SK}}(\underline{\sigma})\bigr\}\Bigr] = \mathsf{P}_*. \tag{1.7}$$

The partial differential equation (1.4) can be solved numerically to high precision, resulting with the numerical evaluation of $\mathsf{P}_* = 0.76321 \pm 0.00003 \approx 0.763166726$ [11, 42], whereas using the replica symmetric bound of [22], it is possible to prove that $\mathsf{P}_* \le \sqrt{2/\pi} \approx 0.797885$.

We next introduce some additional notation necessary for stating our results. Throughout the paper, $O(\cdot)$, $o(\cdot)$, and $\Theta(\cdot)$ stands for the usual $n \to \infty$ asymptotic, while $O_\gamma(\cdot)$, $o_\gamma(\cdot)$ and $\Theta_\gamma(\cdot)$ are used to describe the $\gamma \to \infty$ asymptotic regime. We say that a sequence of events A_n occurs with high probability (w.h.p.) if $\mathbb{P}(A_n) \to 1$ as $n \to \infty$. Finally, for random $\{X_n\}$ and nonrandom $f : \mathbb{R}^+ \to \mathbb{R}^+$, we say that $X_n = o_\gamma(f(\gamma))$ w.h.p. as $n \to \infty$ if there exists nonrandom $g(\gamma) = o_\gamma(f(\gamma))$ such that the sequence $A_n = \{|X_n| \le g(\gamma)\}$ occurs w.h.p. (as $n \to \infty$).

Our first result provides estimates of the minimum and maximum bisection of Erdős–Rényi random graphs in terms of the SK quantity P_* of (1.6).

THEOREM 1.2. *We have*, *w.h.p. as* $n \to \infty$, *that*

$$\frac{\mathsf{mcut}(G(n,[\gamma n]))}{n} = \frac{\gamma}{2} - \mathsf{P}_*\sqrt{\frac{\gamma}{2}} + o_\gamma(\sqrt{\gamma}), \tag{1.8}$$

$$\frac{\mathsf{MCUT}(G(n,[\gamma n]))}{n} = \frac{\gamma}{2} + \mathsf{P}_*\sqrt{\frac{\gamma}{2}} + o_\gamma(\sqrt{\gamma}). \tag{1.9}$$

REMARK 1.3. Recall the Erdős–Rényi random graph $G_I(n, p_n)$, where each edge is independently included with probability p_n. Since the number of edges in $G_I(n, \frac{2\gamma}{n})$ is concentrated around γn, with fluctuations of $O(n^{1/2+\varepsilon})$ w.h.p. for any $\varepsilon > 0$, for the purpose of Theorem 1.2 the random graph $G_I(n, \frac{2\gamma}{n})$ has the same asymptotic behavior as $G(n, [\gamma n])$.

REMARK 1.4. The physics interpretation of Theorem 1.2 is that a zero-magnetization constraint forces a ferromagnet on a random graph to be in a spin glass phase. This phenomenon is expected to be generic for models on nonamenable graphs (whose surface-to-volume ratio is bounded away from zero), in staggering contrast with what happens on amenable graphs (e.g., regular lattices), where such zero magnetization constraint leads to a phase separation.

We next outline the strategy for proving Theorem 1.2 (with the detailed proof provided in Section 2). For graphs $G = (V, E)$, with vertex set $V = [n]$ and n even, we write $\underline{\sigma} \in \Omega_n$ if the assignment of binary variables $\underline{\sigma} = (\sigma_1, \dots, \sigma_n)$, $\sigma_i \in \{-1, +1\}$ to V is such that $\sum_{i \in V} \sigma_i = 0$. We further define the Ising energy function $H_G(\underline{\sigma}) = -\sum_{(i,j)\in E} \sigma_i \sigma_j$, and let $U_-(G) \equiv \min\{H_G(\underline{\sigma}) : \underline{\sigma} \in \Omega_n\}$, $U_+(G) \equiv \max\{H_G(\underline{\sigma}) : \underline{\sigma} \in \Omega_n\}$. It is then clear that

$$\begin{aligned} \mathsf{mcut}(G) &= \tfrac{1}{2}|E| + \tfrac{1}{2}U_-(G), \\ \mathsf{MCUT}(G) &= \tfrac{1}{2}|E| + \tfrac{1}{2}U_+(G). \end{aligned} \tag{1.10}$$

In statistical mechanics $\underline{\sigma}$ is referred to as a "spin configuration" and $U_-(G)$ [resp., $U_+(G)$], its "ferromagnetic (anti-ferromagnetic) ground state energy."

The expected cut size of a random partition is taken care of by the term $\frac{1}{2}|E|$, whereas standard concentration inequalities imply that $U_+(G)$ and $U_-(G)$ are tightly concentrated around their expectation when G is a sparse Erdős–Rényi random graph. Therefore, it suffices to prove that as $n \to \infty$ all limit points of $n^{-1}\mathbb{E}[U_\pm(G)]$ are within $o_\gamma(\sqrt{\gamma})$ of $\pm\mathsf{P}_*\sqrt{2\gamma}$. Doing so is the heart of the whole argument, and it is achieved through the interpolation technique of [23, 24]. Intuitively, we replace the graph G by a complete graph with random edge weights $J_{ij}/\sqrt{n}$ for J_{ij} independent standard normal random variables, and prove that the error induced on $U_\pm(G)$ by this replacement is bounded (in expectation) by $no_\gamma(\sqrt{\gamma})$. Finally, we show that the maximum and minimum cut-width of such weighted complete graph do not change much when optimizing over all partitions $\underline{\sigma} \in \{-1, +1\}^n$ instead of only over the balanced partitions $\underline{\sigma} \in \Omega_n$. Now that the equi-partition constraint has been relaxed, the problem has become equivalent to determining the ground state energy of the SK spin glass model, which is solved by taking the "zero temperature" limit of the Parisi formula (from [44]).

The next result extends Theorem 1.2 to γ-regular random graphs.

THEOREM 1.5. *We have, w.h.p. as $n \to \infty$, that*

$$\frac{\mathsf{mcut}(G^{\mathrm{Reg}}(n,\gamma))}{n} = \frac{\gamma}{4} - \mathsf{P}_*\sqrt{\frac{\gamma}{4}} + o_\gamma(\sqrt{\gamma}), \tag{1.11}$$

$$\frac{\mathsf{MCUT}(G^{\mathrm{Reg}}(n,\gamma))}{n} = \frac{\gamma}{4} + \mathsf{P}_*\sqrt{\frac{\gamma}{4}} + o_\gamma(\sqrt{\gamma}). \tag{1.12}$$

The average degree in an Erdős–Rényi graph $G(n, [\gamma n])$ is 2γ so Theorems 1.2 and 1.5 take the same form in terms of average degree. However, moving from Erdős–Rényi graphs to regular random graphs having the same number of edges is nontrivial, since the fluctuation of the degree of a typical vertex in an Erdős–Rényi graph is $\Theta_\gamma(\sqrt{\gamma})$. Hence, any coupling of these two graph models yield about $n\sqrt{\gamma}$ different edges, and merely bounding the difference in cut-size by the number of different edges, results in the too large $\sqrt{\gamma}$ spread. Instead, as detailed in Section 3, our proof of Theorem 1.5 relies on a delicate construction, similar to that in [20], which "embeds" an Erdős–Rényi graph of average degree slightly smaller than γ, into a γ-regular random graph while establishing that the fluctuations in the contribution of the additional edges is only $no_\gamma(\sqrt{\gamma})$. Our construction starts with the γ-regular graph $\mathcal{G}_1$ and produces an Erdős–Rényi graph $\mathcal{G}_2$, "most" of which is embedded within $\mathcal{G}_1$, whereas [20] go in the converse direction, starting with $\mathcal{G}_2$ and producing $\mathcal{G}_1$ out of it.

Our next result, whose proof is provided in Section 4, shows that up to the first order, the asymptotic of the Max-Cut matches that of the Max bisection for both Erdős–Rényi and random regular graphs.

THEOREM 1.6. (a) *W.h.p. as $n \to \infty$, we have*

$$\frac{\mathsf{MaxCut}(G(n,[\gamma n]))}{n} = \frac{\gamma}{2} + \mathsf{P}_*\sqrt{\frac{\gamma}{2}} + o_\gamma(\sqrt{\gamma}).$$

(b) *W.h.p. as $n \to \infty$, we have*

$$\frac{\mathsf{MaxCut}(G^{\mathrm{Reg}}(n,\gamma))}{n} = \frac{\gamma}{4} + \mathsf{P}_*\sqrt{\frac{\gamma}{4}} + o_\gamma(\sqrt{\gamma}).$$

1.2. *Application to community detection.* As a simple illustration of the potential applications of our results, we consider the problem of detecting communities within the so-called "planted partition model," or stochastic block model. Given parameters $a > b > 0$ and even n, we denote by $G_I(n, a/n, b/n)$ the random graph over vertex set $[n]$, such that given a uniformly random balanced partition $[n] = V_1 \cup V_2$, edges (i, j) are independently present with probability a/n when either both $i, j \in V_1$ or both $i, j \in V_2$, or alternatively present with probability b/n if either $i \in V_1$ and $j \in V_2$, or vice versa. Given a random graph G, the community detection problem requires us to determine whether the null hypothesis $H_0 : G \sim G_I(n, (a+b)/(2n))$ holds, or the alternative hypothesis $H_1 : G \sim G_I(n, a/n, b/n)$ holds.

Under the alternative hypothesis, the cut size of the balanced partition (V_1, V_2) concentrates tightly around $nb/4$. This suggests the optimization-based hypothesis testing

$$T_{\mathsf{cut}}(G;\theta) = \begin{cases} 0, & \text{if } \mathsf{mcut}(G) \le \theta, \\ 1, & \text{otherwise,} \end{cases} \tag{1.13}$$

and we have the following immediate consequence of Theorem 1.2.

COROLLARY 1.7. *Let $\theta_n = (b/4) + \varepsilon_n$ with $\varepsilon_n\sqrt{n} \to \infty$. Then, the test $T_{\mathsf{cut}}(\cdot;\theta_n)$ succeeds w.h.p. as $n \to \infty$, provided $(a-b)^2 \ge 8\mathsf{P}_*^2(a+b) + o(a+b)$.*

Let us stress that we did not provide an efficient algorithm for computing T_{cut} (but see [37] for related work that uses polynomially computable convex relaxations). By contrast, there exist polynomially computable tests that succeed w.h.p. whenever $(a-b)^2 > 2(a+b)$ and no test can succeed below this threshold (see [13, 34, 38]). Nevertheless, the test T_{cut} is so natural that its analysis is of independent interest, and Corollary 1.7 implies that T_{cut} is sub-optimal by a factor of at most $4\mathsf{P}_*^2 \approx 2.33$.

2. Interpolation: Proof of Theorem 1.2. The Erdős–Rényi random graph $G(n, m)$ considers a uniformly chosen element from among all *simple* (i.e., having no loops or double edges), graphs of n vertices and m edges. For $m = [\gamma n]$ and γ bounded, such simple graph differs in only $O(1)$ edges from the corresponding multi-graph which makes a uniform choice while allowing for loops and multiple edges. Hence, the two models are equivalent for our purpose, and letting $G(n,[\gamma n])$ denote hereafter the latter multi-graph, we note that it can be constructed also by sequentially introducing the $[\gamma n]$ edges and independently sampling their end-points from the uniform distribution on $\{1,\dots,n\}$. We further let $G^{\mathrm{Poiss}}_{n,\gamma}$ denote the Poissonized random multi-graph $G(n,N_n)$ having the random number of edges $N_n\sim \mathrm{Pois}(\gamma n)$, independently of the choice of edges. Alternatively, one constructs $G^{\mathrm{Poiss}}_{n,\gamma}$ by generating for $1\le i,j\le n$ the i.i.d. $z_{ij}\sim\mathrm{Pois}(\frac{\gamma}{n})$ and forms the multi-graph on n vertices by taking $(z_{ij}+z_{ji})$ as the multiplicity of each edge $(i,j), i\neq j$ [ending with multiplicity $z_{(i,j)}\sim\mathrm{Pois}(\frac{2\gamma}{n})$ for edge (i,j), $i\neq j$ and the multiplicity $z_{(i,i)}\sim\mathrm{Pois}(\frac{\gamma}{n})$ for each loop (i,i), where $\{z_{(i,j)}, i<j, z_{(i,i)}\}$ are mutually independent]. By the tight concentration of the $\mathrm{Pois}(\gamma n)$ law, it suffices to prove Theorem 1.2 for $G^{\mathrm{Poiss}}_{n,\gamma}$, and in this section we always take for G_n a random multi-graph distributed as $G^{\mathrm{Poiss}}_{n,\gamma}$.

2.1. *Spin models and free energy.* A spin model is defined by the (possibly random) Hamiltonian $H:\{-1,+1\}^n\to\mathbb{R}$ and in this paper we often consider spin models constrained to have zero empirical magnetization, namely from the set $\Omega_n=\{\underline{\sigma}\in\{-1,+1\}^n:\sum_{i=1}^n\sigma_i=0\}$. The constrained partition function is then $Z_n(\beta)=\sum_{\underline{\sigma}\in\Omega_n}e^{-\beta H(\underline{\sigma})}$ with the corresponding constrained free energy density

$$\phi_n(\beta)\equiv\frac{1}{n}\mathbb{E}\big[\log Z_n(\beta)\big]=\frac{1}{n}\mathbb{E}\Big[\log\Big\{\sum_{\underline{\sigma}\in\Omega_n}e^{-\beta H(\underline{\sigma})}\Big\}\Big]. \tag{2.1}$$

The expectation in (2.1) is over the distribution of the function $H(\cdot)$ [i.e., over the collection of random variables $\{H(\underline{\sigma})\}$]. Depending on the model under consideration, the Hamiltonian (or the free energy) might depend on additional parameters which we will indicate, with a slight abuse of notation, as additional arguments of $\phi_n(\cdot)$.

For such spin models, we also consider the expected ground state energy density

$$e_n=\frac{1}{n}\mathbb{E}\Big[\min_{\underline{\sigma}\in\Omega_n}H(\underline{\sigma})\Big], \tag{2.2}$$

which determines the large-β behavior of the free energy density. That is, $\phi_n(\beta)=-\beta e_n+o(\beta)$. We analogously define the maximum energy

$$\widehat{e}_n=\frac{1}{n}\mathbb{E}\Big[\max_{\underline{\sigma}\in\Omega_n}H(\underline{\sigma})\Big], \tag{2.3}$$

which governs the behavior of the free energy density as $\beta\to-\infty$. That is, $\phi_n(\beta)=-\beta\widehat{e}_n+o(\beta)$ (in statistical mechanics it is more customary to change the sign of the Hamiltonian in such a way that β is kept positive). The corresponding Boltzmann measure on Ω_n is

$$\mu_{\beta,n}(\underline{\sigma})=\frac{1}{Z_n(\beta)}\exp\{-\beta H(\underline{\sigma})\}. \tag{2.4}$$

A very important example of a spin model, that is crucial for our analysis is the SK model having the Hamiltonian $H^{\text{SK}}(\cdot)$ of (1.2) on $\{-1,+1\}^n$ and we also consider that model constrained to Ω_n (i.e., subject to zero magnetization constraint).

The second model we consider is the "dilute" ferromagnetic Ising model on $G^{\text{Poiss}}_{n,\gamma} = (V, E)$, corresponding to the Hamiltonian

$$H^{\text{D}}_{\gamma}(\underline{\sigma}) = -\sum_{(i,j)\in E} \sigma_i \sigma_j, \tag{2.5}$$

again restricted to $\underline{\sigma} \in \Omega_n$. We use superscripts to indicate the model to which various quantities refer. For instance $\phi^{\text{SK}}_n(\beta)$ denotes the constrained free energy of the SK model, $\phi^{\text{D}}_n(\beta;\gamma)$ is the constrained free energy of the Ising model on $G^{\text{Poiss}}_{n,\gamma}$, with analogous notations used for the ground state energies e^{SK}_n and $e^{\text{D}}_n(\gamma)$.

The first step in proving Theorem 1.2 is to show that $\mathsf{mcut}(G_n)$ and $\mathsf{MCUT}(G_n)$ are concentrated around their expectations.

LEMMA 2.1. *Fixing $\varepsilon > 0$, we have that*

$$\mathbb{P}\big[\big|\mathsf{mcut}(G_n) - \mathbb{E}\big[\mathsf{mcut}(G_n)\big]\big| > n\varepsilon\big] = O(1/n),$$
$$\mathbb{P}\big[\big|\mathsf{MCUT}(G_n) - \mathbb{E}\big[\mathsf{MCUT}(G_n)\big]\big| > n\varepsilon\big] = O(1/n).$$

PROOF. Recall (1.10) that $\mathsf{mcut}(G_n) = \frac{1}{2}|E_n| + \frac{1}{2}U_-(G_n)$, with $|E_n| = N_n \sim \text{Pois}([\gamma n])$. Therefore,

$$\begin{aligned}
&\mathbb{P}\big[\big|\mathsf{mcut}(G_n) - \mathbb{E}\big[\mathsf{mcut}(G_n)\big]\big| > n\varepsilon\big] \\
&\quad \le \mathbb{P}\big[\big|U_-(G_n) - \mathbb{E}\big[U_-(G_n)\big]\big| > n\varepsilon\big] + \mathbb{P}\big[|N_n - \mathbb{E}N_n| > n\varepsilon\big] \\
&\quad \le \frac{\text{Var}(U_-(G_n))}{n^2\varepsilon^2} + \frac{\text{Var}(N_n)}{n^2\varepsilon^2} \\
&\quad = \frac{\text{Var}(U_-(G_n))}{n^2\varepsilon^2} + O(1/n).
\end{aligned}$$

We complete the proof for $\mathsf{mcut}(G_n)$ by showing that $\text{Var}(U_-(G_n)) \le n\gamma$. Indeed, writing $U_-(G_n) = f(\mathbf{z})$ for $\mathbf{z} = \{z_{ij}, 1 \le i, j \le n\}$ and i.i.d. $z_{ij} \sim \text{Pois}(\gamma/n)$, we let $\mathbf{z}^{(i,j)}$ denote the vector formed when replacing z_{ij} in $\mathbf{z}$ by an i.i.d. copy z'_{ij}. Clearly, $|f(\mathbf{z}) - f(\mathbf{z}^{(i,j)})| \le |z_{ij} - z'_{ij}|$. Hence, by the Efron–Stein inequality [7], Theorem 3.1,

$$\text{Var}\big(U_-(G_n)\big) \le \frac{1}{2}\sum_{i,j} \mathbb{E}\big[\big(f(\mathbf{z}) - f\big(\mathbf{z}^{(i,j)}\big)\big)^2\big] \le \frac{1}{2}\sum_{i,j} \mathbb{E}\big[\big(z_{ij} - z'_{ij}\big)^2\big],$$

yielding the required bound [and the proof for $\mathsf{MCUT}(G_n) = \frac{1}{2}N_n + \frac{1}{2}U_+(G_n)$ proceeds along the same line of reasoning]. □

Next, recall that $e_n^D = n^{-1}\mathbb{E}[U_-(G_n)]$ and $\widehat{e}_n^D = n^{-1}\mathbb{E}[U_+(G_n)]$ [see (2.2) and (2.3), resp.], whereas $|E_n| \sim \text{Pois}(\gamma n)$ has expectation γn. Hence, from the representation (1.10) of $\mathsf{mcut}(G_n)$ and $\mathsf{MCUT}(G_n)$, we conclude that

$$\begin{aligned} \frac{1}{n}\mathbb{E}[\mathsf{mcut}(G_n)] &= \frac{\gamma}{2} + \frac{1}{2}e_n^{\mathrm{D}}(\gamma), \\ \frac{1}{n}\mathbb{E}[\mathsf{MCUT}(G_n)] &= \frac{\gamma}{2} + \frac{1}{2}\widehat{e}_n^{\mathrm{D}}(\gamma). \end{aligned} \tag{2.6}$$

Combining (2.6) with Lemma 2.1, we establish Theorem 1.2, once we show that as $n \to \infty$,

$$e_n^{\mathrm{D}}(\gamma) = -\sqrt{2\gamma}\mathsf{P}_* + o_\gamma(\sqrt{\gamma}) + o(1), \tag{2.7}$$

$$\widehat{e}_n^{\mathrm{D}}(\gamma) = +\sqrt{2\gamma}\mathsf{P}_* + o_\gamma(\sqrt{\gamma}) + o(1). \tag{2.8}$$

Establishing (2.7) and (2.8) is the main step in proving Theorem 1.2, and the key to it is the following proposition of independent interest.

PROPOSITION 2.2. *There exist constants $A_1, A_2 < \infty$ independent of n, β and γ such that*

$$\left|\phi_n^{\mathrm{D}}\left(\frac{\beta}{\sqrt{2\gamma}}, \gamma\right) - \phi_n^{\mathrm{SK}}(\beta)\right| \le A_1\frac{|\beta|^3}{\sqrt{\gamma}} + A_2\frac{\beta^4}{\gamma}. \tag{2.9}$$

We defer the proof of Proposition 2.2 to Section 2.2, where we also apply it to deduce the next lemma, comparing the ground state energy of a dilute Ising ferromagnet to that of the SK model, after both spin models have been constrained to have zero magnetization.

LEMMA 2.3. *There exist $A = A(\gamma_0)$ finite, such that for all $\gamma \ge \gamma_0$ and any n,*

$$\left|\frac{e_n^{\mathrm{D}}(\gamma)}{\sqrt{2\gamma}} - e_n^{\mathrm{SK}}\right| \le A\gamma^{-1/6}, \qquad \left|\frac{\widehat{e}_n^{\mathrm{D}}(\gamma)}{\sqrt{2\gamma}} + e_n^{\mathrm{SK}}\right| \le A\gamma^{-1/6}. \tag{2.10}$$

In view of Lemma 2.3, we get both (2.7) and (2.8) once we control the difference between the ground state energies of the unconstrained and constrained to have zero magnetization SK models. This is essentially established by the following lemma (whose proof is provided in Section 2.3).

LEMMA 2.4. *For any $\delta > 0$, w.h.p. $0 \le U_n^{\mathrm{SK}} - \overline{U}_n^{\mathrm{SK}} \le n^{1/2+\delta}$, where*

$$\overline{U}_n^{\mathrm{SK}} = \min_{\underline{\sigma}\in\{-1,+1\}^n}\{H^{\mathrm{SK}}(\underline{\sigma})\}, \qquad U_n^{\mathrm{SK}} = \min_{\underline{\sigma}\in\Omega_n}\{H^{\mathrm{SK}}(\underline{\sigma})\}. \tag{2.11}$$

Indeed, applying Borel's concentration inequality for the maxima of Gaussian random vectors (see [7], proof of Theorem 5.8), we have that for some $c > 0$, all n and $\delta > 0$,

$$\mathbb{P}[|\overline{U}_n^{\mathrm{SK}} - \mathbb{E}[\overline{U}_n^{\mathrm{SK}}]| > n\delta] \le 2e^{-cn\delta^2}, \tag{2.12}$$

$$\mathbb{P}[|U_n^{\mathrm{SK}} - \mathbb{E}[U_n^{\mathrm{SK}}]| > n\delta] \le 2e^{-cn\delta^2}. \tag{2.13}$$

Recall that $e_n^{\mathrm{SK}} = n^{-1}\mathbb{E}[U_n^{\mathrm{SK}}]$, whereas $n^{-1}\mathbb{E}[\overline{U}_n^{\mathrm{SK}}] \to -\mathsf{P}_*$ by (1.7). Consequently, the bounds of (2.12), (2.13) coupled with Lemma 2.4 imply that $e_n^{\mathrm{SK}} \to -\mathsf{P}_*$ as $n \to \infty$. This, combined with Lemma 2.3 and (2.6), completes the proof of Theorem 1.2.

2.2. *The interpolation argument*. We first deduce Lemma 2.3 out of Proposition 2.2. To this end, we use the inequalities of Lemma 2.5 relating the free energy of a spin model to its ground state energy (these are special cases of general bounds for models with at most c^n configurations, but for the sake of completeness we include their proof).

LEMMA 2.5. *The following inequalities hold for any n, $\beta, \gamma > 0$:*

$$\left|e_n^{\mathrm{D}}(\gamma) + \frac{1}{\beta}\phi_n^{\mathrm{D}}(\beta,\gamma)\right| \le \frac{\log 2}{\beta}, \qquad \left|e_n^{\mathrm{SK}} + \frac{1}{\beta}\phi_n^{\mathrm{SK}}(\beta)\right| \le \frac{\log 2}{\beta}. \tag{2.14}$$

Further, for any n, $\beta < 0$, $\gamma > 0$,

$$\left|\widehat{e}_n^{\mathrm{D}}(\gamma) + \frac{1}{\beta}\phi_n^{\mathrm{D}}(\beta,\gamma)\right| \le \frac{\log 2}{|\beta|}, \qquad \left|\widehat{e}_n^{\mathrm{SK}} + \frac{1}{\beta}\phi_n^{\mathrm{SK}}(\beta)\right| \le \frac{\log 2}{|\beta|}. \tag{2.15}$$

PROOF. Let $H_n(\underline{\sigma})$ be a generic Hamiltonian for $\underline{\sigma} \in \Omega_n$. One then easily verifies that

$$\frac{\partial}{\partial\beta}\left(\frac{\phi_n(\beta)}{\beta}\right) = -\frac{1}{n\beta^2}\mathbb{E}[S(\mu_{\beta,n})] \in \left[-\frac{\log 2}{\beta^2}, 0\right],$$

for the Boltzman measure (2.4) and the nonnegative entropy functional $S(\mu) = -\sum_{\underline{\sigma}\in\Omega_n}\mu(\underline{\sigma})\log\mu(\underline{\sigma})$ which is at most $\log|\Omega_n|$. Further, comparing (2.1) and (2.2) we see that $\beta^{-1}\phi_n(\beta) \to -e_n$ when $\beta \to \infty$ (while n is fixed). Consequently, for any $\beta > 0$,

$$\left|e_n + \frac{\phi_n(\beta)}{\beta}\right| = \left|\int_\beta^\infty \frac{\partial}{\partial u}\left(\frac{\phi_n(u)}{u}\right)\mathrm{d}u\right| \le \frac{\log 2}{\beta}.$$

We apply this inequality separately to the SK model and the diluted Ising model to get the bounds of (2.14). We similarly deduce the bounds of (2.15) upon observing that $\beta^{-1}\phi_n(\beta) \to -\widehat{e}_n$ when $\beta \to -\infty$. $\square$

PROOF OF LEMMA 2.3. Clearly, for any n, $\beta > 0$ and $\gamma > 0$,

$$\begin{aligned}\left|\frac{e_n^{\mathrm{D}}(\gamma)}{\sqrt{2\gamma}} - e_n^{\mathrm{SK}}\right| &\le \left|\frac{1}{\sqrt{2\gamma}} e_n^{\mathrm{D}}(\gamma) + \frac{1}{\beta}\phi_n^{\mathrm{D}}\left(\frac{\beta}{\sqrt{2\gamma}}, \gamma\right)\right| \\ &\quad + \left|\frac{1}{\beta}\phi_n^{\mathrm{SK}}(\beta) - \frac{1}{\beta}\phi_n^{\mathrm{D}}\left(\frac{\beta}{\sqrt{2\gamma}}, \gamma\right)\right| \\ &\quad + \left|e_n^{\mathrm{SK}} + \frac{1}{\beta}\phi_n^{\mathrm{SK}}(\beta)\right|.\end{aligned}$$

In view of (2.14), the first and last terms on the RHS are bounded by $(\log 2)/\beta$. Setting $\beta = \gamma^{1/6}$, we deduce from Proposition 2.2 that the middle term on the RHS is bounded by $A_1\gamma^{-1/6} + A_2\gamma^{-1/2}$, yielding the first (left) bound in (2.10) (for $A = \log 2 + A_1 + A_2\gamma_0^{-1/3}$). In case $\beta < 0$, starting from

$$\begin{aligned}\left|\frac{\widehat{e}_n^{\mathrm{D}}(\gamma)}{\sqrt{2\gamma}} + e_n^{\mathrm{SK}}\right| &\le \left|\frac{1}{\sqrt{2\gamma}} \widehat{e}_n^{\mathrm{D}}(\gamma) + \frac{1}{\beta}\phi_n^{\mathrm{D}}\left(\frac{\beta}{\sqrt{2\gamma}}, \gamma\right)\right| \\ &\quad + \left|\frac{1}{\beta}\phi_n^{\mathrm{SK}}(\beta) - \frac{1}{\beta}\phi_n^{\mathrm{D}}\left(\frac{\beta}{\sqrt{2\gamma}}, \gamma\right)\right| \\ &\quad + \left|e_n^{\mathrm{SK}} - \frac{1}{\beta}\phi_n^{\mathrm{SK}}(\beta)\right|,\end{aligned}$$

and using (2.15), yields the other (right) bound in (2.10), recalling that with $\{H_n^{\mathrm{SK}}(\underline{\sigma})\}$ a zero mean Gaussian process, necessarily $\widehat{e}_n^{\mathrm{SK}} = -e_n^{\mathrm{SK}}$. □

PROOF OF PROPOSITION 2.2. For $t \in [0, 1]$ we consider the interpolating Hamiltonian on Ω_n

$$H_n(\gamma, t, \underline{\sigma}) := \frac{1}{\sqrt{2\gamma}} H^{\mathrm{D}}_{\gamma(1-t)}(\underline{\sigma}) + \sqrt{t}\, H^{\mathrm{SK}}(\underline{\sigma}), \tag{2.16}$$

denoting by $Z_n(\beta, \gamma, t)$, $\phi_n(\beta, \gamma, t)$ and $\mu_{\beta,n}(\cdot; \gamma, t)$, the partition function, free energy density, and Boltzmann measure, respectively, for this interpolating Hamiltonian. Clearly, $\phi_n(\beta, \gamma, 0) = \phi_n^{\mathrm{D}}(\frac{\beta}{\sqrt{2\gamma}}, \gamma)$ and $\phi_n(\beta, \gamma, 1) = \phi_n^{\mathrm{SK}}(\beta)$. Hence,

$$\left|\phi_n^{\mathrm{D}}\left(\frac{\beta}{\sqrt{2\gamma}}, \gamma\right) - \phi_n^{\mathrm{SK}}(\beta)\right| \le \int_0^1 \left|\frac{\partial \phi_n}{\partial t}(\beta, \gamma, t)\right| \mathrm{d}t$$

and it suffices to show that $|\frac{\partial \phi_n}{\partial t}|$ is bounded, uniformly over $t \in [0, 1]$ and n, by the RHS of (2.9). To this end, associate with i.i.d. configurations $\{\underline{\sigma}^j, j \ge 1\}$ from $\mu_{\beta,n}(\cdot; \gamma, t)$ and $\ell \ge 1$, the multi-replica overlaps

$$Q_\ell \equiv \frac{1}{n}\sum_{i=1}^{n}\left(\prod_{j=1}^{\ell} \sigma_i^j\right).$$

Then, denoting by $\langle\cdot\rangle_t$ the expectation over such i.i.d. configurations $\{\underline{\sigma}^j, j \geq 1\}$, and setting $b := \beta/\sqrt{2\gamma}$, it is a simple exercise in spin glass theory (see e.g., [18]), to explicitly express the relevant derivatives as

$$\frac{\partial \phi_n}{\partial t}(\beta,\gamma,t) = \left(\frac{\partial \phi_n}{\partial t}\right)_{\rm SK} + \left(\frac{\partial \phi_n}{\partial t}\right)_{\rm D},$$

$$\left(\frac{\partial \phi_n}{\partial t}\right)_{\rm SK} = \frac{\beta^2}{4}(1 - \mathbb{E}[\langle Q_2^2\rangle_t]), \tag{2.17}$$

$$\left(\frac{\partial \phi_n}{\partial t}\right)_{\rm D} = -\gamma \log(\cosh b) + \gamma \sum_{\ell=1}^{\infty} \frac{(-1)^\ell}{\ell}(\tanh b)^\ell \mathbb{E}[\langle Q_\ell^2\rangle_t]. \tag{2.18}$$

For the reader's convenience, we detail the derivation of (2.17) and (2.18) in Section 2.4, and note in passing that the expressions on their RHS resemble the derivatives of the interpolating free energies obtained in the Gaussian and dilute spin glass models, respectively (see [23, 24]).

Now observe that $|Q_\ell| \leq 1$ for all $\ell \geq 2$ and $Q_1 = 0$ on Ω_n, hence

$$\left|\frac{\partial \phi_n}{\partial t}(\beta,\gamma,t)\right| \leq \gamma\left|\log(\cosh b) - b^2\right| + \frac{\gamma}{2}\left|(\tanh b)^2 - b^2\right| + \gamma \sum_{\ell=3}^{\infty} \frac{1}{\ell}|\tanh b|^\ell.$$

The required uniform bound on $|\frac{\partial \phi_n}{\partial t}|$ is thus a direct consequence of the elementary inequalities

$$\left|\log\cosh x - \tfrac{1}{2}x^2\right| \leq C_1 x^4, \qquad \left|y^2 - x^2\right| \leq C_2 x^4,$$
$$\left|-\log(1-y) - y - \tfrac{1}{2}y^2\right| \leq C_3 |x|^3,$$

which hold for some finite C_1, C_2, C_3 and any $y = |\tanh x|$. □

2.3. *Proof of Lemma* 2.4. Recall that $H^{\rm SK}(\underline{\sigma}) = -\frac{1}{2\sqrt{n}}\underline{\sigma}^T \widetilde{\mathbf{J}}\underline{\sigma}$ where $\widetilde{\mathbf{J}} = \{\widetilde{J}_{ij} = (J_{ij} + J_{ji})/\sqrt{2} : 1 \leq i, j \leq n\}$ is a GOE matrix. Since $\{\widetilde{J}_{ii}\}$ do not affect $U_n^{\rm SK} - \overline{U}_n^{\rm SK}$, we further set all diagonal entries of $\widetilde{\mathbf{J}}$ to zero. By symmetry of the Hamiltonian $H^{\rm SK}(\cdot)$, the configuration $\underline{\sigma}^\star$ that achieves the unconstrained ground state energy $H^{\rm SK}(\underline{\sigma}^\star) = \overline{U}_n^{\rm SK}$ is uniformly random in $\{-1,+1\}^n$. Therefore, $S_n^\star := \frac{1}{2}\sum_{i=1}^n \sigma_i^\star$ is a centered Bin$(n, 1/2)$ random variable, and by the LIL the events $B_n = \{|S_n^\star| \leq b_n\}$ hold w.h.p. for $b_n := \sqrt{n \log n}$. By definition $\overline{U}_n^{\rm SK} \geq -\frac{n}{2}\lambda_{\max}(\widetilde{\mathbf{J}}/\sqrt{n})$, hence the events $C_n = \{\overline{U}_n^{\rm SK} \geq -2n\}$ also hold w.h.p. by the a.s. convergence of the largest eigenvalue $\lambda_{\max}(\cdot)$ for Wigner matrices (see [2], Theorem 2.1.22). Consequently, hereafter our analysis is carried out on the event $\{B_n \cap C_n\}$ and without loss of generality we can and shall further assume that $S_n^\star > 0$ is integer (since n is even).

Since $\underline{\sigma}^\star$ is a global minimizer of the quadratic form $H^{\rm SK}(\underline{\sigma})$ over the hypercube $\{-1,1\}^n$, necessarily $\sigma_i^\star = \mathrm{sign}(f_i^\star)$ for

$$f_i^\star := \frac{1}{2\sqrt{n}} \sum_{j=1}^n \widetilde{J}_{ij}\sigma_j^\star.$$

Consequently, under the event C_n,

$$-2n \le \overline{U}_n^{\rm SK} = H^{\rm SK}(\underline{\sigma}^\star) = -\sum_{i=1}^n \sigma_i^\star f_i^\star = -\sum_{i=1}^n |f_i^\star|,$$

hence $R^\star := \{i \in [n] : |f_i^\star| \le 6\}$ is of size at least $(2/3)n$. Thus, for $n \ge 6b_n$, under the event $B_n \cap C_n$ we can find a collection $W^\star \subseteq \{i \in R^\star : \sigma_i^\star = +1\}$ of size $S_n^\star$ and let $\underline{\widetilde{\sigma}} \in \Omega_n$ be the configuration obtained by setting $\widetilde{\sigma}_i = -\sigma_i^\star = -1$ whenever $i \in W^\star$ while otherwise $\widetilde{\sigma}_i = \sigma_i^\star$. We obviously have then that

$$\overline{U}_n^{\rm SK} = H^{\rm SK}(\underline{\sigma}^\star) \le U_n^{\rm SK} \le H^{\rm SK}(\underline{\widetilde{\sigma}}). \tag{2.19}$$

Further, by our choices of $\underline{\widetilde{\sigma}}$ and $W^\star \subseteq R^\star$, also

$$\begin{aligned} H^{\rm SK}(\underline{\widetilde{\sigma}}) - H^{\rm SK}(\underline{\sigma}^\star) &= \frac{2}{\sqrt{n}} \sum_{i\in W^\star} \sum_{j \in [n]\setminus W^\star} \widetilde{J}_{ij}\sigma_j^\star \\ &\le 4\sum_{i\in W^\star} |f_i^\star| + \frac{4}{\sqrt{n}}\Delta(W^\star) \le 24 S_n^\star + \frac{4}{\sqrt{n}}\Delta(W^\star), \end{aligned} \tag{2.20}$$

where we define, for $W \subseteq [n]$ the corresponding partial sum

$$\Delta(W) := \sum_{i,j\in W, i<j} |\widetilde{J}_{ij}|,$$

of $\binom{|W|}{2}$ i.i.d. variables $\widetilde{J}_{ij}$. Under the event B_n we have that $S_n^\star \le b_n \le y_n := \frac{1}{32} n^{1/2+\delta}$, so by (2.19) and (2.20) it suffices to show that w.h.p. $\{\Delta(W^\star) \le x_n\}$ for $x_n = \sqrt{n} y_n$. To this end, note that by Markov's inequality, for some $c > 0$, all n and any fixed W of size $|W| \le b_n$,

$$\mathbb{P}(\Delta(W) \ge x_n) \le e^{-x_n} \mathbb{E}[e^{|\widetilde{J}|}]^{b_n^2} \le e^{-cx_n}.$$

With at most 2^n such $W \subseteq [n]$, we conclude that

$$\mathbb{P}(\sup\{\Delta(W) : W \subset [n], |W| \le b_n\} \le x_n) \to 1,$$

and in particular w.h.p. $\{\Delta(W^\star) \le x_n\}$ (under $B_n = \{S_n^\star \le b_n\}$).

2.4. *The interpolation derivatives.* Recall the Hamiltonian $H_n(\gamma, t, \underline{\sigma})$ of (2.16), the corresponding partition function $Z_n(\beta, \gamma, t)$ and free energy density $\phi_n(\beta, \gamma, t)$. We view $n^{-1} \log Z_n(\beta, \gamma, t) := \psi_n(t, \mathbf{z}, \mathbf{J})$, as a (complicated) function of the Gaussian couplings $\mathbf{J} = \{J_{ij} : 1 \le i, j \le n\}$ and the Poisson multiplicities $\mathbf{z} = \{z_{ij} : 1 \le i, j \le n\}$. Denoting by $p(t, \cdot)$ the Pois$(\gamma(1-t)/n)$ probability mass function (PMF) of z_{ij}, yields the joint PMF $\mathbf{p}(t, \mathbf{z}) = \prod_{1 \le i, j \le n} p(t, z_{ij})$, and the expression

$$\phi_n(\beta, \gamma, t) = \mathbb{E}[\psi_n(t, \mathbf{z}, \mathbf{J})] = \int \psi_n(t, \mathbf{z}, \mathbf{J}) \mathbf{p}(t, \mathbf{z})\, \mathrm{d}\mu(\mathbf{z}, \mathbf{J}), \tag{2.21}$$

where $\mu = (\nu_{\mathbb{N}})^{n^2} \otimes (\nu_{\mathbb{R}})^{n^2}$ for the counting measure $\nu_{\mathbb{N}}$ on $\mathbb{N}$ and the standard Gaussian measure $\nu_{\mathbb{R}}$ on $\mathbb{R}$. Thus,

$$\begin{aligned} \frac{\partial \phi_n}{\partial t}(\beta, \gamma, t) &= \int \frac{\partial \psi_n}{\partial t}(t, \mathbf{z}, \mathbf{J}) \mathbf{p}(t, \mathbf{z})\, \mathrm{d}\mu(\mathbf{z}, \mathbf{J}) \\ &\quad + \int \psi_n(t, \mathbf{z}, \mathbf{J}) \frac{\partial \mathbf{p}}{\partial t}(t, \mathbf{z})\, \mathrm{d}\mu(\mathbf{z}, \mathbf{J}) \\ &:= \left(\frac{\partial \phi_n}{\partial t}\right)_{\mathrm{SK}} + \left(\frac{\partial \phi_n}{\partial t}\right)_{\mathrm{D}}. \end{aligned} \tag{2.22}$$

Proceeding to verify (2.17), here $\frac{\partial H_n}{\partial t} = \frac{1}{2\sqrt{t}} H^{\mathrm{SK}}$ [since $H^{\mathrm{D}}_{\gamma(1-t)}(\cdot)$ depends on t only through the PMF of $\mathbf{z}$]. Hence,

$$\frac{\partial}{\partial t}[\log Z_n(\beta, \gamma, t)] = -\beta \left\langle \frac{\partial H_n}{\partial t}(\gamma, t, \underline{\sigma}) \right\rangle_t = -\frac{\beta}{2\sqrt{t}} \langle H^{\mathrm{SK}}(\underline{\sigma}) \rangle_t,$$

resulting with

$$\left(\frac{\partial \phi_n}{\partial t}\right)_{\mathrm{SK}} = -\frac{1}{n} \frac{\beta}{2\sqrt{t}} \mathbb{E}_{\mathbf{z}}(\mathbb{E}_{\mathbf{J}}[\langle H^{\mathrm{SK}}(\sigma) \rangle_t]).$$

Applying the Gaussian integration by parts ($\mathbb{E}[J f(J) - f'(J)] = 0$), we arrive at

$$\begin{aligned} \mathbb{E}_{\mathbf{J}}[\langle H^{\mathrm{SK}}(\underline{\sigma}) \rangle_t] &= -\frac{1}{\sqrt{2n}} \sum_{i,j} \mathbb{E}_{\mathbf{J}}[\langle J_{ij} \sigma_i \sigma_j \rangle_t] \\ &= -\frac{1}{\sqrt{2n}} \sum_{i,j} \mathbb{E}_{\mathbf{J}}\left[\frac{\mathrm{d}\langle \sigma_i \sigma_j \rangle_t}{\mathrm{d} J_{ij}}\right] \\ &= -\frac{\beta\sqrt{t}}{2n} \sum_{i,j} \mathbb{E}_{\mathbf{J}}[\langle \sigma_i^2 \sigma_j^2 \rangle_t - \langle \sigma_i \sigma_j \rangle_t^2], \end{aligned}$$

and we get (2.17) from $\langle Q_2^2 \rangle_t = n^{-2} \sum_{i,j} \langle \sigma_i \sigma_j \rangle_t^2$ (cf. [39], Lemma 1.1).

Next, to establish (2.18) let $h_{ij}(z_{ij}) := \mathbb{E}[\psi_n(t, \mathbf{z}, \mathbf{J})|z_{ij}]$, and note that the product form of $\mathbf{p}(t, \mathbf{z})$ and $\mu(\mathbf{z}, \mathbf{J})$, results with

$$\left(\frac{\partial \phi_n}{\partial t}\right)_{\mathrm{D}} = \sum_{i=1}^{n}\sum_{j=1}^{n} \int h_{ij}(z) \frac{\partial p}{\partial t}(t, z)\, \mathrm{d}\nu_{\mathbb{N}}(z). \tag{2.23}$$

The ijth integral on the RHS of (2.23) is merely the value of $(-\gamma/n)g'(\lambda)$, where $g(\lambda) = \mathbb{E}[f(z)]$ for $f = h_{ij}$ and $z \sim \mathrm{Pois}(\lambda)$ at $\lambda = \gamma(1-t)/n$. Differentiating the Pois(λ) PMF one has the identity $g'(\lambda) = \mathbb{E}[f(z+1) - f(z)]$ (under mild regularity conditions on f). This crucial observation transforms (2.23) into

$$\left(\frac{\partial \phi_n}{\partial t}\right)_{\mathrm{D}} = -\frac{\gamma}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{E}\big[h_{ij}(z_{ij}+1) - h_{ij}(z_{ij})\big]. \tag{2.24}$$

Here, $\psi_n(t, \cdot, \cdot) = n^{-1}\log Z_n(\beta, \gamma, t)$ and adding one to z_{ij} corresponds to an extra copy of the edge (i, j) in the dilute Ising model of Hamiltonian $\frac{1}{\sqrt{2\gamma}} H^{\mathrm{D}}_{\gamma(1-t)}(\underline{\sigma})$. Consequently, setting $b := \frac{\beta}{\sqrt{2\gamma}}$,

$$\begin{aligned} &h_{ij}(z_{ij}+1) - h_{ij}(z_{ij}) \\ &\quad = \frac{1}{n}\mathbb{E}\big[\log\langle e^{b\sigma_i\sigma_j}\rangle_t | z_{ij}\big] \\ &\quad = \frac{1}{n}\mathbb{E}\big[\log\{\cosh(b)\big[1 + \tanh(b)\langle\sigma_i\sigma_j\rangle_t\big]\}|z_{ij}\big], \end{aligned} \tag{2.25}$$

since $e^{by} = \cosh(b)[1 + \tanh(b)y]$ for the $\{-1, +1\}$-valued $y = \sigma_i\sigma_j$. Combining (2.24) and (2.25), we obtain by the Taylor series for $-\log(1+x)$ (when $-1 < x < 1$), that

$$\begin{aligned} \left(\frac{\partial \phi_n}{\partial t}\right)_{\mathrm{D}} &= -\frac{\gamma}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{E}\big[\log\{\cosh(b)\big[1 + \tanh(b)\langle\sigma_i\sigma_j\rangle_t\big]\}\big] \\ &= -\gamma\log\cosh(b) + \gamma\sum_{\ell=1}^{\infty}\frac{(-1)^\ell}{\ell}\big(\tanh(b)\big)^\ell \mathbb{E}\left[\frac{1}{n^2}\sum_{i,j=1}^{n}\big(\langle\sigma_i\sigma_j\rangle_t\big)^\ell\right] \\ &= -\gamma\log\cosh(b) + \gamma\sum_{\ell=1}^{\infty}\frac{(-1)^\ell}{\ell}\big(\tanh(b)\big)^\ell \mathbb{E}\big[\langle Q_\ell^2\rangle_t\big], \end{aligned}$$

as stated in (2.18).

3. Graph comparison: Proof of Theorem 1.5. The notion of uniform random γ-regular graph refers to drawing such graph uniformly from among all γ-regular simple graphs on n-vertices, provided, as we assume throughout, that $n\gamma$ is even. We instead denote by $G^{\mathrm{Reg}}(n, \gamma)$ the more tractable configuration model, where each vertex is equipped with γ half-edges and a multi-graph (of possible self-loops and multiple edges) is formed by a uniform random matching of the collection of all γn half-edges. Indeed, as mentioned in the context of Erdős–Rényi graphs (see the start of Section 2), for γ bounded the matching in $G^{\text{Reg}}(n, \gamma)$ produces a simple graph with probability bounded away from zero, and conditional on being simple this graph is uniformly random. Consequently, any property that holds w.h.p. for the configuration model multi-graph $G^{\text{Reg}}(n, \gamma)$ must also hold w.h.p. for the simple uniform random γ-regular graph.

Our strategy for proving Theorem 1.5 is to start from the random regular multi-graph $G_1 \sim G^{\text{Reg}}(n, \gamma)$, deleting some edges and "rewiring" some of the existing ones to obtain a new graph G_2 which is approximately an Erdős–Rényi random graph of $n\gamma_-/2$ edges, where $\gamma_- := \gamma - \sqrt{\gamma}\log\gamma$. Then, with Theorem 1.2 providing us with the typical behavior of extreme bisections of G_2, the main challenge is to control the effect of our edge transformations well enough to handle the minimum and maximum bisections of G_1.

Specifically, drawing i.i.d. $X_i \sim \text{Pois}(\gamma_-)$, we let $Z_i := (\gamma - X_i)_+$ and color Z_i of the γ half-edges of each vertex $i \in [n]$ by BLUE (B). All other half-edges are colored RED (R). Matching the half-edges uniformly, without regard to their colors, we obtain a graph $G_1 \sim G^{\text{Reg}}(n, \gamma)$. Our coloring decomposes G_1 to the sub-graph G_{RR} consisting of all the RR edges and $G_{\mathsf{RB}} \cup G_{\mathsf{BB}}$ having all other edges, which we in turn decompose to the sub-graph G_{BB} consisting of the BB edges and G_{RB} having all the multi-color edges (i.e., RB and BR). To transform G_1 to G_2, we first delete all edges of G_{BB}, disconnect all the multi-colored RB edges and delete all the B half-edges that as a result became unmatched. We then form a new sub-graph $\widetilde{G}_{\mathsf{RR}}$ by uniformly re-matching all the free R half-edges (in case there is an odd number of such half-edges we leave one of them free as a self-loop). The graph G_2 has the vertex set $[n]$ and $E(G_2) = E(G_{\mathsf{RR}}) \cup E(\widetilde{G}_{\mathsf{RR}})$.

We represent by Ω_n the collection of all bisections for a graph G having n vertices, denoting by $\text{cut}_G(\underline{\sigma})$ the cut size for the partition between $\{i \in [n] : \sigma_i = -1\}$ and its complement. Then, for any $\underline{\sigma} \in \Omega_n$ we have

$$\text{cut}_{G_1}(\underline{\sigma}) = \text{cut}_{G_2}(\underline{\sigma}) - \text{cut}_{\widetilde{G}_{\mathsf{RR}}}(\underline{\sigma}) + \text{cut}_{G_{\mathsf{RB}} \cup G_{\mathsf{BB}}}(\underline{\sigma}). \tag{3.1}$$

We control the LHS of (3.1) by three key lemmas, starting with the following consequence of Theorem 1.2, proved in Section 3.1 that gives sharp estimates on the dominant part, namely $\text{cut}_{G_2}(\underline{\sigma})$.

LEMMA 3.1. *We have, w.h.p. as $n \to \infty$,*

$$\frac{\mathsf{mcut}(G_2)}{n} = \frac{\gamma_-}{4} - P_*\sqrt{\frac{\gamma}{4}} + o_\gamma(\sqrt{\gamma}), \tag{3.2}$$

$$\frac{\mathsf{MCUT}(G_2)}{n} = \frac{\gamma_-}{4} + P_*\sqrt{\frac{\gamma}{4}} + o_\gamma(\sqrt{\gamma}). \tag{3.3}$$

Our next lemma, proved in Section 3.2, shows that while both the B half-edge deletions and the R half-edge re-matching that follows, may affect the cut size, on the average (with respect to our random matching), at the scale of interest to us they cancel out each other.

LEMMA 3.2. *Uniformly over all* $\underline{\sigma} \in \Omega_n$,

$$\mathbb{E}[\mathrm{cut}_{G_{\mathsf{RB}}}(\underline{\sigma})] = n\left(\frac{\sqrt{\gamma}\log\gamma}{2} + O_\gamma(1)\right) + o(n), \tag{3.4}$$

$$\mathbb{E}[\mathrm{cut}_{\widetilde{G}_{\mathsf{RR}}}(\underline{\sigma})] = n\left(\frac{\sqrt{\gamma}\log\gamma}{4} + O_\gamma(1)\right) + o(n), \tag{3.5}$$

$$\mathbb{E}[\mathrm{cut}_{G_{\mathsf{BB}}}(\underline{\sigma})] = n\left(\frac{(\log\gamma)^2}{4} + o_\gamma(1)\right) + o(n). \tag{3.6}$$

The last result we need, is the following uniform bound on the fluctuations, proved in Section 3.3, that allows us to control the effect of the edge rewiring on the extremal bisections.

LEMMA 3.3. *There exists* C *sufficiently large*, *independent of* n *and* γ, *such that*

$$\mathbb{P}\Big[\sup_{\underline{\sigma}\in\Omega_n} \big|\mathrm{cut}_{\mathcal{A}}(\underline{\sigma}) - \mathbb{E}[\mathrm{cut}_{\mathcal{A}}(\underline{\sigma})]\big| > Cn\gamma^{1/4}\sqrt{\log\gamma}\Big] = o(1), \tag{3.7}$$

where $\mathcal{A}$ *may be distributed as* $G_{\mathsf{RB}} \cup G_{\mathsf{BB}}$ *or* $\widetilde{G}_{\mathsf{RR}}$.

Turning to prove Theorem 1.5, we have from (3.1) and Lemma 3.3 that w.h.p. as $n \to \infty$,

$$\begin{aligned}\sup_{\underline{\sigma}\in\Omega_n} &\big|\mathrm{cut}_{G_1}(\underline{\sigma}) - \mathrm{cut}_{G_2}(\underline{\sigma}) + \mathbb{E}[\mathrm{cut}_{\widetilde{G}_{\mathsf{RR}}}(\underline{\sigma})] - \mathbb{E}[\mathrm{cut}_{G_{\mathsf{RB}}\cup G_{\mathsf{BB}}}(\underline{\sigma})]\big| \\ &= no_\gamma(\sqrt{\gamma}).\end{aligned} \tag{3.8}$$

In view of Lemma 3.2, we deduce from (3.8) that w.h.p. as $n \to \infty$,

$$\sup_{\underline{\sigma}\in\Omega_n}\left|\mathrm{cut}_{G_1}(\underline{\sigma}) - \mathrm{cut}_{G_2}(\underline{\sigma}) - n\frac{\sqrt{\gamma}\log\gamma}{4}\right| = no_\gamma(\sqrt{\gamma}) + o(n).$$

This in turn implies that w.h.p.

$$\mathsf{mcut}(G_1) = \mathsf{mcut}(G_2) + n\frac{\sqrt{\gamma}\log\gamma}{4} + no_\gamma(\sqrt{\gamma}) + o(n),$$

$$\mathsf{MCUT}(G_1) = \mathsf{MCUT}(G_2) + n\frac{\sqrt{\gamma}\log\gamma}{4} + no_\gamma(\sqrt{\gamma}) + o(n),$$

and Theorem 1.5 thus follows from Lemma 3.1 (recall that $\gamma = \gamma_- + \sqrt{\gamma}\log\gamma$).

3.1. *Proof of Lemma* 3.1. Let G_n^{int} be the random graph generated from the configuration model with i.i.d. $X_i \sim \text{Pois}(\gamma_-)$ degrees. We denote by $G^{\mathsf{clon}}(n, \gamma_-)$ the sub-graph obtained by independently deleting each half-edge of G_n^{int} with probability $1/n$, before matching them. By the thinning property of the Pois law, $G^{\mathsf{clon}}(n, \gamma_-)$ has the law of the Poisson–Cloning model, where one first generates i.i.d. $\zeta_i \sim \text{Pois}(\frac{n-1}{n}\gamma_-)$, then draws a random graph from the configuration model with ζ_i half-edges at vertex i. Recall [32] that the $G_I(n, \frac{\gamma_-}{n})$ and $G^{\mathsf{clon}}(n, \gamma_-)$ models are mutually contiguous. Further, $\gamma_-/\gamma \to 1$, and so by Theorem 1.2, w.h.p.

$$\frac{\mathsf{mcut}(G^{\mathsf{clon}}(n, \gamma_-))}{n} = \frac{\gamma_-}{4} - P_*\sqrt{\frac{\gamma}{4}} + o_\gamma(\sqrt{\gamma}), \tag{3.9}$$

$$\frac{\mathsf{MCUT}(G^{\mathsf{clon}}(n, \gamma_-))}{n} = \frac{\gamma_-}{4} + P_*\sqrt{\frac{\gamma}{4}} + o_\gamma(\sqrt{\gamma}). \tag{3.10}$$

Next, note that for any two graphs $\mathcal{G}_1, \mathcal{G}_2$ on n vertices,

$$\big|\mathsf{MCUT}(\mathcal{G}_1) - \mathsf{MCUT}(\mathcal{G}_2)\big| \le \big|E(\mathcal{G}_1)\Delta E(\mathcal{G}_2)\big|,$$
$$\big|\mathsf{mcut}(\mathcal{G}_1) - \mathsf{mcut}(\mathcal{G}_2)\big| \le \big|E(\mathcal{G}_1)\Delta E(\mathcal{G}_2)\big|.$$

W.h.p. our coupling has $\sum_i (X_i - \zeta_i) = O(1)$ half-edges from G_n^{int} not also in $G^{\mathsf{clon}}(n, \gamma_-)$. Hence, $|E(G_n^{\text{int}})\Delta E(G^{\mathsf{clon}}(n, \gamma_-))| = O(1)$ and (3.9)–(3.10) extend to $\mathsf{mcut}(G_n^{\text{int}})$ and $\mathsf{MCUT}(G_n^{\text{int}})$, respectively.

We proceed to couple G_n^{int} and G_2 such that $|E(G_n^{\text{int}})\Delta E(G_2)| \le n o_\gamma(\sqrt{\gamma})$ w.h.p. thereby yielding the desired conclusion. To this end, G_2 could have alternatively been generated by *one* uniform random matching of only the $X_i' := \min\{X_i, \gamma\}$ RED half-edges that each vertex i has in G_1 (for completeness, we prove this statement in Lemma 3.4). We can thus couple G_2 and G_n^{int} by first forming G_n^{int}, then independently for $i = 1, \dots, n$ color in RED uniformly at random X_i' of the X_i half-edges of vertex i, with all remaining half-edges colored BROWN. Now, to get G_2 we delete all BB edges, disconnect all RB edges and delete the resulting B half-edges, then uniformly re-match all the free R half edges (for Lemma 3.4 applies again in this setting). The claimed bound on $|E(G_n^{\text{int}})\Delta E(G_2)|$ follows since the total number of B half-edges in G_n^{int} is w.h.p. at most

$$2n\mathbb{E}(X_1 - X_1') = 2n\mathbb{E}[(X_1 - \gamma)_+] = nO_\gamma(1), \tag{3.11}$$

where the RHS follows by normal approximation to $\text{Pois}(\gamma_-)$ and our choice of $\gamma_- = \gamma - \sqrt{\gamma}\log\gamma$.

3.2. *Proof of Lemma* 3.2. We first prove (3.4), utilizing the fact that the distribution of $\text{cut}_{G_{\mathsf{RB}}}(\underline{\sigma})$ is the same for all $\underline{\sigma} \in \Omega_n$. Hence,

$$\mathbb{E}[\text{cut}_{G_{\mathsf{RB}}}(\underline{\sigma})] = \mathbb{E}[\mathbb{E}_{\underline{\sigma}^\star}[\text{cut}_{G_{\mathsf{RB}}}(\underline{\sigma}^\star)]] \tag{3.12}$$

for $\underline{\sigma}^\star$ chosen uniformly from Ω_n. Given the graph G_1, we have

$$\mathbb{E}_{\underline{\sigma}^\star}[\text{cut}_{G_{\mathsf{RB}}}(\underline{\sigma}^\star)] = \frac{|E_{\mathsf{RB}}|}{2(1 - 1/n)}, \tag{3.13}$$

where E_{RB} denotes the set of RB edges in G_1 excluding self-loops. Next, noting that the expected number of edges in G_1 excluding self-loops is $\frac{n(n-1)\gamma^2}{2(n\gamma-1)}$ and the probability that an edge connecting two distinct vertices is colored RB is $2\frac{\mathbb{E}[Z_1]}{\gamma}(1-\frac{\mathbb{E}[Z_1]}{\gamma})$, we have

$$\mathbb{E}\big[|E_{\mathsf{RB}}|\big]=\frac{n(n-1)\gamma^2}{n\gamma-1}\frac{\mathbb{E}[Z_1]}{\gamma}\left(1-\frac{\mathbb{E}[Z_1]}{\gamma}\right), \tag{3.14}$$

where $Z_1 \sim (\gamma - X_1)_+$ and $X_1 \sim \mathrm{Pois}(\gamma_-)$. We get (3.4) out of (3.12) and (3.14) upon observing that

$$\begin{aligned}\mathbb{E}[Z_1]&=\gamma-\mathbb{E}[X_1]+\mathbb{E}\big[(\gamma-X_1)_-\big]\\&=\gamma-\gamma_-+\mathbb{E}\big[(X_1-\gamma)_+\big]=\sqrt{\gamma}\log\gamma+O_\gamma(1)\end{aligned} \tag{3.15}$$

[see (3.11) for the right-most identity]. By an analogous calculation, we find that for all $\underline{\sigma}\in\Omega_n$,

$$\begin{aligned}\mathbb{E}\big[\mathrm{cut}_{G_{\mathsf{BB}}}(\underline{\sigma})\big]&=\frac{n(n-1)\gamma^2}{2(n\gamma-1)}\left(\frac{\mathbb{E}[Z_1]}{\gamma}\right)^2\frac{1}{2(1-1/n)}\\&=n\left[\frac{1}{4}(\log\gamma)^2+o_\gamma(1)\right]+o(n).\end{aligned}$$

Turning to (3.5), the same argument as in (3.12) implies that

$$\mathbb{E}\big[\mathrm{cut}_{\widetilde{G}_{\mathsf{RR}}}(\underline{\sigma})\big]=\mathbb{E}\big[\mathbb{E}_{\underline{\sigma}^\star}\big[\mathrm{cut}_{\widetilde{G}_{\mathsf{RR}}}(\underline{\sigma}^\star)\big]\big],$$

for $\underline{\sigma}^\star$ chosen uniformly from Ω_n. Further, similarly to (3.13) we find that given the graph $\widetilde{G}_{\mathsf{RR}}$,

$$\mathbb{E}_{\underline{\sigma}^\star}\big[\mathrm{cut}_{\widetilde{G}_{\mathsf{RR}}}(\underline{\sigma}^\star)\big]=\frac{|E_2|}{2(1-1/n)}, \tag{3.16}$$

where E_2 denotes the set of edges in $\widetilde{G}_{\mathsf{RR}}$ excluding self-loops and $1/(2(1-1/n))$ is the probability that $\underline{\sigma}^\star$ induces different signs on the end points of a fixed edge. Recall that $|E(G_{\mathsf{RB}})|-|E_{\mathsf{RB}}|$ and $\frac{1}{2}|E(G_{\mathsf{RB}})|-|E_2|$ count the number of self-loops in G_{RB} and $\widetilde{G}_{\mathsf{RR}}$, respectively. The expected number of such self-loops is $O(1)$ as $n\to\infty$, hence $\mathbb{E}[|E_2|]=\frac{1}{2}\mathbb{E}[|E_{\mathsf{RB}}|]+O(1)$, which upon comparing (3.13) to (3.16) yields the required expression of (3.5).

3.3. *Proof of Lemma* 3.3. Starting with $\mathcal{A}=G_{\mathsf{RB}}\cup G_{\mathsf{BB}}$, clearly for any $x_n>0$,

$$\mathbb{P}\Big[\sup_{\underline{\sigma}\in\Omega_n}\big|\mathrm{cut}_{\mathcal{A}}(\underline{\sigma})-\mathbb{E}\big[\mathrm{cut}_{\mathcal{A}}(\underline{\sigma})\big]\big|\geq 2x_n\Big]\leq p_1(n)+p_2(n), \tag{3.17}$$

where $\mathbf{Z} = (Z_1, \dots, Z_n)$ count the number of BLUE half-edges at each vertex of G_1 and

$$p_1(n) = \mathbb{P}\Big[\sup_{\underline{\sigma} \in \Omega_n} \big|\mathrm{cut}_{\mathcal{A}}(\underline{\sigma}) - c(\underline{\sigma}, \mathbf{Z})\big| \ge x_n\Big], \tag{3.18}$$

$$p_2(n) = \mathbb{P}\Big[\sup_{\underline{\sigma} \in \Omega_n} \big|c(\underline{\sigma}, \mathbf{Z}) - \mathbb{E}\big[\mathrm{cut}_{\mathcal{A}}(\underline{\sigma})\big]\big| \ge x_n\Big], \tag{3.19}$$

for $c(\underline{\sigma}, \mathbf{Z}) := \mathbb{E}[\mathrm{cut}_{\mathcal{A}}(\underline{\sigma})|\mathbf{Z}]$. Letting $S_n(\mathbf{Z}) = \sum_{i=1}^n Z_i$, note that w.h.p. $\mathbf{Z} \in \mathcal{E}_n$ for $\mathcal{E}_n = \{\mathbf{z} : |S_n(\mathbf{z}) - n\mathbb{E}[Z_1]| \le b_n\}$ and $b_n = \sqrt{n \log n}$. Hence, by a union bound over $\underline{\sigma} \in \Omega_n$ we get that

$$p_1(n) \le 2^n \max_{\mathbf{z} \in \mathcal{E}_n} \max_{\underline{\sigma} \in \Omega_n} \mathbb{P}\big[\big|\mathrm{cut}_{\mathcal{A}}(\underline{\sigma}) - c(\underline{\sigma}, \mathbf{Z})\big| \ge x_n | \mathbf{Z} = \mathbf{z}\big] + o(1). \tag{3.20}$$

We next apply Azuma–Hoeffding inequality to control the RHS of (3.20). To this end, fixing $\mathbf{z} \in \mathcal{E}_n$ and half-edge colors such that $\{\mathbf{Z} = \mathbf{z}\}$, we form G_1 by sequentially pairing a candidate half-edge to uniformly chosen second half-edge, using first BLUE half-edges as candidates for the pairing (until all of them are exhausted). Then, fixing $\underline{\sigma} \in \Omega_n$, we consider Doob's martingale $M_k = \mathbb{E}[\mathrm{cut}_{\mathcal{A}}(\underline{\sigma})|\mathscr{F}_k]$, for the sigma-algebra $\mathscr{F}_k$ generated by all half-edge colors and the first $k \ge 0$ edges to have been paired. This martingale starts at $M_0 = c(\underline{\sigma}, \mathbf{Z})$, has differences $|M_k - M_{k-1}|$ uniformly bounded by some universal finite nonrandom constant κ (independent on n, $\underline{\sigma}$ and $\mathbf{z}$), while $M_\ell = \mathrm{cut}_{\mathcal{A}}(\underline{\sigma})$ for all $\ell \ge S_n(\mathbf{z})$ [since the sub-graph $\mathcal{A} = G_{\mathsf{RB}} \cup G_{\mathsf{BB}}$ is completely formed within our sequential matching first $S_n(\mathbf{z})$ steps]. The bounded difference property of M_k follows easily from the "switching" argument in [45], Theorem 2.19. Thus, from Azuma–Hoeffding inequality we get that for $\mathbf{z} \in \mathcal{E}_n$,

$$\begin{aligned} &\mathbb{P}\big[\big|\mathrm{cut}_{\mathcal{A}}(\underline{\sigma}) - c(\underline{\sigma}, \mathbf{Z})\big| \ge x_n | \mathbf{Z} = \mathbf{z}\big] \\ &\quad \le 2 \exp\Big(-\frac{x_n^2}{8\kappa^2 S_n(\mathbf{z})}\Big) \\ &\quad \le 2 \exp\Big(-\frac{x_n^2}{8\kappa^2 (n\mathbb{E}[Z_1] + b_n)}\Big). \end{aligned} \tag{3.21}$$

Recall (3.15) that $\mathbb{E}[Z_1] = \sqrt{\gamma} \log \gamma + O_\gamma(1)$, hence upon choosing $x_n = Cn\gamma^{1/4}\sqrt{\log \gamma}$ for some $C^2 > 8\kappa^2 \log 3$, we find that the RHS of (3.20) decays to zero as $n \to \infty$.

Turning to control $p_2(n)$, for $i \in [n]$ and $1 \le j \le Z_i$, let $I_{ij}(\underline{\sigma}) = 1$ if the jth B half-edge of vertex i is matched to some half-edge from the opposite side of the partition induced by $\underline{\sigma}$, and $I_{ij}(\underline{\sigma}) = 0$ otherwise. Then

$$\mathrm{cut}_{\mathcal{A}}(\underline{\sigma}) = \sum_{i=1}^n \sum_{j=1}^{Z_i} I_{ij}(\underline{\sigma}) - \mathrm{cut}_{G_{\mathsf{BB}}}(\underline{\sigma}). \tag{3.22}$$

For i such that $\sigma_i = 1$ and $1 \le j \le Z_i$ we similarly set $I'_{ij}(\underline{\sigma}) = 1$ if the jth B half-edge of vertex i is matched to a B half-edge of a vertex from the opposite side, and $I'_{ij} = 0$ otherwise. Clearly, then

$$\mathrm{cut}_{G_{\mathsf{BB}}}(\underline{\sigma}) = \sum_{\{i:\sigma_i=1\}} \sum_{j=1}^{Z_i} I'_{ij}(\underline{\sigma}),$$

so setting $S_n^+(\underline{\sigma}, \mathbf{Z}) := \sum_{\{i:\sigma_i=1\}} Z_i$, we have from (3.22) that

$$\begin{aligned} c(\underline{\sigma}, \mathbf{Z}) &= \sum_{i=1}^{n} \sum_{j=1}^{Z_i} \mathbb{P}[I_{ij}(\underline{\sigma}) = 1|\mathbf{Z}] - \sum_{\{i:\sigma_i=1\}} \sum_{j=1}^{Z_i} \mathbb{P}[I'_{ij}(\underline{\sigma}) = 1|\mathbf{Z}] \\ &= S_n(\mathbf{Z}) \frac{(n\gamma)/2}{n\gamma - 1} - S_n^+(\underline{\sigma}, \mathbf{Z}) \frac{S_n(\mathbf{Z}) - S_n^+(\underline{\sigma}, \mathbf{Z})}{n\gamma - 1}. \end{aligned} \tag{3.23}$$

Considering the extreme values of the RHS of (3.23) yields that for all $\underline{\sigma} \in \Omega_n$,

$$\frac{1}{2} S_n(\mathbf{Z}) \left(1 - \frac{S_n(\mathbf{Z})}{2n\gamma}\right) \le c(\underline{\sigma}, \mathbf{Z}) \left(1 - \frac{1}{n\gamma}\right) \le \frac{1}{2} S_n(\mathbf{Z}),$$

from which we deduce that

$$\begin{aligned} & n \frac{\mathbb{E}[Z_1]}{2} \left(1 - \frac{\mathbb{E}[Z_1]}{2\gamma}\right) + o(n) \\ &\quad \le \inf_{\mathbf{Z} \in \mathcal{E}_n} \inf_{\underline{\sigma} \in \Omega_n} \{c(\underline{\sigma}, \mathbf{Z})\} \\ &\quad \le \sup_{\mathbf{Z} \in \mathcal{E}_n} \sup_{\underline{\sigma} \in \Omega_n} \{c(\underline{\sigma}, \mathbf{Z})\} \le n \frac{\mathbb{E}[Z_1]}{2} + o(n). \end{aligned} \tag{3.24}$$

Further, while proving Lemma 3.2 we have shown that

$$\mathbb{E}[\mathrm{cut}_{\mathcal{A}}(\underline{\sigma})] = n \frac{\mathbb{E}[Z_1]}{2} \frac{n\gamma}{(n\gamma - 1)} \left(1 - \frac{\mathbb{E}[Z_1]}{2\gamma}\right),$$

hence from (3.24) and (3.15) it follows that

$$\sup_{\mathbf{Z} \in \mathcal{E}_n} \sup_{\underline{\sigma} \in \Omega_n} \left|c(\underline{\sigma}, \mathbf{Z}) - \mathbb{E}[\mathrm{cut}_{\mathcal{A}}(\underline{\sigma})]\right| \le n \frac{\mathbb{E}[Z_1]^2}{4\gamma} + o(n) \le n(\log\gamma)^2 + o(n),$$

and since w.h.p. $\mathbf{Z} \in \mathcal{E}_n$, we conclude that $p_2(n) = o(1)$.

Next, we consider the graph $\mathcal{A} = \widetilde{G}_{\mathsf{RR}}$ and proceeding in a similar manner we have the decomposition (3.17), except for replacing in this case $\mathbf{Z}$ in (3.18)–(3.19) by the count $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ of the number of R half-edges at each vertex at the initiation of the second step in forming $\widetilde{G}_{\mathsf{RR}}$. The total number $S_n(\mathbf{Y})$ of R half-edges to be matched in that second step is less than the initial number $S_n(\mathbf{Z})$ of B half-edges. Consequently, if $\mathbf{Z} \in \mathcal{E}_n$ then $\mathbf{Y} \in \mathcal{E}_n^+ := \{\mathbf{y} : S_n(\mathbf{y}) \le n\mathbb{E}[Z_1] + b_n\}$, and we have again a bound of the type (3.20) on $p_1(n)$, just taking here the maximum over $\mathbf{y} \in \mathcal{E}_n^+$ instead of $\mathbf{z} \in \mathcal{E}_n$. Further, we repeat the martingale construction that resulted with the RHS of (3.21). Specifically, here $\mathcal{F}_0$ is the sigma-algebra of $\mathbf{Y}$, namely knowing the degrees of vertices in $\widetilde{G}_{\mathsf{RR}}$, and we expose in $\mathcal{F}_k$ the first k edges to have been paired en-route to the uniform matching that forms $\widetilde{G}_{\mathsf{RR}}$. As before the Doob's martingale $M_k = \mathbb{E}[\mathrm{cut}_{\widetilde{G}_{\mathsf{RR}}}(\underline{\sigma})|\mathcal{F}_k]$ has uniformly bounded differences, starting at $M_0 = c(\underline{\sigma}, \mathbf{Y})$ and with the same choice of x_n the desired bound on $p_1(n)$ follows upon observing that $M_\ell = \mathrm{cut}_{\widetilde{G}_{\mathsf{RR}}}(\underline{\sigma})$ for all $\ell \geq S_n(\mathbf{y})$, and in particular when $\ell = n\mathbb{E}[Z_1] + b_n$. Turning to deal with $p_2(n)$ in this context, by the same reasoning that led to (3.23) we find that for $S = S_n(\mathbf{y}) \geq 2$ and $S^+(\underline{\sigma}) = S_n^+(\underline{\sigma}, \mathbf{y})$,

$$\begin{aligned} c(\underline{\sigma}, \mathbf{y}) &= \mathbb{E}[\mathrm{cut}_{\widetilde{G}_{\mathsf{RR}}}(\underline{\sigma})|\mathbf{Y} = \mathbf{y}] \\ &= \frac{S^+(\underline{\sigma})(S - S^+(\underline{\sigma}))}{S-1} = \frac{S^2 - (2S^+(\underline{\sigma}) - S)^2}{4(S-1)}. \end{aligned} \tag{3.25}$$

While proving Lemma 3.2, we have shown that

$$\overline{c}_n := 4\mathbb{E}[\mathrm{cut}_{\widetilde{G}_{\mathsf{RR}}}(\underline{\sigma})] = n[\mathbb{E}[Z_1] + O_\gamma(1)] + o(n)$$

and that $\overline{c}_n$ is constant over $\underline{\sigma} \in \Omega_n$. As $\overline{c}_n \geq 6x_n$ for n and γ large, while $|S^2/(S-1) - S|$ is uniformly bounded, we deduce from (3.25) that for any $y_n \leq x_n$,

$$\begin{aligned} &\{|2S^+(\sigma) - S'| < x_n\} \cap \{|S - S'| < y_n\} \cap \{|S' - \overline{c}_n| < x_n\} \\ &\quad \Longrightarrow \quad \{|4c(\underline{\sigma}, \mathbf{y}) - \overline{c}_n| < 4x_n\}. \end{aligned}$$

Now, w.h.p. $S' = S_n(\mathbf{Z})$ is in $\mathcal{I}_n := [n\mathbb{E}[Z_1] - b_n, n\mathbb{E}[Z_1] + b_n]$ with $|S' - \overline{c}_n| < x_n$, and taking the union over $\underline{\sigma} \in \Omega_n$, we get similarly to the derivation of (3.20) that

$$\begin{aligned} p_2(n) &\leq 2^n \max_{s_n \in \mathcal{I}_n} \max_{\underline{\sigma} \in \Omega_n} \mathbb{P}[|2S^+(\underline{\sigma}) - s_n| \geq x_n, S > s_n - y_n | S' = s_n] \\ &\qquad + \max_{s_n \in \mathcal{I}_n} \mathbb{P}(S \leq s_n - y_n | S' = s_n) + o(1) \\ &=: p_3(n) + p_4(n) + o(1). \end{aligned} \tag{3.26}$$

Starting with $2N = n\gamma$ half-edges of G_1 of whom $S' = s_n$ colored B (while all others are colored R), the nonnegative number $S' - S$ of half-edges in G_{BB} is stochastically dominated by a $\mathrm{Bin}(s_n, s_n/(2N - s_n))$ random variable. For $s_n \in \mathcal{I}_n$ the latter Binomial has mean $n\mathbb{E}[Z_1]^2/(\gamma - \mathbb{E}[Z_1]) + o(n)$, hence in view of (3.15), $p_4(n) = o(1)$ provided $y_n \geq 3n(\log\gamma)^2$. For bounding $p_3(n)$, we assume w.l.o.g. that $\sigma_i = 1$ iff $i \leq n/2$, with $S^+(\underline{\sigma})$ the total number of R half-edges for vertices $i \leq n/2$, which are matched to B half-edges by the uniform matching in the first step of forming $\widetilde{G}_{\mathsf{RR}}$. Fixing the total number s_n of B half-edges in G_1, clearly $S^+(\underline{\sigma})$ is stochastically decreasing in the number $S'_+ = \sum_{i=1}^{n/2} Z_i$ of B half-edges among vertices $i \leq n/2$. Thus, it suffices to bound $p_3(n)$ in the extreme cases, of $S'_+ = 0$, and of $S'_+ = s_n$. The uniform matching of the first step induces a sampling without replacement with $S^+(\underline{\sigma})$ denoting the number of marked balls when drawing a sample of (random) size $S \in (s_n - y_n, s_n]$, uniformly without replacement from an urn containing $2N - s_n$ balls, of which either N or $N - s_n$ balls are marked. By stochastic monotonicity, it suffices to consider the relevant tails of $S^+(\underline{\sigma}) - s_n/2$ only in the extreme cases of $S = s_n - y_n$ and $S = s_n$. As $2N/n = \gamma$, $s_n/n \ll \gamma$ and $y_n \ll x_n$, standard tail bounds for the hyper-geometric distribution [8] imply that $p_3(n) = o(1)$ for γ sufficiently large, thereby completing the proof.

3.4. *A pairing lemma.* We include here, for completeness, the formal proof of the fidelity of the two stage pairing procedure (which was used in our preceding arguments).

LEMMA 3.4. *Given* 2ℓ *labeled balls of color* R *and* $2m$ *labeled balls of color* B *for some* $m \leq \ell$, *we get a uniform random pairing of the* R *balls by the following two-step procedure*:

(a) *First match the* $2(m+\ell)$ *balls uniformly at random to obtain some* RR, RB *and* BB *pairs*.

(b) *Remove all* B *balls and uniformly re-match the* R *balls which were left unmatched due to the removal of the* B *balls*.

PROOF. We use the notation $(2k-1)!! = (2k-1)(2k-3)\cdots 1$ and $[m]_k = m(m-1)\cdots(m-k+1)$ and let $\mathscr{P}$ denote the random pairing of the 2ℓ R balls by our two-step procedure [which first generated $2s$ pairs of type RB, $(m-s)$ of type BB and $(\ell - s)$ of type RR]. We then have that for any fixed final pairing P of the R balls,

$$\begin{aligned}\mathbb{P}[\mathscr{P} = P] &= \sum_{s=0}^{m} \binom{\ell}{s} \frac{[2m]_{2s}(2(m-s)-1)!!}{(2m+2\ell-1)!!(2s-1)!!} \\ &= \frac{\ell!(2m)!2^{\ell}}{(2m+2\ell)!} \sum_{s=0}^{m} 2^{2s} \binom{m+\ell}{m-s, \ell-s, 2s} = \frac{1}{(2\ell-1)!!},\end{aligned}$$

where the last identity follows upon observing that $\sum_{s=0}^{m} 2^{2s}\binom{m+\ell}{m-s,\ell-s,2s} = \binom{2(m+\ell)}{2\ell}$. □

4. From bisection to cut: Proof of Theorem 1.6. Let $\mathcal{I}^{\pm}(\underline{\sigma}) := \{i : \sigma_i = \pm 1\}$ be the partition of $[n]$ induced by $\underline{\sigma}$ and $m(\underline{\sigma}) := \frac{1}{2}\sum_{i=1}^{n} \sigma_i$ the difference in size of its two sides. Note that by the invariance of $\mathrm{cut}_G(\underline{\sigma})$ under the symmetry $\underline{\sigma} \to -\underline{\sigma}$, it suffices to compare the cuts in $\mathcal{S}_n^+ = \{\underline{\sigma} \in \{-1,+1\}^n : m(\underline{\sigma}) \geq 0\}$ to those in Ω_n. To this end, define the map $T : \mathcal{S}_n^+ \to \Omega_n$ where we flip the spins at the subset $V(\underline{\sigma})$ of smallest $m(\underline{\sigma})$ indices within $\mathcal{I}^+(\underline{\sigma})$, thereby moving all those indices to $\mathcal{I}^-(T(\underline{\sigma}))$. Let $X(\underline{\sigma})$, $Y(\underline{\sigma})$ and $Z(\underline{\sigma})$ count the number of edges from $V(\underline{\sigma})$ to $\mathcal{I}^-(\underline{\sigma})$, $\mathcal{I}^-(T(\underline{\sigma}))$ and $\mathcal{I}^+(T(\underline{\sigma})) = \mathcal{I}^+(\underline{\sigma}) \setminus V(\underline{\sigma})$, respectively. Fixing $0 < \delta < 1/4$, let

$$\mathcal{S}^\star = \{\underline{\sigma} \in \mathcal{S}_n^+ : m(\underline{\sigma}) \le \gamma^{-\delta} n\}. \tag{4.1}$$

Then, for $\underline{\sigma}^\star \in \mathcal{S}_n^+$ such that $\mathsf{MaxCut}(G_n) = \mathrm{cut}_{G_n}(\underline{\sigma}^\star)$ we have

$$\begin{aligned}\mathsf{MaxCut}(G_n) &= \mathrm{cut}_{G_n}(T(\underline{\sigma}^\star)) + X(\underline{\sigma}^\star) - Z(\underline{\sigma}^\star)\\ &\le \mathsf{MCUT}(G_n) + Y(\underline{\sigma}^\star) - Z(\underline{\sigma}^\star).\end{aligned}$$

Considering the union over $\underline{\sigma} \in \mathcal{S}^\star$, we get that

$$\begin{aligned}&\mathbb{P}[\mathsf{MaxCut}(G_n) > \mathsf{MCUT}(G_n) + \Delta_n]\\ &\quad \le 2^n \max_{\underline{\sigma}\in\mathcal{S}^\star} \mathbb{P}[Y(\underline{\sigma}) - Z(\underline{\sigma}) > \Delta_n] + \mathbb{P}[\underline{\sigma}^\star \notin \mathcal{S}^\star]\\ &\quad =: q_1(n) + q_2(n).\end{aligned} \tag{4.2}$$

In proving part (a) of Theorem 1.6, we consider w.l.o.g. the Erdős–Rényi random graphs $G_n \sim G_I(n, \frac{\gamma}{n})$ as in Remark 1.3. For fixed $\underline{\sigma} \in \mathcal{S}_n^+$ the independent variables $Y(\underline{\sigma})$ and $Z(\underline{\sigma})$ are $\mathrm{Bin}(N', \gamma/n)$ and $\mathrm{Bin}(N, \gamma/n)$, respectively, for $N' \le N$ [specifically, $N' = m(\underline{\sigma})(n - m(\underline{\sigma}) - 1)/2$ and $N = m(\underline{\sigma})(n/2)$]. Thus, $Y(\underline{\sigma}) - Z(\underline{\sigma})$ is stochastically dominated by $Z'(\underline{\sigma}) - Z(\underline{\sigma})$ and computing the m.g.f. of the latter variable we get by Markov's inequality that for any $\theta > 0$,

$$\mathbb{P}[Y(\underline{\sigma}) - Z(\underline{\sigma}) > \Delta_n] \le e^{-2\theta\Delta_n}\left[1 + \frac{4\gamma}{n}\sinh^2(\theta)\right]^N.$$

Setting $\Delta_n = n\gamma^{\psi/2}$ for some $\psi \in (1-\delta, 1)$ fixed and the maximal $N = \frac{1}{2}n^2\gamma^{-\delta}$ for $\underline{\sigma} \in \mathcal{S}^\star$, we deduce that

$$\begin{aligned}&\limsup_{n\to\infty} n^{-1}\log \mathbb{P}[Y(\underline{\sigma}) - Z(\underline{\sigma}) > \Delta_n]\\ &\quad \le -2[\theta\gamma^{\psi/2} - \gamma^{1-\delta}\sinh^2(\theta)] =: -J.\end{aligned} \tag{4.3}$$

Since $\psi > 1 - \delta$, we have that $\gamma^{1-\delta}\sinh^2(\gamma^{-\psi/2}) \to 0$, so taking $\theta = \gamma^{-\psi/2}$ results with $J > 1$ for all γ large enough, in which case $q_1(n) = o(1)$ [see (4.2)]. As for controlling $q_2(n)$, recall Theorem 1.2 that w.h.p. $\mathsf{MaxCut}(G_n) \ge \mathsf{MCUT}(G_n) \ge n\gamma/4$. Hence, considering the union over $\underline{\sigma} \notin \mathcal{S}^\star$ we have that

$$q_2(n) \le 2^n \max_{\underline{\sigma}\notin\mathcal{S}^\star} \mathbb{P}\left(\mathrm{cut}_{G_n}(\underline{\sigma}) \ge \frac{n\gamma}{4}\right).$$

For our Erdős–Rényi graphs $\mathrm{cut}_{G_n}(\underline{\sigma}) \sim R_k := \mathrm{Bin}(k(n-k), \frac{\gamma}{n})$ with $k = \frac{n}{2} - m(\underline{\sigma})$. Taking the maximal $k^\star := \frac{n}{2} - n\gamma^{-\delta}$ for $\underline{\sigma} \notin \mathcal{S}^\star$ and computing the relevant m.g.f. yields, similarly to (4.3), that for $f_1(\theta) = e^\theta - 1$, $f_2(\theta) = e^\theta - \theta - 1$ and any $\theta > 0$,

$$\limsup_{n\to\infty} n^{-1} \log \mathbb{P}\left(R_{k^\star} \geq \frac{n\gamma}{4}\right) \leq \frac{\gamma}{4} f_2(\theta) - \gamma^{1-2\delta} f_1(\theta) := -J'. \tag{4.4}$$

Since $\gamma f_2(\gamma^{-1/2})$ is uniformly bounded while $\gamma^{1-2\delta} f_1(\gamma^{-1/2}) = O_\gamma(\gamma^{1/2-2\delta})$ diverges (due to our choice of $\delta < 1/4$), it follows that for $\theta = \gamma^{-1/2}$ and γ large enough, $J' \geq 1$ hence $q_2(n) = o(1)$, thereby completing the proof.

The Erdős–Rényi nature of the graph G_n is only used for deriving the large deviation bounds (4.3) and (4.4). While slightly more complicated, similar computations apply also for $G^{\mathrm{Reg}}(n,\gamma)$. Indeed, in this case $Y(\underline{\sigma}) - Z(\underline{\sigma})$ corresponds to the sum of spins in a random sample of size $\gamma m(\underline{\sigma})$ taken without replacement from a balanced population of γn spins [so by standard tail estimates for the hypergeometric law, here too the LHS of (4.3) is at most -1 for any γ large enough]. Similarly, now $R_{k^\star}$ counts the pairs formed by uniform matching of γn items, between a fixed set of $\gamma k^\star$ items and its complement [so by arguments similar to those we used when proving Lemma 3.3, the LHS of (4.4) is again at most -1 for large γ]. With the rest of the proof unchanged, we omit its details.

Acknowledgment. We thank the anonymous referee for pointing out reference [20] and numerous comments, which helped improve the presentation of the paper.

REFERENCES

[1] ALON, N. (1997). On the edge-expansion of graphs. *Combin. Probab. Comput.* **6** 145–152. MR1447810

[2] ANDERSON, G. W., GUIONNET, A. and ZEITOUNI, O. (2009). *An Introduction to Random Matrices*. Cambridge Univ. Press, Cambridge.

[3] AUFFINGER, A. and CHEN, W. (2015). The Parisi formula has a unique minimizer. *Comm. Math. Phys.* **335** 1429–1444. MR3320318

[4] BAYATI, M., GAMARNIK, D. and TETALI, P. (2013). Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. *Ann. Probab.* **41** 4080–4115. MR3161470

[5] BOLLOBÁS, B. (1988). The isoperimetric number of random regular graphs. *European J. Combin.* **9** 241–244. MR0947025

[6] BOLLOBÁS, B. (2001). *Random Graphs*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **73**. Cambridge Univ. Press, Cambridge. MR1864966

[7] BOUCHERON, S., LUGOSI, G. and MASSART, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford Univ. Press, Oxford. MR3185193

[8] CHVÁTAL, V. (1979). The tail of the hypergeometric distribution. *Discrete Math.* **25** 285–287. MR0534946

[9] COJA-OGHLAN, A. (2007). On the Laplacian eigenvalues of $G_{n,p}$. *Combin. Probab. Comput.* **16** 923–946. MR2351691

[10] COPPERSMITH, D., GAMARNIK, D., HAJIAGHAYI, M. T. and SORKIN, G. B. (2004). Random MAXSAT, random MAXCUT, and their phase transitions. *Random Structures Algorithms* **24** 502–545. MR2060633

[11] CRISANTI, A. and RIZZO, T. (2002). Analysis of the ∞-replica symmetry breaking solution of the Sherrington–Kirkpatrick model. *Phys. Rev. E* (3) **65** 046137, 9. MR1918001
[12] DAUDÉ, H., MARTÍNEZ, C., RASENDRAHASINA, V. and RAVELOMANANA, V. (2012). The MAX-CUT of sparse random graphs. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms* 265–271. ACM, New York. MR3205214
[13] DECELLE, A., KRZAKALA, F., MOORE, C. and ZDEBOROVÁ, L. (2011). Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Phys. Rev. E* (3) **84** 066106.
[14] DÍAZ, J., PETIT, J. and SERNA, M. J. (2002). A survey on graph layout problems. *ACM Comput. Surv.* **34** 313–356.
[15] DÍAZ, J., SERNA, M. J. and WORMALD, N. C. (2004). Computation of the bisection width for random d-regular graphs. In *LATIN* 2004: *Theoretical Informatics*. *Lecture Notes in Computer Science* **2976** 49–58. Springer, Berlin. MR2094779
[16] FEIGE, U. and KRAUTHGAMER, R. (2000). A polylogarithmic approximation of the minimum bisection. In *Foundations of Computer Science* 105–115. Redondo Beach, CA.
[17] FEIGE, U. and OFEK, E. (2005). Spectral techniques applied to sparse random graphs. *Random Structures Algorithms* **27** 251–275. MR2155709
[18] FRANZ, S. and LEONE, M. (2003). Replica bounds for optimization problems and diluted spin systems. *J. Stat. Phys.* **111** 535–564. MR1972121
[19] FRIEDMAN, J. (2003). A proof of Alon's second eigenvalue conjecture. In *Proceedings of the* 35*th Symposium on Theory of Computing* 720–724. San Diego, CA.
[20] FRIEZE, A. M. and ŁUCZAK, T. (1992). On the independence and chromatic numbers of random regular graphs. *J. Combin. Theory Ser. B* **54** 123–132. MR1142268
[21] GAMARNIK, D. and LI, Q. (2014). On the Max-Cut over sparse random graph. Available at arXiv:1411.1698v1.
[22] GUERRA, F. (2003). Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.* **233** 1–12. MR1957729
[23] GUERRA, F. and TONINELLI, F. L. (2002). The thermodynamic limit in mean field spin glass models. *Comm. Math. Phys.* **230** 71–79. MR1930572
[24] GUERRA, F. and TONINELLI, F. L. (2004). The high temperature region of the Viana–Bray diluted spin glass model. *J. Stat. Phys.* **115** 531–555. MR2070106
[25] HALPERIN, E. and ZWICK, U. (2002). A unified framework for obtaining improved approximation algorithms for maximum graph bisection problems. *Random Structures Algorithms* **20** 382–402. MR1900614
[26] HASTAD, J. (1999). Some optimal inapproximability results. In *STOC '97* (*el Paso*, *TX*) 1–10. ACM, New York. MR1715618
[27] HOFSTAD, R. (2009). Random graphs and complex networks. Preprint. Available at http://www.win.tue.nl/~rhofstad/.
[28] JAGANNATH, A. and TOBASCO, I. (2016). Dynamic programming approach to the Parisi variational problem. *Proc. Amer. Math. Soc.* **144** 3135–3150.
[29] JANSON, S., ŁUCZAK, T. and RUCINSKI, A. (2000). *Random Graphs*. Wiley, New York. MR1782847
[30] KHORUNZHIY, O., KIRSCH, W. and MÜLLER, P. (2006). Lifshitz tails for spectra of Erdős–Rényi random graphs. *Ann. Appl. Probab.* **16** 295–309. MR2209343
[31] KHOT, S. (2004). Ruling out PTAS for graph min-bisection, densest subgraph and bipartite clique. In *Foundations of Computer Science* 136–145. Roma, Italy.
[32] KIM, J. H. (2006). Poisson cloning model for random graphs. In *International Congress of Mathematicians*. *Vol. III* 873–897. Eur. Math. Soc., Zürich. MR2275710
[33] LUCZAK, M. J. and MCDIARMID, C. (2001). Bisecting sparse random graphs. *Random Structures Algorithms* **18** 31–38. MR1799801

[34] MASSOULIÉ, L. (2014). Community detection thresholds and the weak Ramanujan property. In *STOC'14—Proceedings of the* 2014 *ACM Symposium on Theory of Computing* 694–703. ACM, New York. MR3238997

[35] MÉZARD, M. and MONTANARI, A. (2009). *Information*, *Physics*, *and Computation*. Oxford Univ. Press, Oxford. MR2518205

[36] MÉZARD, M., PARISI, G. and VIRASORO, M. (1986). *Spin Glass Theory and Beyond*: *An Introduction to the Replica Method and Its Applications*. *World Scientific Lecture Notes in Physics* **9**. World Scientific, Singapore.

[37] MONTANARI, A. and SEN, S. (2015). Semidefinite programs on sparse random graphs and their application to community detection. Available at arXiv:1504.05910.

[38] MOSSEL, E., NEEMAN, J. and SLY, A. (2013). A proof of the block model threshold conjecture. Available at arXiv:1311.4115.

[39] PANCHENKO, D. (2013). *The Sherrington–Kirkpatrick Model*. Springer, New York. MR3052333

[40] PERCUS, A. G., ISTRATE, G., GONÇALVES, B., SUMI, R. Z. and BOETTCHER, S. (2008). The peculiar phase structure of random graph bisection. *J*. *Math*. *Phys*. **49** 125219, 13. MR2484350

[41] POLJAK, S. and TUZA, Z. (1995). Maximum cuts and large bipartite subgraphs. In *Combinatorial Optimization* (*New Brunswick*, *NJ*, 1992–1993). *DIMACS Ser. Discrete Math*. *Theoret*. *Comput*. *Sci*. **20** 181–244. Amer. Math. Soc., Providence, RI. MR1338615

[42] SCHMIDT, M. J. (2008). Replica symmetry breaking at low temperatures. Ph.D. thesis, Univ. Würzburg.

[43] TALAGRAND, M. (2003). *Spin Glasses*: *A Challenge for Mathematicians*: *Cavity and Mean Field Models*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete*. 3. *Folge*. *A Series of Modern Surveys in Mathematics* **46**. Springer, Berlin. MR1993891

[44] TALAGRAND, M. (2006). The Parisi formula. *Ann*. *of Math*. (2) **163** 221–263. MR2195134

[45] WORMALD, N. C. (1999). Models of random regular graphs. In *Surveys in Combinatorics*, 1999 (*Canterbury*). *London Mathematical Society Lecture Note Series* **267** 239–298. Cambridge Univ. Press, Cambridge. MR1725006

[46] ZDEBOROVÁ, L. and BOETTCHER, S. (2010). A conjecture on the maximum cut and bisection width in random regular graphs. *J*. *Stat*. *Mech*. *Theory Exp*. **2010** P02020.

A. DEMBO
DEPARTMENT OF MATHEMATICS AND STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305
USA
E-MAIL: amir@math.stanford.edu

A. MONTANARI
DEPARTMENT OF ELECTRICAL ENGINEERING AND STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305
USA
E-MAIL: montanari@stanford.edu

S. SEN
DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305
USA
E-MAIL: ssen90@stanford.edu