NOISE SENSITIVITY OF CRITICAL RANDOM GRAPHS

EYAL LUBETZKY AND YUVAL PELED

ABSTRACT. We study noise sensitivity of properties of the largest components $(\mathcal{C}_j)_{j\geq 1}$ of the random graph $\mathcal{G}(n,p)$ in its critical window $p = (1+\lambda n^{-1/3})/n$. For instance, is the property "$|\mathcal{C}_1|$ exceeds its median size" noise sensitive? Roberts and Şengül (2018) proved that the answer to this is yes if the noise ε is such that $\varepsilon \gg n^{-1/6}$, and conjectured the correct threshold is $\varepsilon \gg n^{-1/3}$. That is, the threshold for sensitivity should coincide with the critical window— as shown for the existence of long cycles by the first author and Steif (2015).

We prove that for $\varepsilon \gg n^{-1/3}$ the pair of vectors $n^{-2/3}(|\mathcal{C}_j|)_{j\geq 1}$ before and after the noise converges in distribution to a pair of *i.i.d.* random variables, whereas for $\varepsilon \ll n^{-1/3}$ the ℓ^2-distance between the two goes to 0 in probability. This confirms the above conjecture: any Boolean function of the vector of rescaled component sizes is sensitive in the former case and stable in the latter.

We also look at the effect of the noise on the metric space $n^{-1/3}(\mathcal{C}_j)_{j\geq 1}$. E.g., for $\varepsilon \geq n^{-1/3+o(1)}$, we show that the joint law of the spaces before and after the noise converges to a product measure, implying noise sensitivity of any property seen in the limit, e.g., "the diameter of $\mathcal{C}_1$ exceeds its median."

1. INTRODUCTION

A sequence of Boolean functions f_N on a product space of N Bernoulli variables is *noise sensitive*, a notion introduced in the pioneering paper of Benjamini, Kalai and Schramm [4], if the resampling of an ε-fraction of the inputs, for a fixed $\varepsilon > 0$, results in an asymptotically independent copy of f_N in the limit as $N \to \infty$.

More precisely, consider i.i.d. Bernoulli(p_N) random variables $\{\omega(i)\}_{i=1}^N$ and functions $f_N : \{0,1\}^N \to \{0,1\}$, such that $\mathbb{P}(f_N(\omega) = 1)$ is uniformly bounded away from 0 and 1. Let ω^ε be the result of applying the noise operator T_ε on ω, whereby every ω_i is resampled (replaced by a new i.i.d. Bernoulli(p) variable) with probability ε, independently of the other coordinates. The functions f_N are said to be ε_N-noise sensitive for $\varepsilon = \varepsilon_N > 0$ if $\mathrm{Cov}(f_N(\omega), f_N(\omega^\varepsilon)) \to 0$ as $N \to \infty$.

A remarkable argument of [4] (see [9, §1]) showed that if $p_N \equiv \frac{1}{2}$ and f_N is increasing w.r.t. the partial order on the hypercube $\{0,1\}^N$, then noise sensitivity w.r.t. any (and every) $0 < \varepsilon < 1$ occurs if and only if $\lim_{N\to\infty} \sum_{i=1}^N \hat{f}_N(\{i\})^2 = 0$, where $\hat{f}_N(\{i\})$ is the Fourier coefficient of f_N corresponding to the singleton i. This ceases to apply when $p_N \to 0$ fast with N, as Benjamini et al. [4, §6.4] noted: "*When p tends to zero with n, new phenomena occur. Consider, for example, random graphs on n vertices with edge probability $p = n^{-a}$.*"

Consider the Erdős–Rényi random graph $\mathcal{G}(n,p)$—the undirected graph on n vertices where each edge is independently present with probability p—in the critical window of its celebrated phase transition, $p = (1 + \lambda n^{-1/3})/n$ for fixed $\lambda \in \mathbb{R}$. Aldous [3] famously found the limit law of its vector $\mathbf{X}_n$ of rescaled component sizes. Here we show a dichotomy for noise sensitivity of functions in the critical window: if the noise parameter ε is such that $\varepsilon \gg n^{-1/3}$ then any non-degenerate function of $\mathbf{X}_n$ is noise sensitive, whereas if $\varepsilon = o(n^{-1/3})$ it is noise stable (an opposite notion). Analogous results are obtained for functions of the (rescaled) metric space of G_n.

Note that the threshold $\varepsilon \gg n^{-1/3}$ to achieve noise sensitivity exactly coincides with the width of the critical window ($p = (1 + O(n^{-1/3})p_c$ where $p_c = 1/n$). As part of their study of noise sensitivity in $\mathcal{G}(n,p)$, the first author and Steif [12] demonstrated this for the property "containing a cycle of length $\ell \in (an^{1/3}, bn^{1/3})$," and gave a heuristic explanation of why it should occur for a "global" function f_n: Viewing the operator T_ε as a composition of sprinkling and percolation, the latter induces a subcritical graph where f_n is degenerate, thus decorrelated (see [12, §1.2]).

Roberts and Şengül [17] recently established that, for the largest component $\mathcal{C}_1$ of $\mathcal{G}(n,p)$ at $p = 1/n$, the property that $|\mathcal{C}_1| \geq an^{2/3}$ ($a > 0$ fixed) is noise sensitive (N.B. the median of $|\mathcal{C}_1|$ has order $n^{2/3}$ at criticality). More precisely, they showed that this property is noise sensitive whenever $\varepsilon \gg n^{-1/6}$, and conjectured that the correct threshold for noise sensitivity is $\varepsilon \gg n^{-1/3}$, in line with the above intuition.

Let $\mathcal{C}_i = \mathcal{C}_i(G)$ be the i-th largest component of a graph G, and write

$$\mathbf{X}(G) = |V(G)|^{-2/3}(|\mathcal{C}_1(G)|, |\mathcal{C}_2(G)|, \ldots),$$

viewing it as an element of $\ell^2_\downarrow$, the space of sequences $(x_i)_{i \geq 1}$ wherein the $x_i \in \mathbb{R}_+$ are in decreasing order and $\sum x_i^2 < \infty$. Aldous [3] showed that when $G_n \sim \mathcal{G}(n,p)$ for $p = (1 + \lambda n^{-1/3})/n$ with $\lambda \in \mathbb{R}$ fixed, $\mathbf{X}(G_n) \xrightarrow{\mathrm{d}} \boldsymbol{\xi}(\lambda)$ in the ℓ^2-topology, where $\boldsymbol{\xi}(\lambda)$ is the sequence of excursion sizes, in decreasing order, of Brownian motion reflected at 0 with a parabolic drift as given by $\widehat{W}_t^{(\lambda)} = W_t^{(\lambda)} - \min_{0 \leq s \leq t} W_s^{(\lambda)}$ where $W_t^{(\lambda)} = B_t + \lambda t - t^2/2$ for standard Brownian motion B_t.

Our first result confirms the above conjecture of Roberts and Şengül, showing moreover that the "noise sensitivity exponent" (dubbed so in [9, §12]) is indeed $1/3$, coinciding with the width of the critical window, for any Boolean function of $\mathbf{X}(G_n)$.

Theorem 1. *Let $G_n \sim \mathcal{G}(n,p)$ with $p = (1 + \lambda n^{-1/3})/n$ for fixed $\lambda \in \mathbb{R}$, and let $G_n^\varepsilon = T_\varepsilon(G_n)$, where T_ε is the (Bonami–Beckner) noise operator for $0 < \varepsilon < 1$.*

  (i) If $\varepsilon = \varepsilon(n)$ satisfies $\varepsilon^3 n \to \infty$ as $n \to \infty$ then $(\mathbf{X}(G_n), \mathbf{X}(G_n^\varepsilon))$ converges in distribution in $(\ell^2_\downarrow, \|\cdot\|_2)$ to a pair of independent copies of $\boldsymbol{\xi}(\lambda)$.

  (ii) If $\varepsilon = \varepsilon(n)$ satisfies $\varepsilon^3 n \to 0$ as $n \to \infty$ then $\|\mathbf{X}(G_n) - \mathbf{X}(G_n^\varepsilon)\|_2 \xrightarrow{\mathrm{P}} 0$.

Part (i) implies that, for $\varepsilon \gg n^{-1/3}$, any function of $\mathbf{X}(G_n)$ is noise sensitive, whereas Part (ii) implies that, for $\varepsilon \ll n^{-1/3}$, any such function is noise stable in the sense that $\mathbb{P}(f_n(G_n) \neq f_n(G_n^\varepsilon)) \to 0$. The joint law of $\mathbf{X}(G_n), \mathbf{X}(G_n^\varepsilon)$ in the case where $\varepsilon = tn^{-1/3}$ for a fixed $t > 0$ was recently obtained by Rossignol [18] (a special case of the scaling limit of the pair of metric spaces, as mentioned below), whence the limit is nontrivial—neither two i.i.d. variables nor two identical ones.

The key to the proof of Part (i) is a short argument giving rigor to the heuristic that $f_n(G_n), f_n(G_n^\varepsilon)$ decorrelate iff the noise ε "draws" G_n outside of the critical window, where f_n is degenerate. Suppose $\mathbb{P}(f_n(\omega) = 1) \to \alpha$ for some $0 < \alpha < 1$. By constructing ω via sprinkling on a subcritical ω_0 comprised of i.i.d. $\mathrm{Be}(p_0)$ variables for $p_0 = (1 - \varepsilon)p/(1 - \varepsilon p)$ (the percolation portion of the noise), one finds (see Cor. 2.2 and 2.3) that f_n is noise sensitive if and only if $\mathbb{P}(f_n(\omega) = 1 \mid \omega_0) \xrightarrow{\mathrm{P}} \alpha$. This reduces our problem to finding a set of graphs $\mathcal{H}_n$ such that $\mathbb{P}(H_n \in \mathcal{H}_n) \to 1$ for the subcritical $H_n \sim \mathcal{G}(n, p_0)$, and for every $H_n \in \mathcal{H}_n$, the conditional law of $\mathbf{X}(G_n)$ given H_n converges to $\boldsymbol{\xi}(\lambda)$ in distribution. We then utilize a proposition of Aldous—driving his convergence result in [3]—which states that if the initial state of the multiplicative coalescent satisfies certain conditions, then its limit law is that of $\boldsymbol{\xi}(\lambda)$. Verifying these conditions for almost every H_n concludes Theorem 1(i).

We also study the exponent in the noise sensitivity threshold for properties of a critical $\mathcal{G}(n,p)$ graph that are read, instead of from $\mathbf{X}(G_n)$, rather from the rescaled metric spaces arising from the connected components of G_n. For a graph G, let

$$\mathbf{M}(G) = (\mathsf{M}_1(G), \mathsf{M}_2(G), \dots),$$

with $\mathsf{M}_j(G)$ the measured metric space on $\mathcal{C}_j(G)$ equipped with the graph metric scaled by $|V(G)|^{-1/3}$ and counting measure on vertices multiplied by $|V(G)|^{-2/3}$. A breakthrough result of Addario-Berry, Broutin and Goldschmidt [1] established the scaling limit of the rescaled metric spaces in Gromov–Hausdorff distance (d_{GH}). This was extended to convergence in Gromov–Hausdorff-Prokhorov distance (d_{GHP}) by those authors and Miermont in [2, Sec. 4.1], and we let $\mathfrak{M}(\lambda) = (\mathfrak{M}_j(\lambda) : j \geq 1)$ denote the sequence of random measured metric spaces that describe the limit, where every $\mathfrak{M}_j$ is an element of $\mathscr{M}$, the space of isometry-equivalence classes of compact measured metric spaces. Further let $d_{\mathrm{GHP}}^p(\mathbf{M}, \mathbf{M}') = (\sum_i d_{\mathrm{GHP}}(\mathsf{M}_i, \mathsf{M}_i')^p)^{\frac{1}{p}}$ and set $\mathbb{L}_p = \{\mathbf{M} \in \mathscr{M}^{\mathbb{N}} : \sum_i d_{\mathrm{GHP}}(\mathsf{M}_i, \mathsf{Z})^p < \infty\}$ where Z is the zero metric space.

With this notation, our second result shows that the noise sensitivity exponent is $1/3$ for every Boolean function of $\mathbf{M}(G_n)$ when G_n is a critical $\mathcal{G}(n,p)$ graph.

Theorem 2. *Let $G_n \sim \mathcal{G}(n,p)$ with $p = (1 + \lambda n^{-1/3})/n$ for fixed $\lambda \in \mathbb{R}$, and let $G_n^\varepsilon = T_\varepsilon(G_n)$, where T_ε is the (Bonami–Beckner) noise operator for $0 < \varepsilon < 1$.*

(i) Fix $\delta > 0$. If $\varepsilon(n) \geq n^{-1/3+\delta}$ then, for every j, the pair $(\mathsf{M}_j(G_n), \mathsf{M}_j(G_n^\varepsilon))$ converges in distribution to a pair of independent copies of $\mathfrak{M}_j(\lambda)$ in $(\mathscr{M}, d_{\mathrm{GHP}})$. Moreover, the pair $(\mathbf{M}(G_n), \mathbf{M}(G_n^\varepsilon))$ converges in distribution to a pair of independent copies of $\mathfrak{M}(\lambda)$ in $(\mathbb{L}_4, d_{\mathrm{GHP}}^4)$.

(ii) If $\varepsilon(n) = o(n^{-1/3})$ then $d_{\mathrm{GHP}}^4(\mathbf{M}(G_n), \mathbf{M}(G_n^\varepsilon)) \xrightarrow{\mathrm{P}} 0$ in $\mathbb{L}_4$.

As a consequence of Part (i) of this theorem, in the regime $\varepsilon \geq n^{-1/3+\delta}$ we have that any nondegenerate function of $\mathbf{M}(G_n)$ is noise sensitive. For example, this includes properties such as "the diameter of $\mathcal{C}_1(G_n)$ exceeds $an^{1/3}$," (or its special case asking if this diameter exceeds its median), as well as the property "G_n contains a cycle of length between $an^{1/3}$ and $bn^{1/3}$" whose noise sensitivity was established in [12]. As mentioned there, the heuristic for the $1/3$ noise sensitivity exponent may fail for a "local" property (see Remark 3.5 in that work). Theorem 2 thus shows that properties seen in the scaling limit $\mathfrak{M}(\lambda)$ qualify as "global" to that end.

The proof of Theorem 2(i) uses the same approach described above: one wishes to show that when constructing G_n via sprinkling edges on a suitable H_n, the conditional law of $\mathbf{M}(G_n)$ given H_n weakly converges to $\mathfrak{M}(\lambda)$ for almost every H_n. Instead of the key proposition of Aldous on the convergence of the multiplicative coalescent from an initial state satisfying certain conditions, we employ an analog of it for metric spaces by Bhamidi, Broutin, Sen and Wang [5]. To apply this powerful general result, of the various conditions that H_n should satisfy, one must control the mean and variance of the average distance between all pairs of connected vertices.

As mentioned in the context of Theorem 1, the joint law of $(\mathbf{M}(G_n), \mathbf{M}(G_n^\varepsilon))$ in the case where $\varepsilon = tn^{-1/3}$ for fixed $t > 0$ was recently derived by Rossignol [18], as a special case of a convergence result of $\{\mathbf{M}(G_n(t))\}_{t \geq 0}$ where $\{G_n(t)\}_{t \geq 0}$ is the dynamical percolation process on the critical $\mathcal{G}(n,p)$ graph with intensity $n^{-1/3}$. In principal, one could build on this to recover Part (ii) of Theorems 1 and 2 (which corresponds to $t = o(1)$), yet given that the analysis of the case $t = o(1)$ is considerably simpler, we include the short proofs of stability for completeness.

This paper is organized as follows. Section 2 describes the short reduction of noise sensitivity to controlling the conditional variance of $f_n(\omega)$ given a subcritical ω_0. Section 3 includes properties of the subcritical graph $\mathcal{G}(n, p_0)$ which are needed for the proofs. Sections 4 and 5 then establish Theorems 1 and 2, resp., via these tools.

2. CONDITIONAL VARIANCE GIVEN A COMMON CORE

For some sequences $0 < p(n) < 1$, $0 < \varepsilon(n) < 1$ and $N(n)$, where N is an integer going to ∞ with n, let $\Omega_n = \{0,1\}^N$ and let $\omega \sim \mathrm{Be}(p)^{\otimes N}$, the random variable taking values in Ω_n whose N coordinates are i.i.d. $\mathrm{Be}(p)$ random variables.

For a given ω, let $\omega^\varepsilon = T_\varepsilon \omega$, where T_ε is the (Bonami–Beckner) noise operator: $\omega^\varepsilon(i) = \omega(i)$ with probability $1-\varepsilon$ whereas $\omega(i)$ is a new independent $\mathrm{Be}(p)$ variable with probability ε, independently between the N coordinates.

Next, define ω_0, ω_1 and ω_1' as follows. Let

$$p_0 := \frac{p(1-\varepsilon)}{1 - \varepsilon p}\,, \tag{2.1}$$

and let $\omega_0 \sim \mathrm{Be}(p_0)^{\otimes N}$. Conditional on ω_0, let ω_1 and ω_1' be independent outcomes of sprinkling on ω_0:

$$\omega_1(i) = \begin{cases} 1 & \omega_0(i) = 1 \\ \mathrm{Be}(p_1) & \omega_0(i) = 0 \end{cases} \qquad \text{where} \qquad p_1 := \varepsilon p\,, \tag{2.2}$$

independently among the $N - |\omega_0|$ coordinates where $\omega_0(i) = 0$, and similarly—via another set of $N - |\omega_0|$ independent $\mathrm{Be}(p_1)$ variables—for ω_1'.

Observation 2.1. *For any $0 < \varepsilon < 1$ and $0 < p < 1$ we have $(\omega_1, \omega_1') \stackrel{\mathrm{d}}{=} (\omega_1, \omega_1^\varepsilon)$.*

Proof. Note that ω_1 is distributed as $\mathrm{Be}(p)^{\otimes N}$ (as is ω_1') by our choice of p_0, p_1, as the N variables are independent and $\mathbb{P}_n(\omega_1(i) = 1) = p_0 + (1 - p_0)p_1 = p$. Hence, such is also the distribution of ω_1^ε (invariant under the noise operator). It thus remains to verify that the covariances of a given coordinate match. Indeed, for every i,

$$\mathbb{P}_n(\omega_1(i) = \omega_1'(i) = 1) = p_0 + (1 - p_0)p_1^2 = p\frac{1 - \varepsilon + (1 - p)\varepsilon^2 p}{1 - \varepsilon p} = p(1 - \varepsilon + \varepsilon p)\,,$$

whereas

$$\mathbb{P}_n(\omega_1(i) = \omega_1^\varepsilon(i) = 1) = p(1 - \varepsilon + \varepsilon p)\,,$$

implying that the covariances indeed match, as required. ∎

Corollary 2.2. *Let $f_n : \Omega_n \to \{0,1\}$. Then*

$$\mathrm{Cov}(f_n(\omega_1), f_n(\omega_1^\varepsilon)) = \mathrm{Var}\left(\mathbb{P}(f_n(\omega_1) = 1 \mid \omega_0)\right)\,. \tag{2.3}$$

In particular, if $\mathbb{P}_n(f_n(\omega_1) = 1) \to \alpha$ as $n \to \infty$ for some $0 < \alpha < 1$, then the sequence (f_n) is ε-noise sensitive if and only if $\mathbb{P}_n(f_n(\omega_1) = 1 \mid \omega_0) \xrightarrow{\mathrm{P}} \alpha$.

Proof. Eq. (2.3) follows from Observation 2.1, as the term $\mathrm{Cov}(f_n(\omega_1), f_n(\omega_1') \mid \omega_0)$ in the law of total covariance for $\mathrm{Cov}(f_n(\omega_1), f_n(\omega_1'))$ vanishes by the conditional independence of ω_1 and ω_1' given ω_0. Setting $X_n = \mathbb{P}(f_n(\omega_1) = 1 \mid \omega_0)$ for brevity, it thus follows by definition that if the (uniformly bounded) X_n satisfies $X_n \xrightarrow{\mathrm{P}} \alpha$ then $\mathrm{Var}(X_n) \to 0$, whence (f_n) is ε-noise sensitive.

Conversely, if (f_n) is ε-noise sensitive then $\mathrm{Var}(X_n) \to 0$, so by Chebyshev's inequality $X_n - \mathbb{E}X_n \xrightarrow{\mathrm{P}} 0$, where $\mathbb{E}X_n = \mathbb{P}(f_n(\omega_1) = 1) \to \alpha$ by assumption. ∎

If we further suppose that X_n $(n \geq 1)$ is a measurable map from Ω_n to a common topological space $\mathbb{S}$, which converges weakly to some limit X_∞—denoting the measure $\mathbb{P}(X_\infty \in \cdot)$ on $\mathbb{S}$ is denoted μ_∞ for brevity—then we have the following.

Corollary 2.3. *Suppose that for each n there exists $W_n \subseteq \Omega_n$ with $\mathbb{P}_n(W_n) \to 1$, the variables $X_n : \Omega_n \to \mathbb{S}$ have $X_n \xrightarrow{\mathrm{d}} X_\infty$, and for every μ_∞-continuity set $S \subseteq \mathbb{S}$,*

$$\lim_{n \to \infty} \mathbb{E}\left[\left|\mathbb{P}_n(X_n(\omega_1) \in S \mid \omega_0) - \mathbb{P}(X_\infty \in S)\right|^2 \mathbf{1}_{\{\omega_0 \in W_n\}}\right] = 0. \qquad (2.4)$$

Then $(X_n(\omega_1), X_n(\omega_1^\varepsilon)) \xrightarrow{\mathrm{d}} (X_\infty, X'_\infty)$ as $n \to \infty$, where X_∞ and X'_∞ are i.i.d.

Proof. Let $f_n(\omega) = \mathbf{1}_{\{X_n(\omega) \in S_1\}}$ and $g_n(\omega) = \mathbf{1}_{\{X_n(\omega) \in S_2\}}$ for two μ_∞-continuity sets $S_1, S_2 \subseteq \mathbb{S}$, and let $\alpha_1 = \mathbb{P}(X_\infty \in S_1)$. By the same argument that led to (2.3),

$$\mathrm{Cov}(f_n(\omega_1), g_n(\omega_1^\varepsilon)) = \mathrm{Cov}\left(\mathbb{P}(f_n(\omega_1) = 1 \mid \omega_0), \mathbb{P}(g_n(\omega_1) = 1 \mid \omega_0)\right),$$

which vanishes as $n \to \infty$ if $Y_n := \mathbb{P}(f_n(\omega_1) = 1 \mid \omega_0)$ has $\mathrm{Var}(Y_n) \to 0$. Since

$$\left|\mathrm{Var}(Y_n) - \mathbb{E}\left[|Y_n - \alpha_1|^2 \mathbf{1}_{\{\omega_0 \in W_n\}}\right]\right| \leq 3|\mathbb{E}Y_n - \alpha_1| + \mathbb{P}_n(\omega_0 \notin W_n)$$

for the random variables $0 \leq Y_n \leq 1$, and each term on the right vanishes as $n \to \infty$ by the hypotheses $X_n \xrightarrow{\mathrm{d}} X_\infty$ and $\mathbb{P}_n(W_n) \to 1$, resp., the proof is concluded. ∎

Remark 2.4. *When applying Corollary 2.3, we will establish the stronger inequality*

$$\lim_{n \to \infty} \max_{\omega \in W_n} |\mathbb{P}_n(X_n(\omega_1) \in S \mid \omega_0 = \omega) - \mathbb{P}(X_\infty \in S)| = 0 \quad \forall \mu_\infty\text{-cont. set } S \subseteq \mathbb{S}.$$

Compare this with the definition of 1-strong noise sensitivity of $f_n : \Omega_n \to \{0, 1\}$: a 1-witness to f_n is a minimal subset W such that $\omega|_W = 1$ implies $f_n(\omega) = 1$ deterministically; if $\mathcal{W}_n$ is the set of all witnesses, strong noise sensitivity says that

$$\lim_{n \to \infty} \max_{W \in \mathcal{W}_n} |\mathbb{P}_n(f_n(\omega^\varepsilon) = 1 \mid \omega|_W = 1) - \mathbb{P}_n(f_n = 1)| = 0.$$

This stronger notion implies, and is not equivalent to, noise sensitivity (see [12]). Observe that strong noise sensitivity addresses the effect of conditioning that a specific subset W of the variables is open in a critical configuration ω. On the other hand, here we examine the bias due to the entire configuration of the subcritical ω_0.

3. Component structure of the subcritical random graph

Let $p = (1 + \lambda n^{-1/3})/n$ for a fixed $\lambda \in \mathbb{R}$, and $\varepsilon = \varepsilon_n \in (0, 1)$ such that $\varepsilon^3 n \to \infty$. Letting $G \sim \mathcal{G}(n, p)$ be the analog of ω_1 in in §2 for the $N = \binom{n}{2}$ edge variables, define G_0 analogously to ω_0; that is, $G_0 \sim \mathcal{G}(n, p_0)$ where p_0, as per (2.1), is set to

$$p_0 := \frac{(1 - \varepsilon)(1 + \lambda n^{-1/3})}{n - \varepsilon(1 + \lambda n^{-1/3})} = \frac{(1 - \varepsilon)(1 + \lambda n^{-1/3})}{n} + O(n^{-2}). \qquad (3.1)$$

It will be convenient to further denote

$$\theta := 1 - n p_0 = (1 - o(1))\varepsilon, \qquad \text{whence } \theta^3 n \to \infty. \qquad (3.2)$$

Our proofs will rely on several concentration estimates for features of the subcritical random graph G_0, which are known to determine the scaling limit of the component sizes, and moreover, that of the full metric spaces of connected components of G_1 conditioned on G_0. In particular, we will need the following results.

Theorem 3.1 ([10, Thms. 3.3,3.4,4.1]). *For G_0 be as above and $r = 2$ or $r = 3$,*

$$\mathbb{E}\Big[\sum_i |\mathcal{C}_i(G_0)|^r\Big] = \big(1 + O\big(\tfrac{1-\theta}{\theta^3 n}\big)\big)\,\theta^{3-2r} n\,, \tag{3.3}$$

$$\mathrm{Var}\Big(\sum_i |\mathcal{C}_i(G_0)|^r\Big) = O\left(\theta^{3-4r} n\right)\,. \tag{3.4}$$

Theorem 3.2 ([7,13]; [8, Cor. 5.11],[11, Thm. 5.6]). *For G_0 as above,*

$$|\mathcal{C}_1(G_0)| = \frac{\log(\theta^3 n) - 5\log\log(\theta^3 n) + O_{\mathrm{P}}(1)}{-\theta - \log(1-\theta)}\,.$$

Theorem 3.3 ([14, Thm. 11]). *For G_0 as above,*

$$\max_i \mathrm{diam}(\mathcal{C}_i(G_0)) \leq \frac{\log(\theta^3 n) + O_{\mathrm{P}}(1)}{-\log(1-\theta)}\,.$$

We further need estimates on the average distance between all pairs of vertices that share a connected component, established by the next result.

Lemma 3.4. *For G_0 as above, the random variable*

$$Z := \sum_i Z_i \qquad where \qquad Z_i := \sum_{u,v \in \mathcal{C}_i(G_0)} \mathrm{dist}_{G_0}(u,v)$$

satisfies

$$\mathbb{E}[Z] = \big(1 + O\big(\tfrac{1}{\theta^3 n}\big)\big)\frac{1-\theta}{\theta^2} n\,, \tag{3.5}$$

$$\mathrm{Var}(Z) = O\left(\theta^{-7} n\right)\,. \tag{3.6}$$

Proof. We may rewrite $Z = \sum_u \sum_v^* \mathrm{dist}_{G_0}(u,v)$, where the sum $\sum^*$ goes over every vertex v in the same connected component of u. With this in mind, let the random variable $X_k = X_k(u)$ be the number of vertices at distance k in G_0 from a given vertex u. Then by the symmetry of the vertices in G_0 playing the role of u,

$$\mathbb{E}[Z] = n \sum_{k=0}^{\infty} k\,\mathbb{E}[X_k].$$

The upper bound on $\mathbb{E}[Z]$ will be immediate: we claim $\mathbb{E}[X_k] \leq (1-\theta)^k$, and so

$$\mathbb{E}[Z] \leq n \sum_k k(1-\theta)^k = \frac{(1-\theta)n}{\theta^2}\,. \tag{3.7}$$

Indeed, for each v, we have that $\mathbb{P}(\mathrm{dist}_{G_0}(u,v) = k) \leq n^{k-1} p_0^k$ as there are less than n^{k-1} paths of length k from u to v; summing over v gives $\mathbb{E}[X_k] \leq (1-\theta)^k$.

For the lower bound on $\mathbb{E}[Z]$, we will show that

$$\mathbb{E}[X_k] \geq (1-\theta)^k - 3k^2\frac{(1-\theta)^{k+1}}{\theta n}\,, \tag{3.8}$$

which, together with (3.7), will yield (3.5) since $\sum k^3(1-\theta)^k = (\theta^2 + 6(1-\theta))\frac{1-\theta}{\theta^4}$.

To this end, we will need to first estimate $\mathbb{E}[X_l X_k]$ for $l \leq k$; if (v_1, v_2) are such that $\mathrm{dist}_{G_0}(u,v_1) = l$ and $\mathrm{dist}_{G_0}(u,v_2) = k$, then there must be a pair of paths P_1 from u to v_1 and P_2 from u to v_2, of lengths l and k respectively, which intersect only in their first j edges for some $0 \leq j \leq l$. (This is due to the fact that any sub-path of a shortest path is itself a shortest path, thus from P_1, P_2 without this

property we may generate P_1', P_2' as above by replacing the j-prefix of P_1 with that of P_2, where j is the maximal index such that the j-th edge of P_1 is also that of P_2). For each value of $0 \leq j \leq l$ there are at most n^{k+l-j} such pairs of paths P_1, P_2, and the probability that two given such paths both appear in G_0 is at most p_0^{k+l-j}. Hence,

$$\mathbb{E}[X_l X_k] \leq \sum_{j=0}^{l} (1-\theta)^{k+j} \leq \frac{(1-\theta)^k}{\theta} . \tag{3.9}$$

Next, let $(\mathcal{F}_k)$ be the filtration corresponding to revealing the first k levels of the breadth-first-search tree of u in G_0, wherein the variable X_k corresponds to the number of vertices in level k of the tree. Conditional on $\mathcal{F}_{k-1}$, the random variable X_k is distributed as $\mathrm{Bin}(n - \sum_{l=0}^{k-1} X_l, \eta)$ for $\eta = 1 - (1-p_0)^{X_{k-1}}$. We use

$$1 - (1-p_0)^{X_{k-1}} \geq X_{k-1} p_0 - \binom{X_{k-1}}{2} p_0^2$$

to find that

$$\mathbb{E}[X_k] \geq np_0 \mathbb{E}[X_{k-1}] - p_0 \sum_{l=0}^{k-1} \mathbb{E}\left[X_l X_{k-1}\right] - \frac{1}{2} np_0^2 \mathbb{E}[X_{k-1}^2]$$

$$\geq (1-\theta)\mathbb{E}[X_{k-1}] - \frac{(1-\theta)^{k+1}}{\theta}(k/n + p_0),$$

which, by induction on k starting from $X_0 = 1$, satisfies (3.8), as claimed.

We turn to the bound on $\mathrm{Var}(Z)$. Let us denote by $\mathcal{C}_A$, for every subset $A \subseteq [n]$, the event that A is the vertex set of a connected component in G_0, and let

$$Z_A := \mathbf{1}_{\mathcal{C}_A} \sum_{u,v \in A} \mathrm{dist}(u,v) .$$

Clearly, $\sum_{i \neq j} Z_i Z_j = \sum_{A \neq B} Z_A Z_B$, and $Z = \sum_A Z_A$. In addition, $Z_A Z_B = 0$ for every A, B such that $A \cap B \neq \emptyset$; hence,

$$\sum_{A \neq B} Z_A Z_B = \sum_A Z_A \sum_{B \subseteq [n] \setminus A} Z_B.$$

The the random variables Z_A and $\sum_{B \subseteq [n] \setminus A} Z_B$ are conditionally independent on the event $\mathcal{C}_A$, and the latter is the analog of Z for a random graph $F \sim \mathcal{G}(n - |A|, p_0)$ whence its expectation is bounded from above by the right-hand side of (3.7). Therefore,

$$\mathbb{E}\Big[\sum_{A \neq B} Z_A Z_B \Big] = \sum_A \mathbb{E}[Z_A \mid \mathcal{C}_A] \mathbb{P}(\mathcal{C}_A) \mathbb{E}\Big[\sum_{B \subseteq [n] \setminus A} Z_B \mid \mathcal{C}_A \Big]$$

$$\leq \mathbb{E}[Z] \frac{(1-\theta)n}{\theta^2} .$$

(Although $\mathbb{E}[\sum_B Z_B \mathbf{1}_{\mathcal{C}_A}] \leq \mathbb{E}[Z]$, we could not have increased $\mathbb{E}[\sum_{B \subseteq A^c} Z_B \mid \mathcal{C}_A]$ to $\mathbb{E}[Z]$ due to the conditioning, and instead applied the upper bound in (3.7) on F.) We use this inequality and (3.5) (twice) to find that

$$\mathbb{E}\Big[\sum_{i \neq j} Z_i Z_j \Big] - \mathbb{E}[Z]^2 \leq \mathbb{E}[Z] \frac{(1-\theta)n}{\theta^2} O\Big(\frac{1}{\theta^3 n}\Big) \leq O(\theta^{-7} n) .$$

The proof of (3.6) will thus be concluded once we show that

$$\mathbb{E}\Big[\sum_i Z_i^2\Big] = O\left(\theta^{-7}n\right) . \tag{3.10}$$

By Cauchy–Schwartz,

$$Z_i^2 \leq |\mathcal{C}_i(G_0)|^2 \sum_{u,v \in \mathcal{C}_i(G_0)} \mathrm{dist}(u,v)^2 .$$

Hence, again by symmetry of the vertices in G_0 playing the role of u,

$$\mathbb{E}\Big[\sum_i Z_i^2\Big] = n\mathbb{E}\Big[\Big(\sum_{k \geq 0} X_k\Big)^2 \Big(\sum_{k \geq 0} k^2 X_k\Big)\Big] = n \sum_{k,l,m \geq 0} k^2 \mathbb{E}\left[X_k X_l X_m\right] . \tag{3.11}$$

We will treat $\mathbb{E}[X_r X_s X_t]$ for $r \leq s \leq t$ similarly to (3.9), by enumerating over the triples of vertices v_1, v_2, v_3 at respective distances r, s, t from a fixed vertex u. Consider shortest paths P_i from u to v_i for $i = 1, 2, 3,$. If $0 \leq j \leq s$ is the number of common edges between P_2 and P_3, and $0 \leq i \leq r$ is the number of common edges between P_1 and P_2, then given P_3 one can construct P_2 followed by P_1 such that the j-prefix of P_2 coincides with that of P_3, and then the i-prefix of P_1 coincides with that of P_2. Thus, there are at most $n^{t+(s-j)+(r-i)}$ choices for such P_1, P_2, P_3, and by the same token, at most $n^{t+(s-j)+(r-i)}$ choices for P_1, P_2, P_3 where the i-prefix of P_1 coincides with that of P_3. Altogether,

$$\mathbb{E}[X_r X_s X_t] \leq 2 \sum_{i=0}^{r} \sum_{j=0}^{s} (np_0)^{t+(s-j)+(r-i)} \leq \frac{2(1-\theta)^t}{\theta^2} .$$

Splitting $\sum_k k^2 \sum_{l,m} \mathbb{E}[X_k X_l X_m]$ in of (3.11) into $l, m \leq k$ (where we use the last display) and $k < l \vee m$ (where we enumerate on $l \vee m$ and apply the same bound),

$$\mathbb{E}\Big[\sum_i Z_i^2\Big] \leq 2n \sum_{k=0}^{\infty} k^2 \Big(k^2 \frac{(1-\theta)^k}{\theta^2} + 2 \sum_{l=k+1}^{\infty} \frac{l(1-\theta)^l}{\theta^2}\Big),$$

which is $O(n \sum k^4 \frac{(1-\theta)^k}{\theta^2}) = O(\frac{n}{\theta^7})$, giving (3.10) and thus concluding the proof. ∎

4. Proof of Theorem 1

4.1. **Noise sensitivity: proof of Theorem 1, Part (i).** Let $G_0, G_1, G_1', G_1^\varepsilon$ be the analogs of $\omega_0, \omega_1, \omega_1', \omega^\varepsilon$ as defined above for the $N = \binom{n}{2}$ edge variables that correspond to $\mathcal{G}(n,p)$, using p_0 and p_1 as in (2.1)–(2.2). Conditioned on G_0, the connected components of G_1 are constructed by a multiplicative coalescent process where two components $\mathcal{C}_i(G_0), \mathcal{C}_j(G_0)$ will are in G_1 with probability

$$1 - (1 - p_1)^{|\mathcal{C}_i(G_0)||\mathcal{C}_j(G_0)|} .$$

Aldous [3] considered the following related random graph. For a vector of positive weights $\mathbf{x} = (x_1, \ldots, x_N)$ and $q > 0$, let $\mathfrak{G} \sim \mathscr{W}(\mathbf{x}, q)$ be the random graph on the vertices $\{1, \ldots, N\}$ where each edge ij is present with probability $1 - \exp(-q x_i x_j)$ independently of other edges. The weight of a connected component $\mathcal{C}$ of $\mathfrak{G}$ is defined as $\mathrm{w}(\mathcal{C}) = \sum_{i \in \mathcal{C}} x_i$, and the components of $\mathfrak{G}$ are ordered $(\mathcal{C}_1(\mathfrak{G}), \mathcal{C}_2(\mathfrak{G}), \ldots)$ in decreasing order of weights. With these notations, Aldous showed the following.

Proposition 4.1 ([3, Prop. 4]). *Fix $\lambda \in \mathbb{R}$. For each N, let $\mathbf{x}(N) = (x_1, \dots, x_N)$ be a positive vector and $q(N) > 0$ be so that, if $\sigma_r := \|\mathbf{x}\|_r^r$ and $x_{\max} := \|\mathbf{x}\|_\infty$ then*

$$\frac{\sigma_3}{(\sigma_2)^3} \to 1, \qquad q - \frac{1}{\sigma_2} \to \lambda, \qquad \frac{x_{\max}}{\sigma_2} \to 0 \qquad as \ N \to \infty. \tag{4.1}$$

Then the random graph $\mathfrak{G} \sim \mathcal{W}(\mathbf{x}, q)$ satisfies

$$(\mathrm{w}(\mathcal{C}_j(\mathfrak{G})) \ : \ j \geq 1) \xrightarrow{\mathrm{d}} \boldsymbol{\xi}(\lambda) \quad as \ N \to \infty.$$

Observe that if let $\mathfrak{G} \sim \mathcal{W}(\mathbf{x}, q)$ for $\mathbf{x} = (x_i)_{i \geq 1}$ given by

$$x_i = n^{-2/3}|\mathcal{C}_i(G_0)| \qquad \text{and} \qquad q = n^{4/3}(-\log(1 - p_1)) \tag{4.2}$$

(i.e., the vertices of $\mathfrak{G}$ are in correspondence with the connected component of G_0), then by the definitions of $\mathcal{W}(\mathbf{x}, q)$ and construction of G_1, conditioned on G_0,

$$(\mathrm{w}(\mathcal{C}_j(\mathfrak{G})) \ : \ j \geq 1) \overset{\mathrm{d}}{=} n^{-2/3}(|\mathcal{C}_j(G_1)| \ : \ j \geq 1). \tag{4.3}$$

With this in mind, we move to verify that the three conditions in (4.1) that qualify an application of Proposition 4.1 hold for G_0 w.h.p.

Lemma 4.2. *Let $G_0 \sim \mathcal{G}(n, p_0)$ with p_0 as in (3.1) for $\lambda \in \mathbb{R}$ fixed and $0 < \varepsilon(n) < 1$ such that $\varepsilon^3 n$ goes to ∞ with n. Let $\mathbf{x}, q$ be as in (4.2). Then w.h.p.,*

$$\left| \frac{\sigma_3}{(\sigma_2)^3} - 1 \right| < (\varepsilon^3 n)^{-1/5}, \tag{4.4}$$

$$\left| q - \frac{1}{\sigma_2} - \lambda \right| < (\varepsilon^3 n)^{-1/15}, \tag{4.5}$$

$$\frac{x_{\max}}{\sigma_2} < 4(\varepsilon^3 n)^{-1/3} \log(\varepsilon^3 n). \tag{4.6}$$

Proof. Let $\theta = 1 - np_0$ be as in (3.2), and write $\xi := \theta^3 n$ for brevity (recalling that $\theta = (1 - o(1))\varepsilon$ whence $\theta^3 n \to \infty$). We apply Theorem 3.1 to find that the variables $\sigma_2 = n^{-4/3} \sum_i |\mathcal{C}_i(G_0)|^2$ and $\sigma_3 = n^{-2} \sum_i |\mathcal{C}_i(G_0)|^3$ satisfy

$$\mathbb{E}[\sigma_2] = \xi^{-1/3} + O(\xi^{-4/3}), \qquad \mathbb{E}[\sigma_3] = \xi^{-1} + O(\xi^{-2})$$

by (3.3), whereas by (3.4),

$$\mathrm{Var}(\sigma_2) = O(\xi^{-5/3}), \qquad \mathrm{Var}(\sigma_3) = O(\xi^{-3}).$$

By Chebyshev's inequality, we see that the events

$$\mathcal{E}_2 = \{|\sigma_2 - \xi^{-1/3}| \leq \xi^{-3/4}\}, \qquad \mathcal{E}_3 = \{|\sigma_3 - \xi^{-1}| \leq \xi^{-5/4}\}$$

occur w.h.p., and specifically

$$\mathbb{P}(\mathcal{E}_2^c) \leq O(\xi^{-1/6}) = o(1), \qquad \mathbb{P}(\mathcal{E}_3^c) \leq O(\xi^{-1/2}) = o(1). \tag{4.7}$$

For $G_0 \in \mathcal{E}_2 \cap \mathcal{E}_3$ we immediately see that

$$\frac{\sigma_3}{(\sigma_2)^3} = \frac{(1 + O(\xi^{-1/4}))\xi^{-1}}{(1 + O(\xi^{-5/12}))\xi^{-1}} = 1 + O((\varepsilon^3 n)^{-1/4}),$$

and in particular (4.4) holds for large enough n. For the second (somewhat subtler) inequality (4.5), recall $\theta = \varepsilon - (1 - \varepsilon)\lambda n^{-1/3} + O(1/n)$ and $p_1 = \varepsilon p = \varepsilon(1 + \lambda n^{-1/3})/n$. Thus, for every $G_0 \in \mathcal{E}_2$,

$$q - \sigma_2^{-1} = (1 + O(\varepsilon/n))n^{1/3}\varepsilon(1 + \lambda n^{-1/3}) - (1 + O(\xi^{-5/12}))\theta n^{1/3}$$

$$= \lambda + O(\xi^{-5/12}\theta n^{1/3}) + O(n^{-2/3}) = \lambda + O((\varepsilon^3 n)^{-1/12}),$$

yielding (4.5). Finally, from Theorem 3.2 we have $n^{2/3}x_{\max} \leq 2(\log \xi + O_{\mathrm{P}}(1))/\theta^2$, and so G_0 satisfies both $\mathcal{E}_2$ and $x_{\max} \leq 3\xi^{-2/3} \log \xi$ w.h.p., in which case

$$\frac{x_{\max}}{\sigma_2} \leq (3 + o(1))\xi^{-1/3} \log \xi < 4(\varepsilon^3 n)^{-1/3} \log(\varepsilon^3 n)$$

for large n, as claimed in (4.6). This completes the proof.   ∎

Combining Proposition 4.1 with Lemma 4.2 we find that for each $n \geq 1$ there exists a set of graphs $\mathcal{A}_0^{(n)}$ such that $\mathbb{P}(G_0 \in \mathcal{A}_0^{(n)}) \to 1$, and for every $G \in \mathcal{A}_0^{(n)}$, the conditional law of $\mathbf{X}(G_1)$ given that $G_0 = G$ converges in distribution to $\boldsymbol{\xi}(\lambda)$. Namely, for every continuity set S of $\mathbb{P}(\boldsymbol{\xi}(\lambda) \in \cdot)$,

$$\lim_{n \to \infty} \max_{G \in \mathcal{A}_0^{(n)}} |\mathbb{P}_n (\mathbf{X}_n(G_1) \in S \mid G_0 = G) - \mathbb{P}(\boldsymbol{\xi}(\lambda) \in S)| = 0 \,.$$

In particular (cf. Remark 2.4), the hypothesis (2.4) in Corollary 2.3 holds true, and hence we may conclude from it that the pair $(\mathbf{X}(G_1), \mathbf{X}(G_1^\varepsilon))$ converges in distribution to a product of two i.i.d. copies of $\boldsymbol{\xi}(\lambda)$.   ∎

4.2. **Noise stability: proof of Theorem 1, Part (ii).** Recall the setup of $G_0, G_1, G_1', G_1^\varepsilon$ from the previous sections. The hypothesis $\varepsilon^3 n \to 0$ means that G_0 is now also a critical random graph satisfying $\nu_n^0 = n^{-2/3}(|\mathcal{C}_j(G_0) : j \geq 1) \xrightarrow{\mathrm{d}} \boldsymbol{\xi}(\lambda)$. We need the following two results of Aldous [3] for the random graph model $\mathscr{W}(\mathbf{x}, q)$. Recall $\mathcal{C}_j(\mathfrak{G})$ is the component of a weighted graph $\mathfrak{G}$ with the j-th largest weight.

Lemma 4.3 ([3, Lem. 17]). *If $\mathfrak{G}, \bar{\mathfrak{G}}$ are weighted graphs on N vertices with resp. weights $\mathbf{x} = (x_i)_{i=1}^N, \bar{\mathbf{x}} = (\bar{x}_i)_{i=1}^N$ such that $E(\mathfrak{G}) \subseteq E(\bar{\mathfrak{G}})$ and $x_i \leq \bar{x}_i$ for all i, then*

$$\sum_j \left(\mathrm{w}(\mathcal{C}_j(\bar{\mathfrak{G}})) - \mathrm{w}(\mathcal{C}_j(\mathfrak{G}))\right)^2 \leq \sum_j \mathrm{w}(\mathcal{C}_j(\bar{\mathfrak{G}}))^2 - \sum_j \mathrm{w}(\mathcal{C}_j(\mathfrak{G}))^2 \,.$$

Lemma 4.4 ([3, Lem. 20]). *The weighted random graph $\mathfrak{G} \sim \mathscr{W}(\mathbf{x}, q)$ satisfies*

$$\mathbb{P}\left(\sum_j \mathrm{w}(\mathcal{C}_j(\mathfrak{G}))^2 > s \right) \leq \frac{qs\sigma_2}{s - \sigma_2}$$

for every $s > \sigma_2 = \sum_j x_j^2$.

Corollary 4.5. *If $\varepsilon^3 n \to 0$ and G_0, G_1 are as above, then $\|\mathbf{X}(G_1) - \mathbf{X}(G_0)\|_2 \xrightarrow{\mathrm{P}} 0$.*

Proof. Let $\delta > 0$. We let $\bar{\mathfrak{G}} \sim \mathscr{W}(\mathbf{x}, q)$ for $\mathbf{x}$ and q as in (4.2), noting that

$$q = (1 + o(1))\varepsilon n^{1/3} \,,$$

while recalling that $\mathbf{X}(G_1) \stackrel{\mathrm{d}}{=} n^{-2/3}(\mathrm{w}(\mathcal{C}_j(\bar{\mathfrak{G}})))_{j \geq 1}$ conditioned on G_0 as per (4.3). Further let $\mathfrak{G}$ be the edgeless graph $\mathfrak{G}$ on a vertex set corresponding to the connected components of G_0 and the same weights $\mathbf{x}$ (whereby $\sum x_i^2 = \sum_j \mathrm{w}(\mathcal{C}_j(\mathfrak{G}))^2$). Applying Lemma 4.3 for $\mathfrak{G}, \bar{\mathfrak{G}}$ we deduce that, conditioned on G_0,

$$\mathbb{P}\left(\|\mathbf{X}(G_1) - \mathbf{X}(G_0)\|_2^2 > \delta \mid G_0\right) \leq \mathbb{P}\left(\sum_j \mathrm{w}(\mathcal{C}_j(\bar{\mathfrak{G}}))^2 > \sigma_2 + \delta \right),$$

at which point Lemma 4.4 further shows that

$$\mathbb{P}\left(\|\mathbf{X}(G_1) - \mathbf{X}(G_0)\|_2^2 > \delta \mid G_0\right) \leq \frac{q(\sigma_2 + \delta)\sigma_2}{\delta} \,.$$

Recall that $n^{-2/3}(|\mathcal{C}_j(G_0)| \ : \ j \geq 1)$ converges weakly at $n \to \infty$ to a nontrivial limit in the ℓ^2-topology, whence $\sigma_2 = n^{-4/3} \sum |\mathcal{C}_j(G_0)|^2$ converges in distribution precisely to the ℓ^2-norm, squared, of the random limit $\boldsymbol{\xi}(\lambda)$ in the Polish space $\ell_{\downarrow}^2$. As such, $\|\boldsymbol{\xi}(\lambda)\|_2$ is tight, hence so is σ_2^2, while $q \to 0$ by assumption, and this completes the proof. ∎

Part (ii) of Theorem 1 follows directly from Corollary 4.5, applied to (G_0, G_1) and again to (G_0, G_1') (which is equal in law), and using that $(G_0, G_1^\varepsilon) \stackrel{\mathrm{d}}{=} (G_0, G_1')$. ∎

5. Proof of Theorem 2

5.1. **Noise sensitivity: proof of Theorem 2, Part (i).** Our approach here is similar to the proof of Theorem 1, building on a special case of the metric-space extension of Proposition 4.1 that was proved in [5]. In addition to $\mathbf{x} = (x_1, \ldots, x_N)$ and $q > 0$ as before, let $\mathbf{H} = (H_1, \ldots, H_N)$ be finite connected graphs. Define $\mathfrak{G} \sim \mathscr{W}(\mathbf{x}, q)$ as before, whereby the weight of a connected component $\mathcal{C}$ in it is $\mathrm{w}(\mathcal{C}) = \sum_{i \in \mathcal{C}} x_i$. Let $\mathfrak{G}^{\mathbf{H}} \sim \mathscr{W}(\mathbf{x}, \mathbf{H}, q)$ be the (unweighted) random graph obtained by generating a weighted random graph $\mathfrak{G} \sim \mathscr{W}(\mathbf{x}, q)$, replacing each vertex i of $\mathfrak{G}$ with the graph H_i, and connecting a uniform random vertex of H_i with a uniform random vertex of H_j for every edge $ij \in E(\mathfrak{G})$. We let $\mathcal{C}_j(\mathfrak{G}^{\mathbf{H}})$ be the component of $\mathfrak{G}^{\mathbf{H}}$ corresponding to $\mathcal{C}_j(\mathfrak{G})$, the component of $\mathfrak{G}$ with the j-th largest weight, and view it as a measured metric space $\tilde{\mathsf{M}}_j(\mathfrak{G}^{\mathbf{H}})$ via the scaled counting measure

$$\mu_j(A) = \sum_{i \in \mathcal{C}_j(\mathfrak{G})} x_i \cdot \frac{|A \cap V(H_i)|}{|V(H_i)|} \qquad \text{for all } A \subseteq \mathcal{C}_j(\mathfrak{G}^{\mathbf{H}}),$$

(N.B. $\mathrm{w}(\mathcal{C}_j(\mathfrak{G}))$ coincides with its total measure), and the graph metric scaled by

$$\mathfrak{s}(\mathbf{x}, \mathbf{H}) := \frac{(\sigma_2)^2}{\sigma_2 + \sum_{i=1}^N x_i^2 u_i},$$

where $\sigma_2 = \|\mathbf{x}\|_2^2$ and u_i is the expected distance between two independent uniform random vertices in H_i. (We distinguish the notation $\tilde{\mathsf{M}}_j(\mathfrak{G}^{\mathbf{H}})$ from $\mathsf{M}_j(G)$ since the scaling factor here is a function of $\mathbf{x}, \mathbf{H}$ as opposed to scaling by $\mathfrak{s} = |V(G)|^{-1/3}$).

With this notation, following is the metric space analog due to [5] of Prop. 4.1.

Proposition 5.1 ([5, Thm. 3.4]). *Fix $\lambda \in \mathbb{R}$. For each N, let $\mathbf{x}(N) = (x_1, \ldots, x_N)$ be a positive vector and $q(N) > 0$ so that, if $\sigma_r := \|\mathbf{x}\|_r^r$ and $x_{\max} := \|\mathbf{x}\|_\infty$ then assumption (4.1) holds true. In addition, let $\mathbf{H}(N) = (H_1, \ldots, H_N)$ be a sequence of disjoint finite connected graphs so that, if $x_{\min} := \min_i x_i$, $d_{\max} := \max_i \mathrm{diam}(H_i)$, and u_i is the expected distance of two independent uniform random vertices in H_i then there exist some positive real numbers $r_0 > 0$ and $0 < \eta_0 < \frac{1}{2}$ such that*

$$\frac{x_{\max}}{(\sigma_2)^{3/2+\eta_0}} \to 0 \, , \quad \frac{(\sigma_2)^{r_0}}{x_{\min}} \to 0 \, , \quad \frac{(\sigma_2)^{3/2-\eta_0} d_{\max}}{\sum_{i=1}^N x_i^2 u_i + \sigma_2} \to 0 \, , \quad \frac{\sigma_2 x_{\max} d_{\max}}{\sum_{i=1}^N x_i^2 u_i} \to 0$$
(5.1)

as $N \to \infty$. Then the random graph $\mathfrak{G}^{\mathbf{H}} \sim \mathscr{W}(\mathbf{x}, \mathbf{H}, q)$ satisfies, for every $j \geq 1$,

$$\tilde{\mathsf{M}}_j(\mathfrak{G}^{\mathbf{H}}) \stackrel{\mathrm{d}}{\longrightarrow} \mathfrak{M}_j(\lambda) \quad \text{as } N \to \infty \text{ in } (\mathscr{M}, d_{\mathrm{GHP}}).$$

The proof of Theorem 2, (i) consists of several ingredients. We start by setting

$$x_i = n^{-2/3}|\mathcal{C}_i(G_0)|\,, \quad q = n^{4/3}(-\log(1-p_1))\,, \quad H_i = \mathcal{C}_i(G_0)\,. \qquad (5.2)$$

Noting that the x_i's and q are exactly as they were defined in (4.2), we immediately infer that the conditions in (4.1) hold for $G_0 \sim \mathcal{G}(n, p_0)$ w.h.p. using Lemma 4.2. The following result confirms that the additional conditions in (5.1), required for an application of Proposition 5.1, also hold for the graph G_0 w.h.p.

Lemma 5.2. *Fix $\lambda \in \mathbb{R}$ and $\delta > 0$. Let $G_0 \sim \mathcal{G}(n, p_0)$ with p_0 as in (3.1) for $\varepsilon(n)$ which satisfies $n^{-1/3+\delta} \leq \varepsilon \leq 1 - n^{-1/2}$. Further define (x_i) and (H_i) as in (5.2). Then for every fixed $0 < \eta_0 < \frac{1}{2}$ and $r_0 > 1/\delta$, the following inequalities hold w.h.p.*

$$\frac{x_{\max}}{(\sigma_2)^{3/2+\eta_0}} \leq 5(\varepsilon^3 n)^{-(1-2\eta_0)/6}\log(\varepsilon^3 n)\,, \qquad (5.3)$$

$$\frac{(\sigma_2)^{r_0}}{x_{\min}} \leq 2n^{-1/3}\,, \qquad (5.4)$$

$$\frac{(\sigma_2)^{3/2-\eta_0}d_{\max}}{\sigma_2 + \sum_{i=1}^m x_i^2 u_i} \leq 5(\varepsilon^3 n)^{-(1-2\eta_0)/6}\log(\varepsilon^3 n)\,, \qquad (5.5)$$

$$\frac{\sigma_2 x_{\max} d_{\max}}{\sum_{i=1}^m x_i^2 u_i} \leq 10(\varepsilon^3 n)^{-2/3}\log^2(\varepsilon^3 n) \,\wedge\, n^{-1/8}\,, \qquad (5.6)$$

$$\left|\mathfrak{s}(\mathbf{x}, \mathbf{H})n^{1/3} - 1\right| \leq 4(\varepsilon^3 n)^{-2/5}\,. \qquad (5.7)$$

Proof. Observe that the random variable $\sum_{i=1}^N x_i^2 u_i$ is equal to $n^{-4/3}Z$ for Z as defined in Lemma 3.4. We recall that $\theta = 1 - np_0 = \varepsilon + O(n^{-1/3}) = (1 + o(1))\varepsilon$ as per (3.2), and can therefore infer from that lemma that

$$\mathbb{E}\Big[\sum_i x_i^2 u_i\Big] = \Big(1 + O\big((\varepsilon^3 n)^{-1/3}\big)\Big)\frac{1-\theta}{\theta^2 n^{1/3}}\,,$$

$$\mathrm{Var}\Big(\sum_i x_i^2 u_i\Big) = O\Big(\varepsilon^{-7}n^{-5/3}\Big)\,.$$

By Chebyshev's inequality,

$$\mathbb{P}\bigg(\Big|\sum_i x_i^2 u_i - \frac{1-\theta}{\theta^2 n^{1/3}}\Big| > \frac{1}{(\varepsilon^3 n)^{2/5}\varepsilon^2 n^{1/3}}\bigg) = O\Big((\varepsilon^3 n)^{-1/5}\Big) = o(1)\,. \qquad (5.8)$$

We derive the proof from this bound, along with those of Lemma 4.2 and the fact

$$d_{\max} \leq 2\varepsilon^{-1}\log(\varepsilon^3 n)\,,$$

which holds w.h.p. by Theorem 3.3, as $d_{\max} \leq (\log(\varepsilon^3 n) + O_{\mathrm{P}}(1))/\theta$.

For (5.3), recall from (4.7) that, w.h.p., the event $\mathcal{E}_2$ holds, on which one has that $\sigma_2 = (1 + o(1))(\theta^3 n)^{-1/3} = (1 + o(1))(\varepsilon^3 n)^{-1/3}$. Plugging

$$(\sigma_2)^{\frac{1}{2}+\eta_0} = (1 + o(1))(\varepsilon^3 n)^{-\frac{1}{6}-\frac{\eta_0}{3}}$$

in (4.6), we immediately see that (5.3) holds w.h.p.

For (5.4), on the aforementioned event $\mathcal{E}_2$ we have $(\sigma_2)^{r_0} \leq 2(\varepsilon^3 n)^{-r_0/3} \leq 2n^{-\delta r_0}$ for large enough n, by the assumption $\varepsilon \geq n^{-1/3+\delta}$ (this is the only place in the proof where we use this stronger assumption rather than requiring that $\varepsilon^3 n \to \infty$); since $r_0 \geq 1/\delta$, whereas $x_{\min} \geq n^{-2/3}$ holds deterministically, we arrive at (5.4).

For (5.5), the above bounds on σ_2 and $d_{\max}$ show that, w.h.p.,

$$(\sigma_2)^{3/2-\eta_0} d_{\max} \le (2 + o(1))\varepsilon^{-1}(\varepsilon^3 n)^{-1/2+\eta_0/3} \log(\varepsilon^3 n),$$

while $\sum_i x_i^2 u_i = (1 + o(1))(1 - \varepsilon)\varepsilon^{-2} n^{-1/3}$ and $\sigma_2 = (1 + o(1))\varepsilon^{-1} n^{-1/3}$. If $\varepsilon < \frac{1}{2}$ then $\sum_i x_i^2 u_i \ge (\frac{1}{2} - o(1))\varepsilon^{-2} n^{-1/3}$, while if $\varepsilon \ge \frac{1}{2}$ then $\sigma_2 \ge (\frac{1}{2} - o(1))\varepsilon^{-2} n^{-1/3}$, hence in both cases we can combine this bound with the last display to obtain (5.5).

The bound in (5.6) follows from plugging in the above bounds on σ_2 and $d_{\max}$ in (4.6), along with the asymptotics $\sum_i x_i^2 u_i$ established above, to see that w.h.p.

$$\frac{\sigma_2 x_{\max} d_{\max}}{\sum_i x_i^2 u_i} \le (8 + o(1))\frac{\varepsilon^{-1}(\varepsilon^3 n)^{-1} \log^2(\varepsilon^3 n)}{(1 - \varepsilon)\varepsilon^{-2} n^{-1/3}} = \frac{8 + o(1)}{1 - \varepsilon}(\varepsilon^3 n)^{-2/3} \log^2(\varepsilon^3 n).$$

This gives the first bound in the right-hand of (5.6) if we have $\varepsilon < \frac{1}{10}$, while in the case $\frac{1}{10} \le \varepsilon \le 1 - n^{-1/2}$ we arrive at the bound $n^{-1/6+o(1)} < n^{-1/8}$ for large n.

To obtain the final inequality (5.7), we combine the bound (5.8) with the fact that $|\sigma_2 - (\theta n^{1/3})^{-1}| \le (\theta^3 n)^{-3/4}$ on the event $\mathcal{E}_2$, and deduce that w.h.p.,

$$\left| \sigma_2 + \sum_i x_i^2 u_i - \frac{1}{\theta^2 n^{1/3}} \right| \le \frac{1}{(\theta^3 n)^{3/4}} + \frac{(\varepsilon^3 n)^{-2/5}}{\varepsilon^2 n^{1/3}} \le \frac{2(\varepsilon^3 n)^{-2/5}}{\varepsilon^2 n^{1/3}},$$

where we absorbed the term $(1+o(1))\varepsilon(\varepsilon^3 n)^{-5/12}$ into $(\varepsilon^3 n)^{-2/5}$ for large enough n. The above bound on σ_2 given the event $\mathcal{E}_2$ further implies that

$$\left| \frac{(\sigma_2)^2}{(\theta^3 n)^{-2/3}} - 1 \right| \le (2 + o(1))(\theta^3 n)^{-5/12} < (\varepsilon^3 n)^{-2/5}$$

for a sufficiently large n. Combined, we obtain that $\mathfrak{s}(\mathbf{x}, \mathbf{H})$ is bounded w.h.p. via

$$\left| \frac{(\sigma_2)^2 n^{1/3}}{\sigma_2 + \sum_i x_i^2 u_i} - 1 \right| \le \frac{3(\varepsilon^3 n)^{-2/5}}{1 - 2(\varepsilon^3 n)^{-2/5}} \le 4(\varepsilon^3 n)^{-2/5},$$

concluding the proof. ∎

The second step in the proof is to argue that w.h.p. over the random graph G_0, the graphs G_1 conditioned on G_0 and $\mathfrak{G}^{\mathbf{H}} \sim \mathscr{W}(\mathbf{x}, \mathbf{H}, q)$ can be coupled such that $d_{\text{GHP}}(\mathsf{M}_j(G_1), \tilde{\mathsf{M}}_j(\mathfrak{G}^{\mathbf{H}})) \xrightarrow{\mathrm{P}} 0$. This is the (approximate) metric analogue of the distributional equality (4.3) that was used in the proof of Theorem 1. The reason that such a distributional equality does not hold here is twofold. First, G_1 can contain multiple edges between two components of G_0 as well as additional edges within a component of G_0. Second, the scaling factor $\mathfrak{s}(\mathbf{x}, \mathbf{H})$ used for $\tilde{\mathsf{M}}_j(\mathfrak{G}^{\mathbf{H}})$ differs from the $n^{-1/3}$ scaling of dist_{G_1} for $\mathsf{M}_j(G_1)$. On the other hand, note that for the choice of $\mathbf{x}, \mathbf{H}$ as in (5.2), the measure associated with both spaces is the same, namely, the counting measure multiplied by $n^{-2/3}$ (as the factor $|V(H_i)|$ in the definition of μ_j cancels out with the one in x_i).

Lemma 5.3. *Fix $\lambda \in \mathbb{R}$. Suppose that $p = (1 + \lambda n^{-1/3})/n$, let $0 < \varepsilon < 1$ be such that $\varepsilon^3 n \to \infty$ as $n \to \infty$, and define G_0, G_1 to be as above. Then one can construct a random graph $\mathfrak{G}^{\mathbf{H}}$ such that $G_0 \subseteq \mathfrak{G}^{\mathbf{H}} \subseteq G_1$ and so that the following hold.*

(i) $\mathfrak{G}^{\mathbf{H}} \sim \mathscr{W}(\mathbf{x}, \mathbf{H}, q)$ where $\mathbf{x}$, $\mathbf{H}$ and q are as in (5.2).

(ii) There exist a set of graphs $\mathcal{B}_0^{(n)}$ such that $\mathbb{P}(G_0 \in \mathcal{B}_0^{(n)}) \to 1$ as $n \to \infty$, and for every graph $G \in \mathcal{B}_0^{(n)}$ and $j \ge 1$,

$$\mathbb{E}\left[d_{\text{GHP}}\big(\mathsf{M}_j(G_1), \tilde{\mathsf{M}}_j(\mathfrak{G}^{\mathbf{H}})\big) \mid G_0 = G \right] \to 0 \qquad \text{as } n \to \infty.$$

Proof. Given G_0 and G_1, we call a pair of connected components $\mathcal{C}_i(G_0)$ and $\mathcal{C}_j(G_0)$ *excessive in* G_1 if they are connected in G_1 by more than one edge. Similarly, we call a single component $\mathcal{C}_i(G_0)$ excessive if G_1 contains an edge within $\mathcal{C}_i(G_0)$ (i.e., an edge between two of its vertices) that does not belong to G_0. To generate $\mathfrak{G}^{\mathbf{H}}$ from G_1, we remove from each excessive component $\mathcal{C}_i(G_0)$ every edge that does not belong to G_0, and further remove all the edges between every pair of excessive components $\mathcal{C}_i(G_0), \mathcal{C}_j(G_0)$ except one that is chosen uniformly at random.

The graph $\mathfrak{G}^{\mathbf{H}}$ is a $\mathscr{W}(\mathbf{x}, \mathbf{H}, q)$-distributed subgraph of G_1 with the exact same connected components. If we condition on G_0, the probability of the event $\mathfrak{X}(i, j)$ that a pair $\mathcal{C}_i(G_0)$ and $\mathcal{C}_j(G_0)$ is excessive is at most $(|\mathcal{C}_i(G_0)||\mathcal{C}_j(G_0)|p_1)^2$, and the probability of the event $\mathfrak{X}(i)$ that $\mathcal{C}_i(G_0)$ is excessive is at most $|\mathcal{C}_i(G_0)|^2 p_1$. Therefore, as $p_1 = (1 + o(1))\varepsilon/n$, the variable $X := \sum_i \mathbf{1}_{\mathfrak{X}(i)} + \sum_{i<j} \mathbf{1}_{\mathfrak{X}(i,j)}$ satisfies

$$\mathbb{E}[X \mid G_0] \le (1 + o(1))\Big(\frac{\varepsilon}{n}\sum_i |\mathcal{C}_i(G_0)|^2\Big)^2 + (1 + o(1))\frac{\varepsilon}{n}\sum_i |\mathcal{C}_i(G_0)|^2 . \quad (5.9)$$

For every pair of vertices $u, v \in [n]$, consider a shortest path $P = (x_1, \ldots, x_m)$ between them in G_1, and suppose that P uses an edge between a vertex $x_l \in \mathcal{C}_i(G_0)$ and $x_{l+1} \in \mathcal{C}_j(G_0)$ that does not belong to $\mathfrak{G}^{\mathbf{H}}$. Let r be the largest index such that $x_r \in \mathcal{C}_j(G_0)$. If $i = j$ then we replace the segment $x_l, \ldots, x_r$ in P with a shortest path in G_0 from x_l to x_r. If on the other hand $i \ne j$, denote by yz the edge in $\mathfrak{G}^{\mathbf{H}}$ that connects $\mathcal{C}_i(G_0)$ and $\mathcal{C}_j(G_0)$, and replace the segment $x_l, \ldots, x_r$ by a shortest path from x_l to y in G_0, followed by the edge yz and then by a shortest path from z to x_r in G_0. Repeat this procedure until a path in $\mathfrak{G}^{\mathbf{H}}$ is obtained. The number of detours does not exceed the number of excessive pairs and components, and each such detour increases the length of the path by at most $2d_{\max}(G_0)$. Therefore,

$$0 \le \operatorname{dist}_{\mathfrak{G}^{\mathbf{H}}}(u, v) - \operatorname{dist}_{G_1}(u, v) \le 2d_{\max}(G_0)X .$$

Since this bound holds true for every two vertices u, v, and the measure on $\mathsf{M}_j(G_1)$ identifies with the measure on $\tilde{\mathsf{M}}_j(\mathfrak{G}^{\mathbf{H}})$, we find, by taking the scaling factors of the distances into consideration, that

$$d_{\text{GHP}}(\mathsf{M}_j(G_1), \tilde{\mathsf{M}}_j(\mathfrak{G}^{\mathbf{H}})) \le 2\mathfrak{s}(\mathbf{x}, \mathbf{H})d_{\max}(G_0)X + |\mathfrak{s}(\mathbf{x}, \mathbf{H}) - n^{-1/3}|\operatorname{diam}(\mathcal{C}_j(\mathfrak{G}^{\mathbf{H}})) .$$

Define $\mathcal{B}_0^{(n)}$ to be the set of every graph G such that, for $\mathbf{x}, \mathbf{H}, q$ as in (5.2) with G playing the role of G_0, and σ_r, $x_{\max}$, $x_{\min}$ and $d_{\max}$ defined as in Proposition 5.1, the following conditions are satisfied:

- Inequalities (5.3)–(5.7) of Lemma 5.2 hold;
- $\sigma_2 \le 2/(\varepsilon n^{1/3})$;
- $d_{\max} \le 2\log(\varepsilon^3 n)/\varepsilon$.

By Lemma 5.2 and Theorems 3.1 and 3.3, respectively, we have $\mathbb{P}(G_0 \in \mathcal{B}_0^{(n)}) \to 1$. Moreover, for every $G \in \mathcal{B}_0^{(n)}$, we infer from (5.7) and the condition on $d_{\max}$ that

$$\mathbb{E}\Big[d_{\text{GHP}}(\mathsf{M}_j(G_1), \tilde{\mathsf{M}}_j(\mathfrak{G}^{\mathbf{H}})) \mid G_0 = G\Big] \le (2 + o(1))(\varepsilon^3 n)^{-1/3}\log(\varepsilon^3 n)\mathbb{E}[X \mid G_0]$$

$$+ 4(\varepsilon^3 n)^{-2/5}\mathbb{E}\Big[n^{-1/3}\operatorname{diam}(\mathcal{C}_j(\mathfrak{G}^{\mathbf{H}})) \mid G_0 = G\Big] .$$

For all $G \in \mathcal{B}_0^{(n)}$, Proposition 5.1 implies $\mathbb{E}[n^{-1/3}\operatorname{diam}(\mathcal{C}_j(\mathfrak{G}^{\mathbf{H}})) \mid G_0 = G] = O(1)$, and $\mathbb{E}[X \mid G_0 = G] = O(1)$ using $\sum_i \mathcal{C}_i(G)|^2 \le 2n/\varepsilon$ combined with (5.9). It thus follows that $\mathbb{E}[d_{\text{GHP}}(\mathsf{M}_j(G_1), \tilde{\mathsf{M}}_j(\mathfrak{G}^{\mathbf{H}})) \mid G_0 = G] \to 0$ as $n \to \infty$, as claimed. ∎

To complete the proof of Theorem 2, (i), first consider $n^{-1/3+\delta} < \varepsilon < 1 - n^{-1/2}$. By combining Proposition 5.1 with Lemmas 5.2 and 5.3 we find that for each $n \geq 1$ there exists a set of graph $\tilde{\mathcal{B}}_0^{(n)}$ so that $\mathbb{P}(G_0 \in \tilde{\mathcal{B}}_0^{(n)}) \to 1$, and for every $G \in \tilde{\mathcal{B}}_0^{(n)}$,

(1) The measured metric space $\tilde{\mathsf{M}}_j(\mathfrak{G}^{\mathbf{H}})$ conditioned on $G = G_0$ converges in distribution to $\mathfrak{M}_j(\lambda)$ w.r.t. GHP distance.

(2) The spaces $\tilde{\mathsf{M}}_j(\mathfrak{G}^{\mathbf{H}})$ and $\mathsf{M}_j(G_1)$, conditioned on $G_0 = G$, can be coupled such that their GHP distance vanishes in probability.

Therefore, by Slutsky's Lemma for Polish spaces (see, e.g.,[6, Thm. 3.1]), for every $\mathbb{P}(\mathfrak{M}_j(\lambda) \in \cdot)$-continuity Borel set $S \subseteq \mathscr{M}$,

$$\lim_{n \to \infty} \max_{G \in \tilde{\mathcal{B}}_0^{(n)}} |\mathbb{P}_n \left(\mathsf{M}_j(G_1) \in S \mid G_0 = G \right) - \mathbb{P}(\mathfrak{M}_j(\lambda) \in S)| = 0 \,.$$

By Corollary 2.3, this completes the first assertion in Theorem 2, (i) in this regime.

For $\varepsilon \geq 1 - n^{-1/2}$, this assertion will follow directly from Theorem 2, Part (ii), whose proof is in §5.2 below. Indeed, we establish there that if $\tilde{\varepsilon}$ satisfies $\tilde{\varepsilon}^3 n \to 0$, and $\tilde{G}_0$ and $\tilde{G}_1$ are random graphs distributed as

$$\tilde{G}_0 \sim \mathcal{G}\left(n, \tfrac{1-\tilde{\varepsilon}}{1-\tilde{\varepsilon}p}p\right) \quad , \quad \tilde{G}_2 \sim \mathcal{G}\left(n, \tilde{\varepsilon}p\right) \,,$$

then their union $\tilde{G}_1 = \tilde{G}_0 \cup \tilde{G}_2$ satisfies $d_{\text{GHP}}(\mathsf{M}_j(\tilde{G}_0), \mathsf{M}_j(\tilde{G}_1)) \xrightarrow{\text{P}} 0$ as $n \to \infty$. By Markov's inequality, this holds conditionally on $\tilde{G}_2$ such that $\tilde{G}_2 \in \mathcal{K}$, for some set of graphs $\mathcal{K}$ with $\mathbb{P}(\tilde{G}_2 \in \mathcal{K}) \to 1$. In addition, $\mathsf{M}_j(\tilde{G}_0) \xrightarrow{\text{d}} \mathfrak{M}_j(\lambda)$ as $n \to \infty$ since $(n \tfrac{1-\tilde{\varepsilon}}{1-\tilde{\varepsilon}p}p - 1)n^{1/3} \to \lambda$. It follows that for every $\tilde{G}_2 \in \mathcal{K}$, the conditional law of $\mathsf{M}_j(\tilde{G}_1)$ given $\tilde{G}_2$ weakly converges to that of $\mathfrak{M}_j(\lambda)$. We now flip the roles of the graphs by setting $\tilde{\varepsilon} = \tfrac{1-\varepsilon}{1-\varepsilon p}$. Observe that $\tilde{\varepsilon}^3 n \to 0$, $G_0 = \tilde{G}_2$ and $G_1 = \tilde{G}_0 \cup \tilde{G}_2$. Thus, the conditional law of $\mathsf{M}_j(G_1)$ given G_0 such that $G_0 \in \mathcal{K}$ converges in distribution to $\mathfrak{M}_j(\lambda)$ as $n \to \infty$, and Corollary 2.3 concludes the proof.

The second assertion of Theorem 2, Part (i), follows from the first as shown in [2, Thm. 4.1], by the same argument appearing there. ∎

5.2. **Noise stability: proof of Theorem 2, Part (ii).** Let $\varepsilon^3 n \to 0$ and G_0, G_1 as above. We will prove that $d_{\text{GHP}}(\mathsf{M}_j(G_0), \mathsf{M}_j(G_1)) \xrightarrow{\text{P}} 0$ for every fixed j by showing that, w.h.p., $\mathcal{C}_j(G_1)$ is obtained from $\mathcal{C}_j(G_0)$ by gluing to it components of size $o(n^{2/3})$ and diameter $o(n^{1/3})$, which are negligible in the scaling limit.

Lemma 5.4. *Let $j \geq 1$ be an integer, and let A_j be the event that the component of G_1 containing $\mathcal{C}_j(G_0)$ is $\mathcal{C}_j(G_1)$. Then, $\mathbb{P}(A_j) \to 1$ as $n \to \infty$.*

Proof. Suppose that A_j does not hold, and let $i \leq j$ be the smallest index for which A_i does not hold. Let k be the index such that $\mathcal{C}_k(G_1) \supseteq \mathcal{C}_i(G_0)$; then

$$\|\mathbf{X}(G_1) - \mathbf{X}(G_0)\|_2^2 \geq n^{-4/3}\Delta_k^2 \quad \text{for} \quad \Delta_k := |\mathcal{C}_k(G_1)| - |\mathcal{C}_k(G_0)| \,.$$

If $k < i$ then $|\mathcal{C}_k(G_1)| \geq |\mathcal{C}_k(G_0)| + |\mathcal{C}_i(G_0)|$ by the minimality of i, so $\Delta_k \geq |\mathcal{C}_i(G_0)|$. If $k > i$ then we simply use $|\mathcal{C}_k(G_1)| \geq |\mathcal{C}_i(G_0)|$, whence $\Delta_k \geq |\mathcal{C}_i(G_0)| - |\mathcal{C}_k(G_0)|$. So, if $D_{i,a}^{(n)} = \{|\mathcal{C}_i(G_0)| < an^{2/3}\}$ and $D_{i,k,a}^{(n)} = \{|\mathcal{C}_i(G_0)| - |\mathcal{C}_k(G_0)| < an^{2/3}\}$, then

$$\mathbb{P}(A_j^c) \leq \inf_{a>0} \left\{ \mathbb{P}\left(\|\mathbf{X}(G_1) - \mathbf{X}(G_0)\|_2 \geq a\right) + \sum_{i \leq j} \mathbb{P}(D_{i,a}^{(n)}) + \sum_{i \leq j}\sum_{k>i} \mathbb{P}(D_{i,k,a}^{(n)}) \right\} \,.$$

The first term vanishes for any fixed $a > 0$ as $n \to \infty$ by Corollary 4.5, and the other two vanish as $a \downarrow 0$ since $\mathbf{X}(G_0)$ converges to a continues distribution in $\ell_\downarrow^2$. ∎

We need the next standard fact, which follows immediately from the work of Lyons [16] connecting percolation, random walks and electrical networks on trees.

Fact 5.5. *Let $\mathcal{T}$ be a $\mathrm{Bin}(n, \frac{\mu}{n})$-Galton–Watson tree for $\mu > 0$, and let $\mathrm{ht}(\mathcal{T})$ denote its height. Then for every k one has $1/a_k \leq \mathbb{P}(\mathrm{ht}(\mathcal{T}) \geq k) \leq 2/a_k$, where*

$$a_k = \begin{cases} 1 + (1 - \frac{\mu}{n})\frac{1-\mu^{-k}}{\mu-1} & \text{if } \mu \neq 1\,, \\ 1 + (1 - \frac{\mu}{n})k & \text{if } \mu = 1\,. \end{cases}$$

Proof. Letting $\mathcal{T}_{n,k}$ denote the first levels of the complete n-ary tree, we assign the conductance $\mathfrak{c}_j = (\mu/n)^j/(1 - \mu/n)$ to every edge in between levels $j - 1$ and j in this tree (regarding the level of the root to be 0). Further denoting by $\mathcal{R}_k$ the resistance between the root and level k of $\mathcal{T}_{n,k}$, appealing to [16, Thm. 2.1] shows that $1/(\mathcal{R}_k + 1) \leq \mathbb{P}(\mathrm{ht}(\mathcal{T}) \geq k) \leq 2/(\mathcal{R}_k + 1)$. Contracting the vertices in each level yields a path from 0 to k where $\tilde{\mathfrak{c}}_j = n^j\mathfrak{c}_j = \mu^j/(1 - \mu/n)$ for $1 \leq j \leq k$. Since we have $\mathcal{R}_k = \sum_{j=1}^k \tilde{\mathfrak{c}}_j^{-1}$, the required inequality is established. ∎

Lemma 5.6. *Let $j \geq 1$, $\delta > 0$ and denote by $B_{j,\delta}$ the event that*

$$\sup_{v \in \mathcal{C}_j(G_1)\backslash\mathcal{C}_j(G_0)} \mathrm{dist}_{G_1}(v, \mathcal{C}_j(G_0)) \leq \delta n^{1/3}.$$

Then, with A_j defined as in Lemma 5.4, we have $\mathbb{P}(A_j \setminus B_{j,\delta}) \to 0$ as $n \to \infty$.

Proof. Denote by $r_G(u) := \max_v \mathrm{dist}_G(u, v)$ where the maximum is over every vertex v that lies in the same component in G as u. Note that if H is a subgraph of G and u lies in an acyclic component of G then $r_G(u) \geq r_H(u)$.

Let us define a graph $G_0 \subseteq G^{(j)} \subseteq G_1$ obtained by removing from G_1 all the edges that do not belong to G_0 and have an endpoint in $\mathcal{C}_j(G_0)$. As such, G_1 is obtained from $G^{(j)}$ by connecting every pair of non-adjacent vertices in G_0 that has an endpoint in $\mathcal{C}_j(G_0)$ independently with probability p_1. Therefore, $A_j \setminus B_{j,\delta}$ can hold only if an edge is added in this way to $G^{(j)}$ between $\mathcal{C}_j(G_0)$ and T, where

$$T = \left\{ u \notin \mathcal{C}_j(G_0) \,:\, r_{G^{(j)}}(u) > \delta n^{1/3} \right\}$$

(a shortest path from $\mathcal{C}_j(G_0)$ to v in G_1 starts with such an edge, and all its other edges belong to $G^{(j)}$). Let $E_\kappa = \{|\mathcal{C}_j(G_0)| \leq \kappa n^{2/3}\}$ for some large constant κ; writing $\mathbb{P}(E_\kappa) = 1 - \delta_\kappa$, as mentioned above one has $\delta_\kappa \to 0$ as $\kappa \to \infty$. Hence,

$$\mathbb{P}(A_j \setminus B_{j,\delta}) \leq \delta_\kappa + \mathbb{P}(A_j \setminus B_{j,\delta}\,, E_\kappa) \leq \delta_\kappa + \mathbb{E}\big[\mathbb{E}\left[p_1|\mathcal{C}_j(G_0)||T|\mathbf{1}_{E_\kappa} \mid G_0\right]\big]$$

$$\leq \delta_\kappa + p_1\kappa n^{2/3}\mathbb{E}|T|\,. \tag{5.10}$$

Every $u \in T$ either belongs to a cyclic component of G_1, or has $r_{G_1}(u) > \delta n^{1/3}$. The expected number of vertices in cyclic components is $O(n^{2/3})$ (see [15, Lemma 2.2]), and by Fact 5.5, the probability that $r_{G_1}(u) > \delta n^{1/3}$ is at most $(1/\delta + o(1))n^{-1/3}$ if $\lambda = 0$, and $(\lambda/(1 - e^{-\lambda\delta}) + o(1))n^{-1/3}$ if $\lambda \neq 0$. Thus, $\mathbb{E}|T| = O(n^{2/3}/\delta)$.

Combining this, as well as $p_1 = (1+o(1))\varepsilon/n$, with (5.10), we may now conclude that $\mathbb{P}(A_j \setminus B_{j,\delta}) \leq \delta_\kappa + O(\kappa\delta^{-1}\varepsilon n^{1/3}) = \delta_\kappa + o(1)$, since $\varepsilon n^{1/3} \to 0$. ∎

Lemma 5.7. *Let $j \geq 1$, and denote by C_j the event that $\mathrm{dist}_{G_0}(u, v) = \mathrm{dist}_{G_1}(u, v)$ for every two vertices $u, v \in \mathcal{C}_j(G_0)$. Then, $\mathbb{P}(C_j) \to 1$.*

Proof. First we claim that w.h.p. no two vertices in $\mathcal{C}_j(G_0)$ that are non-adjacent in G_0 are adjacent in G_1. Indeed, the expected number of such pairs of vertices is at most $\mathbb{E}[|\mathcal{C}_j(G_0)|^2]p_1 = O(n^{1/3}\varepsilon) \to 0$ as $n \to \infty$.

Next, consider the random graph $G_0 \subseteq G^{(j)} \subseteq G_1$ that was defined in Lemma 5.6. We claim that w.h.p. no component of $G^{(j)}$ is connected to $\mathcal{C}_j(G_0)$ in G_1 by more than one edge. Together with the first claim, this would imply that the shortest path in G_1 between any pair of vertices in $\mathcal{C}_j(G_0)$ contains only edges from G_0.

Let X be the number of pairs of vertices that lie in the same component of $G^{(j)}$ and are both connected by an edge to $\mathcal{C}_j(G_0)$ in G_1. For $E_\kappa = \{|\mathcal{C}_j(G_0)| \leq \kappa n^{2/3}\}$ as in Lemma 5.6, satisfying $\mathbb{P}(E_\kappa^c) = \delta_k \to 0$ as $\kappa \to \infty$, we have

$$\mathbb{P}(X > 0) \leq \mathbb{P}(E_\kappa^c) + \mathbb{E}[X\mathbf{1}_{E_\kappa}] \leq \delta_k + \mathbb{E}\Big[\mathbb{E}\big[p_1^2|\mathcal{C}_j(G_0)|^2\mathbf{1}_{E_\kappa}\sum_i|\mathcal{C}_i(G^{(j)})|^2 \mid G_0\big]\Big]$$

$$\leq \delta_k + (1+o(1))\big(\kappa\varepsilon n^{1/3}\big)^2\mathbb{E}\Big[n^{-4/3}\sum_i|\mathcal{C}_i(G^{(j)})|^2\Big].$$

By Lemma 4.3, $\sum_i|\mathcal{C}_i(G^{(j)})|^2 \leq \sum_i|\mathcal{C}_i(G_1)|^2$ deterministically, since $G^{(j)} \subseteq G_1$. In addition, we have

$$\mathbb{E}\Big[n^{-4/3}\sum_i|\mathcal{C}_i(G_1)|^2\Big] = O(1)$$

by the convergence of $\mathbf{X}(G_1)$ in $(\ell_\downarrow^2, \|\cdot\|_2)$. Therefore,

$$\mathbb{P}(X > 0) \leq \delta_k + O((\kappa\varepsilon n^{1/3})^2) = \delta_k + o(1)$$

since $\varepsilon n^{1/3} \to 0$. ∎

We may now conclude the proof of Theorem 2, Part (ii). Fix $j \geq 1$, and condition on the events $A_j \cap B_{j,\delta} \cap C_j$ as per Lemmas 5.4, 5.6 and 5.7. On these events, we have that $\mathcal{C}_j(G_0)$ is isometrically embedded in $\mathcal{C}_j(G_1)$, and $\mathcal{C}_j(G_1)$ is contained in a ball of radius $o(n^{1/3})$ about $\mathcal{C}_j(G_0)$. In addition, both graph are equipped with the counting measure multiplied by $n^{-2/3}$ and they differ by only $o(n^{2/3})$ vertices. Therefore, $d_{\text{GHP}}(\mathsf{M}_j(G_0), \mathsf{M}_j(G_1)) = o(1)$.

The extension of convergence in the product topology to convergence in d_{GHP}^4 follows, as it did in Part (i), from the proof of [2, Thm. 4.1]. ∎

References

[1] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probab. Theory Related Fields*, 152(3-4):367–406, 2012.

[2] L. Addario-Berry, N. Broutin, C. Goldschmidt, and G. Miermont. The scaling limit of the minimum spanning tree of the complete graph. *Ann. Probab.*, 45(5):3075–3144, 2017.

[3] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, 25(2):812–854, 1997.

[4] I. Benjamini, G. Kalai, and O. Schramm. Noise sensitivity of Boolean functions and applications to percolation. *Inst. Hautes Études Sci. Publ. Math.*, (90):5–43, 1999.

[5] S. Bhamidi, N. Broutin, S. Sen, and X. Wang. Scaling limits of random graph models at criticality: Universality and the basin of attraction of the Erdős-Rényi random graph. 2014. Preprint. Available at `arXiv:1411.3417`.

[6] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons, 1999.

[7] B. Bollobás. The evolution of random graphs. *Trans. Amer. Math. Soc.*, 286(1):257–274, 1984.

[8] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[9] C. Garban and J. E. Steif. *Noise sensitivity of Boolean functions and percolation*, volume 5 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, New York, 2015.

[10] S. Janson and M. J. Luczak. Susceptibility in subcritical random graphs. *J. Math. Phys.*, 49(12):125207, 23, 2008.

[11] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[12] E. Lubetzky and J. E. Steif. Strong noise sensitivity and random graphs. *Ann. Probab.*, 43(6):3239–3278, 2015.

[13] T. Łuczak. Component behavior near the critical point of the random graph process. *Random Structures Algorithms*, 1(3):287–310, 1990.

[14] T. Łuczak. Random trees and random graphs. In *Proceedings of the Eighth International Conference "Random Structures and Algorithms" (Poznan, 1997)*, volume 13, pages 485–500, 1998.

[15] T. Łuczak, B. Pittel, and J. C. Wierman. The structure of a random graph at the point of the phase transition. *Trans. Amer. Math. Soc.*, 341(2):721–748, 1994.

[16] R. Lyons. Random walks, capacity and percolation on trees. *Ann. Probab.*, 20(4):2043–2088, 1992.

[17] M. I. Roberts and B. Şengül. Exceptional times of the critical dynamical Erdős-Rényi graph. *Ann. Appl. Probab.*, 28(4):2275–2308, 2018.

[18] R. Rossignol. Scaling limit of dynamical percolation on the critical Erdős–Rényi random graph. 2017. Preprint. Available at `arXiv:1710.09101`.

E. Lubetzky

Courant Institute, New York University, 251 Mercer Street, New York, NY 10012.
Email address: `eyal@courant.nyu.edu`

Y. Peled

Courant Institute, New York University, 251 Mercer Street, New York, NY 10012.
Email address: `yuval.peled@courant.nyu.edu`