A Appendix

A.1 Proofs in Section 2

Proof of Lemma 2. To expedite the proof, we express the LR statistics in terms of the sufficient statistics $\mathbf{y}_0 = \frac{1}{|C|}\sum_{i\in C}\mathbf{y}_i \sim N(\beta_0, \sigma_0^2)$ and $\mathbf{y}_1 = \frac{1}{|\bar{C}|}\sum_{i\in\bar{C}}\mathbf{y}_i \sim N(\beta_1, \sigma_1^2)$ for $\sigma_0 = \sigma/\sqrt{|C|}$ and $\sigma_1 = \sigma/\sqrt{|\bar{C}|}$. Then, we obtain

$$2\log\Lambda_C(\mathbf{y}) = \frac{1}{\sigma_0^2}(\mathbf{y}_0 - \hat{\beta})^2 + \frac{1}{\sigma_1^2}(\mathbf{y}_1 - \hat{\beta})^2$$

where $\hat{\beta} = \frac{\sigma_1^2}{\sigma_0^2+\sigma_1^2}\mathbf{y}_0 + \frac{\sigma_0^2}{\sigma_0^2+\sigma_1^2}\mathbf{y}_1$ is the MLE under H_0. (The likelihood under the alternative balances with the normalizing constant of the null likelihood.) Thus,

$$\begin{aligned}
2\log\Lambda_C(\mathbf{y}) &= \frac{1}{\sigma_0^2}\left(\frac{\sigma_0^2}{\sigma_0^2+\sigma_1^2}(\mathbf{y}_0-\mathbf{y}_1)\right)^2 + \frac{1}{\sigma_1^2}\left(\frac{\sigma_1^2}{\sigma_0^2+\sigma_1^2}(\mathbf{y}_0-\mathbf{y}_1)\right)^2 \\
&= \frac{(\mathbf{y}_0-\mathbf{y}_1)^2}{\sigma_0^2+\sigma_1^2} = \frac{1}{\sigma^2}\frac{|C||\bar{C}|}{|V|}(\mathbf{y}_0-\mathbf{y}_1)^2 \\
&= \frac{1}{\sigma^2}\frac{|V|}{|C||\bar{C}|}\left(\frac{|\bar{C}|}{|V|}\sum_{v\in C}\mathbf{y}_v - \frac{|C|}{|V|}\sum_{v\in\bar{C}}\mathbf{y}_v\right)^2 \\
&= \frac{1}{\sigma^2}\frac{|V|}{|C||\bar{C}|}\left(\sum_{v\in C}\mathbf{y}_v - \frac{|C|}{|V|}\sum_{v\in V}\mathbf{y}_v\right)^2 \\
&= \frac{1}{\sigma^2}\frac{|V|}{|C||\bar{C}|}\left(\sum_{v\in C}\tilde{\mathbf{y}}_v\right)^2. \qquad (11)
\end{aligned}$$

Now we let $\mathbf{x} = \mathbf{1}_C$, making the statistic above

$$2\sigma^2\log\Lambda_C(\mathbf{y}) = \frac{\mathbf{x}^\top\tilde{\mathbf{y}}\tilde{\mathbf{y}}\mathbf{x}}{\mathbf{x}^\top\mathbf{K}\mathbf{x}} \text{ and } \frac{|\partial C||V|}{|C||\bar{C}|} = \frac{\mathbf{x}^\top\mathbf{L}\mathbf{x}}{\mathbf{x}^\top\mathbf{K}\mathbf{x}}.$$

The result now follows by considering all the indicator functions corresponding to the sets in $\mathcal{C}$. □

Proof of Remark 4. First we notice that (8) is equivalent to

$$\inf_{\mathbf{x}\in\mathbb{R}} -\mathbf{x}^\top\tilde{\mathbf{y}} \text{ s.t. } \mathbf{x}^\top\mathbf{L}\mathbf{x} \le \rho, \|\mathbf{x}\| \le 1$$

because $\mathbf{x}^\top\mathbf{L}\mathbf{x}$ and $\mathbf{x}^\top\tilde{\mathbf{y}}$ are invariant under changes in $\mathbf{1}^\top\mathbf{x}$. This admits the Lagrangian (for parameters $\nu_0, \nu_1 > 0$),

$$-\mathbf{x}^\top\tilde{\mathbf{y}} + \nu_0(\mathbf{x}^\top\mathbf{L}\mathbf{x} - \rho) + \nu_1(\mathbf{x}^\top\mathbf{x} - 1)$$

which is minimized for fixed ν_0, ν_1 at $\mathbf{x} = -\frac{1}{2}[\nu_0\mathbf{L} + \nu_1\mathbf{I}]^{-1}\tilde{\mathbf{y}}$ (which confirms Slater's condition). Hence, the dual program is

$$\sup_{\nu_0,\nu_1\ge 0} -\nu_0\rho - \nu_1 - \frac{1}{2}\tilde{\mathbf{y}}[\nu_0\mathbf{L}+\nu_1\mathbf{I}]^{-1}\tilde{\mathbf{y}} + \frac{1}{4}\tilde{\mathbf{y}}[\nu_0\mathbf{L}+\nu_1\mathbf{I}]^{-1}\tilde{\mathbf{y}}$$

□

A.2 Proofs in Section 3

Proof of Theorem 5 (1). Let the true $C \in \mathcal{C}$ be known. The performance of the optimal test with C known, which by the Neyman-Pearson Lemma is based on $2\log\Lambda_C(\mathbf{y})$, bounds the performance of that with C unknown. To this end, note that, under H_0, the LR statistic (6) has a χ_1^2, while under the alternative H_1^C it has a $\chi_1^2(\lambda)$ distribution with non-centrality parameter

$$\lambda = \frac{\delta^2}{\sigma^2}\frac{|C||\bar{C}|}{|V|} = \frac{\eta^2}{\sigma^2},$$

which is the square of the SNR. For fixed C, asymptotically indistinguishable of H_0 versus H_C^1 follows by considering any threshold and noticing that the associated type 1 and type 2 errors are non-vanishing under the SNR scaling assumed in the statement. Since the risk of testing H_0 versus H_1 is no smaller than the risk of testing H_0 versus H_C^1, the result follows. □

We remark that the proof of the previous result shows that when distinguishing H_0 from H_1^C, the power of the test is maximal when $|C| = |\bar{C}|$ for a fixed value of the SNR.

Proof of Theorem 5 (2). We will begin by constructing from our set, $\mathcal{C}'$, a new set, $\mathcal{S}$, of clusters which are difficult to distinguish in the sense that the Bayes risk for the uniform prior over those in the alternative is bounded away from 0. Enumerate $\mathcal{C}'$ such that $\mathcal{C}' = \{C_i\}_{i=1}^{|\mathcal{C}'|}$. We will build $\mathcal{S}$ by unioning k elements of $\mathcal{C}'$, then draw S, S' uniformly from $\mathcal{S}$. Specifically, let $k = \lfloor\sqrt{|\mathcal{C}'|}\rfloor$ (recall that $c = |C|, \forall C \in \mathcal{C}'$), and let K, K' be independent uniform samples without replacement of k elements from $\{1, \ldots, |\mathcal{C}'|\}$. Then let $S = \cup_{i\in K} C_i$ and $S' = \cup_{i\in K'} C_i$. Notice that $kc = |S| \le n/2$ for n large enough.

$$\begin{aligned}
\frac{|\partial S|}{|S||\bar{S}|} &\le \frac{k \max_{C\in\mathcal{C}'}|\partial C|}{kc(n-kc)} \\
&\le \frac{n-c}{n-kc}\max_{C\in\mathcal{C}'}\frac{|\partial C|}{c(n-c)} \le 2\frac{\rho}{2} = \rho
\end{aligned}$$

Notice that the risk can be bounded by

$$\begin{aligned}
&\sup_{\boldsymbol{\beta}\in\Theta_0}\mathbb{E}_{\boldsymbol{\beta}}T(\mathbf{y}) + \sup_{\boldsymbol{\beta}\in\Theta_1}\mathbb{E}_{\boldsymbol{\beta}}[1 - T(\mathbf{y})] \\
&\ge \mathbb{E}_{\boldsymbol{\beta}=\mathbf{0}}T(\mathbf{y}) + \frac{1}{|\mathcal{S}|}\sum_{S\in\mathcal{S}}\mathbb{E}_{\boldsymbol{\beta}^S}[1 - T(\mathbf{y})] = R^*
\end{aligned}$$

where $\boldsymbol{\beta}^S = \eta\sqrt{\frac{n}{|S||\bar{S}|}}\mathbf{1}_S$ and $\mathcal{S} \subseteq \mathcal{C}$. Then by Proposition 3.2 in [1],

$$R^* \geq 1 - \frac{1}{2}\sqrt{\mathbb{E}\exp\left\{\frac{\eta^2}{\sigma^2}Z\right\} - 1}$$

where

$$Z = \frac{n|S \cap S'|}{\sqrt{|S||\bar{S}||S'||\bar{S}'|}}$$

for S, S' drawn independently uniformly from $\mathcal{S}$. Notice that

$$\frac{n}{\sqrt{|\bar{S}'||\bar{S}|}} \leq 2$$

Hence,

$$Z \leq 2\frac{|S \cap S'|}{\sqrt{|S||S'|}} = 2\frac{|K \cap K'|}{\sqrt{|K||K'|}}$$

And we have that

$$R^* \geq 1 - \frac{1}{2}\sqrt{\mathbb{E}e^{\frac{2\eta^2}{k\sigma^2}|K\cap K'|} - 1}$$

Hence, we can apply Proposition 3.4 from [1] (by substituting $\mu \leftarrow \eta\sqrt{2}/(\sigma\sqrt{k})$) and determine that $R^* > \delta$ if

$$\frac{\eta\sqrt{2}}{\sigma\sqrt{k}} \leq \sqrt{\log\left(1 + \frac{|\mathcal{C}'|\log(1 + 4(1-\delta)^2)}{k^2}\right)}$$

Because $k^2 \asymp |\mathcal{C}'|$ we have asymptotic indistinguishability if $\eta/\sigma = o(\sqrt{k}) = o(|\mathcal{C}'|^{1/4})$. For some explanation for the choice of k the term $k\log(1 + |\mathcal{C}'|/k^2)$ is largest when $k^2 \asymp |\mathcal{C}'|$. ☐

Proof of Lemma 7. Without loss of generality, let $\mathbf{y} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. We recall that, since G is connected, the combinatorial Laplacian $\mathbf{L}$ is symmetric, its smallest eigenvalue is zero and the remaining eigenvalues are positive. By the spectral theorem, we can write $\mathbf{L} = \mathbf{U}\Lambda\mathbf{U}^\top$, where Λ is a $(n-1)\times(n-1)$ diagonal matrix containing the positive eigenvalues of $\mathbf{L}$ in increasing order and the columns of the $n \times (n-1)$ matrix $\mathbf{U}$ are the associated eigenvectors. Then, since each vector $\mathbf{x} \in \mathbb{R}^n$ with $\mathbf{1}^\top\mathbf{x} = 0$ can be written as $\mathbf{U}\mathbf{z}$ for a unique vector $\mathbf{z} \in \mathbb{R}^{n-1}$, we have

$$\begin{aligned}
\mathcal{X} &= \{\mathbf{x} \in \mathbb{R}^n \colon \mathbf{x}^\top\mathbf{L}\mathbf{x} \leq \rho, \mathbf{x}^\top\mathbf{x} = 1, \mathbf{1}^\top\mathbf{x} \leq 0\}\\
&= \{\mathbf{U}\mathbf{z} \colon \mathbf{z} \in \mathbb{R}^{n-1},\\
&\qquad \mathbf{z}^\top\mathbf{U}^\top\mathbf{L}\mathbf{U}\mathbf{z} \leq \rho, \mathbf{z}^\top\mathbf{U}^\top\mathbf{U}\mathbf{z} \leq 1\}\\
&= \{\mathbf{U}\mathbf{z} \colon \mathbf{z} \in \mathbb{R}^{n-1}, \tfrac{1}{\rho}\mathbf{z}^\top\Lambda\mathbf{z} \leq 1, \mathbf{z}^\top\mathbf{z} \leq 1\},
\end{aligned}$$

where in the third identity we have used the fact that $\mathbf{U}^\top\mathbf{U} = \mathbf{I}_{n-1}$. Letting $\mathcal{Z} = \{\mathbf{z} \in \mathbb{R}^{n-1} \colon \frac{1}{\rho}\mathbf{z}^\top\Lambda\mathbf{z} \leq 1, \mathbf{z}^\top\mathbf{z} \leq 1\}$, we see that

$$\sup_{\mathbf{x}\in\mathcal{X}} \mathbf{x}^\top\mathbf{y} = \sup_{\mathbf{z}\in\mathcal{Z}} \mathbf{z}^\top\mathbf{U}^\top\mathbf{y} \stackrel{d}{=} \sup_{\mathbf{z}\in\mathcal{Z}} \mathbf{z}^\top\boldsymbol{\xi},$$

where $\boldsymbol{\xi} \sim N(0, \mathbf{I}_{n-1})$ and $\stackrel{d}{=}$ denotes equality in distribution.

Next, we show that the set $\mathcal{Z}$, which is the intersection of an ellipsoid with the unit ball in $\mathbb{R}^{n-1}$, is contained in an enlarged ellipsoid. The supremum of the Gaussian process $\mathbf{z}^\top\boldsymbol{\xi}$ over $\mathcal{Z}$ will then be bounded by the supremum of the same process over this larger but simpler set, which we will be able to bound using directly a result from [38] based on chaining. To this end, let $\mathbf{A} = \frac{1}{\rho}\Lambda = \mathrm{diag}\{a_i\}_{i=1}^{n-1}$ and $d = \max\{j : a_j < 1\}$. For for a vector $\mathbf{z} \in \mathbb{R}^{n-1}$ set $\mathbf{z}_1 = \mathbf{z}_{[d]}$, $\mathbf{z}_2 = \mathbf{z}_{[n-1]\setminus[d]}$, and $\mathbf{A}_2 = \mathrm{diag}\{a_i\}_{i>d}$. Then, we observe the following chain of implications, holding for vectors $\mathbf{z} \in \mathbb{R}^{n-1}$:

$$\|\mathbf{z}\| \leq 1, \mathbf{z}^\top\mathbf{A}\mathbf{z} \leq 1 \Rightarrow \|\mathbf{z}_1\| \leq 1, \sum_{i>d} a_i\mathbf{z}_i^2 \leq 1$$

$$\Rightarrow \mathbf{z}_1^\top\mathbf{z}_1 + \mathbf{z}_2^\top\mathbf{A}_2\mathbf{z}_2 \leq 2 \Rightarrow \sum_i \frac{\max\{1, a_i\}}{2}\mathbf{z}_i^2 \leq 1.$$

Hence, we have the bound

$$\mathbb{E}\sqrt{\hat{s}} \leq \mathbb{E}\sup_{\mathbf{z}\in\mathbb{R}^{n-1}} \mathbf{z}^\top\boldsymbol{\xi} \text{ s.t. } \sum_i 2\max\{1, a_i\}\mathbf{x}_i^2 \leq 1.$$

Recalling that $a_i = \frac{\lambda_{i+1}}{\rho}$, for $i = 1, \ldots, n-1$, where λ_{i+1} is the $(i+1)$th eigenvalue of $\mathbf{L}$, by Proposition 2.2.1 in [38] the right hand side of the previous expression is bounded by $\sqrt{2\sum_{i>1}\min\{1, \rho\lambda_i^{-1}\}}$. ☐

Supplement to the proof of Theorem 6. The following property of Gaussian processes effectively reduces the study of their supremum to the study of its expectation. It was established by [7] and [10] and can be found in [22].

Lemma 14. *Consider a Gaussian process $\{Z_t\}_{t\in\mathcal{U}}$ where $\mathcal{U}$ is compact with respect to metric*

$$d(s,t) = (\mathbb{E}(Z_s - Z_t)^2)^{1/2}, \quad s, t, \in \mathcal{U},$$

and let $\sigma^2 \geq \sup_{t\in\mathcal{U}}\mathbb{E}Z_t^2$. We have that with probability at least $1-\delta$

$$\left|\sup_{t\in\mathcal{U}} Z_t - \mathbb{E}\sup_{t\in\mathcal{U}} Z_t\right| < \sqrt{2\sigma^2\log\frac{2}{\delta}}.$$

Notice that the natural distance is given by $d(\mathbf{x}_0, \mathbf{x}_1) = (\mathbb{E}((\mathbf{x}_0 - \mathbf{x}_1)^\top\mathbf{y})^2)^{1/2} = \sigma\|\mathbf{x}_0 - \mathbf{x}_1\|$ for $\mathbf{x}_0, \mathbf{x}_1 \in \mathcal{X}$. ☐

A.3 Proof in Section 4

Proof of Corollary 11 (a). The study of the spectra of trees really began in earnest with the work of [12]. Notably, it became apparent that tree have eigenvalues with high multiplicities, particularly the eigenvalue 1.

[30] gave a tight bound on the algebraic connectivity of balanced binary trees (BBT). They found that for a BBT of depth ℓ, the reciprocal of the smallest eigenvalue ($\lambda_2^{(\ell)}$) is

$$\begin{aligned}\frac{1}{\lambda_2^{(\ell)}} \leq 2^\ell - 2\ell + 2 - \frac{2^\ell - \sqrt{2}(2\ell - 1 - 2^{\ell-1})}{2^\ell - 1 - \sqrt{2}(2^{\ell-1} - 1)} \\ + (3 - 2\sqrt{2}\cos(\frac{\pi}{2\ell - 1}))^{-1} \\ \leq 2^\ell + 105 I\{\ell < 4\}\end{aligned} \quad (12)$$

[32] gave a more exact characterization of the spectrum of a balanced binary tree, providing a decomposition of the Laplacian's characteristic polynomial. Specifically, the characteristic polynomial of $\mathbf{L}$ is given by

$$\begin{aligned}\det(\lambda \mathbf{I} - \mathbf{L}) = p_1^{2^{\ell-2}}(\lambda) p_2^{2^{\ell-3}}(\lambda) \dots \\ p_{\ell-3}^{2^2}(\lambda) p_{\ell-2}^2(\lambda) p_{\ell-1}(\lambda) s_\ell(\lambda)\end{aligned} \quad (13)$$

where $s_\ell(\lambda)$ is a polynomial of degree ℓ and $p_i(\lambda)$ are polynomials of degree i with the smallest root satisfying the bound in (12) with ℓ replaced with i. In [33], they extended this work to more general balanced trees.

By (13) we know that at most $\ell + (\ell - 1) + (\ell - 2)2 + ... + (\ell - j)2^{j-1} \leq \ell 2^j$ eigenvalues have reciprocals larger than $2^{\ell-j} + 105 I\{j < 4\}$. Let $k = \max\{\lceil \frac{\ell}{c} 2^{\ell(1-\alpha)} \rceil, 2^3\}$, then we have ensured that at most k eigenvalues are smaller than ρ. For n large enough

$$\begin{aligned}\sum_{i>1} \min\{1, \rho\lambda_i^{-1}\} \leq k + \rho \sum_{j > \log k}^{\ell} \ell 2^j 2^{\ell - j} \\ = k + \ell(\ell - \log k) n\rho = O(n^{1-\alpha} (\log n)^2)\end{aligned}$$

□

Proof of Corollary 11 (b). We will construct $\mathcal{C}'$ in Theorem 5 (b) from subtrees of size $4cn^\alpha$. Let C be such a subtree, then for n large enough

$$\begin{aligned}1 - 4cn^{\alpha-1} \geq \frac{1 - cn^{\alpha-1}}{2} \\ \Rightarrow \frac{n|\partial C|}{|C||\bar{C}|} = [4cn^\alpha(1 - 4cn^{\alpha-1})]^{-1} \\ \leq \frac{1}{2}[cn^\alpha(1 - cn^{\alpha-1})]^{-1} = \frac{rho}{2}\end{aligned}$$

Hence the conditions of Theorem 5 (b) hold with $|\mathcal{C}'| = n/(4cn^\alpha) \asymp n^{1-\alpha}$ □

Proof of Corollary 12 (a). By a simple Fourier analysis (see [36]), we know that the Laplacian eigenvalues are $2(2 - \cos(2\pi i_1/p) - \cos(2\pi i_2/p))$ for all $i_1, i_2 \in [p]$. Let us denote the p^2 eigenvalues as $\lambda_{(i_1,i_2)}$ for $i_1, i_2 \in [p]$. Notice that for $i \in [p]$, $|\{(i_1, i_2) : i_1 \vee i_2 = i\}| \leq 2i$. For simplicity let p be even. We know that if $i_1 \vee i_2 \leq p/2$ then $\lambda_{(i_1,i_2)} = 2 - \cos(2\pi i_1/p) - \cos(2\pi i_2/p) \geq 1 - \cos(2\pi(i_1 \vee i_2)/p)$. Thus,

$$\begin{aligned}&\sum_{(i_1,i_2)\neq(1,1)\in[p]^2} 1 \wedge \frac{\rho}{\lambda_{(i_1,i_2)}} \\ &\leq 2 \sum_{i\in[p/2]} 2i \left(1 \wedge \frac{\rho}{1 - \cos(2\pi i/p)}\right) \\ &\leq \rho \frac{p^2}{2} \frac{2}{p} \sum_{i\in[p/2]} 2 \frac{i/p}{1 - \cos(2\pi i/p)} \\ &\leq \rho \frac{p^2}{2} \int_{1/p}^{1/2} \frac{x dx}{1 - \cos(2\pi x)} \\ &\leq \rho \frac{p^2}{2} \left. \frac{\log(\sin(\pi x)) - \pi x \cot(\pi x)}{2\pi^2} \right|_{1/p}^{1/2} \\ &= \rho \frac{p^2}{2} \frac{(\pi/p)\cot(\pi/p) - \log(\sin(\pi/p))}{2\pi^2}\end{aligned}$$

While we can use the first order expansion of the terms to obtain the behavior,

$$\begin{aligned}(\pi/p)\cot(\pi/p) = 1 + o(\pi/p) \\ -\log(\sin(\pi/p)) = -\log(\pi/p) - \log(1 + o(1))\end{aligned}$$

so we arrive at the following,

$$\begin{aligned}&\sum_{(i_1,i_2)\neq(1,1)\in[p]^2} 1 \wedge \frac{\rho}{\lambda_{(i_1,i_2)}} \\ &\leq \rho \frac{p^2}{4\pi^2}(1 + \log(p/\pi) + o(1)) \\ &= \frac{C}{4\pi^2} p^{1+\beta}(1 + \log(p/\pi) + o(1)) \\ &= O(n^{(1+\beta)/2} \log(p))\end{aligned}$$

which in conjunction with (9) completes our proof. □

Proof of Corollary 13 (a). The Kronecker product of two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ is defined as $\mathbf{A} \otimes \mathbf{B} \in \mathbb{R}^{(n\times n)\times(n\times n)}$ such that $(\mathbf{A} \otimes \mathbf{B})_{(i_1,i_2),(j_1,j_2)} = A_{i_1,j_1} B_{i_2,j_2}$. Some matrix algebra shows that if H_1 and H_2 are graphs on p vertices with Laplacians $\mathbf{L}_1, \mathbf{L}_2$ then the Laplacian of their Kronecker product, $H_1 \otimes H_2$, is given by $\mathbf{L} = \mathbf{L}_1 \otimes \mathbf{I}_p + \mathbf{I}_p \otimes \mathbf{L}_2$ ([28]). Hence, if $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^p$ are eigenvectors, viz. $\mathbf{L}_1\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ and $\mathbf{L}_2\mathbf{v}_2 = \lambda_2\mathbf{v}_2$, then $\mathbf{L}(\mathbf{v}_1 \otimes \mathbf{v}_2) = (\lambda_1 + \lambda_2)\mathbf{v}_1 \otimes \mathbf{v}_2$, where $\mathbf{v}_1 \otimes \mathbf{v}_2$ is the usual tensor product. This completely characterizes the spectrum of Kronecker products of graphs.

We should argue the choice of $\rho \propto p^{2k-\ell-1}$, by showing that it is the results of cuts at level k. We say that an edge $e = ((i_1, ..., i_\ell), (j_1, ..., j_\ell))$ has scale k if $i_k \neq j_k$.

Furthermore, a cut has scale k if each of its constituent edges has scale at least k. Each edge at scale k has weight $p^{k-\ell}$ and there are $p^{\ell-1}$ such edges, so cuts at scale k have total edge weight bounded by

$$p^{\ell-1}\sum_{i=1}^{k} p^{i-\ell} = p^{k-1}\frac{p-\frac{1}{p^{k-1}}}{p-1} \leq \frac{p^k}{p-1}$$

Cuts at scale k leave components of size $p^{\ell-k}$ intact, meaning that $\rho \propto p^{2k-\ell-1}$ for large enough p.

We now control the spectrum of the Kronecker graph. Let the eigenvalues of the base graph H be $\{\nu_j\}_{j=1}^p$ in increasing order. The eigenvalues of G are precisely the sums

$$\lambda_i = \frac{1}{p^{\ell-1}}\nu_{i_1} + \frac{1}{p^{\ell-2}}\nu_{i_2} + ... + \frac{1}{p}\nu_{i_{\ell-1}} + \nu_{i_\ell}$$

for $i = (i_j)_{j=1}^{\ell} \subseteq [p]$. The eigenvalue distribution $\{\lambda_i\}$ stochastically bounds

$$\lambda_i \geq \sum_{j=1}^{\ell} \frac{1}{p^{\ell-j}}\nu_2 I\{\nu_{i_j} \neq 0\} \geq \frac{\nu_2}{p^{Z(i)}}$$

where $Z(i) = \min\{j : \nu_{i_{\ell-j}} \neq 0\}$. Notice that if i is chosen uniformly at random then $Z(i)$ has a geometric distribution with probability of success $(p-1)/p$. Also $\rho/(\frac{\nu_2}{p^{Z(i)}}) = p^{Z(i)+2k-\ell-1}/\nu_2 \geq 1$ if $Z(i) \geq \ell+1-2k+\log_p \nu_2$, so

$$\begin{aligned}&\frac{1}{p^\ell}\sum_{i\in[p]^\ell}\min\{1,\frac{\rho}{\lambda_i}\} \leq \frac{p^{2k-\ell-1}}{\nu_2}\\ &+ \sum_{Z=1}^{\lfloor \ell+1-2k+\log_p \nu_2\rfloor} \frac{p^{Z+2k-\ell-1}}{\nu_2}\frac{1}{p^Z}\frac{p-1}{p}\\ &\leq \frac{(\ell+2)p^{2k-\ell-1}}{\nu_2}\end{aligned}$$

This followed from the geometric probability mass function. We also know that the algebraic connectivity, ν_2, is bounded from below by $4p^{-2}$, so the following result holds.

□

Proof of Corollary 13 (b). Similarly to the proof of Corollary 11 (b), we form $\mathcal{C}'$ as the connected components of the graph with all the edges at coarseness less than $k-2$. So we have more than quadrupled the size of the clusters without increasing their cut size. Hence, $|\mathcal{C}'| \asymp p^{k-2} \asymp n^{k/\ell}/p^2$.

□