A Proofs

In this supplement, we present proofs for "Local Clustering of Density Upper Level Sets". We begin by providing technical lemmas, before moving on to proving the main results of the paper.

Throughout, we will fix $\mathcal{A} \subset \mathbb{R}^d$ to be an arbitrary set. To simplify expressions, for the σ-expansion $\mathcal{A}_\sigma$, we will write the set difference between $\mathcal{A}_\sigma$ and the $(\sigma + r)$-expansion $\mathcal{A}_{\sigma+r}$ as

$$\mathcal{A}_{\sigma,\sigma+r} := \{x : 0 < \rho(x, \mathcal{A}_\sigma) \leq r\},$$

where $\rho(x, \mathcal{A}) = \min_{x' \in \mathcal{A}} \|x - x'\|$.

For notational ease, we write

$$\mathrm{cut}_{n,r} = \mathrm{cut}(\mathcal{C}_\sigma[\mathbf{X}]; G_{n,r}),\ \mu_K = \mathbb{E}(\mathrm{cut}_{n,r}),\ p_K = \frac{\mu_K}{\binom{n}{2}}$$
$$\mathrm{vol}_{n,r} = \mathrm{vol}(\mathcal{C}_\sigma[\mathbf{X}]; G_{n,r}),\ \mu_V = \mathbb{E}(\mathrm{vol}_{n,r}),\ p_V = \frac{\mu_V}{\binom{n}{2}}$$

for the random variable, mean, and probability of cut size and volume, respectively.

A.1 Technical Lemmas

We state Lemma 1 without proof, as it is trivial. We formally include it mainly to comment on its (potential) suboptimality; for sets $\mathcal{A}$ with diameter much larger than σ, the volume estimate of Lemma 1 will be quite poor.

Lemma 1. *For any $\sigma > 0$ and the σ-expansion $\mathcal{A}_\sigma = \mathcal{A} + \sigma B$,*

$$\sigma B \subset \mathcal{A}_\sigma, \quad \text{and } \nu(\mathcal{A} + \sigma B) \leq \nu((1+\sigma)\mathcal{A}) = (1+\sigma)^d \nu(\mathcal{A}).$$

We will need to carefully control the volume of the expansion set using the above estimate; Lemma 2 serves this purpose.

Lemma 2. *For any $0 \leq x \leq 1/2d$,*

$$(1+x)^d \leq 1 + 2dx.$$

The proof of Lemma 2 is based on approximation via Taylor series, and we omit it.

We will repeatedly employ Lemma 1 and Lemma 2 in tandem. As a first example, in Lemma 3, we use it to bound the ratio of $\nu(\mathcal{A}_\sigma)$ to $\nu(\mathcal{A}_{\sigma-r})$. This will be useful when we bound $\mathrm{vol}(\mathcal{C}_\sigma)$.

Lemma 3. *For σ, $\mathcal{A}_\sigma$ as in Lemma 1, let $r > 0$ satisfy $r \leq \sigma/4d$. Then,*

$$\frac{\nu(\mathcal{A}_\sigma)}{\nu(\mathcal{A}_{\sigma-r})} \leq 2.$$

Proof. Fix $q = \sigma - r$. Then,

$$\begin{aligned}\nu(\mathcal{A}_\sigma) &= \nu(\mathcal{A}_{q+\sigma-q}) = \nu(\mathcal{A}_q + (\sigma - q)B) \\ &\leq \nu(\mathcal{A}_q + \frac{(\sigma - q)}{q}\mathcal{A}_q) = \left(1 + \frac{\sigma - q}{q}\right)^d \nu(\mathcal{A}_q)\end{aligned}$$

where the inequality follows from Lemma 1. Of course, $\sigma - q = r$, and $\frac{r}{q} \leq \frac{1}{2d}$ for $r \leq \frac{1}{4d}$. The claim then follows from Lemma 2. □

As part of the proof of Theorem 2, we will require an estimate of a function $g(t)$ for $t \in [x_0, 1/2]$ for some $x_0 > 0$. For $m > 0$ and $x_0 = t_0 < t_1 < \ldots < t_m = 1/2$, define the *stepwise approximation to g* to be $\bar{g}$, given by

$$\bar{g}(t) = g(t_i), \quad \text{for } t \in [t_{i-1}, t_i] \tag{A.1}$$

Lemma 4.
- *For any function f monotone decreasing in t on the interval $[x_0, 1/2]$, $\bar{f}(t) \leq f(t)$ for all $t \in [x_0, 1/2]$.*
- *Fix*

  $$g(t) = \log\left(\frac{1}{t}\right) \text{ for } x \in [x_0, 1/2]$$

  If for all i in $1, \ldots, m$, $(t_i - t_{i-1}) \leq x_0/2$, then $\bar{g}(t) \geq g(t)/2$.

Proof. The first statement is immediately obvious, and we turn to proving the second.

The upper bound $g(t) \geq \bar{g}(t)$ follows immediately from the fact that $g(t)$ is a decreasing function along with the first statement.

By the concavity of the log function,

$$\bar{g}(t) = \log\left(\frac{1}{t_i}\right) \geq \log\left(\frac{1}{t}\right) - \frac{(t_i - t)}{t}.$$

As a result,

$$\bar{g}(t) - \frac{g(t)}{2} \geq \frac{\log\left(\frac{1}{t}\right)}{2} - \frac{(t_i - t)}{t} \geq 1/2 - 1/2 = 0.$$

□

The proof of Theorem 2 also depends on a parameter – which we term *discrete local spread* – to handle the mixing over very small steps. Formally, the discrete local spread $\pi_1(G)$ is given by

$$\pi_1(G) := \frac{d_{\min}(G)^2}{10\text{vol}(V; G)} \tag{A.2}$$

where $d_{\min}(G) = \min_{v \in V} d(v)$ is the minimum degree in G. Intuitively, the discrete local spread gauges how much the walk given by $\mathbf{W}$ has mixed after one step, starting from any node v. We will denote $\pi_1(G_{n,r}[\mathcal{C}_\sigma[\mathbf{X}]])$ by $\widetilde{\pi}_{1,n}$.

Lemma 5. *For $\mathcal{C}_\sigma$ satisfying the conditions of Theorem 2:*

$$\liminf_{n \to \infty} \ \widetilde{\pi}_{1,n} \geq \frac{\lambda_\sigma^2}{\Lambda_\sigma^2} \frac{r^d}{(2D)^d}$$

Prove Lemma 5.

A.2 Cut and volume estimates

Lemma 6. *Under the conditions of Theorem 1, and for any $r < \sigma/2d$,*

$$\mathbb{P}(\mathcal{C}_{\sigma,\sigma+r}) \leq 2\nu(\mathcal{C}_\sigma) \frac{rd}{\sigma} \left(\lambda_\sigma - \frac{r^\gamma}{\gamma + 1} \right)$$

Proof. Recalling that f is the density function for $\mathbb{P}$, we have

$$\mathbb{P}(\mathcal{C}_{\sigma,\sigma+r}) = \int_{\mathcal{C}_{\sigma,\sigma+r}} f(x) dx \tag{A.3}$$

We partition $\mathcal{C}_{\sigma,\sigma+r}$ into slices, based on distance from $\mathcal{C}_\sigma$, as follows: for $k \in \mathbb{N}$,

$$\mathcal{T}_{i,k} = \left\{ x \in \mathbb{R}^d : t_{i,k} < \frac{\rho(x, \mathcal{C}_\sigma)}{r} \leq t_{i+1,k} \right\}, \quad \mathcal{C}_{\sigma,\sigma+r} = \bigcup_{i=0}^{k-1} \mathcal{T}_{i,k}$$

where $t_i = i/k$ for $i = 0, \ldots, k-1$. As a result,

$$\int_{\mathcal{C}_{\sigma,\sigma+r}} f(x) dx = \sum_{i=0}^{k-1} \int_{\mathcal{T}_{i,k}} f(x) dx \leq \sum_{i=0}^{k-1} \nu(\mathcal{T}_{i,k}) \max_{x \in \mathcal{T}_{i,k}} f(x).$$

We substitute

$$\nu(\mathcal{T}_{i,k}) = \nu(\mathcal{C}_\sigma + rt_{i+1,k}B) - \nu(\mathcal{C}_\sigma + rt_{i,k}B) := \nu_{i+1,k} - \nu_{i,k}.$$

where for simplicity we've written $\nu_{i,k} = \nu(\mathcal{C}_\sigma + rt_{i+1,k}B)$. This, in concert with the upper bound

$$\max_{x \in \mathcal{T}_{i,k}} f(x) \leq \lambda_\sigma - (rt_{i,k})^\gamma,$$

which follows from (A1) and (A2), yields

$$\sum_{i=0}^{k-1} \nu(\mathcal{T}_{i,k}) \max_{x \in \mathcal{T}_{i,k}} f(x) \leq \sum_{i=0}^{k-1} \Big\{ \nu_{i+1,k} - \nu_{i,k} \Big\} \Big(\lambda_\sigma - (rt_{i,k})^\gamma \Big)$$
$$= \sum_{i=1}^{k} \underbrace{\nu_{i,k} \Big([\lambda_\sigma - (rt_{i,k})^\gamma] - [\lambda_\sigma - (rt_{i-1,k})^\gamma] \Big)}_{:=\Sigma_k} + \underbrace{\Big(\nu_{k,k} \left[\lambda_\sigma - r^\gamma\right] - \nu_{1,k}\lambda_\sigma \Big)}_{:=\xi_k} \tag{A.4}$$

We first consider the term Σ_k. Here we use Lemma 1 to upper bound

$$\nu_{i,k} \leq \mathrm{vol}(\mathcal{C}_\sigma) \left(1 + \frac{rt_{i,k}}{\sigma}\right)^d$$

and so we can in turn upper bound Σ_k:

$$\Sigma_k \leq \mathrm{vol}(\mathcal{C}_\sigma) r^\gamma \sum_{i=1}^{k} \left(1 + \frac{rt_{i,k}}{\sigma}\right)^d \Big((t_{i-1,k})^\gamma - (t_{i,k})^\gamma \Big). \tag{A.5}$$

This, of course, is a Riemann sum, and as the inequality holds for all values of k it holds in the limit as well, which we compute to be

$$\lim_{k \to \infty} \sum_{i=1}^{k} \left(1 + \frac{rt_{i,k}}{\sigma}\right)^d \Big((t_{i-1,k})^\gamma - (t_{i,k})^\gamma \Big) = \gamma \int_0^1 \left(1 + \frac{rt}{\sigma}\right)^d t^{\gamma-1} dt$$
$$\overset{(i)}{\leq} \gamma \int_0^1 \left(1 + \frac{2drt}{\sigma}\right) t^{\gamma-1} dt = \left(1 + \frac{\gamma 2dr}{\gamma+1}\right).$$

where (i) follows from Lemma 2. We plug this estimate in to (A.5) and obtain

$$\lim_{k \to \infty} \Sigma_k \leq \mathrm{vol}(\mathcal{C}_\sigma) r^\gamma \left(1 + \frac{\gamma 2dr}{\gamma+1}\right).$$

We now provide an upper bound on ξ_k. It will follow the same basic steps as the bound on Σ_k, but will not involve integration:

$$\xi_k \overset{(ii)}{\leq} \nu(\mathcal{C}_\sigma) \left\{ \left(1 + \frac{r}{\sigma}\right)^d (\lambda - r^\gamma) - \lambda \right\}$$
$$\overset{(iii)}{\leq} \nu(\mathcal{C}_\sigma) \left\{ \left(1 + \frac{2dr}{\sigma}\right) (\lambda - r^\gamma) - \lambda \right\} = \nu(\mathcal{C}_\sigma) \left\{ \frac{2dr}{\sigma} (\lambda - r^\gamma) - r^\gamma \right\}.$$

where (ii) follows from Lemma 1 and (iii) from Lemma 2. The final result comes from adding together the upper bounds on Σ_k and ξ_k and taking the limit as $k \to \infty$. □

Lemma 7. *Under the setup and conditions of Theorem 1, and for any $r < \sigma/2d$,*

$$p_K \leq \frac{4\lambda\nu_d r^{d+1}\nu(\mathcal{C}_\sigma)d}{\sigma}\left(\lambda_\sigma - \frac{r^\gamma}{\gamma+1}\right)$$

Proof. We can write $\text{cut}_{n,r}$ as the sum of indicator functions,

$$\text{cut}_{n,r} = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{1}(x_i \in \mathcal{C}_{\sigma,\sigma+r})\mathbf{1}(x_j \in B(x_i,r) \cap \mathcal{C}_\sigma) \tag{A.6}$$

and by linearity of expectation, we can obtain

$$p_K = \frac{\mu_K}{\binom{n}{2}} = 2 \cdot \mathbb{P}(x_i \in \mathcal{C}_{\sigma,\sigma+r}, x_j \in B(x_i,r) \cap \mathcal{C}_\sigma)$$

Writing this with respect to the density function f, we have

$$\begin{aligned} p_K &= 2\int_{\mathcal{C}_{\sigma,\sigma+r}} f(x)\left\{\int_{B(x,r)\cap\mathcal{C}_\sigma} f(x')dx'\right\}dx \\ &\leq 2\nu_d r^d \lambda \int_{\mathcal{C}_{\sigma,\sigma+r}} f(x)dx \end{aligned}$$

where the inequality follows from Assumption (A3), which implies that the density function $f(x') \leq \lambda$ for all $x' \in \mathcal{C}_\sigma \setminus \mathcal{C}$ (otherwise, x' would be in some $\mathcal{C}' \in \mathbb{C}_f(\lambda)$, which (A3) forbids). Then, upper bounding the integral using Lemma 7 gives the final result. □

Lemma 8. *Under the setup and conditions of Theorem 1,*

$$p_V \geq \lambda_\sigma^2 \nu_d r^d \nu(\mathcal{C}_\sigma)$$

Proof. The proof will proceed similarly to Lemma 7. We begin by writing $\text{vol}_{n,r}$ as the sum of indicator functions,

$$\text{vol}_{n,r} = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{1}(x_i \in \mathcal{C}_\sigma)\mathbf{1}(x_j \in B(x_i,r)) \tag{A.7}$$

and by linearity of expectation we obtain

$$p_V = \frac{\mu_V}{\binom{n}{2}} = 2 \cdot \mathbb{P}(x_i \in \mathcal{C}_\sigma, x_j \in B(x_i,r)).$$

Writing this with respect to the density function f, we have

$$\begin{aligned} p_V &= 2\int_{\mathcal{C}_\sigma} f(x)\left\{\int_{B(x,r)} f(x')dx'\right\}dx \\ &\geq 2\int_{\mathcal{C}_{\sigma-r}} f(x)\left\{\int_{B(x,r)} f(x')dx'\right\}dx \\ &\overset{(i)}{\geq} 2\lambda_\sigma^2 \nu_d r^d \int_{\mathcal{C}_{\sigma-r}} f(x)dx \end{aligned}$$

where (i) follows from the fact that $B(x, r) \subset \mathcal{C}_\sigma$ for all $x \in C_{\sigma-r}$, along with the lower bound in Assumption (A1). The claim then follows from Lemma 3. $\square$

We now convert from bounds on p_K and p_V to probabilistic bounds on $\mathrm{cut}_{n,r}$ and $\mathrm{vol}_{n,r}$ in Lemmas 9 and 10. The key ingredient will be Lemma 11, Hoeffding's inequality for U-statistics; the proofs for both are nearly identical and we give only a proof for Lemma 9.

Lemma 9. *The following statement holds for any $\delta \in (0, 1]$: Under the setup and conditions of Theorem 1,*

$$\frac{\mathrm{cut}_{n,r}}{\binom{n}{2}} \leq p_K + \sqrt{\frac{\log(1/\delta)}{n}} \tag{A.8}$$

with probability at least $1 - \delta$.

Lemma 10. *The following statement holds for any $\delta \in (0, 1]$: Under the setup and conditions of Theorem 1,*

$$\frac{\mathrm{vol}_{n,r}}{\binom{n}{2}} \geq p_V - \sqrt{\frac{\log(1/\delta)}{n}} \tag{A.9}$$

with probability at least $1 - \delta$.

Proof of Lemma 9. From (A.6), we see that $\mathrm{cut}_{n,r}$, properly scaled, can be expressed as an order-2 U-statistic,

$$\frac{\mathrm{cut}_{n,r}}{\binom{n}{2}} = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \phi_K(x_i, x_j)$$

where

$$\phi_K(x_i, x_j) = \mathbf{1}(x_i \in \mathcal{A}_{\sigma,\sigma+r})\mathbf{1}(x_j \in B(x_i, r) \cap \mathcal{A}_\sigma) + \mathbf{1}(x_j \in \mathcal{A}_{\sigma,\sigma+r})\mathbf{1}(x_i \in B(x_j, r) \cap \mathcal{A}_\sigma).$$

From Lemma 11 we therefore have

$$\frac{\mathrm{cut}_{n,r}}{\binom{n}{2}} \leq p_k + \sqrt{\frac{\log(1/\delta)}{n}}$$

with probability at least $1 - \delta$. $\square$

A.3 Proof of Theorem 1

The proof of Theorem 1 is more or less given by Lemmas 7, 8, 9, and 10. All that remains is some algebra, which we take care of below.

Fix $\delta \in (0,1]$ and let $\delta' = \delta/2$. Noting that $\Phi_{n,r}(\mathcal{C}_\sigma[\mathbf{X}]) = \frac{\mathrm{cut}_{n,r}}{\mathrm{vol}_{n,r}}$, some trivial algebra gives us the expression

$$\Phi_{n,r}(\mathcal{C}_\sigma[\mathbf{X}]) = \frac{p_K + \left(\frac{\mathrm{cut}_{n,r}}{\binom{n}{2}} - p_K\right)}{p_V + \left(\frac{\mathrm{vol}_{n,r}}{\binom{n}{2}} - p_V\right)} \tag{A.10}$$

We assume (A.8) and (A.9) hold with respect to δ', keeping in mind that this will happen with probability at least $1-\delta$. Along with (A.10) this means

$$\Phi_{n,r}(\mathcal{C}_\sigma[\mathbf{X}]) \leq \frac{p_K + \mathrm{Err}_n}{p_V - \mathrm{Err}_n}$$

for $\mathrm{Err}_n = \sqrt{\frac{\log(1/\delta')}{n}}$. Now, some straightforward algebraic manipulations yield

$$\begin{aligned}
\frac{p_K + \mathrm{Err}_n}{p_V - \mathrm{Err}_n} &= \frac{p_K}{p_V} + \left(\frac{p_K}{p_V - \mathrm{Err}_n} - \frac{p_K}{p_V}\right) + \frac{\mathrm{Err}_n}{p_V - \mathrm{Err}_n} \\
&= \frac{p_k}{p_V} + \frac{\mathrm{Err}_n}{p_V - \mathrm{Err}_n}\left(\frac{p_K}{p_V} + 1\right) \\
&\leq \frac{p_K}{p_V} + 2\frac{\mathrm{Err}_n}{p_V - \mathrm{Err}_n}.
\end{aligned}$$

By Lemmas 7 and Lemma 8, we have

$$\frac{p_K}{p_V} \leq \frac{4rd}{\sigma}\frac{\lambda}{\lambda_\sigma}\frac{\left(\lambda_\sigma - \frac{r^\gamma}{\gamma+1}\right)}{\lambda_\sigma}$$

Then, the choice of

$$n \geq \frac{9\log(2/\delta)}{\epsilon^2}\left(\frac{1}{\lambda_\sigma^2 \nu(\mathcal{C}_\sigma)\nu_d r^d}\right)^2$$

implies $2\frac{\mathrm{Err}_n}{p_V - \mathrm{Err}_n} \leq \epsilon$.

A.4 Concentration inequalities

Given a symmetric kernel function $k : \mathcal{X}^m \to \mathbb{R}$, and data $\{x_1, \ldots, x_n\}$, we define the *order-m U statistic* to be

$$U := \frac{1}{\binom{n}{m}} \sum_{1 \leq i_1 < \ldots < i_m \leq n} k(x_{i_1}, \ldots, x_{i_m})$$

For both Lemmas 11 and 18, let $X_1, \ldots, X_n \in \mathcal{X}$ be independent and identically distributed. We will additionally assume the order-m kernel function k satisfies the boundedness property $\sup_{x_1,\ldots,x_m} |k(x_1, \ldots, x_m)| \leq 1$.

Lemma 11 (Hoeffding's inequality for U-statistics.)**.** *For any $t > 0$,*

$$\mathbb{P}(|U - \mathbb{E}U| \geq t) \leq 2 \exp\left\{-\frac{2nt^2}{m}\right\}$$

Further, for any $\delta > 0$, we have

$$U \leq \mathbb{E}U + \sqrt{\frac{m \log(1/\delta)}{2n}},$$
$$U \geq \mathbb{E}U - \sqrt{\frac{m \log(1/\delta)}{2n}}$$

each with probability at least $1 - \delta$.

A.5 Mixing time on graphs

For $N \in \mathbb{N}$ and a set V of N vertices, take $G = (V, E)$ to be an undirected and unweighted graph, with associated adjacency matrix $\mathbf{A}$, random walk matrix $\mathbf{W}$, and stationary distribution $\boldsymbol{\pi} = (\pi_u)_{u \in V}$ where $\pi_v = \frac{\mathbf{D}_{vv}}{\mathrm{vol}(V;G)}$. For $v \in V$,

$$q_{vu}^{(m)} = e_v \mathbf{W}^m e_u, \quad \mathbf{q}_v^{(m)} = \left(q_{vu}^{(m)}\right)_{u \in V}, \quad \mathbf{q}_v = (\mathbf{q}_{v\cdot}^{(1)}, \mathbf{q}_{v\cdot}^{(2)}, \ldots), \tag{A.11}$$

denote respectively the m-step transition probability, distribution, and sequence distributions of the random walk over G originating at v. Letting $\mathbf{q} = (\mathbf{q}_v)_{v \in V}$, the relative pointwise mixing time is thus

$$\tau_\infty(\mathbf{q}; G) = \min\left\{ m : \forall u, v \in V, \frac{\left|q_{vu}^{(m)} - \boldsymbol{\pi}_u\right|}{\boldsymbol{\pi}_u} \leq 1/4 \right\}$$

Two key quantities relate the mixing time to the expansion of subsets S of V. The *local spread* is defined to be

$$s(G) := \frac{9D_{\min}}{10}\pi_{\min}$$

for $D_{\min} := \min_{v \in V} \mathbf{D}_{vv}$ and $\pi_{\min} := D_{\min}/\mathrm{vol}(V;G)$.

where $\beta(S) := \inf_{v \in S} \mathbf{q}_v^{(1)}(S^c)$, and by convention we let $\mathbf{p}(S) = \sum_{u \in S} p_u$ for any distribution vector $\mathbf{p} = (p_u)_{u \in V}$ over V. We collect some necessary facts about the local spread in Lemma 12.

Lemma 12.
- *If $\boldsymbol{\pi}(S) \leq s(G)$, then for every $u \in S$, $\mathbf{q}_u^{(1)}(S) \geq 1/10$.*
- *For any $v, u \in V$, and $m \in N$ greater than 0, $q_{vu}^{(m)}/\pi_{\min} \leq 1/s(G)$.*

Proof. If $t = \boldsymbol{\pi}(S) \leq \frac{9D_{\min}}{10}\pi_{\min}$, divide both sides by $\pi_{\min}$ to obtain

$$|S| \leq \frac{9D_{\min}}{10}$$

which implies $\mathbf{q}_v^{(1)}(S^c) \geq 1/10$ for all $v \in S$. This implies the first statement.

The second statement follows from the fact $q_{vu}^{(m)} \leq 1/D_{\min}$ for any m. □

The local spread facilitates conversion between $\tau_\infty(\mathbf{q}_v; G)$ and the more easily manageable *total variation* mixing time, given by

$$\tau_1(\boldsymbol{\rho}; G) = \min\left\{m : \forall v \in V, \|\boldsymbol{\rho}_v - \boldsymbol{\pi}\|_{TV} \leq 1/4\right\}$$

where

$$\boldsymbol{\rho}_v^{(m)} = \frac{1}{m}\sum_{k=1}^{m+1} \mathbf{q}_v^m, \quad \boldsymbol{\rho}_v = \left(\boldsymbol{\rho}_v^{(1)}, \boldsymbol{\rho}_v^{(2)}, \boldsymbol{\rho}_v^{(3)} \dots\right), \quad \boldsymbol{\rho} = (\boldsymbol{\rho}_v)_{v\in V} \tag{A.12}$$

and $\|\mathbf{p} - \boldsymbol{\pi}\|_{TV} = \sum_{v\in V} |p_v - \pi_v|$ is the total variation norm between distributions $\mathbf{p}$ and $\boldsymbol{\pi}$.

Lemma 13. *For* $\mathbf{q}$ *as in* (A.11) *and* $\boldsymbol{\rho}$ *as in* (A.12),

$$\tau_\infty(\mathbf{q}; G) \leq 2752\tau_1(\boldsymbol{\rho}; G)\log\left(4\max\left\{1, \frac{1}{10s(G)}\right\}\right)$$

Proof. Masking dependence on v for the moment, let

$$\Delta_u^{(m)} = q_{vu}^{(m)} - \pi_u, \quad \delta_u^{(m)} = \frac{\Delta_u^{(m)}}{\pi_u}$$

and $\boldsymbol{\Delta}^{(m)} = (\Delta_u^{(m)})_{u\in V}$, $\boldsymbol{\delta}^{(m)} = (\delta_u^{(m)})_{u\in V}$. For a vector $\boldsymbol{\Delta} = (\Delta_u)_{u\in V}$, the $L^p(\boldsymbol{\pi})$ norm is given by

$$\|\boldsymbol{\Delta}\|_{L^p(\boldsymbol{\pi})} = \left(\sum_{u\in V} (\Delta_u)^p\, \pi_u\right)^{1/p}$$

To go between the $L^\infty(\boldsymbol{\pi})$ and $L^1(\boldsymbol{\pi})$ norms, we have

$$\begin{aligned}
\left\|\boldsymbol{\delta}^{(2m)}\right\|_{L^\infty(\pi)} &\overset{(i)}{\leq} \left\|\boldsymbol{\delta}^{(m)}\right\|_{L^2(\pi)}^2 \\
&= \left\|(\boldsymbol{\delta}^{(m)})^2\right\|_{L^1(\pi)} \\
&\overset{(ii)}{\leq} \left\|(\boldsymbol{\delta}^{(m)})\right\|_{L^1(\pi)} \left\|(\boldsymbol{\delta}^{(m)})\right\|_{L^\infty(\pi)}
\end{aligned}$$

where (i) is a result of Benjamini and Morris and (ii) follows from Holder's inequality. Now, we upper bound the second factor on the right hand side by observing

$$\left\|(\boldsymbol{\delta}^{(m)})\right\|_{L^\infty(\pi)} \leq \max\left\{1, \max_{u\in V} \frac{q_{vu}^{(m)}}{\pi_u}\right\}$$
$$\overset{(iii)}{\leq} \max\left\{1, \frac{1}{s(G)}\right\}$$

where (iii) follows from Lemma 12.

Now, we leverage the following well-known fact (PhD thesis of Montenegro): for any $\epsilon > 0$, if $m \geq \tau_1(\mathbf{q}_v^{(m)}; G) \cdot \log(1/\epsilon)$ then

$$\left\|\mathbf{q}_v^{(m)} - \boldsymbol{\pi}\right\|_{TV} \leq \epsilon.$$

But $\left\|\mathbf{q}_v^{(m)} - \boldsymbol{\pi}\right\|_{TV}$ is exactly $\left\|(\boldsymbol{\delta}^{(m)})\right\|_{L^1(\pi)}$. Therefore, picking

$$m_0 = \tau_1(\mathbf{q}_v^{(m)}; G) \cdot \log\left(4 \max\left\{1, \frac{1}{s(G)}\right\}\right)$$

implies $\left\|(\boldsymbol{\delta}^{(m)})\right\|_{L^\infty(\pi)} \leq 1/4$ for all $m \geq 2m_0$. Then,

$$\left\|(\boldsymbol{\delta}^{(m)})\right\|_{L^\infty(\pi)} = \sup_u \left\{\frac{\left|q_{vu}^{(m)} - \boldsymbol{\pi}_u\right|}{\boldsymbol{\pi}_u}\right\}.$$

and since none of the above depended on a specific choice for v, the supremum can be taken over all seed nodes v as well. Thus $\tau_\infty(\mathbf{q}^{(m)}; G) \leq 2m_0$.

Finally, it is known (PhD thesis of Montenegro) that

$$\tau_1(\mathbf{q}^{(m)}; G) \leq 1376\tau_1(\boldsymbol{\rho}^{(m)}; G)$$

and so the desired result holds.

□

The second key quantity is the *conductance function*

$$\Phi(t; G) := \min_{\substack{S\subseteq V,\\ \boldsymbol{\pi}(S)\leq t}} \Phi(S; G) \qquad (\pi_{\min} \leq t < 1)$$

where $\Phi(S; G)$ is the normalized cut of S in G given by (3).

A.6 Proof of Theorem 2

Give proof structure.

We begin by introducing some more notation. For vertex set $V = \{v_1, \ldots, v_N\}$, take $G = (V, E)$ to be an undirected and unweighted graph, with associated adjacency matrix $\mathbf{A}$, random walk matrix $\mathbf{W}$, and stationary distribution $\boldsymbol{\pi}$. We slightly overload notation and let the *conductance function* be given by

$$\Phi(t; G) := \min_{\substack{S \subseteq V \\ \pi(S) \leq t}} \Phi(S; G)$$

We will pay special attention to the conductance function evaluated over the subgraph $G_{n,r}[\mathcal{C}_\sigma[\mathbf{X}]]$. For ease of notation, we therefore introduce $\widetilde{\Phi}_{n,r}(t) = \Phi(t; G_{n,r}[\mathcal{C}_\sigma[\mathbf{X}]])$.

Lemma 14 is adapted from the PhD thesis of Montenegro. It leverages the conductance function to produce an upper bound on the total variation distance between the random walk

$$\boldsymbol{\rho}^t = \frac{1}{t} \sum_{s=1}^{t} e_v \mathbf{W}^s \tag{A.13}$$

and the stationary distribution $\boldsymbol{\pi}$. Consider $\mathbf{q} = (\mathbf{q}^1, \mathbf{q}^2, \ldots)$ a sequence of distributions over V, where $\mathbf{q}^t = (\mathbf{q}_1^t, \ldots, \mathbf{q}_{|V|}^t)$. The *total variation mixing time* of $\mathbf{q}$ to stationary distribution $\boldsymbol{\pi}$ is given by

$$\tau_1(\mathbf{q}; G) = \min_{t_0 \in N} \left\{ t : \left\| \mathbf{q}^t - \boldsymbol{\pi} \right\|_{TV} \leq 1/4 \right\}$$

where $\|\mathbf{q}^t - \boldsymbol{\pi}\|_{TV} = \sum_{v_i \in V} |\mathbf{q}_i - \boldsymbol{\pi}_i|$ is the total variation distance. Then,

Lemma 14. *Let $\mathbf{W}$ be the random walk matrix over a graph $G = (V, E)$, with stationary distribution $\boldsymbol{\pi}$. Then, for any $v \in V$:*

$$\left\| \rho^t - \boldsymbol{\pi} \right\|_{TV} \leq \max \left\{ \frac{1}{4}, \frac{1}{10} + \frac{70}{t} \left(5 + \int_{x = \pi_1(G)}^{1/2} \frac{4}{x \Phi^2(x; G)} dx \right) \right\}$$

where $\pi_1(G)$ is defined as in (A.2).

To prove Lemma 14 we must first introduce the concept of a blocking conductance function.[1]

Definition 1 (Blocking Conductance Function of PhD thesis of Montenegro)**.** *Let x_0 satisfy*

$$x_0 \geq \min_{v \in V} \boldsymbol{\pi}(v)$$

A function $\phi(x; G) : [x_0, 1/2] \to [0, 1]$ is a blocking conductance function *if for all $S \subset V$ with $\boldsymbol{\pi}(S) = x \in [x_0, 1/2]$, either of the following hold:*

[1] For more details, see PhD thesis of Montenegro

1. Exterior inequality. *For all* $y \in \left[\frac{1}{2}x, x\right] : \phi_{int}(S) \geq \phi(\max\{x_0, y\})$

2. Interior inequality. *For all* $y \in \left[x, \frac{3}{2}x\right] : \phi_{ext}(S) \geq \phi(\max\{y, 1-y\})$.

where ϕ_{int} *and* ϕ_{ext} *are defined respectively as*

$$\phi_{int}(S) = \sup_{\lambda \leq \boldsymbol{\pi}(S)} \min_{\substack{B \subset S \\ \boldsymbol{\pi}(B) \leq \lambda}} \frac{\lambda \text{cut}(S \setminus B, S^c; G)}{\text{vol}(V; G)\left[\boldsymbol{\pi}(S)\boldsymbol{\pi}(S^c)\right]^2}$$

$$\phi_{ext}(S) = \sup_{\lambda \leq \boldsymbol{\pi}(S)} \min_{\substack{B \subset S^c \\ \boldsymbol{\pi}(B) \leq \lambda}} \frac{\lambda \text{cut}(S \setminus B, S^c; G)}{\text{vol}(V; G)\left[\boldsymbol{\pi}(S)\boldsymbol{\pi}(S^c)\right]^2}$$

Proof of Lemma 14. Consider the function $\phi(\cdot, G) : [\pi_1(G), 1/2] \to [0, 1]$ defined by

$$\phi(x; G) = \begin{cases} \frac{1}{5}, & x = \pi_1(G) \\ \frac{1}{4}\Phi^2(x; G), & x \in (\pi_1(G), 1/2] \end{cases} \tag{A.14}$$

In the PhD thesis of Montenegro, letting

$$h^t(x_0) = \sup_{S : \pi(S) < x_0} \left(\boldsymbol{\rho}^t(S) - \boldsymbol{\pi}(S)\right)$$

the following theorem is given: if ϕ is a blocking conductance function (meaning it satisifies Definition 1),

$$\begin{aligned} \left\|\boldsymbol{\rho}^t - \boldsymbol{\pi}\right\|_{TV} &\leq \max\left\{\frac{1}{4}, h^t(x_0) + \frac{70}{t}\left(\frac{1}{\phi(x_0; G)} + \int_{x=x_0}^{1/2} \frac{4}{x\phi(x; G)}\right)\right\} \\ &= \max\left\{\frac{1}{4}, h^t(x_0) + \frac{70}{t}\left(5 + \int_{x=x_0}^{1/2} \frac{4}{x\Phi^2(x; G)}\right)\right\}. \end{aligned}$$

We will wait to the end to show the ϕ is in fact a blocking conductance function, and first show that the above expression reduces to the desired result.

It is also observed in the PhD thesis of Montenegro that the function $h^t(x)$ is decreasing for any fixed x, as t increases. All that remains for us to show is that $h^s(\pi_1(G)) \leq 1/10$ for some $s \leq t$. We choose $s = 1$, and note that for any set $S \subset V$ with $\boldsymbol{\pi}(S) \leq \pi_1(G)$,

$$\boldsymbol{\rho}^1(S) \leq \frac{|S|}{d_{\min}(G)} \leq \frac{\boldsymbol{\pi}(S)\text{vol}(V; G)}{d_{\min}(G)^2} \leq \frac{1}{10},$$

where we use the fact that $\boldsymbol{\pi}(S) \geq \frac{|S| d_{\min}(G)}{\text{vol}(V;G)}$.

Now we complete the proof by verifying that ϕ is in fact a blocking conductance function. It is easy to see that $\frac{1}{4}\Phi^2(x; G)$ is a valid blocking conductance function for all $x \in (\pi_1(G), 1/2]$, (as it satisfies the exterior inequality). For $x = \pi_1(G)$

we will instead turn to the interior inequality, and show the following: for any $S \subset V$ such that $\pi(S) = \pi_1(G)$,

$$\frac{1}{5} \leq \phi_{int}(S).$$

For any S such that $\pi(S) \leq \pi_1(G)$, the following statement holds: for every $u \in S$, $\mathrm{cut}(u, S^c; G) \geq 9/10 \cdot \deg(u; G)$. Fixing $\lambda = \pi(S)/2$, we have

$$\begin{aligned}
\phi_{int}(S) &\geq 4 \min_{\substack{B \subset S \\ \boldsymbol{\pi}(B) \leq \lambda}} \frac{\lambda \mathrm{cut}(S \setminus B, S^c; G)}{4\mathrm{vol}(V; G) \left[\lambda(1-\lambda)\right]^2} \\
&\geq \min_{\substack{B \subset S \\ \boldsymbol{\pi}(B) \leq \lambda}} \frac{9\lambda \sum_{u \in S \setminus B} \deg(u; G)}{40\mathrm{vol}(V; G) \left[\lambda(1-\lambda)\right]^2} \\
&\geq \frac{9\lambda^2}{40[\lambda(1-\lambda)^2]} \geq \frac{1}{5}.
\end{aligned}$$

□

We turn to lower bounding $\widetilde{\Phi}_{n,r}(t)$. First, we exhibit a lower bound on a continuous space analogue, over the set $\mathcal{C}_\sigma$. Let $\nu_{\mathbb{P}}(\cdot)$ denote the weighted volume; formally, for $\mathcal{S} \subset \mathbb{R}^d$

$$\nu_{\mathbb{P}}(\mathcal{S}) := \int_{\mathcal{S}} f(x) dx$$

The r-ball walk over $\mathcal{C}_\sigma$ is a Markov chain with transition probability for $x \in \mathcal{C}_\sigma, \mathcal{S}, \mathcal{S}' \subset \mathcal{C}_\sigma$ given by,

$$P_{\mathbb{P},r}(x; \mathcal{S}) := \frac{\nu_{\mathbb{P}}(\mathcal{S} \cap B(x,r))}{\nu_{\mathbb{P}}(\mathcal{C}_\sigma \cap B(x,r))}, \quad Q_{\mathbb{P},r}(\mathcal{S}, \mathcal{S}') := \int_{x \in \mathcal{S}} f(x) P_{\mathbb{P},r}(x; \mathcal{S}') dx.$$

stationary distribution defined by

$$\ell_{\mathbb{P},r}(x) := \frac{\nu_{\mathbb{P}}(\mathcal{C}_\sigma \cap B(x,r))}{\nu_{\mathbb{P}}(B(x,r))}, \quad \pi_{\mathbb{P},r}(\mathcal{S}) := \frac{1}{\int_{\mathcal{C}_\sigma} f(x) \ell_{\mathbb{P},r}(x) dx} \int_{\mathcal{S}} f(x) \ell_{\mathbb{P},r}(x) dx$$

and corresponding conductance function

$$\widetilde{\Phi}_{\mathbb{P},r}(t) := \min_{\substack{\mathcal{S} \subset \mathcal{C}_\sigma, \\ \pi_{\mathbb{P},r}(\mathcal{S}) \leq t}} \frac{Q_{\mathbb{P},r}(\mathcal{S}, \mathcal{C}_\sigma \setminus \mathcal{S})}{\pi_{\mathbb{P},r}(\mathcal{S})}$$

Lemma 15 (a weighted analogue of Theorem 4.6 of Kannan) provides a lower bound on $\widetilde{\Phi}_{\mathbb{P},r}(t)$, as well as a stepwise approximation to $\widetilde{\Phi}_{\mathbb{P},r}$.

Lemma 15. *Under the conditions on $\mathcal{C}_\sigma$ given by Theorem 2, the following bounds hold:*

- *for* $0 < t < 1/2$,

$$\widetilde{\Phi}_{\mathbb{P},r}(t) > \min\left\{\frac{1}{288\sqrt{d}}, \frac{r}{81\sqrt{d}D}\log\left(1 + \frac{\lambda_\sigma^2}{\Lambda_\sigma^2 t}\right)\right\} \cdot \frac{\lambda_\sigma^4}{\Lambda_\sigma^4}$$

- *Let*

$$m = \frac{2^{d+1}D^d\Lambda_\sigma^2}{r^d\lambda_\sigma^2}$$

and $t_i = (i+1)/m$ *for* $i = 0, \ldots, m-1$. *Then, for* $1/m < t < 1/2$

$$\overline{\Phi}_{\mathbb{P},r}(t) > \min\left\{\frac{1}{288\sqrt{d}}, \frac{r}{162\sqrt{d}D}\mathrm{Log}\left(\frac{\Lambda_\sigma^2}{\lambda_\sigma^2 t}\right)\right\} \cdot \frac{\lambda_\sigma^4}{\Lambda_\sigma^4}$$

where $\overline{\Phi}_{\mathbb{P},r}(t)$ *is defined as in* (A.1) *with respect to* $t_0, \ldots t_{m-1}$, *and* $\mathrm{Log}(A/t) = \max\{\log(1+2A), \log(A/t)\}$.

Proof. We begin by stating directly the result of Kannan et al.: for $0 < t < 1/2$, and any S such that $\pi_{\nu,r}(S) = t$,

$$\frac{Q_{\nu,r}(\mathcal{S}, \mathcal{C}_\sigma \setminus \mathcal{S})}{t} > \min\left\{\frac{1}{288\sqrt{d}}, \frac{r}{81\sqrt{d}D}\log\left(1 + \frac{1}{t}\right)\right\}. \tag{A.15}$$

Now, we note that

$$\pi_{\mathbb{P},r}(S) \leq \pi_{\nu,r}(S) \cdot \frac{\Lambda_\sigma^2}{\lambda_\sigma^2}, \quad Q_{\mathbb{P},r}(\mathcal{S}, \mathcal{C}_\sigma \setminus \mathcal{S}) \geq Q_{\nu,r}(\mathcal{S}, \mathcal{C}_\sigma \setminus \mathcal{S}) \cdot \frac{\lambda_\sigma^2}{\Lambda_\sigma^2}$$

and plugging these estimates in to (A.15) gives the final bound on $\widetilde{\Phi}_{\mathbb{P},r}(t)$. The bound on $\overline{\Phi}_{\mathbb{P},r}(t)$ then follows from $\mathrm{Log}(A/t) \leq \log(1 + 1/t)$ for all $0 < t < 1/2$ and application of Lemma 4. ◻

The introduction of the stepwise approximation allows us to make use of Lemma 16, which gives us (pointwise) consistency of the discrete graph functionals $\widetilde{\Phi}_{n,r}(t)$ to the continuous functionals $\widetilde{\Phi}_{\mathbb{P},r}(t)$.

Lemma 16. *Under the conditions on* $\mathcal{C}_\sigma$ *given by Theorem 2, fix any* $0 < t < 1/2$. *Then the following statement holds: with probability one, as* $n \to \infty$,

$$\liminf_{n\to\infty} \widetilde{\Phi}_{n,r}(t) \geq \widetilde{\Phi}_{\mathbb{P},r}(t)$$

As a consequence, for m *and* $(t_i)_{i=1}^m$ *defined as in Lemma 15, we have that*

$$\liminf_{n\to\infty} \overline{\Phi}_{n,r} \geq \overline{\Phi}_{\mathbb{P},r} \tag{A.16}$$

Lemma 16 stems from Garcia Trillos 16, with a pair of notable distinctions: here, we do not allow the radius r to go to zero, but rather set it to be constant, and in the minimization we have an additional constraint on the measure $\pi(\cdot)$, in both the discrete and continuous functionals. As such, we will need to re-prove a number of the statements of Garcia Trillos 16 to work in this context. We defer this work to Section A.7, and here will only show that (A.16) is implied immediately by the pointwise result.

Proof of (A.16). We take as given that for any $0 < t < 1/2$,

$$\liminf_{n \to \infty} \widetilde{\Phi}_{n,r}(t) \geq \widetilde{\Phi}_{\mathbb{P},r}(t).$$

In particular, this will occur for $t_0, t_1, \ldots, t_m$ and therefore

$$\liminf_{n \to \infty} \overline{\Phi}_{n,r} \geq \overline{\Phi}_{\mathbb{P},r}$$

uniformly over $[1/m, 1/2]$. $\square$

Lemma 16 coupled with Lemma 14 will yield a bound on the total variation mixing time of $\boldsymbol{\rho}^t$. Lemma 17 handles the last step, conversion from the total variation mixing time of $\boldsymbol{\rho}^t$ to the relative pointwise mixing time of $e_v \mathbf{W}^t$.

Lemma 17. *Let $\boldsymbol{\rho} = (\boldsymbol{\rho}^1, \boldsymbol{\rho}^2, \ldots)$ be a sequence of distributions, with $\boldsymbol{\rho}^t$ defined as in* (A.13)*. Then, for $\mathbf{q} = (e_v \mathbf{W}^1, e_v \mathbf{W}^2, \ldots)$*

$$\tau_\infty(\mathbf{q}; G) \leq 2752 \tau_1(\boldsymbol{\rho}; G) \log \left(\frac{4}{\min \{1, 10\pi_1(G)\}} \right)$$

Proof. We introduce a few pieces of notation to simplify the analysis. Let

$$\Delta_k^{(t)} = (e_v \mathbf{W}^t)[k] - \boldsymbol{\pi}[k], \quad \delta_k^{(t)} = \frac{\Delta_k^{(t)}}{\boldsymbol{\pi}[k]}$$

and $\boldsymbol{\Delta}^{(t)} = (\Delta_1^{(t)}, \ldots, \Delta_N^{(t)})$, $\boldsymbol{\delta}^{(t)} = (\delta_1^{(t)}, \ldots, \delta_N^{(t)})$. For a vector $\boldsymbol{\Delta} = (\Delta_1, \ldots, \Delta_N)$, the $L^p(\boldsymbol{\pi})$ norm is given by

$$\|\boldsymbol{\Delta}\|_{L^p(\boldsymbol{\pi})} = \left(\sum_{k=1}^N (\Delta_k)^p \, \pi_k \right)^{1/p}$$

The $L^p(\boldsymbol{\pi})$ norms are relevant because

1. If $t > \tau_\infty(\mathbf{q}; G)$, then $\left\|\boldsymbol{\delta}^{(t)}\right\|_{L^\infty(\pi)} < 1/4$.
2. If $t > \tau_1(\mathbf{q}; G)$, then $\left\|\boldsymbol{\delta}^{(t)}\right\|_{L^1(\pi)} < 1/4$.

To go between the two, we have

$$\begin{aligned}
\left\|\boldsymbol{\delta}^{(2t)}\right\|_{L^\infty(\pi)} &\overset{(i)}{\leq} \left\|\boldsymbol{\delta}^{(t)}\right\|^2_{L^2(\pi)} \\
&= \left\|(\boldsymbol{\delta}^{(t)})^2\right\|_{L^1(\pi)} \\
&\overset{(ii)}{\leq} \left\|(\boldsymbol{\delta}^{(t)})\right\|_{L^1(\pi)} \left\|(\boldsymbol{\delta}^{(t)})\right\|_{L^\infty(\pi)}
\end{aligned}$$

where (i) is a result of Benjamini and Morris and (ii) follows from Holder's inequality. Now, we upper bound the second factor on the right hand side by observing

$$\begin{aligned}
\left\|(\boldsymbol{\delta}^{(t)})\right\|_{L^\infty(\pi)} &\leq \max\left\{1, \max_{k=1,\ldots,N} \frac{(e_v \mathbf{W}^t)[k]}{\boldsymbol{\pi}[k]}\right\} \\
&\overset{(iii)}{\leq} \max\left\{1, \frac{\mathrm{vol}(V;G)}{d_{\min}(G)^2}\right\} \\
&= \max\left\{1, \frac{1}{10\pi_1(G)}\right\}
\end{aligned}$$

where (iii) follows from $\boldsymbol{\pi}[k] \geq d_{\min}(G)/\mathrm{vol}(V;G)$ and $(e_v \mathbf{W}^t) \geq 1/d_{\min}$.

As a result, it is sufficient to exhibit $t > 0$ such that

$$\left\|(\boldsymbol{\delta}^{(t)})\right\|_{L^1(\pi)} \leq \min\left\{1/4, \frac{10\pi_1(G)}{4}\right\}$$

Here, we leverage the following well-known fact (PhD thesis of Montenegro): for any $\epsilon > 0$, if $t \geq \tau_1(e_v \mathbf{W}^t; G) \cdot \log(1/\epsilon)$ then

$$\left\|(\boldsymbol{\delta}^{(t)})\right\|_{L^1(\pi)} \leq \epsilon$$

Therefore, picking

$$t_0 = \tau_1(e_v \mathbf{W}^t; G) \cdot \log\left(\frac{1}{\min\left\{1/4, \frac{10\pi_1(G)}{4}\right\}}\right)$$

implies $\left\|(\boldsymbol{\delta}^{(t)})\right\|_{L^\infty(\pi)} \leq 1/4$ for all $t \geq 2t_0$.

Finally, it is known (PhD thesis of Montenegro) that

$$\tau_1(\mathbf{q}; G) \leq 1376\tau_1(\boldsymbol{\rho}; G)$$

and so the desired result holds.

□

Proof of Theorem 2. Fix arbitrary $v = x_i \in \mathcal{C}_\sigma[\mathbf{X}]$, and let

$$\widetilde{q}_n^{(t)} = e_v \mathbf{W}_{\mathcal{C}_\sigma[\mathbf{X}]}^t, \quad \widetilde{\mathbf{q}}_n = (\widetilde{q}_n^{(1)}, \widetilde{q}_n^{(2)}, \dots)$$

be the distribution of a random walk over the subgraph $G_{n,r}[\mathcal{C}_\sigma[\mathbf{X}]]$ after t. For simplicity, we will refer to the subgraph $G_{n,r}[\mathcal{C}_\sigma[\mathbf{X}]]$ as $\widetilde{G}_{n,r}$ hereafter. Our goal is to upper bound $\tau_\infty(\widetilde{\mathbf{q}}_n; \widetilde{G}_{n,r})$.

By Lemma 17,

$$\tau_\infty(\widetilde{\mathbf{q}}_n; \widetilde{G}_{n,r}) \le 2752 \tau_1(\widetilde{\mathbf{q}}_n; \widetilde{G}_{n,r}) \max\left\{2, \log\left(\frac{4}{10\widetilde{\pi}_{1,n}}\right)\right\}$$

We first upper bound $\tau_1(\widetilde{\mathbf{q}}_n; \widetilde{G}_{n,r})$. From Lemma 14, we have that

$$\limsup_{n\to\infty} \tau_1(\widetilde{\mathbf{q}}_n; \widetilde{G}_{n,r}) \le \limsup_{n\to\infty} \frac{1400}{3}\left(5 + \int_{x=\widetilde{\pi}_{1,n}}^{1/2} \frac{4}{\widetilde{\Phi}_{n,r}^2(x)} dx\right)$$

We set aside the constant term for the moment and turn to the integral. By Lemma 5,

$$\limsup_{n\to\infty} \int_{x=\widetilde{\pi}_{1,n}}^{1/2} \frac{4}{\widetilde{\Phi}_{n,r}^2(x)} \le \limsup_{n\to\infty} \int_{x=\pi_{1,\mathbb{P},r}}^{1/2} \frac{4}{\widetilde{\Phi}_{n,r}^2(x)} dx$$

where $\pi_{1,\mathbb{P},r} = \frac{\lambda_\sigma^2}{\Lambda_\sigma^2}\frac{r^d}{(2D)^d}$. We now replace the discrete conductance function $\widetilde{\Phi}_{n,r}$ by the continuous conductance function $\overline{\Phi}_{n,r}$:

$$\begin{aligned}
\limsup_{n\to\infty} \int_{x=\pi_{1,\mathbb{P},r}}^{1/2} \frac{4}{\widetilde{\Phi}_{n,r}^2(x)} dx &\overset{(i)}{\le} \limsup_{n\to\infty} \int_{x=\pi_{1,\mathbb{P},r}}^{1/2} \frac{4}{\overline{\Phi}_{n,r}^2(x)} dx \\
&= \int_{x=\pi_{1,\mathbb{P},r}}^{1/2} \limsup_{n\to\infty} \frac{4}{\overline{\Phi}_{n,r}^2(x)} dx \\
&\overset{(ii)}{\le} \int_{x=\pi_{1,\mathbb{P},r}}^{1/2} \frac{4}{\overline{\Phi}_{\mathbb{P},r}^2(x)} dx
\end{aligned}$$

where (i) follows from Lemma 4 and (ii) from Lemma 16 (along with the continuous mapping theorem). Now, we make use of the lower bound on the continuous conductance function of Lemma 15:

$$\begin{aligned}
&\int_{x=\pi_{1,\mathbb{P},r}}^{1/2} \frac{4}{\overline{\Phi}_{\mathbb{P},r}^2(x)} dx \le \frac{\Lambda_\sigma^8}{\lambda_\sigma^8} \cdot \left(331776 \int_{x=\pi_{1,\mathbb{P},r}}^{1/2} \frac{d}{x} dx + \int_{x=\pi_{1,\mathbb{P},r}}^{1/2} \frac{81 d D^2}{r^2 x \mathrm{Log}(\frac{\Lambda_\sigma^2}{x\lambda_\sigma^2})}\right) dx \\
&\le \frac{\Lambda_\sigma^8}{\lambda_\sigma^8} \cdot \left(331776 \underbrace{\int_{x=\pi_{1,\mathbb{P},r}}^{1/2} \frac{d}{x} dx}_{:=\mathcal{J}_1} + 81 \underbrace{\int_{x=\pi_{1,\mathbb{P},r}}^{\lambda_\sigma^2/(4\Lambda_\sigma^2)} \frac{dD^2}{r^2 x \log(\frac{\Lambda_\sigma^2}{x\lambda_\sigma^2})}}_{:=\mathcal{J}_2} + 81 \underbrace{\int_{x=\lambda_\sigma^2/(4\Lambda_\sigma^2)}^{1/2} \frac{dD^2}{r^2 x \log(1 + \frac{4\lambda_\sigma^2}{\Lambda_\sigma^2})}}_{:=\mathcal{J}_3}\right)
\end{aligned}$$

□

Simple integration yields the following upper bounds on $\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3$:

$$\begin{aligned}
\mathcal{J}_1 &\leq d^2 \log\left(\frac{2D\Lambda_\sigma^2}{r\lambda_\sigma^2}\right) \\
\mathcal{J}_2 &\leq \frac{dD^2}{r^2}\left[\log(2d) + \log\left(\log\left(\frac{2d}{r}\right)\right)\right] \\
\mathcal{J}_3 &\overset{(iii)}{\leq} 2\frac{dD^2}{r^2}\frac{\Lambda_\sigma^2}{\lambda_\sigma^2}\log\left(4\frac{\Lambda_\sigma^2}{\lambda_\sigma^2}\right)
\end{aligned}$$

where (iii) uses the upper bound $\frac{1}{\log(1+x)} \leq \frac{1}{x}$.

A.7 Proof of pointwise consistency of conductance function.

B OTHER STUFF

Lemma 18 (Bernstein's inequality for U-statistics)**.** *Additionally, assume* $\sigma^2 = \mathrm{Var}\left(k(X_1, \ldots, X_m)\right) < \infty$*. Then for any* $\delta > 0$*,*

$$\mathbb{P}(U - \mathbb{E}U \geq t) \leq \exp\left\{-\frac{n}{2m}\frac{t^2}{\sigma^2 + t/3}\right\},$$

Moreover if $\sigma^2 \leq \mu/n$*,*

$$\begin{aligned}
U &\leq \mathbb{E}U \cdot \left(1 + \max\left\{\sqrt{\frac{2m\log(1/\Delta)}{\mu}}, \frac{2m\log(1/\Delta)}{3\mu}\right\}\right), \\
U &\geq \mathbb{E}U \cdot \left(1 - \max\left\{\sqrt{\frac{2m\log(1/\Delta)}{\mu}}, \frac{2m\log(1/\Delta)}{3\mu}\right\}\right)
\end{aligned}$$

each with probability at least $1 - \Delta$*.*

Multiplicative bound: As $\tilde{k}(x_1, x_2)$ is the sum of two Bernoulli random variables with negative covariance (since $\mathbf{1}(x_i \in \mathcal{A}_{\sigma,\sigma+r})\mathbf{1}(x_j \in B(x_i, r) \cap \mathcal{A}_\sigma) = 1$ implies $\mathbf{1}(x_j \in \mathcal{A}_{\sigma,\sigma+r})\mathbf{1}(x_i \in B(x_j, r) \cap \mathcal{A}_\sigma) = 0$ and vice versa), we can upper bound $\mathrm{Var}\left(\tilde{k}(x_1, x_2)\right) \leq \widetilde{p}$, where we recall

$$\widetilde{p} = 2 \cdot \mathbb{P}\left(\mathbf{1}(x_1 \in \mathcal{A}_{\sigma,\sigma+r})\mathbf{1}(x_2 \in B(x_1, r) \cap \mathcal{A}_\sigma)\right)$$

From Lemma 18, we therefore have

$$\frac{\widetilde{\mathcal{E}}}{\binom{n}{2}} \leq \widetilde{p} + \max\left\{\sqrt{\frac{4\log(1/\Delta)\widetilde{p}}{n}}, \frac{4\log(1/\Delta)}{3n}\right\}$$

with probability at least $1 - \Delta$.

Multiplicative bound: The two terms on the right hand side are both distributed Bernoulli$(p/2)$. Moreover, since $\mathbf{1}(x_i \in A_\sigma) = 1$ implies $\mathbf{1}(x_j \in A_\sigma) = 0$, they have negative covariance. We can therefore upper bound $\mathrm{Var}(k'(x_i, x_j)) \le p$, and so from Lemma 18, we have

$$\frac{\mathcal{V}}{\binom{n}{2}} \ge p - \max\left\{\sqrt{\frac{4\log(1/\Delta)p}{n}}, \frac{4\log(1/\Delta)}{3n}\right\}$$

with probability at least $1 - \Delta$.