Supplementary Material
New Resistance Distances with Global Information on Large Graphs

Abstract

The first part is the proofs of Theorems in Section 3. The second part of this supplementary file shows heat maps of distances and their cluster structures for experiment in Sections 4.

1 Proofs

Theorem 3.3. *For connected ϵ-neighborhood random geometric graphs constructed from a valid region X in R^d (von Luxburg et al., 2014), the global part of $E_1(I_2)$ ($E_1^{global}(I_2)$) dominates the local part ($E_1^{local}(I_2)$) almost surely (for any pair (x_s, x_t)) as $n \to \infty$. Concretely, the following statements hold:*

1. *For unwighted graph $w_{ij} = 1$: $\lim_{n\to\infty} \frac{E_1^{global}}{E_1^{local}} \to \infty$ almost surely as $n \to \infty$ and $\epsilon \to 0$.*
2. *For Euclidean weighted graph with $w_{ij} = d(x_i, x_j)$: $\frac{E_1^{global}}{E_1^{local}} \to \infty$ almost surely as $n \to \infty$ and $\epsilon \to 0$.*
3. *For Gaussian weighted graph with $w_{ij} = \exp(\frac{d(x_i,x_j)^2}{\delta^2})$: $\frac{E_1^{global}}{E_1^{local}} \to \infty$ almost surely as $n \to \infty$, $\epsilon \to 0$ and $O(\delta) > O(\frac{\epsilon}{\sqrt{-\ln(\epsilon)}})$.*

Proof of Theorem 3.3 We work with the assumption that every node is connected to at least one another node.

Case 1: unweighted graph, $w_{ij} = 1 \forall d(x_i, x_j) < \epsilon$.

$O(E_1^{local})$:

$$E_1^{local} = \sum_{i,(s,i)\in E} w_{si}|i_{si}| + \sum_{i,(t,i)\in E} w_{ti}|i_{ti}| = 2 \tag{1}$$

because $I_2 = (i_e)_{e\in E}$ is an unit flow.

$O(E_1^{global})$: We construct a set of parallel hyperplanes $P_1, P_2, \cdots$ that: (1) are orthogonal to the line between s and t, (2) intersect with the line segment between s and t, and (3) of ϵ distance apart from each others as in Figure 1. By this way of construction, any edge of the graph intersects at most one hyperplane in the set. Let E^j denote the set of edges that intersect with plane P_j, then: $E^j \subset E$ and $E^j \cap E^l = \emptyset$ for any two different hyperplanes. Hence, $\cup_j E^j \subset E$, therefore, $E_1(I_2) \geq \sum_j \sum_{e\in E^j} |i_e|$.

Since any of these hyperplanes is an $s - t$ cut of the graph, $\sum_{e\in E^j} |i_e| \geq 1$ because the total flow on E^j would not be less than a min cut, which is of size 1.

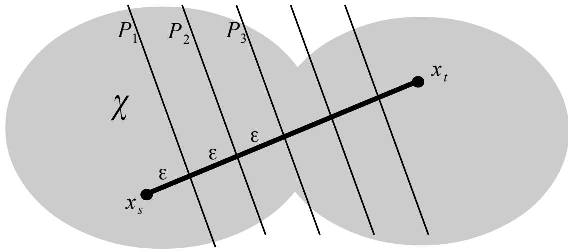

Figure 1: Parallel hyperplanes that partition the set of all edges into disjoint sets of $s-t$ cuts.

The number of such hyperplanes is of order $O(\frac{d(s,t)}{\epsilon})$, then,

$$O(E_1^{global} + E_1^{local}) = O(E_1) \geq O(\frac{d(s,t)}{\epsilon}) = O(\frac{1}{\epsilon}). \tag{2}$$

Hence, from (1) and (2), when $\lim_{n\to\infty} \epsilon = 0$ then $\lim_{n\to\infty} E_1^{global} + E_1^{local} = \infty$ and therefore,

$$\lim_{n\to\infty} \frac{E_1^{global} + E_1^{local}}{E_1^{local}} = \lim_{n\to\infty} \frac{E_1^{global} + 2}{2} = \infty \tag{3}$$

In this case, $E_1^{global} >> E_1^{local}$.

Case 2: Euclidean weighted graph, $w_{ij} = d(x_i, x_j) \forall d(x_i, x_j) < \epsilon$.

$O(E_1^{local})$: $E_1^{local} = \sum_{i,(i,s)\in E} w_{si}|i_{si}| + \sum_{i,(i,t)\in E} w_{ti}|i_{ti}|$. Let $w_{min} = \min_{e\in E} w_e$ be the minimum weight of all edges in the graph. Because $I_2 = (i_e)_{e\in E}$ is an unit flow, therefore

$$2w_{min} \leq E_1^{local} \leq 2\epsilon. \tag{4}$$

$O(E_1^{global})$: Since w_{ij} is Euclidean distance, triangle inequality applies. Therefore, all paths from x_s to x_t, including the shortest path of length $sp(x_s, x_t)$, are not shorter than $d(x_s, x_t)$.

Since $E_1(I_2(x_s, x_t)) \geq sp(x_s, x_t)$,

$$E_1(I_2(x_s, x_t)) = E_1^{global}(I_2(x_s, x_t)) + E_1^{local}(I_2(x_s, x_t)) \geq d(x_s, x_t)(= O(1)). \tag{5}$$

Hence, from (4) and (5), as long as $\epsilon \to 0$, with probability 1:

$$\lim_{n\to\infty} \frac{E_1^{global}}{E_1^{local}} \geq \lim_{n\to\infty} \frac{d(x_s, x_t) - 2\epsilon}{2\epsilon} = \infty. \tag{6}$$

Case 3: For Gaussian weighted graph, $w_{ij} = \exp(\frac{d(x_i,x_j)^2}{\delta^2})$ being the distance corresponding to a similarity graph (equivalent to similarity between x_i and x_j being $\exp(\frac{-d(x_i,x_j)^2}{\delta^2})$).

$O(E_1^{local})$: In ϵ-neighborhood graph, $\forall (i,j) \in E, d(x_i, x_j) < \epsilon$, therefore $w_{si}, w_{tj} < \exp(\frac{\epsilon^2}{\delta^2})$. Then, for unit flow $(i_e)_e = I_2(x_s, x_t)$,

$$\begin{aligned} E_1^{local} &= \sum_{i,(i,s)\in E} w_{si}|i_{si}| + \sum_{i,(i,t)\in E} w_{ti}|i_{ti}| \\ &< \sum_{i,(i,s)\in E} \exp(\frac{\epsilon^2}{\delta^2})|i_{si}| + \sum_{i,(i,t)\in E} \exp(\frac{\epsilon^2}{\delta^2})|i_{ti}| \\ &= 2\exp(\frac{\epsilon^2}{\delta^2}). \end{aligned} \tag{7}$$

$O(E_1^{global})$: $\forall (i,j) \in E, d(x_i, x_j), w_{ij} = \exp(\frac{d(x_i,x_j)^2}{\delta^2}) \geq 1$. Therefore,

$$E_1^{global} \geq sp(x_s, x_t) - E_1^{local} \geq \frac{d(x_s, x_t)}{\epsilon} - E_1^{local}. \tag{8}$$

From (7) and (8), as $\lim_{n\to\infty}$, if $\epsilon \to 0$ and $O(\delta) > O(\frac{\epsilon}{\sqrt{-\ln(\epsilon)}})$, we have:

$$\frac{E_1^{global}}{E_1^{local}} \geq \frac{\frac{1}{\epsilon}}{2\exp(\frac{\epsilon^2}{\delta^2})} - 1.$$

$$\begin{aligned}\ln(\frac{\frac{1}{\epsilon}}{2\exp(\frac{\epsilon^2}{\delta^2})}) &= \ln\frac{1}{\epsilon} - \ln(2\exp(\frac{\epsilon^2}{\delta^2})) \\ &= -\ln(\epsilon) - \ln(2) - \frac{\epsilon^2}{\delta^2}\end{aligned}$$

Since $O(\delta) > O(\frac{\epsilon}{\sqrt{-\ln(\epsilon)}})$, $O(\delta^2) > O(\frac{\epsilon^2}{-\ln(\epsilon)})$ and $O(-\ln(\epsilon)) > O(\frac{\epsilon^2}{\delta^2})$.
As $\lim_{n\to\infty}\epsilon = 0$, $\lim_{n\to\infty} -\ln(\epsilon) = \infty$, therefore $\lim_{n\to\infty} -\ln(\epsilon) - \ln(2) - \frac{\epsilon^2}{\delta^2} = \infty$. Hence,

$$\lim_{n\to\infty}\frac{E_1^{global}}{E_1^{local}} = \infty. \tag{9}$$

Note that, it is advisable to set δ such that w_{ij} are not too small or too large to avoid computing issues. One of the common practice is to have $O(\delta) = O(\epsilon)$, which satisfies the condition of $O(\delta) > O(\frac{\epsilon}{\sqrt{-\ln(\epsilon)}})$ when $\epsilon \to 0$.

Theorem 3.4. *For connected k-nearest neighbor (random geometric) graphs constructed from a valid region X in R^d (von Luxburg et al., 2014), the global part of $E_1(I_2)$ dominates the local part almost surely as $n \to \infty$. Concretely, there exist constants c_1, c_2 that the following statements hold:*

1. *For unwighted graph $w_{ij} = 1$: $\lim_{n\to\infty} \frac{E_1^{global}}{E_1^{local}} \to \infty$ almost surely as $n \to \infty$, $k > \log(n)$ and $\frac{k}{n} \to 0$ with a probability of at least $1 - c_1 n \exp(-c_2\sqrt{nk})$ (converging to 1).*
2. *For Euclidean weighted graph with $w_{ij} = d(x_i, x_j)$: $\frac{E_1^{global}}{E_1^{local}} \to \infty$ almost surely as $n \to \infty$, $k > \log(n)$ and $\frac{k}{n} \to 0$ with a probability of at least $1 - c_1 n \exp(-c_2\sqrt{nk})$ (converging to 1).*
3. *For Gaussian weighted graph with $w_{ij} = \exp(\frac{d(x_i,x_j)^2}{\delta^2})$: $\frac{E_1^{global}}{E_1^{local}} \to \infty$ almost surely as $n \to \infty$, $k > \log(n)$, $\frac{k}{n} \to 0$ and $O(\delta) = (\frac{k}{n})^{\frac{1}{d}}$ with a probability of at least $1 - c_1 n \exp(-c_2 k \cdot \log(\frac{n}{k})^{\frac{d}{2}})$ (converging to 1).*

Recall the definition of k-nearest neighbor radii: $R_k(x) = \max_{i,(i,s)\in E} d(x_s, x_i)$ be the distance of x to its k-nearest neighbor in X. Let $B(x, \eta)$ be the ball centered at x with radius η. Hence, there are only k points from the sampled n points lying in $B(x, R_k(x))$. Let p_{min} and p_{max} be the minimum and maximum probability density in p.

We first prove a lemma that for a fixed point x, $R_k(x) \to 0$ with a high probability.

Lemma 1.1. *For any fixed node $x \in X$ in a random geometric knn graph, any ϵ_0 as a function of n satisfying $O((\frac{k}{n})^{\frac{1}{d}}) < O(\epsilon_0) < O(1)$, then $O(R_k(x)) < O(\epsilon_0) \to 0$ as $n \to \infty$ with a probability at least $1 - c_1 \exp(\frac{-n\gamma(\epsilon_0)}{2})$, converging to 1 when $n, k \to \infty$ and $\frac{k}{n} \to 0$ for some constant c_1.*

Proof. In this case, we want ϵ_0 to play the role of ϵ in Theorem 3.3. The difference is that, in this case, ϵ_0 neighborhoods contain all the neighbors of all points with a high enough probability (converging to 1), as opposed to the case of ϵ neighborhoods that contain all neighbors of all points with probability 1.

The volume of $B(x, \epsilon_0)$ is $c\epsilon_0^d$ for constant $c = \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}$. Let $\gamma(\epsilon_0) = \int_{B(x,\epsilon_0)} p(x)dx$ denote the probability mass of $B(x, \epsilon_0)$, then

$$p_{min} c\epsilon_0^d \leq \gamma(\epsilon_0) \leq p_{max} c\epsilon_0^d, \tag{10}$$

or $O(\gamma(\epsilon_0)) = O(\epsilon_0^d)$.

We prove that $R_k(x) < \epsilon_0$ with a probability converging to 1. This is equivalent to proving that $R_k(x) \geq \epsilon_0$ with a probability converging to 0. The probability of $R_k(x) \geq \epsilon_0$ is the probability that there are less than k points in X (with n points) lying inside of $B(x, \epsilon_0)$ (with probability mass of $\gamma(\epsilon_0)$).

The number of points lying in $B(x, \epsilon_0)$ follows a binomial distribution $B(n, \gamma(\epsilon_0))$. Let $F(k, n, \gamma(\epsilon_0))$ be the cumulative distribution function of $B(n, \gamma(\epsilon_0))$. Let $P(R_k(x) \geq \epsilon_0)$ be the probability that $R_k(x) \geq \epsilon_0$, then, $P(R_k(x) \geq \epsilon_0) = F(k-1, n, \gamma(\epsilon_0)) < F(k, n, \gamma(\epsilon_0))$ (we prove k for simplicity). Chernoff's inequality for binomial distribution gives

$$F(k, n, \gamma(\epsilon_0)) \leq \exp(\frac{-(n\gamma(\epsilon_0) - k)^2}{2n\gamma(\epsilon_0)})$$

As $n, k \to \infty$, $\frac{k}{n} \to 0$, $O(n\gamma(\epsilon_0)) > O(k)$ by the way we choose ϵ_0, hence,

$$O(\exp(\frac{-(n\gamma(\epsilon_0) - k)^2}{2n\gamma(\epsilon_0)})) = O(\exp(\frac{-n\gamma(\epsilon_0)}{2})).$$

Because $n\gamma(\epsilon_0) > k \to \infty$, $\exp(\frac{-n\gamma(\epsilon_0)}{2}) \to 0$. Hence, there exists a constant c_1 that $R_k(x) < \epsilon_0 \to 0$ with a probability $1 - F(k, n, \gamma(\epsilon_0))$, which is at least $1 - c_1 \exp(\frac{-n\gamma(\epsilon_0)}{2})$. □

Lemma 1.2. *For all nodes in X in a random geometric knn graph, $O(R_k(x)) < O(\epsilon_0) \to 0 \forall i = 1 \cdots n$ as $n \to \infty$ with a probability at least $1 - c_1 n \exp(-c_2 n \epsilon_0^d)$ for some constant c_1, c_2, converging to 1 when $n, k \to \infty$, $\frac{k}{n} \to 0$ and $O(\epsilon_0) > (\frac{\log(n)}{n})^{\frac{1}{d}}$. If $k > \log(n)$ then the last condition is already included in the choice of ϵ_0.*

Proof. The Lemma 1.1 shows that for any fixed $x_i \in X$, $P(R_k(x_i) \geq \epsilon_0) \leq c_1 \exp(\frac{-n\gamma(\epsilon_0)}{2})$. Therefore, from 10 the probability that there exists at least one x_i such that $R_k(x_i) \geq \epsilon_0 \forall i = 1 \cdots n$ satisfies (for all i together)

$$P(R_k(x_i) \geq \epsilon_0) \leq c_1 n \exp(\frac{-n\gamma(\epsilon_0)}{2}) \leq c_1 n \exp(-c_2 n \epsilon_0^d) \tag{11}$$

for some constant c_2. Hence, the probability that $R_k(x_i) < \epsilon_0$ for all $i = 1 \cdots n$, as $n, k \to \infty$ and $\frac{k}{n} \to 0$, satisfies

$$P(R_k(x_i) < \epsilon_0) \geq 1 - c_1 n \exp(-c_2 n \epsilon_0^d) \forall i. \tag{12}$$

Now we show that $c_1 n \exp(-c_2 n \epsilon_0^d) \to 0$ as $n, k \to \infty$ and $\frac{k}{n} \to 0$ and $O(\epsilon_0) > O(\frac{\log(n)}{n})^{\frac{1}{d}}$. Since $O(\epsilon_0^d) > O(\frac{\log(n)}{n})$, $O(-c_2 n \epsilon_0^d) > O(\log(n))$, therefore,

$$c_1 n \exp(-c_2 n \epsilon_0^d) = c_1 \exp(\log(n) - c_2 n \epsilon_0^d) \to c_1 \exp(-\infty) = 0.$$

As $O(\epsilon_0) > O((\frac{k}{n})^{\frac{1}{d}})$ by choice and $k > \log(n)$, then the last condition in the lemma is already implied.

□

Proof of Theorem 3.4

In all cases, we choose different ϵ_0 to make the formulations simple and intuitive, even though a range of ϵ_0 would work.

Case 1: unweighted graph, $w_{ij} = 1$ for x_j is one of the k-nearest neighbors of x_i.

$O(E_1^{local})$: For unit flow $I_2 = (i_e)_{e \in E}$:

$$E_1^{local} = \sum_{i,(i,s) \in E} w_{si} |i_{si}| + \sum_{j,(j,t) \in E} w_{tj} |i_{tj}| = 2. \tag{13}$$

$O(E_1^{global})$: Since $E_1(I_2(x_s, x_t)) \geq sp(x_s, x_t)$, we prove that $O(sp(x_s, x_t)) > O(1)$ with a probability converging to 1 as $n \to \infty$.

In this case, we choose $\epsilon_0 = (\frac{k}{n})^{\frac{1}{2d}}$. From Lemma 1.2, with a probability at least $1 - c_1 n \exp(-c_2 \sqrt{nk})$, all $R_k(x) < \epsilon_0$, or equivalently, $d(x_i, x_j) < \epsilon_0 \ \forall (i, j) \in E$. Therefore, $O(sp(x_s, x_t)) \geq O(\frac{d(x_s, x_t)}{\epsilon_0})$. Since $O(\frac{d(x_s, x_t)}{\epsilon_0}) = O((\frac{n}{k})^{\frac{1}{2d}}) \to \infty$. Hence, with a probability of at least $1 - c_1 n \exp(-c_2 \sqrt{nk})$, as $n \to \infty$, $k > \log(n)$ and $\frac{k}{n} \to 0$,

$$E_1(I_2(x_s, x_t)) \to \infty \tag{14}$$

Therefore, from (20) and (14) with a probability of at least $1 - c_1 n \exp(-c_2 \sqrt{nk})$ (converging to 1), as $n \to \infty$, $k \to \infty$ and $\frac{k}{n} \to 0$,

$$\frac{E_1^{global}}{E_1^{local}} \to \infty. \tag{15}$$

Case 2: Euclidean weighted graph, $w_{ij} = d(x_i, x_j)$ for x_j is one of the k-nearest neighbors of x_i. We also choose $\epsilon_0 = (\frac{k}{n})^{\frac{1}{2d}}$ as previous case.

$O(E_1^{local})$: For unit flow $I_2 = (i_e)_{e\in E}$:

$$E_1^{local}(I_2(x_s,x_t)) = \sum_{i,(i,s)\in E} w_{si}|i_{si}| + \sum_{i,(i,t)\in E} w_{ti}|i_{ti}| \leq \max_{i,(i,s)\in E} d(x_s,x_i) + \max_{j,(j,t)\in E} d(x_t,x_j). \quad (16)$$

As $n\to\infty$, $k > log(n)$ and $\frac{k}{n}\to 0$, according to Lemma 1.2, $c_1 n\exp(-c_2\sqrt{nk})\to 0$. Hence, with a probability not smaller than $1 - c_1 n\exp(-c_2\sqrt{nk})$, $R_k(x) < \epsilon_0 \forall x\in X$ or $d(x_i,x_j) < \epsilon_0 \forall (i,j)\in E$. From (16),

$$E_1^{local} < 2\epsilon_0 \to 0. \quad (17)$$

$O(E_1^{global})$:

$$E_1^{local}(I_2(x_s,x_t)) + E_1^{global}(I_2(x_s,x_t)) = E_1(I_2(x_s,x_t)) \geq sp(x_s,x_t) \geq d(x_s,x_t) \in O(1). \quad (18)$$

From (17) and (18), we have, with a probability at least $1 - c_1 n\exp(-c_2\sqrt{nk})$ (converging to 1) as $n\to\infty$, $k > log(n)$ and $\frac{k}{n}\to 0$,

$$\frac{E_1^{global}}{E_1^{local}} \to \infty. \quad (19)$$

Case 3: Gaussian weighted graph, $w_{ij} = \exp(\frac{d(x_i,x_j)^2}{\delta^2})$ for x_j is one of the k-nearest neighbors of x_i. Since $d_(x_i,x_j) \in O((\frac{k}{n})^{\frac{1}{d}})$ for most of (x_i,x_j) pair, it is necessary to choose $O(\delta) = O((\frac{k}{n})^{\frac{1}{d}})$ so that weights of most edges in the graph are of constant range, not going to ∞ nor 0.

We define some notations for simpler formulation. Let $t = \frac{n}{k}$, hence, $t\to\infty$.

We choose $\epsilon_0 = (\frac{k}{n})^{\frac{1}{d}} \cdot \sqrt{\frac{\log(t)}{d+1}}$ in this case to bound E_1^{local} and E_1^{global}. In fact, we just need $O(\epsilon_0) > O((\frac{k}{n})^{\frac{1}{d}})$ and still small enough according to some complicated formula. In this case, we have to choose smaller ϵ_0 compared to previous cases just to show that global energy dominates local one.

With a probability of at least $1 - c_1 n\exp(-c_2 n\epsilon_0^d)$, then all $R_k(x) < \epsilon_0$, meaning that $d(x_i,x_j) < \exp(\frac{\epsilon^2}{\delta^2})$. In this case, we bound E_1^{local} and E_1^{global} as follows.

$O(E_1^{local})$: For unit flow $I_2 = (i_e)_{e\in E}$:

$$\begin{aligned} E_1^{local}(I_2(x_s,x_t)) &= \sum_{i,(i,s)\in E} w_{si}|i_{si}| + \sum_{i,(i,t)\in E} w_{ti}|i_{ti}| \\ &\leq \max_{i,(i,s)\in E} \exp(\frac{d(x_s,x_i)^2}{\delta^2}) + \max_{j,(j,t)\in E} \exp(\frac{d(x_t,x_j)^2}{\delta^2}) \\ &\leq 2\exp(\frac{\epsilon_0^2}{\delta^2}). \end{aligned} \quad (20)$$

Therefore,

$$\begin{aligned} O(E_1^{local}) &\leq O(\exp(\frac{\epsilon_0^2}{\delta^2})) \\ &= O(\exp(\sqrt{\frac{\log(t)}{d+1}}^2)) \\ &= O(t^{\frac{1}{d+1}}). \end{aligned} \quad (21)$$

$O(E_1^{global})$: By the definition of Gaussian weighted graph, $w_i j > 1 \forall (i,j)\in E$. In case that $R_k(x) < \epsilon_0$, meaning that $d(x_i,x_j) < \epsilon_0 \forall (i,j)\in E$ (all edges are of length less than ϵ_0), the number of edges on any path between x_s and x_t must not smaller than $\frac{d(x_s,x_t)}{\epsilon_0}$. Hence,

$$E_1^{global} + E_1^{local} \geq sp(x_s,x_t) \geq \frac{d(x_s,x_t)}{\epsilon_0} \in O(\frac{1}{\epsilon}). \quad (22)$$

Therefore,

$$
\begin{aligned}
O(E_1^{global} + E_1^{local}) &\geq O(\frac{1}{\epsilon}) \\
&= O(\frac{t^{\frac{1}{d}}}{\sqrt{\frac{\log(t)}{d+1}}}) \\
&= O(\frac{t^{\frac{1}{d}}}{\sqrt{\log(t)}}) \\
&= O(t^{\frac{1}{d}} \cdot \log(t)^{\frac{-1}{2}}). \qquad (23)
\end{aligned}
$$

Putting (21) and (23) together, we have

$$
\begin{aligned}
O(\frac{E_1^{global}}{E_1^{local}}) &= O(\frac{E_1^{global} + E_1^{local}}{E_1^{local}} - 1) \\
&\geq O(\frac{t^{\frac{1}{d}} \cdot \log(t)^{\frac{-1}{2}}}{t^{\frac{1}{d+1}}} - 1) \\
&= O(\frac{t^{\frac{1}{d(d+1)}}}{\log(t)^{\frac{1}{2}}} - 1) \\
&= \infty \qquad (24)
\end{aligned}
$$

because $t^{\frac{1}{d(d+1)}} >> \log(t)^{\frac{1}{2}}$ as $t^{\alpha} >> \log(t) \forall \alpha > 0, t \rightarrow \infty$. Hence, $E_1^{global} >> E_1^{local}$ with a probability of at least $1 - c_1 n \exp(-c_2 n \epsilon_0^d)$. Replacing ϵ_0, updating constant c_2, we can have the probability as $1 - c_1 n \exp(-c_2 k \cdot \log(\frac{n}{k})^{\frac{d}{2}})$. According to Lemma 1.2, this probability also converges to 1.

Lemma 3.5. *Let $V^{(i)} = L^{-1}_{.i}$ denote the i-th column of L^{-1}, respectively. Then,*

$$V^{(s)} - V^{(t)} = L^{-1}(e_s - e_t) \tag{25}$$

is a possible potential assignment to the nodes of the network that makes the unit flow from x_s to x_t on the network.

Proof. First, let $V^{'} \in R^n$ be the potential on nodes of the graph with *unit potential difference* between x_s and x_t, namely

$$V'_s - V'_t = V'^T(e_s - e_t) = 1.$$

Kirchhoff's voltage law: the potential assignment in the network minimizes the energy function $V'^T LV'$:

$$V' = \arg\min_{x \in R^n} x^T Lx, \ \ s.t. \ x^t(e_s - e_t) = 1. \tag{26}$$

Lagrange multipliers method gives us

$$V' = \frac{L^{-1}(e_s - e_t)}{(e_s - e_t)^T L^{-1}(e_s - e_t)} + \alpha\mathbf{1}$$

with $\mathbf{1}$ is the vector of all 1 in R^n and any $\alpha \in R$, and the energy

$$E' = \frac{1}{(e_s - e_t)^T L^{-1}(e_s - e_t)}.$$

Using Ohm's law $E' = V'I' = I'^2 R_{st}$, then V' makes the total flow from x_s to x_t of I' as:

$$I' = \frac{E'}{V'} = \frac{1}{(e_s - e_t)^T L^{-1}(e_s - e_t)}.$$

Second, showing $V^{(s)} - V^{(t)}$ is a potential assignment on the graph to make an unit flow from x_s to x_t by rescaling V' (also I'). To have an unit flow ($I = 1$) from x_s to x_t then voltage arrangement in the network, ignoring constant terms , can be

$$V' \cdot \frac{1}{I'} = L^{-1}(e_s - e_t) = V^{(s)} - V^{(t)}. \tag{27}$$

This means that $V^{(s)} - V^{(t)}$ is an valid potential assignment to the nodes of the network that makes an unit flow from x_s to x_t, resulting in the flow $I_2(x_s, x_t)$ on the graph. This gives us the embeddings of graph into *edge space.* □

Theorem 3.6. *The following embedding f of the nodes of graph G into an L^p space:*

$$
\begin{aligned}
&f : X \to R^{|E|} \\
&\quad x_s \to f(x_s) = \{\cdots, \frac{V_i^{(s)} - V_j^{(s)}}{r_{ij}^{(p-1)/p}}, \cdots\}^T_{(i,j)\in E}
\end{aligned}
\tag{28}
$$

makes the p-norm of the space coincide with R_p: $\|f(x_s) - f(x_t)\|_p = R_p(x_s, x_t)$.

Proof. We prove the theorem by explicitly constructing the embedding of the nodes in an L^p space. For simplicity, denote $V = V^{(s)} - V^{(t)}$ as the potential arrangement for an unit flow from s to t. We rewrite $R_p^p = E_p(I_2)$ in potential form using Lemma 3.4:

$$
\begin{aligned}
&R_p^p(x_s, x_t) \\
&= \sum_{e=(i,j)\in E} r_e |I_2(x_s, x_t)_e|^p \\
&= \sum_{e=(i,j)\in E} r_e \frac{|V_i - V_j|^p}{r_e^p} \\
&= \sum_{e=(i,j)\in E} r_e^{1-p} \cdot |(L_{is}^{-1} - L_{it}^{-1}) - (L_{js}^{-1} - L_{jt}^{-1})|^p \\
&= \sum_{e=(i,j)\in E} |\frac{L_{is}^{-1} - L_{js}^{-1}}{r_{ij}^{(p-1)/p}} - \frac{L_{it}^{-1} - L_{jt}^{-1}}{r_{ij}^{(p-1)/p}}|^p \\
&= \sum_{e=(i,j)\in E} |\frac{V_i^{(s)} - V_j^{(s)}}{r_{ij}^{(p-1)/p}} - \frac{V_i^{(t)} - V_t^{(t)}}{r_{ij}^{(p-1)/p}}|^p
\end{aligned}
\tag{29}
$$

We could see that R_p is in the form of p-norm on an m dimensional space. There is a natural embedding f of the nodes as follows, for any s:

$$
\begin{aligned}
&f : X \to R^{|E|} \\
&\quad x_s \to f(x_s) = \{\cdots, \frac{V_i^{(s)} - V_j^{(s)}}{r_{ij}^{(p-1)/p}}, \cdots\}^T_{(i,j)\in E}.
\end{aligned}
\tag{30}
$$

Endowing this $R^{|E|}$ space with p-norm $\|\cdot\|_p$, then using (29)

$$R_p(x_s, x_t) = \|f(x_s) - f(x_t)\|_p$$

Hence, R_p distance induces an embedding of nodes into the L_p space $(R^{|E|}, \|\cdot\|_p)$. We call f an *edge space embedding* as each dimension of the embedding space (of nodes) corresponds to an edge of the graph. This makes R_p a metric induced by the p norm. □

2 Heat Maps of Distances

For subsection 4.2: Data size effect

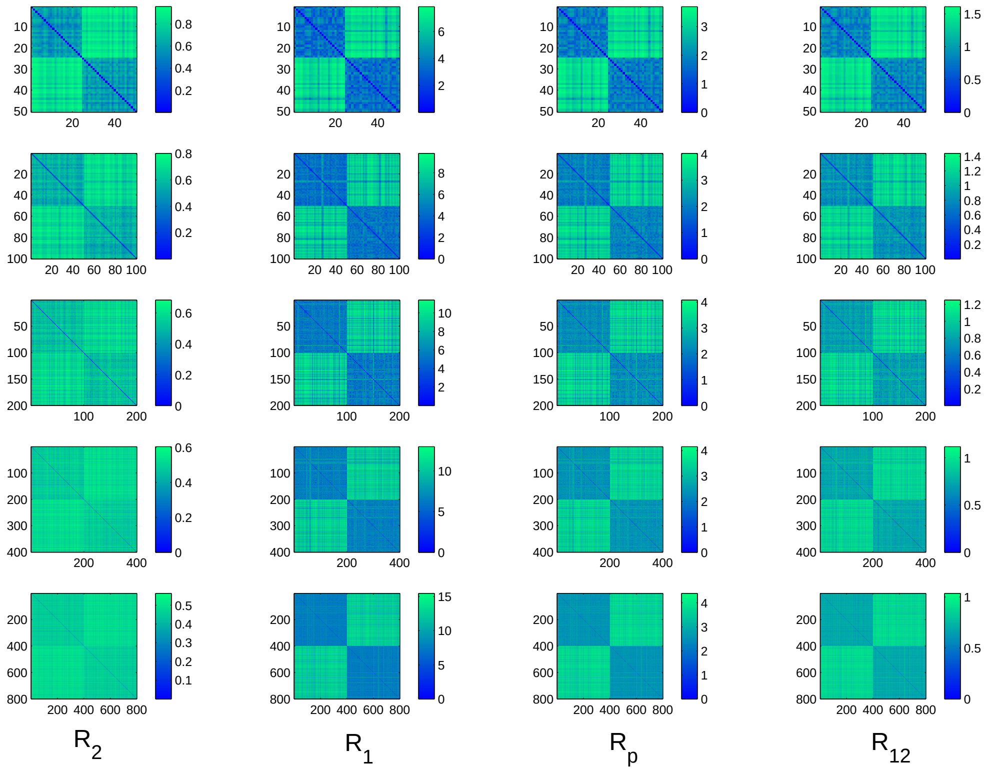

Figure 2: Heat maps of pairwise distances. The four columns are for the four distances. The rows are for our generated data sets of different sizes ranging from 50 to 800.

We could observe that for (the square root of) resistance distances (R_2), the two-blocked structures were clearly observed in small sized data sets. As the data sizes became larger, the structures disappeared and became totally unrecognizable in the last row. On the other hand, the other distances (R_1, R_p and R_{12}) showed consistently the two-blocked structures for all data sizes. This meant that the resistance distance suffered from global information loss problem in large random geometric graphs. Our proposed distances could retain global information and showed cluster structures clearly.

For subection 4.3: Dimensional effect

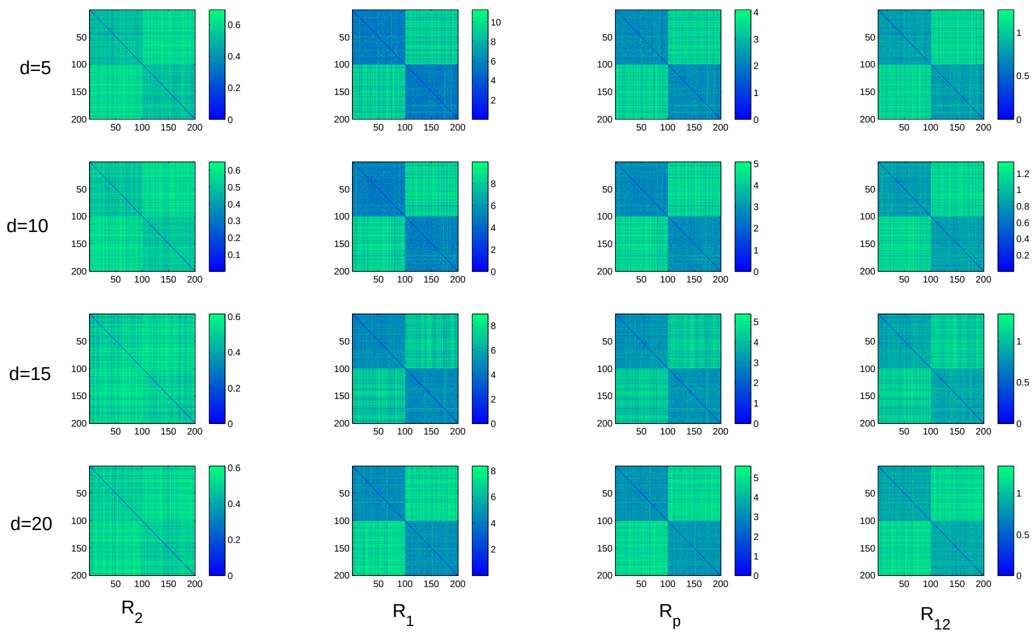

Figure 3: The effect of space dimension in the distances. The rows are for different dimensions of the space $d = 5, 10, 15$ and 20. As the dimension becomes larger, the resistance distance could not recognize clusters. Our proposed distances are robust to the dimensions of the spaces.

We could observe similar behavior. As the dimension of the space increased, resistance distance failed to show two-block structure due to the global information loss problem. Our proposed distances (R_1, R_2, and R_{12}) could still show two-block structures. It meant that our proposed distances could overcome the global information loss problem.