## 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 $\epsilon$-neighborhood random geometric graphs constructed from a valid region $X$ in $R^{d}$ (von Luxburg et al., 2014), the global part of $E_{1}\left(I_{2}\right)\left(E_{1}^{g l o b a l}\left(I_{2}\right)\right)$ dominates the local part $\left(E_{1}^{\text {local }}\left(I_{2}\right)\right)$ almost surely (for any pair $\left.\left(x_{s}, x_{t}\right)\right)$ as $n \rightarrow \infty$. Concretely, the following statements hold:

1. For unwighted graph $w_{i j}=1: \lim _{n \rightarrow \infty} \frac{E_{1}^{\text {global }}}{E_{1}^{\text {local }}} \rightarrow \infty$ almost surely as $n \rightarrow \infty$ and $\epsilon \rightarrow 0$.
2. For Euclidean weighted graph with $w_{i j}=d\left(x_{i}, x_{j}\right): \frac{E_{1}^{\text {glooal }}}{E_{1}^{\text {local }}} \rightarrow \infty$ almost surely as $n \rightarrow \infty$ and $\epsilon \rightarrow 0$.
3. For Gaussian weighted graph with $w_{i j}=\exp \left(\frac{d\left(x_{i}, x_{j}\right)^{2}}{\delta^{2}}\right): \frac{E_{1}^{\text {global }}}{E_{1}^{\text {local }}} \rightarrow \infty$ almost surely as $n \rightarrow \infty$, $\epsilon \rightarrow 0$ and $O(\delta)>O\left(\frac{\epsilon}{\sqrt{-\ln (\epsilon)}}\right)$.
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, $w_{i j}=1 \forall d\left(x_{i}, x_{j}\right)<\epsilon$.
$O\left(E_{1}^{\text {local }}\right)$ :

$$
\begin{equation*}
E_{1}^{l o c a l}=\sum_{i,(s, i) \in E} w_{s i}\left|i_{s i}\right|+\sum_{i,(t, i) \in E} w_{t i}\left|i_{t i}\right|=2 \tag{1}
\end{equation*}
$$

because $I_{2}=\left(i_{e}\right)_{e \in E}$ is an unit flow.
$O\left(E_{1}^{\text {global }}\right)$ : 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 $\epsilon$ 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, $E_{1}\left(I_{2}\right) \geq \sum_{j} \sum_{e \in E^{j}}\left|i_{e}\right|$.

Since any of these hyperplanes is an $s-t$ cut of the graph, $\sum_{e \in E^{j}}\left|i_{e}\right| \geq 1$ 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 $O\left(\frac{d(s, t)}{\epsilon}\right)$, then,

$$
\begin{equation*}
O\left(E_{1}^{\text {global }}+E_{1}^{\text {local }}\right)=O\left(E_{1}\right) \geq O\left(\frac{d(s, t)}{\epsilon}\right)=O\left(\frac{1}{\epsilon}\right) . \tag{2}
\end{equation*}
$$

Hence, from (1) and (2), when $\lim _{n \rightarrow \infty} \epsilon=0$ then $\lim _{n \rightarrow \infty} E_{1}^{\text {global }}+E_{1}^{\text {local }}=\infty$ and therefore,

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{E_{1}^{\text {global }}+E_{1}^{\text {local }}}{E_{1}^{\text {local }}}=\lim _{n \rightarrow \infty} \frac{E_{1}^{\text {global }}+2}{2}=\infty \tag{3}
\end{equation*}
$$

In this case, $E_{1}^{g l o b a l} \gg E_{1}^{\text {local }}$.
Case 2: Euclidean weighted graph, $w_{i j}=d\left(x_{i}, x_{j}\right) \forall d\left(x_{i}, x_{j}\right)<\epsilon$.
$O\left(E_{1}^{\text {local }}\right): E_{1}^{\text {local }}=\sum_{i,(i, s) \in E} w_{s i}\left|i_{s i}\right|+\sum_{i,(i, t) \in E} w_{t i}\left|i_{t i}\right|$. Let $w_{\text {min }}=\min _{e \in E} w_{e}$ be the minimum weight of all edges in the graph. Because $I_{2}=\left(i_{e}\right)_{e \in E}$ is an unit flow, therefore

$$
\begin{equation*}
2 w_{\min } \leq E_{1}^{l o c a l} \leq 2 \epsilon \tag{4}
\end{equation*}
$$

$O\left(E_{1}^{g l o b a l}\right)$ : Since $w_{i j}$ is Euclidean distance, triangle inequality applies. Therefore, all paths from $x_{s}$ to $x_{t}$, including the shortest path of length $s p\left(x_{s}, x_{t}\right)$, are not shorter than $d\left(x_{s}, x_{t}\right)$.

Since $E_{1}\left(I_{2}\left(x_{s}, x_{t}\right)\right) \geq s p\left(x_{s}, x_{t}\right)$,

$$
\begin{equation*}
E_{1}\left(I_{2}\left(x_{s}, x_{t}\right)\right)=E_{1}^{\text {global }}\left(I_{2}\left(x_{s}, x_{t}\right)\right)+E_{1}^{l o c a l}\left(I_{2}\left(x_{s}, x_{t}\right)\right) \geq d\left(x_{s}, x_{t}\right)(=O(1)) \tag{5}
\end{equation*}
$$

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

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{E_{1}^{\text {global }}}{E_{1}^{\text {local }}} \geq \lim _{n \rightarrow \infty} \frac{d\left(x_{s}, x_{t}\right)-2 \epsilon}{2 \epsilon}=\infty \tag{6}
\end{equation*}
$$

Case 3: For Gaussian weighted graph, $w_{i j}=\exp \left(\frac{d\left(x_{i}, x_{j}\right)^{2}}{\delta^{2}}\right)$ being the distance corresponding to a similarity graph (equivalent to similarity between $x_{i}$ and $x_{j}$ being $\left.\exp \left(\frac{-d\left(x_{i}, x_{j}\right)^{2}}{\delta^{2}}\right)\right)$.
$O\left(E_{1}^{\text {local }}\right)$ : In $\epsilon$-neighborhood graph, $\forall(i, j) \in E, d\left(x_{i}, x_{j}\right)<\epsilon$, therefore $w_{s i}, w_{t j}<\exp \left(\frac{\epsilon^{2}}{\delta^{2}}\right)$. Then, for unit flow $\left(i_{e}\right)_{e}=I_{2}\left(x_{s}, x_{t}\right)$,

$$
\begin{align*}
E_{1}^{\text {local }} & =\sum_{i,(i, s) \in E} w_{s i}\left|i_{s i}\right|+\sum_{i,(i, t) \in E} w_{t i}\left|i_{t i}\right| \\
& <\sum_{i,(i, s) \in E} \exp \left(\frac{\epsilon^{2}}{\delta^{2}}\right)\left|i_{s i}\right|+\sum_{i,(i, t) \in E} \exp \left(\frac{\epsilon^{2}}{\delta^{2}}\right)\left|i_{t i}\right| \\
& =2 \exp \left(\frac{\epsilon^{2}}{\delta^{2}}\right) . \tag{7}
\end{align*}
$$

$O\left(E_{1}^{g l o b a l}\right): \forall(i, j) \in E, d\left(x_{i}, x_{j}\right), w_{i j}=\exp \left(\frac{d\left(x_{i}, x_{j}\right)^{2}}{\delta^{2}}\right) \geq 1$. Therefore,

$$
\begin{equation*}
E_{1}^{g l o b a l} \geq s p\left(x_{s}, x_{t}\right)-E_{1}^{l o c a l} \geq \frac{d\left(x_{s}, x_{t}\right)}{\epsilon}-E_{1}^{l o c a l} \tag{8}
\end{equation*}
$$

From $\sqrt[7]{ }$ and 8 , as $\lim _{n \rightarrow \infty}$, if $\epsilon \rightarrow 0$ and $O(\delta)>O\left(\frac{\epsilon}{\sqrt{-\ln (\epsilon)}}\right)$, we have:

$$
\begin{aligned}
\frac{E_{1}^{\text {global }}}{E_{1}^{\text {local }}} & \geq \frac{\frac{1}{\epsilon}}{2 \exp \left(\frac{\epsilon^{2}}{\delta^{2}}\right)}-1 \\
\ln \left(\frac{\frac{1}{\epsilon}}{2 \exp \left(\frac{\epsilon^{2}}{\delta^{2}}\right)}\right) & =\ln \frac{1}{\epsilon}-\ln \left(2 \exp \left(\frac{\epsilon^{2}}{\delta^{2}}\right)\right) \\
& =-\ln (\epsilon)-\ln (2)-\frac{\epsilon^{2}}{\delta^{2}}
\end{aligned}
$$

Since $O(\delta)>O\left(\frac{\epsilon}{\sqrt{-\ln (\epsilon)}}\right), O\left(\delta^{2}\right)>O\left(\frac{\epsilon^{2}}{-\ln (\epsilon)}\right)$ and $O(-\ln (\epsilon))>O\left(\frac{\epsilon^{2}}{\delta^{2}}\right)$.
As $\lim _{n \rightarrow \infty} \epsilon=0, \lim _{n \rightarrow \infty}-\ln (\epsilon)=\infty$, therefore $\lim _{n \rightarrow \infty}-\ln (\epsilon)-\ln (2)-\frac{\epsilon^{2}}{\delta^{2}}=\infty$. Hence,

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{E_{1}^{\text {global }}}{E_{1}^{\text {local }}}=\infty \tag{9}
\end{equation*}
$$

Note that, it is advisable to set $\delta$ such that $w_{i j}$ 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 $O(\delta)>$ $O\left(\frac{\epsilon}{\sqrt{-\ln (\epsilon)}}\right)$ when $\epsilon \rightarrow 0$.

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 $E_{1}\left(I_{2}\right)$ dominates the local part almost surely as $n \rightarrow \infty$. Concretely, there exist constants $c_{1}, c_{2}$ that the following statements hold:

1. For unwighted graph $w_{i j}=1: \lim _{n \rightarrow \infty} \frac{E_{1}^{\text {global }}}{E_{1}^{\text {local }}} \rightarrow \infty$ almost surely as $n \rightarrow \infty, k>\log (n)$ and $\frac{k}{n} \rightarrow 0$ with a probability of at least $1-c_{1} n \exp \left(-c_{2} \sqrt{n k}\right)$ (converging to 1 ).
2. For Euclidean weighted graph with $w_{i j}=d\left(x_{i}, x_{j}\right): \frac{E_{1}^{\text {global }}}{E_{1}^{\text {local }}} \rightarrow \infty$ almost surely as $n \rightarrow \infty$, $k>\log (n)$ and $\frac{k}{n} \rightarrow 0$ with a probability of at least $1-c_{1} n \exp \left(-c_{2} \sqrt{n k}\right)$ (converging to 1 ).
3. For Gaussian weighted graph with $w_{i j}=\exp \left(\frac{d\left(x_{i}, x_{j}\right)^{2}}{\delta^{2}}\right): \frac{E_{1}^{\text {global }}}{E_{1}^{\text {local }}} \rightarrow \infty$ almost surely as $n \rightarrow \infty$, $k>\log (n), \frac{k}{n} \rightarrow 0$ and $O(\delta)=\left(\frac{k}{n}\right)^{\frac{1}{d}}$ with a probability of at least $1-c_{1} n \exp \left(-c_{2} k \cdot \log \left(\frac{n}{k}\right)^{\frac{d}{2}}\right)$ (converging to 1 ).

Recall the definition of $k$-nearest neighbor radii: $R_{k}(x)=\max _{i,(i, s) \in E} d\left(x_{s}, x_{i}\right)$ be the distance of $x$ to its $k$-nearest neighbor in $X$. Let $B(x, \eta)$ be the ball centered at $x$ with radius $\eta$. Hence, there are only $k$ points from the sampled $n$ points lying in $B\left(x, R_{k}(x)\right)$. 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 $x, R_{k}(x) \rightarrow 0$ with a high probability.
Lemma 1.1. For any fixed node $x \in X$ in a random geometric knn graph, any $\epsilon_{0}$ as a function of $n$ satisfying $O\left(\left(\frac{k}{n}\right)^{\frac{1}{d}}\right)<O\left(\epsilon_{0}\right)<O(1)$, then $O\left(R_{k}(x)\right)<O\left(\epsilon_{0}\right) \rightarrow 0$ as $n \rightarrow \infty$ with a probability at least $1-c_{1} \exp \left(\frac{-n \gamma\left(\epsilon_{0}\right)}{2}\right)$, converging to 1 when $n, k \rightarrow \infty$ and $\frac{k}{n} \rightarrow 0$ for some constant $c_{1}$.

Proof. In this case, we want $\epsilon_{0}$ to play the role of $\epsilon$ in Theorem 3.3. The difference is that, in this case, $\epsilon_{0}$ neighborhoods contain all the neighbors of all points with a high enough probability (converging to $1)$, as opposed to the case of $\epsilon$ neighborhoods that contain all neighbors of all points with probability 1.

The volume of $B\left(x, \epsilon_{0}\right)$ is $c \epsilon_{0}^{d}$ for constant $c=\frac{\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}+1\right)}$. Let $\gamma\left(\epsilon_{0}\right)=\int_{B\left(x, \epsilon_{0}\right)} p(x) d x$ denote the probability mass of $B\left(x, \epsilon_{0}\right)$, then

$$
\begin{equation*}
p_{\min } c \epsilon_{0}^{d} \leq \gamma\left(\epsilon_{0}\right) \leq p_{\max } c \epsilon_{0}^{d} \tag{10}
\end{equation*}
$$

or $O\left(\gamma\left(\epsilon_{0}\right)\right)=O\left(\epsilon_{0}^{d}\right)$.
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 $B\left(x, \epsilon_{0}\right)$ (with probability mass of $\left.\gamma\left(\epsilon_{0}\right)\right)$.

The number of points lying in $B\left(x, \epsilon_{0}\right)$ follows a binomial distribution $B\left(n, \gamma\left(\epsilon_{0}\right)\right)$. Let $F\left(k, n, \gamma\left(\epsilon_{0}\right)\right)$ be the cumulative distribution function of $B\left(n, \gamma\left(\epsilon_{0}\right)\right)$. Let $P\left(R_{k}(x) \geq \epsilon_{0}\right)$ be the probability that $R_{k}(x) \geq \epsilon_{0}$, then, $P\left(R_{k}(x) \geq \epsilon_{0}\right)=F\left(k-1, n, \gamma\left(\epsilon_{0}\right)\right)<F\left(k, n, \gamma\left(\epsilon_{0}\right)\right)$ (we prove $k$ for simplicity). Chernoff's inequality for binomial distribution gives

$$
F\left(k, n, \gamma\left(\epsilon_{0}\right)\right) \leq \exp \left(\frac{-\left(n \gamma\left(\epsilon_{0}\right)-k\right)^{2}}{2 n \gamma\left(\epsilon_{0}\right)}\right)
$$

As $n, k \rightarrow \infty, \frac{k}{n} \rightarrow 0, O\left(n \gamma\left(\epsilon_{0}\right)\right)>O(k)$ by the way we choose $\epsilon_{0}$, hence,

$$
O\left(\exp \left(\frac{-\left(n \gamma\left(\epsilon_{0}\right)-k\right)^{2}}{2 n \gamma\left(\epsilon_{0}\right)}\right)\right)=O\left(\exp \left(\frac{-n \gamma\left(\epsilon_{0}\right)}{2}\right) .\right.
$$

Because $n \gamma\left(\epsilon_{0}\right)>k \rightarrow \infty, \exp \left(\frac{-n \gamma\left(\epsilon_{0}\right)}{2}\right) \rightarrow 0$. Hence, there exists a constant $c_{1}$ that $R_{k}(x)<\epsilon_{0} \rightarrow$ 0 with a probability $1-F\left(k, n, \gamma\left(\epsilon_{0}\right)\right)$, which is at least $1-c_{1} \exp \left(\frac{-n \gamma\left(\epsilon_{0}\right)}{2}\right)$.

Lemma 1.2. For all nodes in $X$ in a random geometric knn graph, $O\left(R_{k}(x)\right)<O\left(\epsilon_{0}\right) \rightarrow 0 \forall i=1 \cdots n$ as $n \rightarrow \infty$ with a probability at least $1-c_{1} n \exp \left(-c_{2} n \epsilon_{0}^{d}\right)$ for some constant $c_{1}, c_{2}$, converging to 1 when $n, k \rightarrow \infty, \frac{k}{n} \rightarrow 0$ and $O\left(\epsilon_{0}\right)>\left(\frac{\log (n)}{n}\right)^{\frac{1}{d}}$. If $k>\log (n)$ then the last condition is already included in the choice of $\epsilon_{0}$.

Proof. The Lemma 1.1 shows that for any fixed $x_{i} \in X, P\left(R_{k}\left(x_{i}\right) \geq \epsilon_{0}\right) \leq c_{1} \exp \left(\frac{-n \gamma\left(\epsilon_{0}\right)}{2}\right)$. Therefore, from 10 the probability that there exists at least one $x_{i}$ such that $R_{k}\left(x_{i}\right) \geq \epsilon_{0} \forall i=1 \cdots n$ satisfies (for all $i$ together)

$$
\begin{equation*}
P\left(R_{k}\left(x_{i}\right) \geq \epsilon_{0}\right) \leq c_{1} n \exp \left(\frac{-n \gamma\left(\epsilon_{0}\right)}{2}\right) \leq c_{1} n \exp \left(-c_{2} n \epsilon_{0}^{d}\right) \tag{11}
\end{equation*}
$$

for some constant $c_{2}$. Hence, the probability that $R_{k}\left(x_{i}\right)<\epsilon_{0}$ for all $i=1 \cdots n$, as $n, k \rightarrow \infty$ and $\frac{k}{n} \rightarrow 0$, satisfies

$$
\begin{equation*}
P\left(R_{k}\left(x_{i}\right)<\epsilon_{0}\right) \geq 1-c_{1} n \exp \left(-c_{2} n \epsilon_{0}^{d}\right) \forall i . \tag{12}
\end{equation*}
$$

Now we show that $c_{1} n \exp \left(-c_{2} n \epsilon_{0}^{d}\right) \rightarrow 0$ as $n, k \rightarrow \infty$ and $\frac{k}{n} \rightarrow 0$ and $O\left(\epsilon_{0}\right)>O\left(\frac{\log (n)}{n}\right)^{\frac{1}{d}}$. Since $O\left(\epsilon_{0}^{d}\right)>O\left(\frac{\log (n)}{n}\right), O\left(-c_{2} n \epsilon_{0}^{d}\right)>O(\log (n))$, therefore,

$$
c_{1} n \exp \left(-c_{2} n \epsilon_{0}^{d}\right)=c_{1} \exp \left(\log (n)-c_{2} n \epsilon_{0}^{d}\right) \rightarrow c_{1} \exp (-\infty)=0
$$

As $O\left(\epsilon_{0}\right)>O\left(\left(\frac{k}{n}\right)^{\frac{1}{d}}\right)$ 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 $\epsilon_{0}$ to make the formulations simple and intuitive, even though a range of $\epsilon_{0}$ would work.

Case 1: unweighted graph, $w_{i j}=1$ for $x_{j}$ is one of the $k$-nearest neighbors of $x_{i}$.
$O\left(E_{1}^{\text {local }}\right)$ : For unit flow $I_{2}=\left(i_{e}\right)_{e \in E}$ :

$$
\begin{equation*}
E_{1}^{l o c a l}=\sum_{i,(i, s) \in E} w_{s i}\left|i_{s i}\right|+\sum_{j,(j, t) \in E} w_{t j}\left|i_{t j}\right|=2 . \tag{13}
\end{equation*}
$$

$O\left(E_{1}^{g l o b a l}\right)$ : Since $E_{1}\left(I_{2}\left(x_{s}, x_{t}\right)\right) \geq s p\left(x_{s}, x_{t}\right)$, we prove that $O\left(s p\left(x_{s}, x_{t}\right)\right)>O(1)$ with a probability converging to 1 as $n \rightarrow \infty$.

In this case, we choose $\epsilon_{0}=\left(\frac{k}{n}\right)^{\frac{1}{2 d}}$. From Lemma 1.2 , with a probability at least $1-c_{1} n \exp \left(-c_{2} \sqrt{n k}\right)$, all $R_{k}(x)<\epsilon_{0}$, or equivalently, $d\left(x_{i}, x_{j}\right)<\epsilon_{0} \forall(i, j) \in E$. Therefore, $O\left(s p\left(x_{s}, x_{t}\right)\right) \geq O\left(\frac{d\left(x_{s}, x_{t}\right)}{\epsilon_{0}}\right)$. Since $O\left(\frac{d\left(x_{s}, x_{t}\right)}{\epsilon_{0}}\right)=O\left(\left(\frac{n}{k}\right)^{\frac{1}{2 d}}\right) \rightarrow \infty$. Hence, with a probability of at least $1-c_{1} n \exp \left(-c_{2} \sqrt{n k}\right)$, as $n \rightarrow \infty, k>\log (n)$ and $\frac{k}{n} \rightarrow 0$,

$$
\begin{equation*}
E_{1}\left(I_{2}\left(x_{s}, x_{t}\right)\right) \rightarrow \infty \tag{14}
\end{equation*}
$$

Therefore, from 20) and 14 with a probability of at least $1-c_{1} n \exp \left(-c_{2} \sqrt{n k}\right)$ (converging to $1)$, as $n \rightarrow \infty, k \rightarrow \infty$ and $\frac{k}{n} \rightarrow 0$,

$$
\begin{equation*}
\frac{E_{1}^{\text {global }}}{E_{1}^{\text {local }}} \rightarrow \infty \tag{15}
\end{equation*}
$$

Case 2: Euclidean weighted graph, $w_{i j}=d\left(x_{i}, x_{j}\right)$ for $x_{j}$ is one of the $k$-nearest neighbors of $x_{i}$. We also choose $\epsilon_{0}=\left(\frac{k}{n}\right)^{\frac{1}{2 d}}$ as previous case.

$$
O\left(E_{1}^{l o c a l}\right): \text { For unit flow } I_{2}=\left(i_{e}\right)_{e \in E}:
$$

$$
\begin{equation*}
E_{1}^{\text {local }}\left(I_{2}\left(x_{s}, x_{t}\right)\right)=\sum_{i,(i, s) \in E} w_{s i}\left|i_{s i}\right|+\sum_{i,(i, t) \in E} w_{t i}\left|i_{t i}\right| \leq \max _{i,(i, s) \in E} d\left(x_{s}, x_{i}\right)+\max _{j,(j, t) \in E} d\left(x_{t}, x_{j}\right) . \tag{16}
\end{equation*}
$$

As $n \rightarrow \infty, k>\log (n)$ and $\frac{k}{n} \rightarrow 0$, according to Lemma 1.2, $c_{1} n \exp \left(-c_{2} \sqrt{n k}\right) \rightarrow 0$. Hence, with a probability not smaller than $1-c_{1} n \exp \left(-c_{2} \sqrt{n k}\right), R_{k}(x)<\epsilon_{0} \forall x \in X$ or $d\left(x_{i}, x_{j}\right)<\epsilon_{0} \forall(i, j) \in E$. From (16),

$$
\begin{equation*}
E_{1}^{\text {local }}<2 \epsilon_{0} \rightarrow 0 . \tag{17}
\end{equation*}
$$

$$
\begin{align*}
& O\left(E_{1}^{\text {global }}\right): \\
& \quad E_{1}^{\text {local }}\left(I_{2}\left(x_{s}, x_{t}\right)\right)+E_{1}^{\text {global }}\left(I_{2}\left(x_{s}, x_{t}\right)\right)=E_{1}\left(I_{2}\left(x_{s}, x_{t}\right)\right) \geq s p\left(x_{s}, x_{t}\right) \geq d\left(x_{s}, x_{t}\right) \in O(1) . \tag{18}
\end{align*}
$$

From 17) and 18, we have, with a probability at least $1-c_{1} n \exp \left(-c_{2} \sqrt{n k}\right)$ (converging to 1) as $n \rightarrow \infty, k>\log (n)$ and $\frac{k}{n} \rightarrow 0$,

$$
\begin{equation*}
\frac{E_{1}^{\text {global }}}{E_{1}^{\text {local }}} \rightarrow \infty \tag{19}
\end{equation*}
$$

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

We define some notations for simpler formulation. Let $t=\frac{n}{k}$, hence, $t \rightarrow \infty$.
We choose $\epsilon_{0}=\left(\frac{k}{n}\right)^{\frac{1}{d}} \cdot \sqrt{\frac{\log (t)}{d+1}}$ in this case to bound $E_{1}^{\text {local }}$ and $E_{1}^{\text {global }}$. In fact, we just need $O\left(\epsilon_{0}\right)>O\left(\left(\frac{k}{n}\right)^{\frac{1}{d}}\right)$ and still small enough according to some complicated formula. In this case, we have to choose smaller $\epsilon_{0}$ compared to previous cases just to show that global energy dominates local one.

With a probability of at least $1-c_{1} n \exp \left(-c_{2} n \epsilon_{0}^{d}\right)$, then all $R_{k}(x)<\epsilon_{0}$, meaning that $d\left(x_{i}, x_{j}\right)<$ $\exp \left(\frac{\epsilon^{2}}{\delta^{2}}\right)$. In this case, we bound $E_{1}^{\text {local }}$ and $E_{1}^{\text {global }}$ as follows.
$O\left(E_{1}^{\text {local }}\right)$ : For unit flow $I_{2}=\left(i_{e}\right)_{e \in E}:$

$$
\begin{align*}
E_{1}^{l o c a l}\left(I_{2}\left(x_{s}, x_{t}\right)\right) & =\sum_{i,(i, s) \in E} w_{s i}\left|i_{s i}\right|+\sum_{i,(i, t) \in E} w_{t i}\left|i_{t i}\right| \\
& \leq \max _{i,(i, s) \in E} \exp \left(\frac{d\left(x_{s}, x_{i}\right)^{2}}{\delta^{2}}\right)+\max _{j,(j, t) \in E} \exp \left(\frac{d\left(x_{t}, x_{j}\right)^{2}}{\delta^{2}}\right) \\
& \leq 2 \exp \left(\frac{\epsilon_{0}^{2}}{\delta^{2}}\right) \tag{20}
\end{align*}
$$

Therefore,

$$
\begin{align*}
O\left(E_{1}^{\text {local }}\right) & \leq O\left(\exp \left(\frac{\epsilon_{0}^{2}}{\delta^{2}}\right)\right) \\
& =O\left(\exp \left(\sqrt{\frac{\log (t)}{d+1}}^{2}\right)\right) \\
& =O\left(t^{\frac{1}{d+1}}\right) . \tag{21}
\end{align*}
$$

$O\left(E_{1}^{g l o b a l}\right)$ : 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 $d\left(x_{i}, x_{j}\right)<\epsilon_{0} \forall(i, j) \in E$ (all edges are of length less than $\epsilon_{0}$ ), the number of edges on any path between $x_{s}$ and $x_{t}$ must not smaller than $\frac{d\left(x_{s}, x_{t}\right)}{\epsilon_{0}}$. Hence,

$$
\begin{equation*}
E_{1}^{g l o b a l}+E_{1}^{\text {local }} \geq s p\left(x_{s}, x_{t}\right) \geq \frac{d\left(x_{s}, x_{t}\right)}{\epsilon_{0}} \in O\left(\frac{1}{\epsilon}\right) . \tag{22}
\end{equation*}
$$

Therefore,

$$
\begin{align*}
O\left(E_{1}^{\text {global }}+E_{1}^{\text {local }}\right) & \geq O\left(\frac{1}{\epsilon}\right) \\
& =O\left(\frac{t^{\frac{1}{d}}}{\sqrt{\frac{\log (t)}{d+1}}}\right) \\
& =O\left(\frac{t^{\frac{1}{d}}}{\sqrt{\log (t)}}\right) \\
& =O\left(t^{\frac{1}{d}} \cdot \log (t)^{\frac{-1}{2}}\right) . \tag{23}
\end{align*}
$$

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

$$
\begin{align*}
O\left(\frac{E_{1}^{\text {global }}}{E_{1}^{\text {local }}}\right) & =O\left(\frac{E_{1}^{\text {global }}+E_{1}^{\text {local }}}{E_{1}^{\text {local }}}-1\right) \\
& \geq O\left(\frac{t^{\frac{1}{d}} \cdot \log (t)^{\frac{-1}{2}}}{t^{\frac{1}{d+1}}}-1\right) \\
& =O\left(\frac{t^{\frac{1}{d(d+1)}}}{\log (t)^{\frac{1}{2}}}-1\right) \\
& =\infty \tag{24}
\end{align*}
$$

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

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

$$
\begin{equation*}
V^{(s)}-V^{(t)}=L^{-1}\left(e_{s}-e_{t}\right) \tag{25}
\end{equation*}
$$

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 $V^{\prime} \in R^{n}$ be the potential on nodes of the graph with unit potential difference between $x_{s}$ and $x_{t}$, namely

$$
V_{s}^{\prime}-V_{t}^{\prime}=V^{\prime T}\left(e_{s}-e_{t}\right)=1
$$

Kirchhoff's voltage law: the potential assignment in the network minimizes the energy function $V^{T} L V^{\prime}$ :

$$
\begin{equation*}
V^{\prime}=\arg \min _{x \in R^{n}} x^{T} L x, \text { s.t. } x^{t}\left(e_{s}-e_{t}\right)=1 . \tag{26}
\end{equation*}
$$

Lagrange multipliers method gives us

$$
V^{\prime}=\frac{L^{-1}\left(e_{s}-e_{t}\right)}{\left(e_{s}-e_{t}\right)^{T} L^{-1}\left(e_{s}-e_{t}\right)}+\alpha \mathbf{1}
$$

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

$$
E^{\prime}=\frac{1}{\left(e_{s}-e_{t}\right)^{T} L^{-1}\left(e_{s}-e_{t}\right)} .
$$

Using Ohm's law $E^{\prime}=V^{\prime} I^{\prime}=I^{\prime 2} R_{s t}$, then $V^{\prime}$ makes the total flow from $x_{s}$ to $x_{t}$ of $I^{\prime}$ as:

$$
I^{\prime}=\frac{E^{\prime}}{V^{\prime}}=\frac{1}{\left(e_{s}-e_{t}\right)^{T} L^{-1}\left(e_{s}-e_{t}\right)}
$$

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 $V^{\prime}$ (also $I^{\prime}$ ). To have an unit flow $(I=1)$ from $x_{s}$ to $x_{t}$ then voltage arrangement in the network, ignoring constant terms, can be

$$
\begin{equation*}
V^{\prime} \cdot \frac{1}{I^{\prime}}=L^{-1}\left(e_{s}-e_{t}\right)=V^{(s)}-V^{(t)} \tag{27}
\end{equation*}
$$

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 $I_{2}\left(x_{s}, x_{t}\right)$ 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:

$$
\begin{align*}
f: X & \rightarrow R^{|E|} \\
x_{s} & \rightarrow f\left(x_{s}\right)=\left\{\cdots, \frac{V_{i}^{(s)}-V_{j}^{(s)}}{r_{i j}^{(p-1) / p}}, \cdots\right\}_{(i, j) \in E}^{T} \tag{28}
\end{align*}
$$

makes the p-norm of the space coincide with $R_{p}:\left\|f\left(x_{s}\right)-f\left(x_{t}\right)\right\|_{p}=R_{p}\left(x_{s}, x_{t}\right)$.
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 $R_{p}^{p}=E_{p}\left(I_{2}\right)$ in potential form using Lemma 3.4:

$$
\begin{align*}
& R_{p}^{p}\left(x_{s}, x_{t}\right) \\
= & \sum_{e=(i, j) \in E} r_{e}\left|I_{2}\left(x_{s}, x_{t}\right)_{e}\right|^{p} \\
= & \sum_{e=(i, j) \in E} r_{e} \frac{\left|V_{i}-V_{j}\right|^{p}}{r_{e}^{p}} \\
= & \sum_{e=(i, j) \in E} r_{e}^{1-p} \cdot\left|\left(L_{i s}^{-1}-L_{i t}^{-1}\right)-\left(L_{j s}^{-1}-L_{j t}^{-1}\right)\right|^{p} \\
= & \sum_{e=(i, j) \in E} \left\lvert\, \frac{L_{i s}^{-1}-L_{j s}^{-1}}{r_{i j}^{(p-1) / p}}-\frac{L_{i t}^{-1}-L_{j t}^{-1}}{\left.r_{i j}^{(p-1) / p}\right|^{p}}\right. \\
= & \sum_{e=(i, j) \in E}\left|\frac{V_{i}^{(s)}-V_{j}^{(s)}}{r_{i j}^{(p-1) / p}}-\frac{V_{i}^{(t)}-V_{t}^{(t)}}{r_{i j}^{(p-1) / p}}\right|^{p} \tag{29}
\end{align*}
$$

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 $s$ :

$$
\begin{align*}
f: X & \rightarrow R^{|E|} \\
x_{s} & \rightarrow f\left(x_{s}\right)=\left\{\cdots, \frac{V_{i}^{(s)}-V_{j}^{(s)}}{r_{i j}^{(p-1) / p}}, \cdots\right\}_{(i, j) \in E}^{T} \tag{30}
\end{align*}
$$

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

$$
R_{p}\left(x_{s}, x_{t}\right)=\left\|f\left(x_{s}\right)-f\left(x_{t}\right)\right\|_{p}
$$

Hence, $R_{p}$ distance induces an embedding of nodes into the $L_{p}$ space $\left(R^{|E|},\|\cdot\|_{p}\right)$. 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 $\left(R_{2}\right)$, 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. Shortest path distances ( $s p$ ), as expected, showed non-smooth distance both within and between clusters. 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. Shortest path distances ( $s p$ ), as expected, showed non-smooth distance both within and between clusters. It meant that our proposed distances could overcome the global information loss problem.

