SUPPLEMENTARY MATERIAL

Getting lost in space: Large sample analysis of the commute distance

---

Ulrike von Luxburg **Agnes Radl**
Max Planck Institute for Biological Cybernetics, Tübingen, Germany
`{ulrike.luxburg,agnes.radl}@tuebingen.mpg.de`

Matthias Hein
Saarland University, Saarbrücken, Germany
`hein@cs.uni-sb.de`

This supplement is devoted to the proof of our main results: Theorems 2 and 3 of the main paper. For convenience, we formulate all proofs in terms of the effective resistance between two vertices. To convert effective resistance to the commute distance, one just has to multiply by the factor $\mathrm{vol}(G)$. We rely on the notation that has been introduced in our main paper.

1 Lower bound on the resistance distance for arbitrary graphs

It is easy to prove that the resistance distance between two points is lower bounded by the sum of the inverse degrees.

Proposition 4 (Lower bound) *Let G be a weighted, undirected, connected graph and consider two vertices s and t, $s \neq t$. Assume that G remains connected if we remove s and t. Then the effective resistance between s and t is bounded by*

$$R_{st} \geq \frac{Q_{st}}{1 + w_{st} Q_{st}}$$

where $Q_{st} = 1/(d_s - w_{st}) + 1/(d_t - w_{st})$. Note that if s and t are not connected by a direct edge (that is, $w_{st} = 0$), then the rhs simplifies to $1/d_s + 1/d_t$.

Proof. The proof is based on Rayleigh's monotonicity principle that states that increasing edge weights in the graph can never increase the effective resistance between two vertices (cf. Corollary 7 in Section IX.2 of Bollobás, 1998). Given our original graph G, we build a new graph G' by setting the weight of all edges to infinity, except the edges that are adjacent to s or t (setting the weight of an edge to infinity means that this edge has no resistance any more). This can also be interpreted as taking all vertices except s and t and merging them to one super-node a. Now our graph G' consists of three vertices s, a, t, with several parallel edges from s to a, several parallel edges from a to t. Exploiting the laws in electrical networks (resistances add along edges in series, conductances add along edges in parallel; see Section 2.3 in Lyons and Peres (2010) for detailed instructions and examples) leads to the desired result. ☺

2 Upper bound on the resistance in deterministic geometric graphs

This is the part that requires the hard work. Our proof is based on a theorem that shows how the resistance between two points in the graph can be computed in terms of flows on the graph. The following result is taken from Corollary 6 in Section IX.2 of Bollobás (1998).

Theorem 5 (Resistance in terms of flows, cf. Bollobás, 1998) *Let $G = (V, E)$ be an unweighted graph. The effective resistance R_{st} between two fixed vertices s and t can be expressed as*

$$R_{st} = \inf \left\{ \sum_{e \in E} u_e^2 \;\middle|\; u = (u_e)_{e \in E} \text{ unit flow from } s \text{ to } t \right\}. \tag{1}$$

Note that evaluating the formula in the above theorem for any fixed flow leads to an upper bound on the effective resistance. The key to obtaining a tight bound is to distribute the flow as widely and uniformly over the graph as possible.

For the case of geometric graphs (that is, graphs whose vertices correspond to points in some underlying space $\mathbb{R}^d$), we are going to use a grid on the underlying space to construct an efficient flow between two vertices. Let $X_1, ..., X_n$ be a fixed set of points in $\mathbb{R}^d$ and consider a geometric graph G with vertices $X_1, ..., X_n$. Fix any two of them, say $s := X_1$ and $t := X_2$. Let $\mathcal{X} \subset \mathbb{R}^d$ be a connected set that contains both s and t. Consider a regular grid with grid width g on $\mathcal{X}$. We say that grid cells are neighbors of each other if they touch each other in at least one point.

Definition 6 (Valid grid) *We call the* grid valid *if the following properties are satisfied:*

1. *Each cell of the grid contains at least one of the points $X_1, ..., X_n$.*
2. *Points in the same or neighboring cells of the grid are always connected in the graph G.*
3. *The grid width g is small enough: Define the bottleneck h of the region $\mathcal{X}$ (see Definition 1 in the main paper) as the largest u such that the set $\{x \in \mathcal{X} \mid dist(x, \partial\mathcal{X}) > u/2\}$ is connected. We require that there exists a piecewise linear path between s and t which has distance at least $h/2$ from $\partial\mathcal{X}$. Denote the length of this path by $d(s,t)$. We require that $\sqrt{d-1}g \le h$ (a cube of side length g should fit in the bottleneck) and $d(s,t) > 4\sqrt{d-1}g$ (the grid should be small enough such that the cubes containing s and t do not overlap).*

Below we construct a flow from s to t with the help of the underlying grid. The main idea is to first distribute the flow from s to all points in the grid cell $C(s)$ that contains s. Then we follow a path of grid cells that goes from the cell $C(s)$ to the corresponding cell of $C(t)$, and in the last step we go to t itself.

At this point note a general concept that we will use over and over again. Assume we have N_1 points in grid cell C_1 and N_2 points in grid cell C_2, and assume that C_1 and C_2 are neighboring cells. Then, if the grid is valid, all points in C_1 are connected to all points in C_2, that is we have $N_1 N_2$ different edges between the two cells.

We now prove the following general proposition that gives an upper bound on the resistance distance between vertices in a fixed geometric graph.

Proposition 7 (Resistance on a fixed geometric graph) *Consider a fixed set of points $X_1, ..., X_n$ in some connected region $\mathcal{X} \subset \mathbb{R}^d$ with bottleneck h (where the bottleneck is defined as in the definition of a valid region in the main paper). Denote $s = X_1$ and $t = X_2$. Assume that s and t have distance at least $h/2$ from $\partial\mathcal{X}$. Let $d(s,t)$ be be the length of a piecewise linear path between s and t which has distance at least $h/2$ from $\partial\mathcal{X}$. Consider a geometric graph on $X_1, ..., X_n$ and let g be the width of a valid grid on $\mathcal{X}$. By $N_{\min}$ denote the minimal number of points in each grid cell, and define a as*

$$a \;:=\; \left\lfloor h/(2g\sqrt{d-1}) \right\rfloor. \tag{2}$$

Then the effective resistance between s and t can be bounded as follows:

$$R_{st} \leq \frac{1}{d_s} + \frac{1}{d_t} + \frac{1}{N_{\min}}\left(\frac{1}{d_s} + \frac{1}{d_t}\right) + \begin{cases} \frac{1}{N_{\min}^2}\left(\frac{1}{d-1}(\log(a)+1) + \frac{2d(s,t)}{g(2a-1)^{d-1}}\right) & \textit{if } d = 3 \\ \frac{1}{N_{\min}^2}\left(\frac{2}{d-1} + \frac{2d(s,t)}{g(2a-1)^{d-1}}\right) & \textit{if } d > 3 \end{cases} \tag{3}$$

The actual proof is rather lengthy and a bit tedious. The rest of this section is devoted to it.

Construction of the flow — overview.

Without loss of generality we assume that there exists a straight line connecting s and t which is along the first dimension of the space.

Step 0: We start a unit flow in vertex s.

Step 1: We make a step to all neighbors $\mathrm{Neigh}(s)$ of s and distribute the flow uniformly over all edges. That is, we traverse d_s edges and send flow $1/d_s$ over each edge.

Step 2: In one step we distribute the flow from the points in $\mathrm{Neigh}(s)$ uniformly to all points in the grid cell $C(s)$ that contains s. This is necessary as some of the neighbors of s are outside of $C(s)$, and for the next step we need that all flow sits in cell $C(s)$.

Step 3: We now distribute the flow from $C(s)$ to a larger region, namely to a $(d-1)$-dim hypercube $H(s)$ of side length h that is perpendicular to the linear path from s to t and centered at $C(s)$, see left plot of Figure 1. This can be achieved in several substeps that will be defined below.

Step 4: We now traverse from $H(s)$ to an analogous hypercube $H(t)$ located at t using parallel paths, see right plot of Figure 1.

Step 5: From the hypercube $H(t)$ we send the flow to the neighborhood $\mathrm{Neigh}(t)$ (this is the "reverse" of steps 2 and 3).

Step 6: From $\mathrm{Neigh}(t)$ we finally send the flow to the destination t ("reverse" of step 1).

Details of the flow construction and computation of the resistance beween s and t.

We now describe the individual steps and their contribution to the bound on the resistance. By the "contribution of a step" we mean the part of the sum in Theorem 5 that goes over the edges considered in the current step.

Step 1 We start with a unit flow at s that we send over all d_s adjacent edges. This leads to flow $1/d_s$ over d_s edges. According to the formula in Theorem 5 this contributes

$$r_1 = d_s \cdot \frac{1}{d_s^2} = \frac{1}{d_s}$$

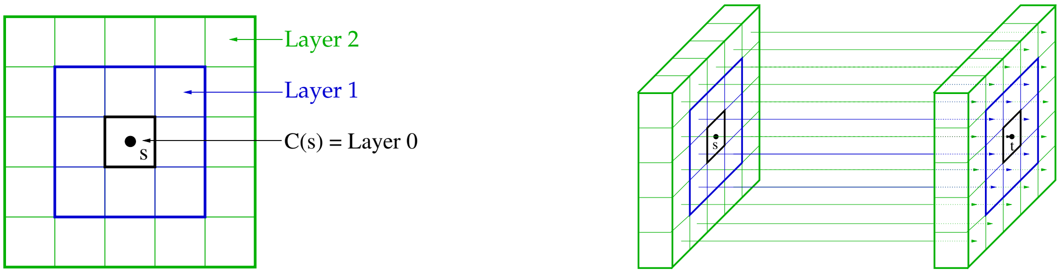

Figure 1: Left plot: front view of a $(d-1)$-dimensional hypercube with three layers, $d = 3$. Right plot: step 3 of the flow, $d = 3$.

to the overall resistance R_{st}.

Step 2: We now redistribute all flow from the neighbors of s to all points in the cell $C(s)$. We want to use as many edges as possible in this step. In our graph, all neighbors of s are connected to all points in $C(s)$ by construction. We now distribute the flow of each of the neighbors of s uniformly to all points in $C(s)$. That is, we use at least $d_s N_{\min}$ edges to distribute our flow, that is each edge gets flow at most $1/(d_s N_{\min})$. This leads to a contribution of

$$r_2 = d_s N_{\min} \left(\frac{1}{d_s N_{\min}} \right)^2 = \frac{1}{d_s N_{\min}}$$

Step 3: To distribute the flow to the whole hypercube we divide the hypercube into layers. Layer 0 consists of the cell $C(s)$ itself, the first layer consists of all cells adjacent to $C(s)$, and so on (see Figure 1, left plot). Layer i has side length $(2i+1)g$. The number l_i of grid cells in Layer i, $i \geq 1$, can be lower bounded by

$$l_i \geq 2(d-1)(2i-1)^{d-2}.$$

All in all we consider $a = \lfloor h/(2g\sqrt{d-1}) \rfloor \leq \lfloor h/(2(g-1)\sqrt{d-1}) \rfloor$ layers, so that the final layer has diameter just a bit smaller than the bottleneck h. We now distribute the flow stepwise through all layers, starting with unit flow in Layer 0.

To send the flow from Layer $i-1$ to Layer i we want to use as many edges as possible. We can lower bound the number of edges between these two layers by $l_i N_{\min}^2$: each of the l_i cells in Layer i is adjacent to at least one of the cells in Layer $i-1$, and each cell contains at least $N_{\min}$ points. Consequently, we can upper bound the contribution of the edges between Layer i and $i-1$ to the resistance by

$$r_{3,i} \leq l_i N_{\min}^2 \cdot \left(\frac{1}{l_i N_{\min}^2} \right)^2 = \frac{1}{l_i N_{\min}^2}.$$

All in all we have a layers. Thus the overall contribution of Step 3 to the resistance can be bounded by

$$r_3 = \sum_{i=1}^{a} r_{3,i} \ \leq \ \frac{1}{N_{\min}^2} \sum_{i=1}^{a} \frac{1}{l_i}.$$

Using the bounds on the l_i and assumption $d \geq 3$ we get

$$\begin{aligned} r_3 &\leq \frac{1}{N_{\min}^2} \sum_{i=1}^{a} \frac{1}{l_i} \\ &= \frac{1}{2(d-1)N_{\min}^2} \sum_{i=1}^{a} \frac{1}{(2i-1)^{d-2}} \\ &\leq \frac{1}{2(d-1)N_{\min}^2} \sum_{i=1}^{2a} \frac{1}{(i)^{d-2}} \end{aligned}$$

In case $d = 3$, the sum on the right hand side is the finite harmonic series and is upper bounded by $\log(a) + 1$. If $d > 3$, the over-harmonic series on the right hand side converges to a constant smaller than 2. All in all we get

$$r_3 \leq \begin{cases} \frac{1}{2(d-1)N_{\min}^2}\left(\log(a)+1\right) & \text{if } d=3 \\ \frac{1}{2(d-1)N_{\min}^2} & \text{if } d>3 \end{cases}$$

Step 4: Now we transfer all flow in "parallel cell paths" from $H(s)$ to $H(t)$. We have $(2a-1)^{d-1}$ parallel rows of cells going from $H(s)$ to $H(t)$, each of them contains $d(s,t)/g$ cells. Thus all in all we traverse $(2a-1)^{d-1}N_{\min}^2 d(s,t)/g$ edges, and each edge carries flow $1/((2a-1)^{d-1}N_{\min}^2)$. Thus step 4 contributes

$$r_4 \leq (2a-1)^{d-1}N_{\min}^2 \frac{d(s,t)}{g} \cdot \left(\frac{1}{(2a-1)^{d-1}N_{\min}^2}\right)^2 = \frac{d(s,t)}{g(2a-1)^{d-1}N_{\min}^2}$$

Step 5 is completely analogous to steps 2 and 3, with the analogous contribution $r_5 = \frac{1}{d_t N_{\min}} + r_3$.

Step 6 is completely analogous to step 1 with overall contribution of $r_6 = 1/d_t$.

Summing up all these contributions leads to the following overall bound on the resistance:

$$R_{st} \leq \frac{1}{d_s} + \frac{1}{d_t} + \frac{1}{N_{\min}}\left(\frac{1}{d_s} + \frac{1}{d_t}\right) + \begin{cases} \frac{1}{N_{\min}^2}\left(\frac{1}{d-1}(\log(a)+1) + \frac{2d(s,t)}{g(2a-1)^{d-1}}\right) & \text{if } d=3 \\ \frac{1}{N_{\min}^2}\left(\frac{2}{d-1} + \frac{2d(s,t)}{g(2a-1)^{d-1}}\right) & \text{if } d>3 \end{cases}$$

with a as defined in Eq. (2). This concludes the proof of Proposition 7. ☺

Let us make a couple of technical remarks about this proof. For the ease of presentation we simplified the proof in a couple of respects.

Strictly speaking, we do not need to distribute the whole unit flow to the outmost Layer a. The reason is that in each layer, a fraction of the flow already "branches off" in direction of t. We simply ignore this leaving flow when bounding the flow in Step 2, our construction leads to an upper bound. It is not so difficult to take the outbound flow into account, but it does not change the order of magnitude of the final result. So for the ease of presentation we drop this additional complication and stick to our rough upper bound.

When we consider Step 2 and the first couple of layers in Step 3 together, it turns out that some edges are used twice and might lead to "loops" in the flow. To ensure that we have a proper flow these loops could be removed. This would then just reduce the contribution of Steps 2 and 3, so that our current estimate is an overestimation of the whole resistance.

In step 3 one has to take care that the flow is always uniformly distributed in each layer, that is each grid cell of the layer gets the same amount of flow. The easiest way to achieve this is to introduce a "redistribution phase" for each layer. In this phase, the arriving flow from Layer $i-1$ is redistributed in Layer i to achieve a uniform distribution. We omitted this step for simplicity, but it does not change the final result apart from constants depending on the dimension.

The proof as it is spelled out above considers the case where s and t are connected by a straight line. It can be generalized to the case where they are connected by a piecewise linear path. This does not change the result in the end, but adds some technicality at the corners of the paths.

The construction of the flow only works if the bottleneck of $\mathcal{X}$ is not smaller than the diameter of one grid cell, if s and t are at least a couple of grid cells apart from each other, and if s and t are not too close to the boundary of $\mathcal{X}$. We took care of these conditions in Part 3 of the definition of a valid grid.

3 General properties of random geometric graphs

In this subsection we prove some basic results on random geometric graphs. These results are well-known in the random geometric graph community, but not so much in the machine learning community. Unfortunately, we did not find any reference where the material is presented in the way we need it (often the results are used implicitly or are tailored towards particular applications). Hence we present the proofs in this section. We encourage the reader to skip this section for the first reading, it is meant as a reference.

In the following, assume that $\mathcal{X}$ is a valid region according to Definition 1. Given any probability density p on $\mathbb{R}^d$, we now restrict the density to the region $\mathcal{X}$.

An important tool for dealing with random geometric graphs is the following well-known concentration inequality for binomial random variables that has first appeared in Angluin and Valiant (1977).

Proposition 8 (Concentration inequalities) *Let N be a $Bin(n, p)$-distributed random variable. Then, for all $\delta \in]0, 1]$,*

$$\begin{aligned} P\Big(N \;\leq\; (1-\delta)np\Big) \;&\leq\; \exp(-\frac{1}{3}\delta^2 np) \\ P\Big(N \;\geq\; (1+\delta)np\Big) \;&\leq\; \exp(-\frac{1}{3}\delta^2 np). \end{aligned}$$

We will see below that computing expected, minimum and maximum degrees in random geometric graphs always boils down to counting the number of data points in certain balls in the space. The following proposition is a straightforward application of the concentration inequality above and serves as "template" for all later proofs.

Proposition 9 (Counting sample points) *Consider a sample $X_1, \ldots, X_n$ drawn i.i.d. according to density p on $\mathcal{X}$. Let $B_1, \ldots, B_K$ be a fixed collection of subsets of $\mathcal{X}$ (the B_i do not need to be disjoint). Denote by $b_{\min} := \min_{i=1,\ldots,K} \int_{B_i} p(x)dx$ the minimal probability mass of the sets B_i (similarly by $b_{\max}$ the maximal probability mass), and by $N_{\min}$ and $N_{\max}$ the minimal (resp. maximal) number of sample points in the sets B_i. Then for all $\delta \in]0, 1]$*

$$\begin{aligned} P\Big(N_{\max} \geq (1+\delta)nb_{\max}\Big) \;&\leq\; K \cdot \exp(-\delta^2 nb_{\max}/3) \\ P\Big(N_{\min} \leq (1-\delta)nb_{\min}\Big) \;&\leq\; K \cdot \exp(-\delta^2 nb_{\min}/3). \end{aligned}$$

Proof. This is a straightforward application of Proposition 8 using the union bound. ☺

When working with ε-graphs or kNN-graphs, we often need to know the degrees of the vertices. As a rule of thumb, the expected degree of a vertex in the ε-graph is of the order $\Theta(n\varepsilon^d)$, the expected degree of a vertex in both the symmetric and mutual kNN graph is of the order $\Theta(k)$. The expected kNN-distance is of the order $\Theta((k/n)^{1/d})$. All these rules of thumb also apply to the minimal and maximal values of these quantities in the graph, provided the graph is "sufficiently connected". The following propositions make these rules of thumb explicit.

Proposition 10 (Degrees in the ε-graph) *Consider an ε-graph on a valid region $\mathcal{X} \subset \mathbb{R}^d$.*

1. *Then, for all $\delta \in]0, 1]$, the minimal and maximal degrees in the ε-graph satisfy*

$$\begin{aligned} P\Big(d_{\max} \geq (1+\delta)n\varepsilon^d p_{\max}\eta_d\Big) \;&\leq\; n \cdot \exp(-\delta^2 n\varepsilon^d p_{\max}\eta_d/3) \\ P\Big(d_{\min} \leq (1-\delta)n\varepsilon^d p_{\min}\eta_d\alpha\Big) \;&\leq\; n \cdot \exp(-\delta^2 n\varepsilon^d p_{\min}\eta_d\alpha/3). \end{aligned}$$

   In particular, if $n\varepsilon^d/\log n \to \infty$, then these probabilities converge to 0 as $n \to \infty$.

2. *If $n \to \infty, \varepsilon \to 0$ and $n\varepsilon^d/\log n \to \infty$, and the density p is continuous, then for each interior point $X_i \in \mathcal{X}$ the degree is a consistent density estimate: $d_i/(n\varepsilon^d\eta_d) \longrightarrow p(X_i)$ a.s.*

Proof. Part 1 follows by applying Proposition 9 to the balls of radius ε centered at the data points. Note that for the bound on $d_{\min}$, we need to take into account boundary effects as only a part of the

ε-ball around a boundary point is contained in $\mathcal{X}$. This is where the constant α comes in (recall the definition of α from the definition of a valid region). *Part 2* is a standard density estimation argument: the expected degree of X_i is the expected number of points in the ε-ball around X_i. For ε small enough, the density is approximately constant in this ball because we assumed the density to be continous. Then the expected number of points is approximately $n\varepsilon^d \eta_d p(X_i)$ where η_d denotes the volume of a d-dimensional unit ball. By concentration arguments it is easy to see that the actual number of points is close to this expectation, and that convergence holds under the conditions stated. ☺

Recall the definitions of the k-nearest neighbor radii: $R_k(x)$ denotes the distance of x to its k-nearest neighbor among the X_i, and the maximum and minimum values are denoted $R_{k,\max} := \max_{i=1,\dots,n} R_k(X_i)$ and $R_{k,\min} := \max_{i=1,\dots,n} R_k(X_i)$.

Proposition 11 (Degrees in the kNN-graph) *Consider a valid region $\mathcal{X} \subset \mathbb{R}^d$.*

1. *With probability at least $1 - n\exp(-c_1 k)$ the minimal and maximal kNN-radii satisfy*
$$R_{k,\min} \geq c_2 (k/n)^{1/d} \qquad \text{and} \qquad R_{k,\max} \leq c_3 (k/n)^{1/d}.$$
2. *Moreover, with probability at least $1 - n\exp(-c_4 k)$ the minimal and maximal degree in both the symmetric and mutual kNN graph are of the order $\Theta(k)$ (the constants differ).*
3. *If the density is continuous, $n \to \infty$, $k/n \to 0$ and $k/\log n \to \infty$, then in both the symmetric and the mutual kNN graph, $k/d_i \to 1$.*

Proof. *Part 1.* Define the constant $a = 1/(2p_{\max})$ and the radius $r := a\,(k/n)^{1/d}$, fix a sample point x, and denote by μ the probability mass of the ball around x with radius r. Set $\mu_{\max} := r^d \eta_d p_{\max} \geq \max_{x \in \mathcal{X}} \mu$. Note that if k/n is small enough, then $\mu_{\max} < 1$. The main idea is that $R_k(x) \leq r$ if and only if there are at least k data points in the ball of radius r around x. Let $M \sim Bin(n, \mu)$ and $V \sim Bin(n, \mu_{\max})$. Note that by the choices of a and r we have $E(V) = k/2$. All this leads to

$$P\Big(R_k(x) \leq r\Big) \;\leq\; P\Big(M \geq k\Big) \;\leq\; P\Big(V \geq k\Big) \;=\; P\Big(V \geq 2E(V)\Big).$$

Applying the concentration inequality of Proposition 8 and using a union bound leads to the following result for the minimal kNN radius:

$$P\Big(R_{k,\min} \leq a\left(\frac{k}{n}\right)^{1/d}\Big) \;\leq\; n\exp(-k/6).$$

By a similar approach we can prove the analogous statement for the maximal kNN radius. Note that for the bound on $R_{k,\max}$ we additionally need to take into account boundary effects: at the boundary of $\mathcal{X}$, only part of the ball around a point is contained in $\mathcal{X}$, which affects the value of $\mu_{\min}$. We thus define $\mu_{\min} = r^d \eta_d p_{\min} \alpha$ where α is the constant defined in the general assumptions. Then we continue similarly to above and get

$$P\Big(R_{k,\max} \geq \tilde{a}\left(\frac{k}{n}\right)^{1/d}\Big) \;\leq\; n\exp(-k/6).$$

Part 2. In the directed kNN graph, the degree of each vertex is exactly k. Thus, in the mutual kNN graph, the maximum degree over all vertices is upper bounded by k, in the symmetric kNN graph the minimum degree over all vertices is lower bounded by k.

For the symmetric graph, observe that the maximal degree in the graph is bounded by the maximal number of points in the balls of radius $R_{k,\max}$ centered at the data points. We know that with high probability, a ball of radius $R_{k,\max}$ contains of the order $\Theta(nR^d_{k,\max})$ points. Using Part 1 we know that with high probability, $R_{k,\max}$ is of the order $(k/n)^{1/d}$. Thus the maximal degree in the symmetric kNN graph is of the order $\Theta(k)$, with high probability.

In the mutual graph, observe that the minimal degree in the graph is bounded by the minimal number of points in the balls of radius $R_{k,\min}$ centered at the data points. Then the statement follows analogously to the last one.

Part 3, proof sketch. Consider a fixed point x in the interior of $\mathcal{X}$. We know that both in the symmetric and mutual kNN graph, two points cannot be connected if their distance is larger than $R_{k,\max}$. As we know that $R_{k,\max}$ is of the order $(k/n)^{1/d}$, under the growth conditions on n and k this radius becomes arbitrarily small. Thus, because of the continuity of the density, if n is large enough we can assume that the density in the ball $B(x, R_{k,\max})$ of radius $R_{k,\max}$ around x is approximately constant. Thus, all points $y \in B(x, R_{k,\max})$ have approximately the same expected k-nearest neighbor radius $R := (k/(n \cdot p(x)\eta_d))^{1/d}$. Moreover, by concentration arguments it is easy to see that the actual kNN radii only deviate by a factor $1 \pm \delta$ from their expected values.

Then, with high probability, all points inside of $B(x, R(1-\delta))$ are among the k nearest neighbors of x, and all k nearest neighbors of x are inside $B(x, R(1+\delta))$. On the other hand, with high probability x is among the k nearest neighbors of all points $y \in B(x, R(1-\delta))$, and not among the k nearest neighbors of any point outside of $B(x, R(1+\delta))$. Hence, in the mutual kNN graph, with high probability x is connected exactly to all points $y \in B(x, R(1-\delta))$. In the symmetric kNN graph, x might additionally be connected to the points in $B(x, R(1+\delta)) \setminus B(x, R(1-\delta))$. By construction, with high probability the number of sample points in these balls is $(1+\delta)k$ and $(1-\delta)k$. Driving δ to 0 leads to the result. ☺

4 Finally completing the proof of Theorems 2 and 3

First of all, note that by Rayleigh's principle (cf. Corollary 7 in Section IX.2 of Bollobás, 1998) the effective resistance between vertices cannot decrease if we delete edges from the graph. So given a sample from the underlying density p, some random geometric graph based on this sample, and some valid region $\mathcal{X}$, we first delete all points that are not in $\mathcal{X}$. Then we consider the remaining geometric graph. The effective resistances on this graph are upper bounds on the resistances of the original graph. Then we conclude the proofs with the following arguments:

Proof of Theorem 3. The lower bound on the deviation follows immediately from Proposition 4. The upper bound is a consequence of Proposition 7 and the properties of random geometric graphs presented above. In particular, note that we can choose the grid width $g := \frac{\varepsilon}{2\sqrt{d-1}}$ to obtain a valid grid. The quantity $N_{\min}$ can be bounded as stated in Proposition 9 and the degrees behave as described in Proposition 10 (we use $\delta = 1/2$ in these results for simplicity). Plugging all these results together leads to the following statement:

$$
\begin{aligned}
&\left| n\varepsilon^d R_{ij} - \left(\frac{n\varepsilon^d}{d_i} + \frac{n\varepsilon^d}{d_j} \right) \right| \\
&\leq \frac{2^{d+2}(d-1)^{d/2}}{p_{\min}^2 \eta_d \alpha} \cdot \frac{1}{n\varepsilon^d} + \begin{cases} \frac{2^{2d+2}(d-1)^d}{p_{\min}^2} \cdot \frac{1}{n\varepsilon^d} \left(\frac{\log(h/\varepsilon)+1}{d-1} + \frac{\sqrt{d-1}\; d(s,t)}{2^{d-3}h^{d-1}} \cdot \varepsilon^{d-2} \right) & \text{if } d = 3 \\ \frac{2^{2d+2}(d-1)^d}{p_{\min}^2} \cdot \frac{1}{n\varepsilon^d} \left(\frac{2}{d-1} + \frac{\sqrt{d-1}\; d(s,t)}{2^{d-3}h^{d-1}} \cdot \varepsilon^{d-2} \right) & \text{if } d > 3 \end{cases} \\
&\approx \begin{cases} c_5 \frac{1+\log(1/\varepsilon)+\varepsilon}{n\varepsilon^3} & \text{if } d = 3 \\ c_6 \frac{1}{n\varepsilon^d} & \text{if } d > 3 \end{cases}
\end{aligned}
$$

☺

The constants in this result look terrible with respect to the dimension d. However, we believe that this is due to the fact that our setup is an unfortunate mix between balls and cubes (the constants come from the fact that we have to ensure that the diagonals of cubes still fit into certain balls). We hope to tidy up the proof in future and provide better constants. The same goes for the next result which we prove now.

Proof of Theorem 2. This proof is completely analogous to the ε-graph. As grid width g we choose $g = R_{k,\min}/(2\sqrt{d-1})$ where $R_{k,\min}$ is the minimal k-nearest neighbor distance (note that this works for both the symmetric and the mutual kNN-graph). Then the statements of the theorem follow by combining Propositions 7, 9 and 11 (by $c := 1/(2p_{\max})$ we denote the constant hidden in the O-notation of Proposition 11 for $R_{k,\min}$). The result is

$$\begin{aligned}
&\left| kR_{ij} - \left(\frac{k}{d_i} + \frac{k}{d_j} \right) \right| \\
&\leq \frac{2^{d+2}(d-1)^{d/2}}{p_{\min}^2 \; c \; c^d} \cdot \frac{1}{k} + \begin{cases} \frac{2^{2d+2}(d-1)^d}{p_{\min}^2 c^{2d}} \cdot \frac{1}{k} \left(\frac{\log(h/(c(k/n)^{1/d})+1}{d-1} + \frac{\sqrt{d-1}\; d(s,t) c^{d-2}}{2^{d-3} h^{d-1}} \cdot (k/n)^{(d-2)/d} \right) & \text{if } d = 3 \\ \frac{2^{2d+2}(d-1)^d}{p_{\min}^2 c^{2d}} \cdot \frac{1}{k} \left(\frac{2}{d-1} + \frac{\sqrt{d-1}\; d(s,t) c^{d-2}}{2^{d-3} h^{d-1}} \cdot (k/n)^{(d-2)/d} \right) & \text{if } d > 3 \end{cases} \\
&\approx \begin{cases} c_4 \frac{1}{k} \left(1 + \log(n/k) + (k/n)^{1/3} \right) & \text{if } d = 3 \\ c_5 \frac{1}{k} & \text{if } d > 3 \end{cases}
\end{aligned}$$

☺

References

D. Angluin and L. G. Valiant. Fast probabilistic algorithms for Hamiltonian circuits and matchings. In *STOC*, 1977.

B. Bollobás. *Modern Graph Theory*. Springer, 1998.

R. Lyons and Y. Peres. Probability on trees and networks. *Book in preparation, available online on the webpage of Yuval Peres*, 2010.