

# Probabilistic Analysis of Optimization Problems on Sparse Random Shortest Path Metrics

**Stefan Klootwijk**

Department of Applied Mathematics, University of Twente, Enschede, The Netherlands  
<https://people.utwente.nl/s.klootwijk>  
s.klootwijk@utwente.nl

**Bodo Manthey**

Department of Applied Mathematics, University of Twente, Enschede, The Netherlands  
<https://wwwhome.ewi.utwente.nl/~mantheyb/>  
b.manthey@utwente.nl

---

## Abstract

---

Simple heuristics for (combinatorial) optimization problems often show a remarkable performance in practice. Worst-case analysis often falls short of explaining this performance. Because of this, “beyond worst-case analysis” of algorithms has recently gained a lot of attention, including probabilistic analysis of algorithms.

The instances of many (combinatorial) optimization problems are essentially a discrete metric space. Probabilistic analysis for such metric optimization problems has nevertheless mostly been conducted on instances drawn from Euclidean space, which provides a structure that is usually heavily exploited in the analysis. However, most instances from practice are not Euclidean. Little work has been done on metric instances drawn from other, more realistic, distributions. Some initial results have been obtained in recent years, where random shortest path metrics generated from dense graphs (either complete graphs or Erdős–Rényi random graphs) have been used so far.

In this paper we extend these findings to sparse graphs, with a focus on grid graphs. A random shortest path metric is constructed by drawing independent random edge weights for each edge in the graph and setting the distance between every pair of vertices to the length of a shortest path between them with respect to the drawn weights. For such instances generated from a grid graph, we prove that the greedy heuristic for the minimum distance maximum matching problem, and the nearest neighbor and insertion heuristics for the traveling salesman problem all achieve a constant expected approximation ratio. Additionally, for instances generated from an arbitrary sparse graph, we show that the 2-opt heuristic for the traveling salesman problem also achieves a constant expected approximation ratio.

**2012 ACM Subject Classification** Mathematics of computing → Approximation algorithms; Mathematics of computing → Combinatorial optimization; Theory of computation → Approximation algorithms analysis; Theory of computation → Random network models

**Keywords and phrases** Random shortest paths, Random metrics, Approximation algorithms, First-passage percolation

**Digital Object Identifier** 10.4230/LIPIcs.AofA.2020.19

**Funding** This research was supported by NWO grant 613.001.402.

## 1 Introduction

Large-scale optimization problems, such as the traveling salesman problem (TSP), are relevant for many applications. Often it is not possible to solve these problems to optimality within a reasonable amount of time, especially when instances get larger. Therefore, in practice these kind of problems are tackled by using approximation algorithms or ad-hoc heuristics. Even though the worst-case performance of these, often simple, heuristics is usually rather bad, they often show a remarkably good performance in practice.



© Stefan Klootwijk and Bodo Manthey;  
licensed under Creative Commons License CC-BY

31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2020).

Editors: Michael Drmota and Clemens Heuberger; Article No. 19; pp. 19:1–19:16



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

In order to find theoretical results that are closer to the practical observations, probabilistic analysis has been a useful tool over the last decades. One of the main challenges here is to choose a probability distribution on the set of possible instances of the problem: on the one hand this distribution should be sufficiently simple in order to make the probabilistic analysis possible, but on the other hand the distribution should somehow reflect realistic instances.

In the “early days” of probabilistic analysis, random instances were either generated by using independent random edge lengths or embedded in Euclidean space (e.g. [2, 11]). Although these models have some nice mathematical properties that enable the probabilistic analysis, they have shortcomings regarding their realism: in practice, instances are often metric, but not Euclidean, and independent random edge lengths are not even metric.

Recently, Bringmann et al. [6] widened the scope of models for generating random instances by using the following model, already proposed by Karp and Steele in 1985 [17]: given an undirected complete graph, draw edge weights independently at random and then define the distance between any two vertices as the total weight of the shortest path between them, measured with respect to those random weights. Even though this model broadens the scope of random metric spaces, the resulting instances from this model are not very realistic.

In this paper we adapt this model in the sense that we start with a sparse graph instead of a complete graph. We believe that this yields instances that are more realistic, for instance since in practice the underlying (road, communication, etc.) networks are almost always sparse.

## Related Work

The model described above is known by two different names: *random shortest path metrics* and *first-passage percolation*. It was introduced by Hammersley and Welsh under the latter name as a model for fluid flow through a (random) porous medium [12, 14]. A lot of studies have been conducted on first-passage percolation, mostly on this model defined on the lattice  $\mathbb{Z}^d$ .

For first-passage percolation on complete graphs many structural results exist. We know for instance that the expected distance between two arbitrary fixed vertices is approximately  $\ln(n)/n$  and that the distance from a fixed vertex to the vertex that is farthest away from it is approximately  $2 \ln(n)/n$  [6, 15]. We also know that the diameter in this case is approximately  $3 \ln(n)/n$  [13, 15]. Bringmann et al. used this model to analyze heuristics for matching, TSP and  $k$ -median [6].

There has been a lot of interest in first-passage percolation on the integer lattice  $\mathbb{Z}^d$ . Although very few precise results are known for this model, there are many existential results available. For instance, the distance between the origin and  $n\mathbf{e}_1$  (where  $\mathbf{e}_1$  is the unit vector in the first coordinate direction) is known to be  $\Theta(n)$  [12]. Also, the set of vertices within distance  $t$  from the origin grows linearly in  $t$  and, after rescaling, converges to some convex domain [20]. The survey by Auffinger et al. [1] contains a thorough overview.

## Our Results

This paper aims at extending the results of Bringmann et al. [6] and Klootwijk et al. [18] to the more realistic setting of random shortest path metrics generated from sparse graphs. For simplicity, most of the results in this paper assume that these sparse graphs are (finite square) grid graphs. We believe that the probabilistic analysis of simple heuristics in different random models will enhance the understanding of the performance of these heuristics, which are used in many applications.

In this paper we consider two different types of simple heuristics. In Section 4 we conduct a probabilistic analysis of three greedy-like heuristics: the greedy heuristic for the minimum-distance perfect matching problem, and the nearest neighbor heuristic and insertion heuristic for the TSP. In Section 5 we conduct a probabilistic analysis of a local search heuristic: the 2-opt heuristic for the TSP. We show that all four heuristics yield a constant approximation ratio for random shortest path metrics generated from (finite square) grid graphs (greedy-like in Section 4) or arbitrary sparse graphs (local search in Section 5). We are aware that our results are mostly purely theoretical, because, e.g., cheapest insertion already achieves an approximation ratio of 2 and is often used to initialize 2-opt [10, 21]. However, they are non-trivial results about practically used algorithms, beyond the classical worst-case analysis.

## 2 Notation and Model

For  $n \in \mathbb{N}$ , we use  $[n]$  as shorthand notation for  $\{1, \dots, n\}$ . Sometimes we use  $\exp(\cdot)$  to denote the exponential function. We denote by  $X \sim P$  that a random variable  $X$  is distributed according to a probability distribution  $P$ , where in particular  $\text{Exp}(\lambda)$  denotes the exponential distribution with parameter  $\lambda$ . We write  $X \sim \sum_{i=1}^n \text{Exp}(\lambda_i)$  if  $X$  is the sum of  $n$  independent exponentially distributed random variables having parameters  $\lambda_1, \dots, \lambda_n$ . In particular,  $X \sim \sum_{i=1}^n \text{Exp}(\lambda)$  denotes an Erlang distribution with parameters  $n$  and  $\lambda$ . If a random variable  $X_1$  is stochastically dominated by a random variable  $X_2$ , i.e., we have  $F_1(x) \geq F_2(x)$  for all  $x$  (where  $F_i$  is the distribution function of  $X_i$ ), we denote this by  $X_1 \preceq X_2$ .

### Random Shortest Path Metrics

Given an undirected connected graph  $G = (V, E)$ , the corresponding random shortest path metric is constructed as follows. First, for each edge  $e \in E$ , draw a random edge weight  $w(e)$  independently according to the exponential distribution with parameter 1. (Exponential distributions are technically easiest to handle since they are memoryless.) Then, define the distance function  $d : V \times V \rightarrow \mathbb{R}_{\geq 0}$  as follows: for each  $u, v \in V$ ,  $d(u, v)$  is the total weight of a lightest  $u, v$ -path in  $G$  (w.r.t. to the random weights  $w(\cdot)$ ). Observe that this definition immediately implies that  $d(v, v) = 0$  for all  $v \in V$ , that  $d(u, v) = d(v, u)$  for all  $u, v \in V$ , and that  $d(u, v) \leq d(u, s) + d(s, v)$  for all  $u, s, v \in V$ . We call the distance function  $d$  obtained by this process a random shortest path metric generated from  $G$ .

We use the following notation to denote properties of these random shortest path metrics.  $\Delta_{\max} := \max_{u, v} d(u, v)$  denotes the diameter of the random metric. The  $\Delta$ -ball around a vertex  $v$ ,  $B_{\Delta}(v) := \{u \in V \mid d(u, v) \leq \Delta\}$ , is the set of vertices within distance  $\Delta$  of  $v$ . Let  $\pi_k(v)$  denote the  $k$ th closest vertex from  $v$  (including  $v$  itself and breaking ties arbitrarily). Note that  $v = \pi_1(v)$  for all  $v \in V$ . The distance from a vertex  $v$  to this  $k$ th closest vertex from it is denoted by  $\tau_k(v) := d(v, \pi_k(v)) = \min\{\Delta \mid |B_{\Delta}(v)| \geq k\}$ . Slightly abusing notation, we let  $B_{\tau_k(v)}(v) := \{\pi_i(v) \mid i = 1, \dots, k\}$  denote the set of the  $k$  closest vertices to  $v$  (including  $v$  itself). The size of the cut in  $G$  induced by this set, which plays an important role in our analysis, is denoted by  $\chi_k(v) := |\delta(B_{\tau_k(v)}(v))|$ , where  $\delta(U) := \{\{u, v\} \in E \mid u \in U, v \notin U\}$  denotes the cut induced by  $U$ .

During our analysis in Sections 4 and 5 it is convenient to describe (partial) solutions to the minimum-distance perfect matching problem and the TSP as sets of “edges”. In order to emphasize that such “edges” in principle do not coincide with edges from  $G$ , we use quotation marks to distinguish them.

### Sparse Graphs

Throughout this paper,  $G$  is a sparse connected undirected simple graph on  $n$  vertices, i.e., we have  $|E| = \Theta(|V|) = \Theta(n)$ . Most of the results in this paper are restricted to a specific class of sparse graphs, namely the finite square grid graph. This graph has vertex set  $V = [N]^2$ , and two vertices  $(x_1, y_1), (x_2, y_2) \in V$  are connected if and only if  $|x_1 - x_2| + |y_1 - y_2| = 1$ . It is easy to see that for these graphs we have  $|V| = n = N^2$  and  $|E| = 2N^2 - 2N = \Theta(n)$ .

For practical reasons we assume in this paper that  $n$  is even. Note that the results concerning the minimum-distance perfect matching problem are only valid when  $n$  is even. All results concerning the travelling salesman problem can easily be extended to odd  $n$ .

## 3 Structural Properties

In this section, we provide some structural properties regarding sparse random shortest path metrics that are used later on in our probabilistic analyses of the greedy heuristic for maximum matching and the 2-opt heuristic for the TSP in such random metric spaces. We start of with some technical lemmas from known literature and some results regarding sums of lightest edge weights in  $G$  (which hold for arbitrary sparse graphs). After that, we consider a random growth process that is closely related to the special case of random shortest path metrics generated from a finite square grid graph and use it to derive a clustering result and a tail bound on the diameter  $\Delta_{\max}$ .

### Technical Lemmas

► **Lemma 1** ([16, Thm. 5.1(i,iii)]). *Let  $X \sim \sum_{i=1}^m \text{Exp}(a_i)$ . Let  $\mu = \mathbb{E}[X] = \sum_{i=1}^m 1/a_i$  and  $a_* = \min_i a_i$ .*

(i) *For any  $\lambda \geq 1$ ,*

$$\mathbb{P}(X \geq \lambda\mu) \leq \lambda^{-1} \exp(-a_*\mu(\lambda - 1 - \ln(\lambda))).$$

(iii) *For any  $\lambda \leq 1$ ,*

$$\mathbb{P}(X \leq \lambda\mu) \leq \exp(-a_*\mu(\lambda - 1 - \ln(\lambda))).$$

► **Corollary 2.** *Let  $X \sim \sum_{i=1}^m \text{Exp}(a_i)$ . Let  $\mu = \mathbb{E}[X] = \sum_{i=1}^m 1/a_i$  and  $a_* = \min_i a_i$ . For any  $x$ ,*

$$\mathbb{P}(X \leq x) \leq \exp(a_*\mu(1 + \ln(x/\mu))).$$

**Proof.** Let  $\lambda := x/\mu$ . If  $\lambda \leq 1$ , the result is a weaker version of Lemma 1(iii). If  $\lambda > 1$ , then  $1 + \ln(x/\mu) > 0$  and hence  $\mathbb{P}(X \leq x) \leq 1 < \exp(a_*\mu(1 + \ln(x/\mu)))$ . ◀

► **Lemma 3** ([5, Thm. 2(ii)]). *Let  $X \sim \sum_{i=1}^m \text{Exp}(\lambda_i)$  and  $Y \sim \sum_{i=1}^m \text{Exp}(\eta)$ . Then*

$$X \succsim Y \quad \text{if and only if} \quad \prod_{i=1}^m \lambda_i \leq \eta^m.$$

### Sums of Lightest Edge Weights in $G$

All main results in this paper make use of some observations relating sums of the  $m$  lightest edge weights in a sparse graph  $G$ . The lemmas and corollary below summarize some structural properties concerning these sums. They hold for arbitrary sparse graphs  $G$ .

► **Lemma 4.** *Let  $S_m$  denote the sum of the  $m$  lightest edge weights in  $G$ . Then*

$$\sum_{i=0}^{m-1} \text{Exp}\left(\frac{e|E|}{m}\right) \lesssim S_m \lesssim \sum_{i=0}^{m-1} \text{Exp}\left(\frac{|E|}{m}\right).$$

**Proof.** Let  $\sigma_k$  denote the  $k$ th lightest edge weight in  $G$ . Since all edge weights are independent and standard exponentially distributed, we have  $\sigma_1 = S_1 \sim \text{Exp}(|E|)$ . Using the memorylessness property of the exponential distribution, it follows that  $\sigma_2 \sim \sigma_1 + \text{Exp}(|E| - 1)$ , i.e., the second lightest edge weight is equal to the lightest edge weight plus the minimum of  $|E| - 1$  standard exponential distributed random variables. In general, we get  $\sigma_{k+1} \sim \sigma_k + \text{Exp}(|E| - k)$ . The definition  $S_m = \sum_{k=1}^m \sigma_k$  yields

$$S_m \sim \sum_{i=0}^{m-1} (m-i) \cdot \text{Exp}(|E| - i) \sim \sum_{i=0}^{m-1} \text{Exp}\left(\frac{|E| - i}{m - i}\right).$$

Now, the first stochastic dominance relation follows from Lemma 3 by observing that

$$\prod_{i=0}^{m-1} \frac{|E| - i}{m - i} = \frac{|E|!}{m!(|E| - m)!} = \binom{|E|}{m} \leq \left(\frac{e|E|}{m}\right)^m,$$

where the inequality follows from applying the well-known inequality  $\binom{n}{k} \leq (en/k)^k$ .

The second stochastic dominance relation follows by observing that  $|E| \geq m$ , which implies that  $(|E| - i)/(m - i) \geq |E|/m$  for all  $i = 0, \dots, m - 1$ . ◀

► **Corollary 5.** *Let  $S_m$  denote the sum of the  $m$  lightest edge weights in  $G$ . Then  $\mathbb{E}[S_m] = \Theta(m^2/n)$ .*

**Proof.** From Lemma 4 we can immediately see that

$$\mathbb{E}\left[\sum_{i=0}^{m-1} \text{Exp}\left(\frac{e|E|}{m}\right)\right] \leq \mathbb{E}[S_m] \leq \mathbb{E}\left[\sum_{i=0}^{m-1} \text{Exp}\left(\frac{|E|}{m}\right)\right].$$

The result follows by observing that

$$\mathbb{E}\left[\sum_{i=0}^{m-1} \text{Exp}\left(\frac{e|E|}{m}\right)\right] = \frac{m^2}{e|E|} \quad \text{and} \quad \mathbb{E}\left[\sum_{i=0}^{m-1} \text{Exp}\left(\frac{|E|}{m}\right)\right] = \frac{m^2}{|E|},$$

and recalling that  $|E| = \Theta(n)$  by our restrictions imposed on  $G$ . ◀

► **Lemma 6.** *Let  $S_m$  denote the sum of the  $m$  lightest edge weights in  $G$ . Then we have*

$$\mathbb{P}(S_m \leq cn) \leq \exp\left(m \left(2 + \ln\left(\frac{c|E|n}{m^2}\right)\right)\right).$$

**Proof.** First of all, Lemma 4 yields

$$S_m \lesssim \sum_{i=0}^{m-1} \text{Exp}\left(\frac{e|E|}{m}\right).$$

Now, we apply Corollary 2 with  $\mu = m^2/e|E|$ ,  $a_* = e|E|/m$ , and  $x = cn$  to obtain

$$\mathbb{P}(S_m \leq cn) \leq \mathbb{P}\left(\sum_{i=0}^{m-1} \text{Exp}\left(\frac{e|E|}{m}\right) \leq cn\right) \leq \exp\left(m \left(1 + \ln\left(\frac{ce|E|n}{m^2}\right)\right)\right).$$

The result follows immediately. ◀

► **Lemma 7.** *Let  $S_m$  denote the sum of the  $m$  lightest edge weights in  $G$ . Then we have  $\text{TSP} \geq \text{MM} \geq S_{n/2}$ , where  $\text{TSP}$  and  $\text{MM}$  are the total distance of a shortest TSP tour and a minimum-distance perfect matching, respectively.*

**Proof.** The first inequality is trivial. For the second one, consider a minimum-distance perfect matching in  $G$ , and take the union of the shortest paths between each pair of matched vertices. This union  $H$  must contain at least  $n/2$  different edges of  $G$ , since  $H$  is a forest with  $n$  vertices, all of which are non-isolated in  $H$ . These  $n/2$  different edges together have a weight of at least  $S_{n/2}$  and at most  $\text{MM}$ . So, the second inequality follows. ◀

### A Random Growth Process

In this subsection, and the following ones, we assume that  $G$  is a finite square grid graph with  $n = N^2$  vertices. It is possible to obtain qualitatively similar results for more general classes of sparse graphs, but in order to improve readability we chose to stick with grid graphs.

In order to understand sparse random shortest path metrics it is important to get a feeling for the distribution  $\tau_k(v)$ . However, this distribution depends heavily on the exact position of  $v$  within  $G$ , which makes it rather complicated to derive it. In order to overcome this, we derive instead a stochastic upper bound on  $\tau_k(v)$  which holds for any vertex  $v \in V$ . The following lemma and corollary establish this.

► **Lemma 8** ([4, Thm. 3]). *Let  $U \subseteq V$ . Then we have*

$$|\delta(U)| \geq \begin{cases} 2\sqrt{|U|} & \text{if } |U| \leq n/4, \\ \sqrt{n} & \text{if } n/4 \leq |U| \leq 3n/4, \\ 2\sqrt{n - |U|} & \text{if } |U| \geq 3n/4. \end{cases}$$

► **Remark.** Bollobás and Leader [4] proved (a more general version of) this result for all  $|U| \leq n/2$ . The result for  $|U| > n/2$  follows immediately from their result by observing that  $\delta(U) = \delta(V \setminus U)$  and  $|V \setminus U| = n - |U|$  for all  $U \subseteq V$ .

► **Corollary 9.** *For any  $v \in V$  and any  $k \leq n/4$  we have*

$$\tau_k(v) \lesssim \sum_{i=1}^{k-1} \text{Exp}(2\sqrt{i}).$$

**Proof.** The values of  $\tau_k(v)$  are generated by a birth process as follows. (Similar birth processes have been analysed before (e.g. [6, 8, 18]).) For  $k = 1$  we have  $\tau_k(v) = 0$  and also  $\sum_{i=1}^{k-1} \text{Exp}(2\sqrt{i}) = 0$ . For  $k \geq 2$ , we can obtain  $\tau_k(v)$  from  $\tau_{k-1}(v)$  by looking at all edges that “leave”  $B_{\tau_{k-1}(v)}(v)$ , i.e., edges  $(u, x)$  with  $u \in B_{\tau_{k-1}(v)}(v)$  and  $x \notin B_{\tau_{k-1}(v)}(v)$ . By definition there are  $\chi_{k-1}(v)$  such edges, and from Lemma 8 it follows that  $\chi_{k-1}(v) \geq 2\sqrt{k-1}$  (since  $k \leq n/4$ ). Moreover, by definition of  $\tau_{k-1}(v)$  these edges are conditioned to have a length of at least  $\tau_{k-1}(v) - d(v, u)$ . Using the memorylessness of the exponential distribution, it follows that  $\tau_k(v) - \tau_{k-1}(v)$  is the minimum of  $\chi_{k-1}(v)$  exponential random variables (with parameter 1), or, equivalently,  $\tau_k(v) - \tau_{k-1}(v) \sim \text{Exp}(\chi_{k-1}(v)) \lesssim \text{Exp}(2\sqrt{k-1})$ , where the stochastic dominance follows since  $\chi_{k-1}(v) \geq 2\sqrt{k-1}$ . The result follows by induction. ◀

Now we use this stochastic upper bound on  $\tau_k(v)$  that holds for any  $v \in V$  to derive some bounds on the cumulative distribution functions of  $\tau_k(v)$  and  $|B_\Delta(v)|$ . The final bound is a crucial ingredient for the construction of clusterings in the next section.

► **Lemma 10.** For any  $\Delta > 0$ ,  $v \in V$ , and  $k \in [n]$  such that  $k \leq \min\{n/4, \Delta^2 + 1\}$ , we have

$$\mathbb{P}(\tau_k(v) \leq \Delta) \geq 1 - \frac{\sqrt{k-1}}{\Delta} \cdot \exp\left(-2(\sqrt{k}-1) \left(\frac{\Delta}{\sqrt{k-1}} - 1 - \ln\left(\frac{\Delta}{\sqrt{k-1}}\right)\right)\right).$$

**Proof.** From Corollary 9 we can see that

$$\mathbb{P}(\tau_k(v) \leq \Delta) \geq \mathbb{P}\left(\sum_{i=1}^{k-1} \text{Exp}(2\sqrt{i}) \leq \Delta\right) = 1 - \mathbb{P}\left(\sum_{i=1}^{k-1} \text{Exp}(2\sqrt{i}) \geq \Delta\right).$$

Next, we want to apply the result of Lemma 1(i). For this purpose, set

$$\mu := \mathbb{E}\left[\sum_{i=1}^{k-1} \text{Exp}(2\sqrt{i})\right] = \sum_{i=1}^{k-1} \frac{1}{2\sqrt{i}} \quad \text{and} \quad \lambda := \frac{\Delta}{\mu},$$

and observe that  $\sqrt{k}-1 \leq \mu \leq \sqrt{k-1}$ . Also note that  $\lambda = \Delta/\mu \geq \Delta/\sqrt{k-1} \geq 1$  since  $k \leq \Delta^2 + 1$ . Lemma 1(i) now yields

$$1 - \mathbb{P}\left(\sum_{i=1}^{k-1} \text{Exp}(2\sqrt{i}) \geq \Delta\right) \geq 1 - \lambda^{-1} \exp(-2\mu(\lambda - 1 - \ln(\lambda))).$$

It can now be seen that this final expression is increasing in both  $\mu$  and  $\lambda$ . Therefore, we may apply the inequalities  $\mu \geq \sqrt{k}-1$  and  $\lambda \geq \Delta/\sqrt{k-1}$  to obtain the desired result. ◀

► **Lemma 11.** For any  $\Delta > 0$ ,  $v \in V$ , and  $k \in [n]$  such that  $k \leq \min\{n/4, \Delta^2 + 1\}$ , we have

$$\mathbb{P}(|B_\Delta(v)| \geq k) \geq 1 - \frac{\sqrt{k-1}}{\Delta} \cdot \exp\left(-2(\sqrt{k}-1) \left(\frac{\Delta}{\sqrt{k-1}} - 1 - \ln\left(\frac{\Delta}{\sqrt{k-1}}\right)\right)\right).$$

**Proof.** This lemma follows immediately from Lemma 10 by observing that  $|B_\Delta(v)| \geq k$  if and only if  $\tau_k(v) \leq \Delta$ . ◀

► **Corollary 12.** Let  $n \geq 9$ . There exists a constant  $c_1 \geq 4$  such that for any  $\Delta > 0$  and  $v \in V$  we have

$$\mathbb{P}\left(|B_\Delta(v)| < \min\left\{\frac{\Delta^2}{4}, \frac{n}{4}\right\}\right) \leq \frac{c_1}{\Delta^2}.$$

**Proof.** First of all, observe that for  $\Delta \leq 2$  the statement is trivial since in that case we have  $c_1/\Delta^2 \geq 1$ . Therefore, from now on assume w.l.o.g. that  $\Delta > 2$ . Let  $s_\Delta := \min\{\Delta^2/4, n/4\}$ . Using Lemma 11 with  $k = s_\Delta$  we obtain

$$\mathbb{P}(|B_\Delta(v)| < s_\Delta) \leq \frac{\sqrt{s_\Delta-1}}{\Delta} \cdot \exp\left(-2(\sqrt{s_\Delta}-1) \left(\frac{\Delta}{\sqrt{s_\Delta-1}} - 1 - \ln\left(\frac{\Delta}{\sqrt{s_\Delta-1}}\right)\right)\right).$$

So, it remains to show that there exists a constant  $c_1$  such that for any  $\Delta > 2$  we have

$$\Delta\sqrt{s_\Delta-1} \cdot \exp\left(-2(\sqrt{s_\Delta}-1) \left(\frac{\Delta}{\sqrt{s_\Delta-1}} - 1 - \ln\left(\frac{\Delta}{\sqrt{s_\Delta-1}}\right)\right)\right) \leq c_1.$$

The tedious computations needed to show that this is true can be found in Appendix A. ◀

### Clustering

The following theorem shows that we can partition the vertices of sparse random shortest path metrics into a suitably small number of clusters with a given maximum diameter. Its proof follows closely the ideas of Bringmann et al. [6], albeit with a different value of  $s_\Delta$ .

► **Theorem 13.** *Let  $G$  be a finite square grid graph on  $n = N^2$  vertices, consider a (sparse) random shortest path metric generated using this graph, and let  $\Delta > 0$ . There exists a partition of vertices into clusters, each of diameter at most  $4\Delta$ , such that the expected number of clusters needed is bounded from above by  $O(1 + n/\Delta^2)$ , where constant of the  $O$ -term is uniform with respect to  $\Delta$ .*

**Proof.** Let  $n$  be sufficiently large ( $n \geq 9$  suffices) and let  $s_\Delta := \min\{\Delta^2/4, n/4\}$ , as in Corollary 12. We call vertex  $v$   $\Delta$ -dense if  $|B_\Delta(v)| \geq s_\Delta$  and  $\Delta$ -sparse otherwise. Using Corollary 12 we can bound the expected number of  $\Delta$ -sparse vertices by  $O(n/\Delta^2)$ . We put each  $\Delta$ -sparse vertex in its own cluster (of size 1), which has diameter  $0 \leq 4\Delta$ .

Now, only the  $\Delta$ -dense vertices remain. We cluster them according to the following process. Consider an auxiliary graph  $H$  whose vertices are the  $\Delta$ -dense vertices and where two vertices  $u, v$  are connected by an edge if and only if  $B_\Delta(u) \cap B_\Delta(v) \neq \emptyset$ . Consider an arbitrary maximal independent set  $S$  in  $H$ , and observe that  $|S| \leq n/s_\Delta$  by construction of  $H$ . We create the initial clusters  $C_1, \dots, C_{|S|}$ , each of which equals  $B_\Delta(v)$  for some vertex  $v \in S$ . These initial clusters have diameter at most  $2\Delta$ .

Next, consider an arbitrary  $\Delta$ -dense vertex  $v$  that is not yet part of any cluster. By the maximality of  $S$ , we know that there must exist a vertex  $u \in S$  such that  $A := B_\Delta(u) \cap B_\Delta(v) \neq \emptyset$ . Let  $x \in A$  be arbitrarily chosen, and observe that  $d(v, u) \leq d(v, x) + d(x, u) \leq \Delta + \Delta = 2\Delta$ . We add  $v$  to the initial cluster corresponding to  $u$ , and repeat this step until all  $\Delta$ -dense vertices have been added to some initial cluster. By construction, the diameter of all these clusters is now at most  $4\Delta$ : consider two arbitrary vertices  $w, y$  in a cluster that initially corresponded to  $u \in S$ ; then we have  $d(w, y) \leq d(w, u) + d(u, y) \leq 2\Delta + 2\Delta = 4\Delta$ .

So, now we have in expectation at most  $O(n/\Delta^2)$  clusters containing one ( $\Delta$ -sparse) vertex each, and at most  $n/s_\Delta = O(1 + n/\Delta^2)$  clusters containing at least  $s_\Delta$  ( $\Delta$ -dense) vertices each, all with diameter at most  $4\Delta$ . The result follows. ◀

### A Tail Bound for $\Delta_{\max}$

Recall that  $\Delta_{\max} = \max_{u,v} d(u, v)$  is the diameter of the random metric. The following lemma shows that  $\Delta_{\max} \leq O(\sqrt{n})$  with high probability. Due to space constraints, its proof can be found in Appendix B.

► **Lemma 14.** *Let  $x \geq 9\sqrt{n}$ . Then we have  $\mathbb{P}(\Delta_{\max} \geq x) \leq ne^{-x}$ .*

## 4 Analyses of Greedy-like Heuristics for Matching and TSP

In this section, we show that three greedy-like heuristics (greedy for minimum-distance perfect matching, and nearest neighbor and insertion for TSP) achieve a constant expected approximation ratio on sparse random shortest path metrics generated from a finite square grid graph. The three proofs are very alike, and the ideas behind them are built upon ideas by Bringmann et al. [6]: we divide the steps of the greedy-like heuristics into bins, depending on the value which they add to the total distance of our matching or TSP tour. Using the clustering (Theorem 13) we bound the total contribution of these bins by  $O(n)$ , and using our observation regarding sums of lightest edge weights (Lemmas 6 and 7) we show that the optimal matching or TSP tour has a distance of  $\Omega(n)$  with sufficiently high probability.



### Greedy Heuristic for Minimum-Distance Perfect Matching

The first problem that we consider is the minimum-distance perfect matching problem. Even though solving the minimum-distance perfect matching problem to optimality is not very difficult (it can be done in  $O(n^3)$  time), in practice this is often too slow, especially if the number of vertices is large. Therefore, people often rely on (simple) heuristics to solve this problem in practical situations. The greedy heuristic is arguably the simplest one among these heuristics. It starts with an empty matching and iteratively adds a pair of currently unmatched vertices (an “edge”) to the matching such that the distance between them is minimal. Let GR denote the total distance of the matching computed by the greedy heuristic, and let MM denote the total distance of an optimal matching.

It is known that the worst-case approximation ratio for this heuristic on metric instances is  $O(n^{\log_2(3/2)})$  [19]. Moreover, for random Euclidean instances, the greedy heuristic has an approximation ratio of  $O(1)$  with high probability [2], and for random shortest path metrics generated from complete graphs or Erdős–Rényi random graphs the expected approximation ratio of the greedy heuristic is  $O(1)$  as well [6, 18]. We show that a similar result holds for sparse random shortest path metrics generated from a finite square grid graph.

► **Theorem 15.**  $\mathbb{E}[\text{GR}] = O(n)$ .

**Proof.** We put “edges” that are being added to the greedy matching into bins according to their distance: bin  $i$  receives all “edges”  $\{u, v\}$  with  $d(u, v) \in (4(i-1), 4i]$ . Let  $X_i$  denote the number of “edges” that end up in bin  $i$  and set  $Y_i := \sum_{k=i}^{\infty} X_k$ , i.e.,  $Y_i$  is the number of “edges” in the greedy matching with distance at least  $4(i-1)$ . Observe that  $Y_1 = n/2$ . For  $i > 1$ , by Theorem 13, we can partition the vertices in an expected number of  $O(1 + n/(i-1)^2)$  clusters (where the constant of the  $O$ -term is uniform with respect to  $i$ ), each of diameter at most  $4(i-1)$ . Just before the greedy heuristic adds for the first time an “edge” of distance more than  $4(i-1)$ , it must be the case that each of these clusters contains at most one unmatched vertex (otherwise the greedy heuristic would have chosen an “edge” between two vertices in the same cluster). Therefore, we can conclude that  $\mathbb{E}[Y_i] \leq O(1 + n/(i-1)^2)$  for  $i > 1$ . On the other hand, for  $4(i-1) \geq 9\sqrt{n}$ , it follows from Lemma 14 that  $\mathbb{E}[Y_i] \leq n \cdot \mathbb{P}(\Delta_{\max} \geq 4(i-1)) \leq n^2 e^{-4(i-1)}$ .

Now we sum over all bins, bound the length of each “edge” in bin  $i$  by  $4i$ , and subsequently use Fubini’s theorem and the derived bounds on  $\mathbb{E}[Y_i]$ . This yields

$$\begin{aligned} \mathbb{E}[\text{GR}] &\leq \sum_{i=1}^{\infty} 4i \cdot \mathbb{E}[X_i] = \sum_{i=1}^{\infty} 4 \cdot \mathbb{E}[Y_i] = 2n + \sum_{i=2}^{3\sqrt{n}} 4 \cdot \mathbb{E}[Y_i] + \sum_{i=3\sqrt{n}}^{\infty} 4 \cdot \mathbb{E}[Y_i] \\ &\leq 2n + \sum_{i=2}^{3\sqrt{n}} O\left(1 + \frac{n}{(i-1)^2}\right) + \sum_{i=3\sqrt{n}}^{\infty} 4n^2 e^{-4(i-1)} = O(n) + O(n) + o(1) = O(n), \end{aligned}$$

which finishes the proof. ◀

► **Theorem 16.** For random shortest path metrics generated from a finite grid graph we have  $\mathbb{E}\left[\frac{\text{GR}}{\text{MM}}\right] = O(1)$ .

**Proof.** Let  $c > 0$  be a sufficiently small constant. Then the approximation ratio of the greedy heuristic on random shortest path metrics generated from a finite grid graph is

$$\mathbb{E}\left[\frac{\text{GR}}{\text{MM}}\right] \leq \mathbb{E}\left[\frac{\text{GR}}{cn}\right] + \mathbb{P}(\text{MM} < cn) \cdot O\left(n^{\log_2(3/2)}\right),$$

since the worst-case approximation ratio of the greedy heuristic on metric instances is  $O(n^{\log_2(3/2)})$  [19]. By Theorem 15 the first term is  $O(1)$ . Combining Lemmas 6 and 7, the second term can be bounded from above by  $\exp(n(1 + \frac{1}{2} \ln(c \cdot \Theta(1)))) \cdot O(n^{\log_2(3/2)}) = o(1)$  since  $c$  is sufficiently small. ◀

### Nearest Neighbor Heuristic for TSP

One of the most intuitive heuristics for the TSP is the nearest neighbor heuristic. This greedy-like heuristic starts with an arbitrary vertex as its current vertex and iteratively builds a TSP tour by traveling from its current vertex to the closest unvisited vertex and adding the corresponding “edge” to the tour (and closing the tour by going back to its first vertex after all vertices have been visited). Let  $\text{NN}$  denote the total distance of the TSP tour computed by the nearest neighbor heuristic, and let  $\text{TSP}$  denote the total distance of an optimal TSP tour.

It is known that the worst-case approximation ratio for this heuristic on metric instances is  $O(\ln(n))$  [21]. Moreover, for random Euclidean instances, the nearest neighbor heuristic has an approximation ratio of  $O(1)$  with high probability [3], and for random shortest path metrics generated from complete graphs or Erdős–Rényi random graphs the expected approximation ratio of the nearest neighbor heuristic is  $O(1)$  as well [6, 18]. We show that a similar result holds for sparse random shortest path metrics generated from a finite square grid graph.

► **Theorem 17.**  $\mathbb{E}[\text{NN}] = O(n)$ .

**Proof.** We put “edges” that are being added to the nearest neighbor TSP tour into bins according to their distance: bin  $i$  receives all “edges”  $\{u, v\}$  with  $d(u, v) \in (4(i-1), 4i]$ . Let  $X_i$  and  $Y_i$  be defined as in the proof of Theorem 15. Now we have  $Y_1 = n$ . For  $i > 1$ , by Theorem 13, we can partition the vertices in an expected number of  $O(1 + n/(i-1)^2)$  clusters (where the constant of the  $O$ -term is uniform with respect to  $i$ ), each of diameter at most  $4(i-1)$ . Every time the nearest neighbor heuristic adds an “edge” of distance more than  $4(i-1)$ , this must be an “edge” from a vertex in some cluster  $C_k$  to a vertex in another cluster  $C_\ell$ , and the tour must have already visited all other vertices in  $C_k$  (otherwise the nearest neighbor heuristic would have chosen an “edge” to an unvisited vertex in  $C_k$ ). Therefore, we can conclude that  $\mathbb{E}[Y_i] \leq O(1 + n/(i-1)^2)$  for  $i > 1$ . On the other hand, for  $4(i-1) \geq 9\sqrt{n}$ , it follows from Lemma 14 that  $\mathbb{E}[Y_i] \leq n \cdot \mathbb{P}(\Delta_{\max} \geq 4(i-1)) \leq n^2 e^{-4(i-1)}$ .

Note that we have derived exactly the same bounds as in the proof of Theorem 15. So, using the same calculations as in that proof, it follows now that  $\mathbb{E}[\text{NN}] = O(n)$ . ◀

► **Theorem 18.** *For random shortest path metrics generated from a finite grid graph we have  $\mathbb{E} \left[ \frac{\text{NN}}{\text{TSP}} \right] = O(1)$ .*

The proof of this theorem is similar to that of Theorem 16, with the worst-case approximation ratio of the nearest neighbor heuristic on metric instances being  $O(\ln(n))$  [21].

### Insertion Heuristics for TSP

Another group of greedy-like heuristics for the TSP are the insertion heuristics. An insertion heuristic starts with an initial optimal tour on a few vertices that are selected according to some predefined rule  $R$ , and iteratively chooses (according to the same rule  $R$ ) a vertex that is not in the tour yet and inserts this vertex in the current tour such that the total distance of the tour increases the least. An example of such a rule  $R$  would be to start with an initial (optimal) tour on three arbitrary vertices, and then use farthest insertion, i.e., at each step insert the vertex whose minimal distance to a vertex already in the tour is maximal.

Let  $IN_R$  denote the total distance of the TSP tour computed by the insertion heuristic using rule  $R$ , and let  $TSP$  denote the total distance of an optimal TSP tour. It is known that the worst-case approximation ratio for this heuristic for any rule  $R$  on metric instances is  $O(\ln(n))$  [21]. Moreover, for random shortest path metrics generated from complete graphs or Erdős–Rényi random graphs the expected approximation ratio of the nearest neighbor heuristic is  $O(1)$  for any rule  $R$  [6, 18]. We show that a similar result holds for sparse random shortest path metrics generated from a finite square grid graph.

► **Theorem 19.**  $\mathbb{E}[IN_R] = O(n)$ .

**Proof.** We put the steps of the insertion heuristic into bins according to the distance they add to the tour: bin  $i$  receives all steps with a contribution in the range  $(8(i-1), 8i]$ . Let  $X_i$  and  $Y_i$  be defined as in the proof of Theorem 15. Again we have  $Y_1 = n$ . For  $i > 1$ , by Theorem 13, we can partition the vertices in an expected number of  $O(1 + n/(i-1)^2)$  clusters (where the constant of the  $O$ -term is uniform with respect to  $i$ ), each of diameter at most  $4(i-1)$ . Every time the contribution of a step of the insertion heuristic is more than  $8(i-1)$ , this step must add a vertex to the tour that is part of a cluster  $C_k$  of which no other vertex is in the tour yet (otherwise the contribution of this step would have been less than  $8(i-1)$ ). Therefore, we can conclude that  $\mathbb{E}[Y_i] \leq O(1 + n/(i-1)^2)$  for  $i > 1$ . On the other hand, for  $8(i-1) \geq 9\sqrt{n}$ , it follows from Lemma 14 that  $\mathbb{E}[Y_i] \leq 2n \cdot \mathbb{P}(\Delta_{\max} \geq 8(i-1)) \leq 2n^2 e^{-8(i-1)}$ .

Using the same method as in the proof of Theorem 15 (i.e., summing over all bins, bounding the contribution of each step in bin  $i$  by  $8i$  and using Fubini's theorem and the derived bounds on  $\mathbb{E}[Y_i]$ ), and adding the expected contribution  $\mathbb{E}[T_R]$  of the initial tour, yields

$$\begin{aligned} \mathbb{E}[IN_R] &\leq \mathbb{E}[T_R] + \sum_{i=1}^{\infty} 8i \cdot \mathbb{E}[X_i] = \mathbb{E}[T_R] + \sum_{i=1}^{\infty} 8 \cdot \mathbb{E}[Y_i] \\ &= \mathbb{E}[T_R] + 8n + \sum_{i=2}^{2\sqrt{n}} 8 \cdot \mathbb{E}[Y_i] + \sum_{i=2\sqrt{n}}^{\infty} 8 \cdot \mathbb{E}[Y_i] \\ &\leq O(n) + 8n + \sum_{i=2}^{2\sqrt{n}} O\left(1 + \frac{n}{(i-1)^2}\right) + \sum_{i=2\sqrt{n}}^{\infty} 16n^2 e^{-8(i-1)} = O(n), \end{aligned}$$

where we used Theorem 17 to bound the expected contribution of the initial tour by  $\mathbb{E}[T_R] \leq \mathbb{E}[TSP] \leq \mathbb{E}[NN] = O(n)$ . Observe that this proof is independent of the choice of rule  $R$ . ◀

► **Theorem 20.** *For random shortest path metrics generated from a finite grid graph we have  $\mathbb{E}\left[\frac{IN_R}{TSP}\right] = O(1)$ .*

The proof of this theorem is similar to that of Theorem 16, with the worst-case approximation ratio of the insertion heuristic on metric instances being  $O(\ln(n))$  [21].

## 5 Analysis of 2-opt for TSP

In this section, we consider the probably most famous local search heuristic for the TSP, the 2-opt heuristic, and show that it achieves a constant expected approximation ratio as well. Since we do not make use of any of the lemmas that have been tailored to random shortest path metrics generated from finite square grid graphs, the results in this section hold for random shortest path metrics generated from any sparse graph.

The 2-opt heuristic starts with an arbitrary initial solution and iteratively improves this solution by applying so-called 2-exchanges until no improvement is possible anymore. In a 2-exchange, two “edges”  $\{u_1, v_1\}$  and  $\{u_2, v_2\}$  that are visited in this order in the current solution are removed from it and replaced by the two “edges”  $\{u_1, u_2\}$  and  $\{v_1, v_2\}$  to obtain a new solution. The improvement of this 2-exchange is  $\delta = d(u_1, v_1) + d(u_2, v_2) - d(u_1, u_2) - d(v_1, v_2)$ . A solution is called 2-optimal if  $\delta \leq 0$  for all possible 2-exchanges.

The actual performance of the 2-opt heuristic strongly depends on the choice of the initial solution and the sequence of improvements. In this paper we look at the worst possible outcome of the 2-opt heuristic, as others have been doing before (see e.g. [7, 9]), since this decouples the actual heuristic from the initialization and therefore keeps the analysis tractable. Let WLO denote the total distance of the worst 2-optimal TSP tour, and let TSP denote the total distance of an optimal TSP tour.

It is known that the worst-case approximation ratio for this heuristic on metric instances is  $O(\sqrt{n})$  [7]. Moreover, for Euclidean instances, the 2-opt heuristic has an expected approximation ratio of  $O(1)$  [7]. For random shortest path metrics on complete graphs, there is a trivial upper bound of  $O(\ln(n))$  for the expected approximation ratio of the 2-opt heuristic, but it is an open problem whether this can be improved or not [6]. We show that for random shortest path metrics generated from sparse graphs, the expected approximation ratio of the 2-opt heuristic is  $O(1)$ .

The crucial observation that enables us to show this result is the fact that for any 2-optimal solution for the TSP it holds that each edge  $e \in E$  can appear at most twice in the disjoint union of all shortest paths that correspond to this solution. In other words, the total distance of any 2-optimal solution can be bounded by twice the sum of all edge weights in  $G$ . The following lemma and theorems formalize this observation and its consequences.

► **Lemma 21.** *Consider a 2-optimal solution for the TSP. For each  $i, j \in V$ , let  $P_{ij}$  denote the set of all (directed) edges in the shortest  $i, j$ -path. Moreover, let  $x_{ij} = 1$  if the solution travels directly from vertex  $i$  to vertex  $j$ , and  $x_{ij} := 0$  otherwise. Then, for any  $i, j, k, l \in V$  with  $x_{ij} = x_{kl} = 1$  we have either  $P_{ij} \cap P_{kl} = \emptyset$  or  $(i, j) = (k, l)$ .*

**Proof.** Let  $i, j, k, l \in V$  such that  $x_{ij} = x_{kl} = 1$ , and suppose that  $(i, j) \neq (k, l)$ . Set  $A := \{i, j, k, l\}$  and observe that  $|A|$  equals either 3 or 4. ( $|A| = 2$  would imply  $(i, j) = (k, l)$ .)

First, suppose that  $|A| = 4$ . Suppose, by way of contradiction, that  $P_{ij} \cap P_{kl} \neq \emptyset$ . Take  $e = (s, t) \in P_{ij} \cap P_{kl}$ . Then  $d(i, j) = d(i, s) + w(e) + d(t, j)$  and  $d(k, l) = d(k, s) + w(e) + d(t, l)$ . Moreover, using the triangle inequality, we can see that  $d(i, k) \leq d(i, s) + d(s, k)$  and  $d(j, l) \leq d(j, t) + d(t, l)$ . Let  $\delta = \delta(i, j, k, l)$  denote the improvement of the 2-exchange where  $\{i, j\}$  and  $\{k, l\}$  are replaced by  $\{i, k\}$  and  $\{j, l\}$ , and note that  $\delta \leq 0$  since we are considering a 2-optimal solution for the TSP. It follows that

$$\begin{aligned} 0 &\geq \delta = d(i, j) + d(k, l) - d(i, k) - d(j, l) \\ &\geq d(i, s) + w(e) + d(t, j) + d(k, s) + w(e) + d(t, l) - d(i, s) - d(s, k) - d(j, t) - d(t, l) \\ &= 2w(e) > 0, \end{aligned}$$

which is clearly a contradiction. Therefore we must have  $P_{ij} \cap P_{kl} = \emptyset$  in this case.

Now, suppose that  $|A| = 3$ . Since the  $x$  variables describe a solution to the TSP, this implies that either  $j = k$  or  $i = l$ . These cases are analogously, so w.l.o.g. we assume that  $j = k$ . The proof that  $P_{ij} \cap P_{kl} = \emptyset$  in this case is similar to the proof for  $|A| = 4$ , with the exception that here we have  $\delta = d(i, j) + d(j, l) - d(i, j) - d(j, l) = 0$  (instead of  $\delta \leq 0$ ). The result follows. ◀

► **Theorem 22.**  $\mathbb{E}[\text{WLO}] = O(n)$ .

**Proof.** For each  $i, j \in V$ , let  $P_{ij}$  denote the set of all (directed) edges in the shortest  $i, j$ -path. Moreover, let  $x_{ij} = 1$  if WLO travels directly from vertex  $i$  to vertex  $j$ , and  $x_{ij} = 0$  otherwise.

From Lemma 21 we know that each edge  $e \in E$  can appear at most twice in the disjoint union of all shortest  $i, j$ -paths that form a 2-optimal tour (at most once per direction). This yields

$$\text{WLO} = \sum_{i,j \in V} d(i,j)x_{ij} = \sum_{\substack{i,j \in V \\ x_{ij}=1}} \sum_{e \in P_{ij}} w(e) \leq \sum_{e \in E} 2w(e) = 2S_{|E|},$$

where  $S_m$  denotes the sum of the  $m$  lightest edge weights in  $G$  as before. Combining this with Corollary 5, it follows that

$$\mathbb{E}[\text{WLO}] \leq \mathbb{E}[2S_{|E|}] = O\left(\frac{|E|^2}{n}\right) = O(n),$$

where the last equality follows by recalling that  $|E| = \Theta(n)$  for sparse graphs. ◀

► **Theorem 23.** *For random shortest path metrics generated from a sparse graph we have  $\mathbb{E}\left[\frac{\text{WLO}}{\text{TSP}}\right] = O(1)$ .*

The proof of this theorem is similar to that of Theorem 16, with the worst-case approximation ratio of the 2-opt heuristic on metric instances being  $O(\sqrt{n})$  [7].

## 6 Concluding Remarks

We have analyzed simple heuristics for matching and TSP on random shortest path metrics generated from sparse graphs, since we believe that these models yield more realistic metric spaces than random shortest path metrics generated from dense or even complete graphs. However, for the greedy-like heuristics we had to restrict ourselves to (finite square) grid graphs. It is possible to adapt our proofs for all classes of sparse graphs that have sufficiently fast growing cut sizes  $|\delta(U)|$  (as long as  $|U|$  is not too large). It seems to be sufficient to have  $|\delta(U)| \geq \Omega(|U|^\varepsilon)$  if  $|U| \leq c'n$  for some constants  $\varepsilon, c' \in (0, 1)$ . Sparse graphs that have this property include  $d$ -dimensional grid graphs and other lattice graphs. We raise the question whether it is possible to extend our findings for these heuristics to arbitrary sparse graphs.

On the other hand, especially if we consider random shortest path metrics generated from grid graphs, in our view the model could be improved by using only a (possibly random) subset of the vertices of  $G$  for defining the random metric space, i.e., restricting the distance function  $d$  of the metric to some sub-domain  $V' \times V'$ , where  $V' \subset V$ . It would be interesting to see whether this model could be analyzed as well.

Finally, in our analysis of the 2-opt local search heuristic, we had to decouple the actual heuristic from the initialization in order to make the analysis tractable. We leave it as an open problem to prove rigorous results about hybrid heuristics that consist of an initialization and a local search algorithm.

## References

- 1 Antonio Auffinger, Michael Damron, and Jack Hanson. 50 years of first passage percolation. *arXiv e-prints*, 2015. [arXiv:1511.03262](https://arxiv.org/abs/1511.03262).
- 2 David Avis, Burgess Davis, and John Michael Steele. Probabilistic analysis of a greedy heuristic for Euclidean matching. *Probability in the Engineering and Informational Sciences*, 2(2):143–156, 1988. [doi:10.1017/S0269964800000711](https://doi.org/10.1017/S0269964800000711).
- 3 Jon Louis Bentley and James Benjamin Saxe. An analysis of two heuristics for the Euclidean traveling salesman problem. In *Proceedings of the Eighteenth Annual Allerton Conference on Communication, Control, and Computing*, pages 41–49, 1980.
- 4 Béla Bollobás and Imre Leader. Edge-isoperimetric inequalities in the grid. *Combinatorica*, 11(4):299–314, 1991. [doi:10.1007/BF01275667](https://doi.org/10.1007/BF01275667).
- 5 Jean-Louis Bon and Eugen Păltănea. Ordering properties of convolutions of exponential random variables. *Lifetime Data Analysis*, 5(2):185–192, 1999. [doi:10.1023/A:1009605613222](https://doi.org/10.1023/A:1009605613222).
- 6 Karl Bringmann, Christian Engels, Bodo Manthey, and B. V. Raghavendra Rao. Random shortest paths: Non-Euclidean instances for metric optimization problems. *Algorithmica*, 73(1):42–62, 2015. [doi:10.1007/s00453-014-9901-9](https://doi.org/10.1007/s00453-014-9901-9).
- 7 Barun Chandra, Howard Karloff, and Craig Tovey. New results on the old  $k$ -opt algorithm for the traveling salesman problem. *SIAM Journal on Computing*, 28(6):1998–2029, 1999. [doi:10.1137/S0097539793251244](https://doi.org/10.1137/S0097539793251244).
- 8 Robert Davis and Armand Prieditis. The expected length of a shortest path. *Information Processing Letters*, 46(3):135–141, 1993. [doi:10.1016/0020-0190\(93\)90059-I](https://doi.org/10.1016/0020-0190(93)90059-I).
- 9 Christian Engels and Bodo Manthey. Average-case approximation ratio of the 2-opt algorithm for the TSP. *Operations Research Letters*, 37(2):83–84, 2009. [doi:10.1016/j.orl.2008.12.002](https://doi.org/10.1016/j.orl.2008.12.002).
- 10 Matthias Englert, Heiko Röglin, and Berthold Vöcking. Worst case and probabilistic analysis of the 2-opt algorithm for the TSP. *Algorithmica*, 68(1):190–264, 2014. [doi:10.1007/s00453-013-9801-4](https://doi.org/10.1007/s00453-013-9801-4).
- 11 Alan M. Frieze and Joseph E. Yukich. Probabilistic analysis of the TSP. In Gregory Gutin and Abraham P. Punnen, editors, *The Traveling Salesman Problem and Its Variations*, chapter 7, pages 257–307. Springer, Boston, MA, 2007. [doi:10.1007/0-306-48213-4\\_7](https://doi.org/10.1007/0-306-48213-4_7).
- 12 John Michael Hammersley and Dominic James Anthony Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In Jerzy Neyman and Lucien Marie Le Cam, editors, *Bernoulli 1713 Bayes 1763 Laplace 1813, Anniversary Volume, Proceedings of an International Research Seminar Statistical Laboratory, University of California, Berkeley 1963*, pages 61–110. Springer Berlin Heidelberg, 1965. [doi:10.1007/978-3-642-49750-6\\_7](https://doi.org/10.1007/978-3-642-49750-6_7).
- 13 Refael Hassin and Eitan Zemel. On shortest paths in graphs with random weights. *Mathematics of Operations Research*, 10(4):557–564, 1985. [doi:10.1287/moor.10.4.557](https://doi.org/10.1287/moor.10.4.557).
- 14 C. Douglas Howard. Models of first-passage percolation. In Harry Kesten, editor, *Probability on Discrete Structures*, pages 125–173. Springer Berlin Heidelberg, 2004. [doi:10.1007/978-3-662-09444-0\\_3](https://doi.org/10.1007/978-3-662-09444-0_3).
- 15 Svante Janson. One, two and three times  $\log n/n$  for paths in a complete graph with random weights. *Combinatorics, Probability and Computing*, 8(4):347–361, 1999. [doi:10.1017/S0963548399003892](https://doi.org/10.1017/S0963548399003892).
- 16 Svante Janson. Tail bounds for sums of geometric and exponential variables. *Statistics & Probability Letters*, 135:1–6, 2018. [doi:10.1016/j.spl.2017.11.017](https://doi.org/10.1016/j.spl.2017.11.017).
- 17 Richard Manning Karp and John Michael Steele. Probabilistic analysis of heuristics. In Eugene Leighton Lawler, Jan Karel Lenstra, Alexander Hendrik George Rinnooy Kan, and David Bernard Shmoys, editors, *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*, pages 181–205. John Wiley & Sons Ltd., 1985.

- 18 Stefan Klootwijk, Bodo Manthey, and Sander K. Visser. Probabilistic analysis of optimization problems on generalized random shortest path metrics. In Gautam K. Das, Partha S. Mandal, Krishnendu Mukhopadhyaya, and Shin-ichi Nakano, editors, *WALCOM: Algorithms and Computation, 13th International Conference, WALCOM 2019, Guwahati, India, February 27 – March 2, 2019, Proceedings*, pages 108–120. Springer Nature Switzerland AG, 2019. doi:10.1007/978-3-030-10564-8\_9.
- 19 Edward M. Reingold and Robert Endre Tarjan. On a greedy heuristic for complete matching. *SIAM Journal on Computing*, 10(4):676–681, 1981. doi:10.1137/0210050.
- 20 Daniel Richardson. Random growth in a tessellation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 74(3):515–528, 1973. doi:10.1017/S0305004100077288.
- 21 Daniel J. Rosenkrantz, Richard Edwin Stearns, and Philip M. Lewis II. An analysis of several heuristics for the traveling salesman problem. *SIAM Journal on Computing*, 6(3):563–581, 1977. doi:10.1137/0206041.

## A Computations for the Proof of Corollary 12

In this appendix we show that there exists a constant  $c_1$  such that for any  $\Delta > 2$  we have

$$\Delta\sqrt{s_\Delta - 1} \cdot \exp\left(-2(\sqrt{s_\Delta} - 1)\left(\frac{\Delta}{\sqrt{s_\Delta} - 1} - 1 - \ln\left(\frac{\Delta}{\sqrt{s_\Delta} - 1}\right)\right)\right) \leq c_1,$$

where  $s_\Delta := \min\{\Delta^2/4, n/4\}$  and  $n \geq 9$ . We consider two cases:  $\Delta^2 \leq n$  and  $\Delta^2 \geq n$ .

For the first case, suppose that  $\Delta^2 \leq n$ . Then we have  $s_\Delta = \Delta^2/4$ , and we need to show that

$$f(\Delta) := \Delta\sqrt{\Delta^2/4 - 1} \cdot \exp\left(-(\Delta - 2)\left(\frac{\Delta}{\sqrt{\Delta^2/4 - 1}} - 1 - \ln\left(\frac{\Delta}{\sqrt{\Delta^2/4 - 1}}\right)\right)\right) \leq c_1.$$

Now observe that  $\lambda - 1 - \ln(\lambda)$  is an increasing function of  $\lambda$  for  $\lambda \geq 1$ . Combining this with the observation that  $\sqrt{\Delta^2/4 - 1} \leq \sqrt{\Delta^2/4} = \Delta/2$  (for any  $\Delta \geq 2$ ), it follows now that

$$f(\Delta) \leq \frac{1}{2}\Delta^2 e^{-(\Delta-2)(1-\ln(2))}.$$

So,  $f(\Delta)$  is upper bounded by a function  $g(\Delta)$  of the form  $g(\Delta) = c_2\Delta^2 e^{-c_3\Delta}$  for some constants  $c_2, c_3 \geq 0$ . It is well-established that such a function has a finite global maximum (which can be shown to be equal to  $\frac{1}{4}c_2c_3^2 e^{-c_3^2/2}$ ). Therefore, we can conclude that in this case there exists a constant  $c_1$  such that  $f(\Delta) \leq c_1$  for all  $\Delta > 2$ .

For the second case, suppose that  $\Delta^2 \geq n$ . Then we have  $s_\Delta = n/4$ , and we need to show that we have

$$h(\Delta, n) := \Delta\sqrt{n/4 - 1} \cdot \exp\left(-(\sqrt{n} - 2)\left(\frac{\Delta}{\sqrt{n/4 - 1}} - 1 - \ln\left(\frac{\Delta}{\sqrt{n/4 - 1}}\right)\right)\right) \leq c_1,$$

for all pairs  $(\Delta, n)$  satisfying  $\Delta^2 \geq n \geq 9$ . The first step of the proof is to show that  $h(\Delta, n) \leq h(\sqrt{n}, n)$  for all  $\Delta \geq \sqrt{n}$ . To do so, we compute the partial derivative of  $h(\Delta, n)$  with respect to  $\Delta$ , and show that it is non-positive for all  $\Delta \geq \sqrt{n}$ . The partial derivative equals

$$\begin{aligned} \frac{\partial h(\Delta, n)}{\partial \Delta} &= \left(\sqrt{n/4 - 1} - (\sqrt{n} - 2)(\Delta - \sqrt{n/4 - 1})\right) \\ &\quad \cdot \exp\left(-(\sqrt{n} - 2)\left(\frac{\Delta}{\sqrt{n/4 - 1}} - 1 - \ln\left(\frac{\Delta}{\sqrt{n/4 - 1}}\right)\right)\right). \end{aligned}$$



## 19:16 Prob. Analysis of Optimization Problems on Sparse Random Shortest Path Metrics

Now observe that for all  $n \geq 9$  we have

$$\sqrt{n/4-1} \cdot \frac{\sqrt{n}-1}{\sqrt{n}-2} \leq \sqrt{n/4} \cdot 2 = \sqrt{n} \leq \Delta.$$

This inequality can be rewritten into  $\sqrt{n/4-1} - (\sqrt{n}-2)(\Delta - \sqrt{n/4-1}) \leq 0$ , which (together with the fact that  $e^x > 0$  for all  $x \in \mathbb{R}$ ) shows that the partial derivative of  $h(\Delta, n)$  with respect to  $\Delta$  is indeed non-positive for all  $\Delta \geq \sqrt{n}$ . So, we may now conclude that  $h(\Delta, n) \leq h(\sqrt{n}, n)$  for all  $\Delta \geq \sqrt{n}$ .

Next, notice that  $h(\sqrt{n}, n) = f(\sqrt{n})$ . In the first case we have already shown that there exists a constant  $c_1$  such that  $f(\Delta) \leq c_1$  for all  $\Delta > 2$ . So, it follows immediately that  $h(\Delta, n) \leq h(\sqrt{n}, n) = f(\sqrt{n}) \leq c_1$  for all pairs  $(\Delta, n)$  satisfying  $\Delta^2 \geq n \geq 9$ .

Combining both cases, we can now see that indeed there exists a constant  $c_1$  such that for any  $\Delta > 2$  we have

$$\Delta \sqrt{s_\Delta - 1} \cdot \exp\left(-2(\sqrt{s_\Delta} - 1) \left(\frac{\Delta}{\sqrt{s_\Delta - 1}} - 1 - \ln\left(\frac{\Delta}{\sqrt{s_\Delta - 1}}\right)\right)\right) \leq c_1,$$

where  $s_\Delta := \min\{\Delta^2/4, n/4\}$  and  $n \geq 9$ .

► **Remark.** Numerical computations show that  $c_1 \geq 4.0647$  is sufficient for this result to hold.

### B Proof of Lemma 14

► **Lemma 14.** *Let  $x \geq 9\sqrt{n}$ . Then we have  $\mathbb{P}(\Delta_{\max} \geq x) \leq ne^{-x}$ .*

**Proof.** Fix an arbitrary  $v \in V$  and recall that we assume  $n$  to be even (the proof for odd  $n$  is similar, but requires some more care with the bounds of the summations). We first show that  $\mathbb{P}(\tau_n(v) \geq x) \leq e^{-x}$ . Using a similar argument as in the proof of Corollary 9, we can derive from Lemma 8 that

$$\tau_n(v) \lesssim \sum_{i=1}^{\frac{1}{4}n-1} \text{Exp}(2\sqrt{i}) + \sum_{i=\frac{1}{4}n}^{\frac{3}{4}n} \text{Exp}(\sqrt{n}) + \sum_{i=\frac{3}{4}n+1}^{n-1} \text{Exp}(2\sqrt{n-i}).$$

From this, we can see that

$$\mathbb{P}(\tau_n(v) \geq x) \leq \mathbb{P}\left(\sum_{i=1}^{\frac{1}{4}n-1} \text{Exp}(2\sqrt{i}) + \sum_{i=\frac{1}{4}n}^{\frac{3}{4}n} \text{Exp}(\sqrt{n}) + \sum_{i=\frac{3}{4}n+1}^{n-1} \text{Exp}(2\sqrt{n-i}) \geq x\right).$$

In order to bound this probability, we once more use Lemma 1(i). For this purpose, set

$$\mu := \mathbb{E}\left[\sum_{i=1}^{\frac{1}{4}n-1} \text{Exp}(2\sqrt{i}) + \sum_{i=\frac{1}{4}n}^{\frac{3}{4}n} \text{Exp}(\sqrt{n}) + \sum_{i=\frac{3}{4}n+1}^{n-1} \text{Exp}(2\sqrt{n-i})\right] = \frac{n+2}{2\sqrt{n}} + \sum_{i=1}^{\frac{1}{4}n-1} \frac{1}{\sqrt{i}},$$

and  $\lambda := x/\mu$ , and observe that  $\mu \leq \frac{1}{2}\sqrt{n} + \sqrt{n-4} + 1/\sqrt{n} \leq \frac{3}{2}\sqrt{n}$ . Together with  $x \geq 9\sqrt{n}$  this implies  $\lambda \geq 6$ . Lemma 1(i) now yields

$$\mathbb{P}(\tau_n(v) \geq x) \leq \lambda^{-1} e^{-2\mu(\lambda-1-\ln(\lambda))} \leq e^{-2\mu(\lambda/2)} = e^{-\lambda\mu} = e^{-x},$$

where we used  $\lambda-1-\ln(\lambda) \geq \lambda/2$  (which holds for all  $\lambda \geq 5.36$ ) for the second inequality. The final result follows from observing that  $\Delta_{\max} = \max_v \tau_n(v)$  and applying the appropriate union bound. ◀