



Probabilistic Analysis of Optimization Problems on Sparse Random Shortest Path Metrics

Stefan Klootwijk¹ · Bodo Manthey¹

Received: 7 February 2022 / Accepted: 6 August 2023 / Published online: 26 August 2023
© The Author(s) 2023

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 sparse graphs with ‘fast growing cut sizes’, i.e. graphs for which $|\delta(U)| = \Omega(|U|^\epsilon)$ for some constant $\epsilon \in (0, 1)$ for all subsets U of the vertices, where $\delta(U)$ is the set of edges connecting U to the remaining vertices. 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 sparse graph with fast growing cut sizes, 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

An earlier version of this work has appeared in the *Proceedings of the 31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2020)*. This research was supported by NWO Grant 613.001.402.

✉ Bodo Manthey
b.manthey@utwente.nl

¹ University of Twente, Enschede, The Netherlands

for the traveling salesman problem also achieves a constant expected approximation ratio.

Keywords Random shortest path metrics · First-passage percolation · Approximation algorithms · Probabilistic analysis · Average-case analysis

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.

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. [3, 14]). 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. [8] widened the scope of models for generating random instances by using the following model, already proposed by Karp and Steele in 1985 [21]: 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.

1.1 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

[15, 18]. 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 defined 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 $2 \ln(n)/n$ [8, 19]. We also know that the diameter in this case is approximately $3 \ln(n)/n$ [16, 19]. Bringmann et al. used this model to analyze heuristics for matching, TSP and k -median [8].

A lot of studies have been conducted on the model of random shortest path metrics (or first-passage percolation). Many of these studies focused on first-passage percolation defined 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 $\Theta(n)$. Also, the set of vertices within distance t from a given vertex grows linearly in t and, after rescaling, converges to some convex domain [26]. The survey by Auffinger et al. [1] contains a thorough overview.

1.2 Our Results

This paper aims at extending the results of Bringmann et al. [8] and Klootwijk et al. [22] to the more realistic setting of random shortest path metrics generated from sparse graphs, i.e., graphs $G = (V, E)$ for which $|E| = O(|V|)$. 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 provide a probabilistic analysis of some simple heuristics in the model of random shortest path metrics generated from sparse graphs. For most of the results in this paper we need to restrict ourselves to classes of sparse graphs that have ‘fast growing cut sizes’. The following definition formalizes the notion of fast growing cut sizes

Definition 1 Let \mathcal{G} be a family of sparse connected undirected simple graphs. We say that \mathcal{G} has fast growing cut sizes if there exist constants $c > 0$ and $\varepsilon, c' \in (0, 1)$ such that for any $G = (V, E) \in \mathcal{G}$ and any $U \subseteq V$ with $|U| \leq c'|V|$ we have

$$|\delta(U)| \geq c|U|^\varepsilon,$$

where $\delta(U) := \{\{u, v\} \in E \mid u \in U, v \notin U\}$ denotes the cut induced by U .

In the remainder of this paper, whenever we say that a family of sparse graphs has fast growing cut sizes, we implicitly assume that it satisfies Definition 1 for some constants c, ε, c' .

Intuitively, this definition implies that a family of sparse graphs with fast growing cut sizes cannot have too many ‘bottlenecks’. Loosely speaking, a bottleneck is given by two relatively large sets of vertices with only relatively few edges between them. Even though Definition 1 might seem rather restrictive at first glance, many graph classes

actually have fast growing cut sizes. Examples include d -dimensional grid graphs (see Example 2 in Sect. 2.2 for a proof), other lattice graphs and random geometric graphs (with high probability). Empirically, also many (real-life) network graphs have fast growing cut sizes. In particular, Definition 1 can be seen as a generalization of the notion of expander graphs, since setting $c' = 1/2$ and $\varepsilon = 1$ yields the definition of a family of expander graphs [17].¹

In Sect. 3 we provide some structural properties of random shortest path metrics generated from sparse graphs with fast growing cut sizes. Partially, these properties can be seen as a generalization of some of the structural properties found by Bringmann et al. for random shortest path metrics generated from complete graphs [8].

For the probabilistic analyses in this paper we consider two different types of simple heuristics. In Sect. 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 Sect. 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 sparse graphs with fast growing cut sizes (greedy-like in Sect. 4) or arbitrary sparse graphs (local search in Sect. 5). We are aware that our results regarding the 2-opt heuristic are mostly purely theoretical, because, e.g., cheapest insertion already achieves an approximation ratio of 2 and is often used to initialize 2-opt [12, 27]. 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 the set $\{1, \dots, n\}$. Sometimes we use $\exp(\cdot)$ to denote the exponential function with base e . We denote by $X \sim P$ that a random variable X is distributed according to a probability distribution P . $\text{Exp}(\lambda)$ denotes the exponential distribution with parameter λ . 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 λ . If a random variable X_1 is stochastically dominated by another random variable X_2 , i.e., we have $\mathbb{P}(X_1 \leq x) \geq \mathbb{P}(X_2 \leq x)$ for all x , we denote this by $X_1 \preceq X_2$.

Furthermore, we use $H_n^{(m)}$ as shorthand notation for the n th generalized harmonic number of order m , i.e., $H_n^{(m)} = \sum_{i=1}^n 1/i^m$. Observe that for $m \in (0, 1)$ we can view the generalized harmonic numbers as Riemann sums for $\int 1/x^m dx$ and bound them as follows:

$$\frac{(n+1)^{1-m} - 1}{1-m} \leq H_n^{(m)} \leq \frac{n^{1-m}}{1-m}. \quad (1)$$

¹ Actually, we do not allow $\varepsilon = 1$ in Definition 1. The main reason for this is the fact that $\varepsilon = 1$ would cause exceptions in our analysis since $\int 1/x^\varepsilon dx = x^{1-\varepsilon}/(1-\varepsilon)$ only holds if $\varepsilon \neq 1$. However, observe that $c|U| \geq c|U|^\varepsilon$ for any $\varepsilon \in (0, 1)$. Hence, if some family of graphs would satisfy Definition 1 with $\varepsilon = 1$, then it also satisfies this definition for any $\varepsilon \in (0, 1)$.

In particular, for any $y > 1$ (and $m \in (0, 1)$) this implies that

$$\frac{y^{1-m} - 1}{1 - m} \leq H_{\lceil y \rceil - 1}^{(m)} \leq \frac{y^{1-m}}{1 - m}. \tag{2}$$

2.1 Random Shortest Path Metrics

Given an undirected simple connected graph $G = (V, E)$, the corresponding random shortest path metric is constructed as follows. First, for each edge $e \in E$, we draw a random edge weight $w(e)$ independently according to the exponential distribution² with parameter 1. Then, we 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 (with respect 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 . Note that even though the graph G is not a complete graph, the metric $d(\cdot, \cdot)$ is complete in the sense that between each pair of vertices $u, v \in V$ it has a direct connection of distance $d(u, v)$. It is tempting to refer to these direct connections in the metric space as ‘edges’ (with weight/length/distance equal to $d(u, v)$). In order to avoid potential confusion with the edges of the graph G that is used to generate the metric space, we write quotation marks around the ‘edges’ of the metric space.

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

2.2 Sparse Graphs

Throughout this paper, we consider random shortest path metrics generated from sparse connected undirected simple graphs on n vertices. We have $|E| = O(|V|) = O(n)$ for any sparse graph $G = (V, E)$. The probabilistic analysis of the 2-opt heuristic for the TSP in Sect. 5 works for any such graph. However, for the probabilistic analyses of the greedy-like heuristics in Sect. 4 we need to restrict ourselves to classes of sparse graphs that have ‘fast growing cut sizes’ as defined in Definition 1.

² Exponential distributions are technically easiest to handle since they are memoryless. A (continuous, non-negative) probability distribution of a random variable X is said to be memoryless if and only if $\mathbb{P}(X > s + t \mid X > t) = \mathbb{P}(X > s)$ for all $s, t \geq 0$ [28, p. 294].

Looking at this definition, note that $c|U|^\varepsilon$ is a subadditive function. Hence, when checking whether a family of sparse graphs has fast growing cut sizes, we can restrict ourselves to *connected* subsets $U \subseteq V$ with $|U| \leq c'n$: if $|\delta(U)| \geq c|U|^\varepsilon$ for all such connected subsets $U \subseteq V$, then it follows for any unconnected subset $\tilde{U} = U_1 \cup \dots \cup U_k$ (where the U_i are the maximal connected subsets of \tilde{U}) that

$$|\delta(\tilde{U})| = \sum_{i=1}^k |\delta(U_i)| \geq \sum_{i=1}^k c|U_i|^\varepsilon \geq c \left(\sum_{i=1}^k |U_i| \right)^\varepsilon = c|\tilde{U}|^\varepsilon.$$

We end this section with showing that d -dimensional grid graphs have fast growing grid sizes. A d -dimensional grid graph has vertex set $V = [N]^d$, and two vertices $(u_1, \dots, u_d), (v_1, \dots, v_d) \in V$ are connected by an edge if and only if $\sum_{i=1}^d |u_i - v_i| = 1$. For such graphs we have $|V| = n = N^d$ and $|E| = dN^{d-1}(N - 1) = O(n)$.

Example 2 For any integer $d > 1$, the family of d -dimensional grid graphs has fast growing cut sizes. To see this, let $G = (V, E)$ be a d -dimensional grid graph with $n = N^d$ vertices. It is known (cf. [6, Thm. 3]) that for any $U \subset V$ with $|U| \leq n/2$ we have

$$|\delta(U)| \geq \min_{r \in [d]} \left\{ r n^{\frac{1}{r} - \frac{1}{d}} |U|^{1 - \frac{1}{r}} \right\}.$$

Exploiting the inequality $|U| \leq n/2$ (or, equivalently, $n \geq 2|U|$) we now obtain that

$$|\delta(U)| \geq \min_{r \in [d]} \left\{ r (2|U|)^{\frac{1}{r} - \frac{1}{d}} |U|^{1 - \frac{1}{r}} \right\} = \min_{r \in [d]} \left\{ r 2^{\frac{1}{r} - \frac{1}{d}} \right\} \cdot |U|^{1 - \frac{1}{d}} = 2^{1 - \frac{1}{d}} \cdot |U|^{1 - \frac{1}{d}}.$$

hence for any $d > 1$, the family of d -dimensional grid graphs satisfies Definition 1 for $c = 2^{1-1/d}$, $\varepsilon = 1 - 1/d$ and $c' = 1/2$.

3 Structural Properties

In this section, we provide some structural properties regarding random shortest path metrics generated from sparse graphs 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 on sparse graphs with fast growing cut sizes and use it to derive a clustering result and a tail bound on the diameter Δ_{\max} for random shortest path metrics generated using these graphs.

3.1 Technical Lemmas

Lemma 3 ([20, Thm. 5.1(i,iii)]). *Let $X \sim \sum_{i=1}^m X_i$ with $X_i \sim \text{Exp}(a_i)$ independently. Let $\mu = \mathbb{E}[X] = \sum_{i=1}^m 1/a_i$ and $a_* = \min_i a_i$.*

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

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

(ii) For any $\lambda \leq 1$,

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

Corollary 4 Let $X \sim \sum_{i=1}^m X_i$ with $X_i \sim \text{Exp}(a_i)$ independently. Let $\mu = \mathbb{E}[X] = \sum_{i=1}^m 1/a_i$ and $a_* = \min_i a_i$. For any x ,

$$\mathbb{P}(X \leq x) \leq \exp(a_*\mu(1 + \ln(x/\mu))) = \left(\frac{ex}{\mu}\right)^{a_*\mu}.$$

Proof Let $\lambda := x/\mu$. If $\lambda \leq 1$, the result is a weaker version of Lemma 3(ii). If $\lambda > 1$, then $1 + \ln(x/\mu) > 0$ and hence $\mathbb{P}(X \leq x) \leq 1 < \exp(a_*\mu(1 + \ln(x/\mu)))$. \square

Lemma 5 ([7, 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

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

3.2 Sums of Lightest Edge Weights in G

All main results in this paper make use of some observations related to 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 6 Let S_m denote the sum of the m lightest edge weights in G . Then

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

Proof Let σ_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

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

Now, the first stochastic dominance relation follows from Lemma 5 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 $i = 0, \dots, m - 1$. \square

Corollary 7 *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 6 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 . \square

Lemma 8 *Let S_m denote the sum of the m lightest edge weights in G . Then we have*

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

Proof First of all, Lemma 6 yields

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

Now, we apply Corollary 4 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. \square

Lemma 9 *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 TSP and 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 inequality, consider a minimum-distance perfect matching in G , and take the union of the shortest paths between each matched pair of vertices. This union must contain at least $n/2$ different edges of G . These edges have a total weight of at least $S_{n/2}$ and at most MM . So, $MM \geq S_{n/2}$. \square

3.3 A Random Growth Process

In this subsection, and the following one, we assume that G is a sparse graph with fast growing cut sizes.

In order to understand the structure of sparse random shortest path metrics it is important to get a feeling for the distribution of the distances in the random metric, in particular the distribution of $\tau_k(v)$. 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 derivation of this result is a generalization of the case in which G is a complete graph, which has been analysed before (e.g. [8, 10, 19]). The (proof of the) following lemma shows this generalization.

Lemma 10 *Let \mathcal{G} be a family of sparse graphs with fast growing cut sizes. Then, for any $G = (V, E) \in \mathcal{G}$, any $v \in V$ and any $k \leq c'n$ we have*

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

Proof The values of $\tau_k(v)$ are generated by a birth process as follows. For $k = 1$ we have $\tau_k(v) = 0$ and also $\sum_{i=1}^{k-1} \text{Exp}(ci^\varepsilon) = 0$. 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 Definition 1 it follows that $\chi_{k-1}(v) \geq c(k-1)^\varepsilon$ for $k \leq c'n$. Moreover, conditioned on the first $k-1$ phases of the birth process, these edges must have a weight of at least $\tau_{k-1}(v) - d(v, u)$ (otherwise we would have $d(v, x) \leq d(v, u) + d(u, x) < \tau_{k-1}(v)$). 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, $\tau_k(v) - \tau_{k-1}(v) \sim \text{Exp}(\chi_{k-1}(v))$. We also know that $\text{Exp}(\chi_{k-1}(v)) \lesssim \text{Exp}(c(k-1)^\varepsilon)$ since $\chi_{k-1} \geq c(k-1)^\varepsilon$ for $k \leq c'n$. So, we obtain $\tau_k(v) - \tau_{k-1}(v) \lesssim \text{Exp}(c(k-1)^\varepsilon)$ for such k . The result follows using induction. \square

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 on $|B_\Delta(v)|$ is a crucial ingredient for the construction of clusterings in the next section.

Lemma 11 *Let \mathcal{G} be a family of sparse graphs with fast growing cut sizes. Then, for any $G = (V, E) \in \mathcal{G}$, any $\Delta > 0$, any $v \in V$ and any $k \in [n]$ with $k \leq$*

$\min\{c'n, (c(1 - \varepsilon)\Delta)^{1/(1-\varepsilon)}\}$ we have

$$\mathbb{P}(\tau_k(v) \leq \Delta) \geq 1 - \frac{H_{k-1}^{(\varepsilon)}}{c\Delta} \cdot \exp\left(-H_{k-1}^{(\varepsilon)}\left(\frac{c\Delta}{H_{k-1}^{(\varepsilon)}} - 1 - \ln\left(\frac{c\Delta}{H_{k-1}^{(\varepsilon)}}\right)\right)\right).$$

Proof From Lemma 10 we can see that

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

Next, we want to apply the result of Lemma 3(i). For this purpose, set

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

and recall from (1) that

$$\frac{k^{1-\varepsilon} - 1}{1 - \varepsilon} \leq H_{k-1}^{(\varepsilon)} \leq \frac{(k - 1)^{1-\varepsilon}}{1 - \varepsilon} < \frac{k^{1-\varepsilon}}{1 - \varepsilon}.$$

Moreover, since $k \leq (c(1 - \varepsilon)\Delta)^{1/(1-\varepsilon)}$, we can derive that $\lambda = c\Delta/H_{k-1}^{(\varepsilon)} \geq c\Delta(1 - \varepsilon)/k^{1-\varepsilon} \geq 1$. Lemma 3(i) now yields

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

Finally, we substitute the values of μ and λ to obtain the desired result. □

By observing that $|B_\Delta(v)| \geq k$ if and only if $\tau_k(v) \leq \Delta$, we can immediately derive the following corollary.

Corollary 12 *Let \mathcal{G} be a family of sparse graphs with fast growing cut sizes. Then, for any $G = (V, E) \in \mathcal{G}$, any $\Delta > 0$, any $v \in V$ and any $k \in [n]$ with $k \leq \min\{c'n, (c(1 - \varepsilon)\Delta)^{1/(1-\varepsilon)}\}$ we have*

$$\mathbb{P}(|B_\Delta(v)| < k) \leq \frac{H_{k-1}^{(\varepsilon)}}{c\Delta} \cdot \exp\left(-H_{k-1}^{(\varepsilon)}\left(\frac{c\Delta}{H_{k-1}^{(\varepsilon)}} - 1 - \ln\left(\frac{c\Delta}{H_{k-1}^{(\varepsilon)}}\right)\right)\right).$$

We now use this bound to derive a bound on the probability distribution of $|B_\Delta(v)|$ that is a crucial ingredient for the construction of clusterings in the next section.

Lemma 13 *Let \mathcal{G} be a family of sparse graphs with fast growing cut sizes. Then, there exists a constant c_1 such that for any $\Delta > 0$, any $G = (V, E) \in \mathcal{G}$ with n sufficiently large, and any $v \in V$ we have*

$$\mathbb{P} \left(|B_\Delta(v)| < \min \left\{ c'(c(1 - \varepsilon)\Delta)^{1/(1-\varepsilon)}, c'n \right\} \right) \leq \frac{c_1}{\Delta^{1/(1-\varepsilon)}}.$$

Proof For ease of notation, define $\xi := c'(c(1 - \varepsilon))^{1/(1-\varepsilon)}$ and assume w.l.o.g. that $c_1 \geq 1/\xi$. Now observe that for $\Delta \leq 1/\xi^{1-\varepsilon}$, the statement is trivial since in that case we have $c_1/\Delta^{1/(1-\varepsilon)} \geq c_1\xi \geq 1$. So, we are left with the case where $\Delta > 1/\xi^{1-\varepsilon}$.

Let $s_\Delta := \min\{\xi\Delta^{1/(1-\varepsilon)}, c'n\}$ and observe that $s_\Delta > 1$ (since $\Delta > 1/\xi^{1-\varepsilon}$ and since n is sufficiently large). Using Corollary 12 with $k = \lceil s_\Delta \rceil$ we obtain

$$\begin{aligned} \mathbb{P}(|B_\Delta(v)| < s_\Delta) &= \mathbb{P}(|B_\Delta(v)| < \lceil s_\Delta \rceil) \\ &\leq \frac{H_{\lceil s_\Delta \rceil - 1}^{(\varepsilon)}}{c\Delta} \cdot \exp \left(-H_{\lceil s_\Delta \rceil - 1}^{(\varepsilon)} \left(\frac{c\Delta}{H_{\lceil s_\Delta \rceil - 1}^{(\varepsilon)}} - 1 - \ln \left(\frac{c\Delta}{H_{\lceil s_\Delta \rceil - 1}^{(\varepsilon)}} \right) \right) \right). \end{aligned}$$

So, it remains to show that there exists a constant c_1 such that for any $\Delta > 1/\xi^{1-\varepsilon}$ we have

$$\frac{H_{\lceil s_\Delta \rceil - 1}^{(\varepsilon)} \Delta^{1/(1-\varepsilon)}}{c\Delta} \cdot \exp \left(-H_{\lceil s_\Delta \rceil - 1}^{(\varepsilon)} \left(\frac{c\Delta}{H_{\lceil s_\Delta \rceil - 1}^{(\varepsilon)}} - 1 - \ln \left(\frac{c\Delta}{H_{\lceil s_\Delta \rceil - 1}^{(\varepsilon)}} \right) \right) \right) \leq c_1.$$

To do so, we consider two cases: $\xi\Delta^{1/(1-\varepsilon)} \leq c'n$ and $\xi\Delta^{1/(1-\varepsilon)} \geq c'n$.

For the first case, suppose that $\xi\Delta^{1/(1-\varepsilon)} \leq c'n$. Then it follows that $s_\Delta = \xi\Delta^{1/(1-\varepsilon)}$, and we need to show that the function

$$\begin{aligned} f(\Delta) &:= \frac{H_{\lceil \xi\Delta^{1/(1-\varepsilon)} \rceil - 1}^{(\varepsilon)} \Delta^{1/(1-\varepsilon)}}{c\Delta} \\ &\quad \times \exp \left(-H_{\lceil \xi\Delta^{1/(1-\varepsilon)} \rceil - 1}^{(\varepsilon)} \left(\frac{c\Delta}{H_{\lceil \xi\Delta^{1/(1-\varepsilon)} \rceil - 1}^{(\varepsilon)}} - 1 - \ln \left(\frac{c\Delta}{H_{\lceil \xi\Delta^{1/(1-\varepsilon)} \rceil - 1}^{(\varepsilon)}} \right) \right) \right) \end{aligned}$$

is bounded from above by a constant. Now, observe that $\lambda - 1 - \ln(\lambda)$ is an increasing function of λ for $\lambda \geq 1$. Combining this with the observation, following from (2), that

$$\frac{c\Delta}{H_{\lceil \xi\Delta^{1/(1-\varepsilon)} \rceil - 1}^{(\varepsilon)}} \geq \frac{c(1 - \varepsilon)\Delta}{(\xi\Delta^{1/(1-\varepsilon)})^{1-\varepsilon}} = \frac{c(1 - \varepsilon)}{\xi^{1-\varepsilon}} = \left(\frac{1}{c'} \right)^{1-\varepsilon} > 1,$$

and setting $\gamma := (1/c')^{1-\varepsilon}$ for ease of notation, it follows that

$$f(\Delta) \leq \Delta^{1/(1-\varepsilon)} \cdot \exp \left(-H_{\lceil \xi\Delta^{1/(1-\varepsilon)} \rceil - 1}^{(\varepsilon)} \cdot (\gamma - 1 - \ln(\gamma)) \right)$$

$$\leq \Delta^{1/(1-\varepsilon)} \cdot \exp\left(-\frac{\xi^{1-\varepsilon} \Delta - 1}{1 - \varepsilon} \cdot (\gamma - 1 - \ln(\gamma))\right),$$

where the second inequality follows by using a bound for the generalized harmonic number (cf. (2)). It is well-established that the function on the right-hand side has a finite global maximum (since $\gamma > 1$ implies $\gamma - 1 - \ln(\gamma) > 0$). Therefore, we can conclude that in this case there exists a constant c_1 such that $f(\Delta) \leq c_1$ for all $\Delta > 1/\xi^{1-\varepsilon}$.

For the second case, suppose that $\xi \Delta^{1/(1-\varepsilon)} \geq c'n$. Then it follows that $s_\Delta = c'n$, and we need to show that the function

$$g(\Delta, n) := \frac{H_{\lceil c'n \rceil - 1}^{(\varepsilon)} \Delta^{1/(1-\varepsilon)}}{c\Delta} \cdot \exp\left(-H_{\lceil c'n \rceil - 1}^{(\varepsilon)} \left(\frac{c\Delta}{H_{\lceil c'n \rceil - 1}^{(\varepsilon)}} - 1 - \ln\left(\frac{c\Delta}{H_{\lceil c'n \rceil - 1}^{(\varepsilon)}}\right)\right)\right)$$

is bounded from above by a constant as long as $\xi \Delta^{1/(1-\varepsilon)} \geq c'n$ and n is sufficiently large. Observe that we can rewrite the inequality $\xi \Delta^{1/(1-\varepsilon)} \geq c'n$ as $c(1-\varepsilon)\Delta \geq n^{1-\varepsilon}$. The first step of the proof is to show that $g(\Delta, n) \leq g(n^{1-\varepsilon}/(c(1-\varepsilon)), n)$ for all $\Delta \geq n^{1-\varepsilon}/(c(1-\varepsilon))$. To do so, we compute the partial derivative of $g(\Delta, n)$ with respect to Δ , and show that it is non-positive for all $\Delta \geq n^{1-\varepsilon}/(c(1-\varepsilon))$. The partial derivative equals

$$\begin{aligned} \frac{\partial g(\Delta, n)}{\partial \Delta} &= \frac{H_{\lceil c'n \rceil - 1}^{(\varepsilon)} \Delta^{1/(1-\varepsilon)}}{c(1-\varepsilon)\Delta^2} \cdot \left(\varepsilon - c(1-\varepsilon)\Delta + (1-\varepsilon)H_{\lceil c'n \rceil - 1}^{(\varepsilon)}\right) \\ &\quad \times \exp\left(-H_{\lceil c'n \rceil - 1}^{(\varepsilon)} \left(\frac{c\Delta}{H_{\lceil c'n \rceil - 1}^{(\varepsilon)}} - 1 - \ln\left(\frac{c\Delta}{H_{\lceil c'n \rceil - 1}^{(\varepsilon)}}\right)\right)\right). \end{aligned}$$

Now observe that for sufficiently large n we have

$$c(1-\varepsilon)\Delta \geq n^{1-\varepsilon} \geq \varepsilon + (c'n)^{1-\varepsilon} \geq \varepsilon + (1-\varepsilon)H_{\lceil c'n \rceil - 1}^{(\varepsilon)},$$

where we subsequently used the bound on Δ for this case, the fact that n is sufficiently large, and (2) to bound the generalized harmonic number. Together with the facts that $e^x > 0$ for all $x \in \mathbb{R}$ and $H_{\lceil c'n \rceil - 1}^{(\varepsilon)} \Delta^{1/(1-\varepsilon)}/(c(1-\varepsilon)\Delta^2) \geq 0$, this shows that the partial derivative of $g(\Delta, n)$ with respect to Δ is indeed non-positive for all $\Delta \geq n^{1-\varepsilon}/(c(1-\varepsilon))$.

Next, notice that $g(n^{1-\varepsilon}/(c(1-\varepsilon)), n) = f(n^{1-\varepsilon}/(c(1-\varepsilon)))$. In the first case we have already shown that there exists a constant c_1 such that $f(\Delta) \leq c_1$ for all $\Delta > 1/\xi^{1-\varepsilon}$. So, it follows immediately that $g(\Delta, n) \leq g(n^{1-\varepsilon}/(c(1-\varepsilon)), n) = f(n^{1-\varepsilon}/(c(1-\varepsilon))) \leq c_1$ as long as $\xi \Delta^{1/(1-\varepsilon)} \geq c'n$ and n is sufficiently large.

Combining both cases yields the desired result. □

3.4 Clustering and a Tail Bound for Δ_{\max}

The following theorem shows that we can partition the vertices of random shortest path metrics generated from sparse graphs from fast growing cut sizes into a suitably small number of clusters with a given maximum diameter. Its proof follows closely the ideas of Bringmann et al. [8], albeit with a different value of s_Δ .

Theorem 14 *Let \mathcal{G} be a family of sparse graphs with fast growing cut sizes. Then, there exists a constant c_1 such that for any $\Delta > 0$ and any $G \in \mathcal{G}$ there exists a partition of the vertices of a random shortest path metric generated from G into clusters, each of diameter at most 4Δ , such that the expected number of clusters needed is bounded from above by*

$$\frac{1}{c'} + \frac{(c_1\xi + 1)n}{\xi \Delta^{1/(1-\varepsilon)}} = O\left(1 + \frac{n}{\Delta^{1/(1-\varepsilon)}}\right),$$

where $\xi = c'(c(1 - \varepsilon))^{1/(1-\varepsilon)}$.

Proof Let $G \in \mathcal{G}$ with n be sufficiently large, and let $s_\Delta := \min\{\xi \Delta^{1/(1-\varepsilon)}, c'n\}$. Consider a random shortest path metric generated from G . We call vertex v Δ -dense if $|B_\Delta(v)| \geq s_\Delta$ and Δ -sparse otherwise. Using Lemma 13 we can bound the expected number of Δ -sparse vertices by $c_1 n / \Delta^{1/(1-\varepsilon)}$. We put each Δ -sparse vertex in its own cluster (of size 1), which has diameter $0 \leq 4\Delta$.

Now, only the Δ -dense vertices remain. We cluster them according to the following process. Consider an auxiliary graph H whose vertices are the Δ -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 $v \in S$. Observe that these initial clusters have diameter at most 2Δ .

Next, consider an arbitrary Δ -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 Δ -dense vertices have been added to some initial cluster. By construction, the diameter of all these clusters is now at most 4Δ : 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 $c_1 n / \Delta^{1/(1-\varepsilon)}$ clusters containing one (Δ -sparse) vertex each, and at most $n/s_\Delta \leq 1/c' + n/\xi \Delta^{1/(1-\varepsilon)}$ clusters containing at least s_Δ (Δ -dense) vertices each, all with diameter at most 4Δ . The result follows. \square

This clustering result is useful as long as Δ is not too large. However, for large values of Δ , in particular $\Delta \geq \Delta_{\max}/4$, a ‘partition’ always requires only one cluster. Recall that $\Delta_{\max} = \max_{u,v} d(u, v)$ is the diameter of the random metric space.

For random shortest path metrics generated from complete graphs we know that $\Delta_{\max} \leq O(\log(n)/n)$ with high probability [19]. For random shortest path metrics

generated from sparse graphs the diameter is significantly larger. Intuitively this follows from the fact that in a sparse graph there are significantly fewer different paths between most pairs of vertices compared to the number of different paths in a complete graph. Hence, it becomes significantly less likely to have a really short path between every pair of vertices.

For random shortest path metrics generated from arbitrary graphs, the best possible general bound is $\Delta_{\max} \leq O(n)$ with high probability. Note that for random shortest path metrics generated from a path graph on n vertices, we can easily derive that $\mathbb{E}[\Delta_{\max}] = \Theta(n)$ (this follows from Corollary 7). Hence, the bound in the following lemma is tight.

Lemma 15 *Let $G = (V, E)$ be an arbitrary connected graph on n vertices and consider a random shortest path metric generated from G . For any $x \geq 6n$ we have $\mathbb{P}(\Delta_{\max} \geq x) \leq ne^{-x/2}$.*

Proof Fix an arbitrary $v \in V$ and let $x \geq 6n$. We first show that $\mathbb{P}(\tau_n(v) \geq x) \leq e^{-x/2}$. Since G is connected, we know that $|\delta(U)| \geq 1$ for all $\emptyset \neq U \subset V$, and hence in particular $\chi_k(v) \geq 1$ for all $k \in [n]$. Using the same approach as in the proof of Lemma 10, we can derive that

$$\tau_n(v) \lesssim \sum_{i=1}^{n-1} \text{Exp}(1).$$

From this, we can see that

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

In order to bound this probability, we once more use Lemma 3(i). For this purpose, set

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

and $\lambda := x/\mu$, and observe that $\lambda \geq 6$ (since $x \geq 6n$). Lemma 3(i) now yields

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

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. □

4 Analysis 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 sparse graphs with fast growing cut sizes. The three proofs are very alike, and the ideas behind them are built upon ideas by Bringmann et al. [8]: we divide the steps of the greedy-like heuristics into bins, depending on the value which they add to the total distance of our (partial) matching or TSP tour. Using the clustering (Theorem 14) we bound the total contribution of these bins by $O(n)$, and using our observations regarding sums of lightest edge weights (Lemmas 8 and 9) we show that the optimal matching or TSP tour has a value of $\Omega(n)$ with sufficiently high probability.

4.1 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 [23]), 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 $O(n^{\log_2(3/2)})$ [25]. Moreover, for random Euclidean instances, the greedy heuristic has an approximation ratio of $O(1)$ with high probability [3]. For instances with independent edge lengths (thus not necessarily metric), the greedy heuristic returns a matching with an expected distance of $\Theta(\ln(n))$ [2] and the optimal matching has a total distance of $\Theta(1)$ with high probability [30], which gives an approximation ratio of $O(\ln(n))$. For random shortest path metrics generated from complete graphs or Erdős–Rényi random graphs the expected approximation ratio of the greedy heuristic is $O(1)$ [8, 22]. We show that a similar result holds for random shortest path metrics generated from sparse graphs with fast growing cut sizes.

Theorem 16 *For random shortest path metrics generated from sparse graphs with fast growing cut sizes we have $\mathbb{E}[\text{GR}] = O(n)$.*

Proof We put ‘edges’ that are being added to the greedy matching into bins according to their distance: bin i receives all ‘edges’ $\{u, v\}$ satisfying $d(u, v) \in (4(i - 1), 4i]$. 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 denotes 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 14, we can partition the vertices in an expected number of at most $O(1 + n/(i - 1)^{1/(1-\varepsilon)})$ clusters (where the constant hidden by the O -notation does not depend on 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 $4(i - 1)$ to the matching, it must be the case that each of these clusters contains at most one unmatched vertex (otherwise the greedy heuristic could have chosen a shorter ‘edge’ between two vertices in the same cluster). Therefore, we can conclude that $\mathbb{E}[Y_i] \leq O(1 + n/(i - 1)^{1/(1-\varepsilon)})$ for $i > 1$. On the other hand, for values of i such that $4(i - 1) \geq 6n$, it follows from Lemma 15 that $\mathbb{E}[Y_i] \leq (n/2) \cdot \mathbb{P}(\Delta_{\max} \geq 4(i - 1)) \leq n^2 e^{-2(i-1)}$.

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

$$\begin{aligned} \mathbb{E}[\text{GR}] &\leq \sum_{i=1}^{\infty} 4i \cdot \mathbb{E}[X_i] = \sum_{i=1}^{\infty} 4 \cdot \mathbb{E}[Y_i] = 4 \cdot \mathbb{E}[Y_1] + \sum_{i=2}^{2n-1} 4 \cdot \mathbb{E}[Y_i] + \sum_{i=2n}^{\infty} 4 \cdot \mathbb{E}[Y_i] \\ &\leq 2n + \sum_{i=2}^{2n-1} O\left(1 + \frac{n}{(i - 1)^{1/(1-\varepsilon)}}\right) + \sum_{i=2n}^{\infty} 4n^2 e^{-2(i-1)} \\ &= O(n) + O(n) + o(1) = O(n), \end{aligned}$$

which finishes the proof. □

Theorem 17 *For random shortest path metrics generated from sparse graphs with fast growing cut sizes we have $\mathbb{E}[\frac{\text{GR}}{\text{MM}}] = O(1)$.*

Proof Let $\hat{c} > 0$ be a sufficiently small constant. Then the approximation ratio of the greedy heuristic on random shortest path metrics generated from sparse graphs with fast growing cut sizes can be bounded by

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

since the worst-case approximation ratio of the greedy heuristic on metric instances is known to be $O(n^{\log_2(3/2)})$ [25]. By Theorem 16 the first term is $O(1)$. Combining Lemmas 8 and 9, the second term can be bounded from above by $(\hat{c} \cdot \Theta(1))^{n/2} \cdot O(n^{\log_2(3/2)}) = o(1)$ since \hat{c} is sufficiently small. □

4.2 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 NN denote the total distance of the TSP tour computed by the nearest neighbor heuristic, 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 $O(\ln(n))$ [27]. Moreover, for random Euclidean instances, the nearest neighbor heuristic has an approximation ratio of $O(1)$ with high probability [5]. For

instances with independent edge lengths (thus not necessarily metric), the nearest neighbor heuristic returns a TSP tour with an expected length of $H_{n-1} + n/(n - 1) = \Theta(\ln(n))$,³ while the optimal TSP tour has a total length of $\Theta(1)$ with high probability [13], which gives an approximation ratio of $O(\ln(n))$. 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 [8, 22]. We show that a similar result holds for random shortest path metrics generated from sparse graphs with fast growing cut sizes.

Theorem 18 *For random shortest path metrics generated from sparse graphs with fast growing cut sizes we have $\mathbb{E}[\text{NN}] = O(n)$.*

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’ $\{u, v\}$ satisfying $d(u, v) \in (4(i - 1), 4i]$. Let X_i and Y_i be defined as in the proof of Theorem 16. Observe that $Y_1 = n$. For $i > 1$, by Theorem 14, we can partition the vertices in an expected number of at most $O(1 + n/(i - 1)^{1/(1-\varepsilon)})$ clusters (where the constant hidden by the O -notation does not depend on 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_ℓ , and the tour must have already visited all other vertices in C_k (otherwise the nearest neighbor heuristic could have chosen a shorter ‘edge’ to an unvisited vertex in C_k). Therefore, we can conclude that $\mathbb{E}[Y_i] \leq O(1 + n/(i - 1)^{1/(1-\varepsilon)})$ for $i > 1$. On the other hand, for values of i such that $4(i - 1) \geq 6n$, it follows from Lemma 15 that $\mathbb{E}[Y_i] \leq n \cdot \mathbb{P}(\Delta_{\max} \geq 4(i - 1)) \leq n^2 e^{-2(i-1)}$.

Note that (except for Y_1) we have derived exactly the same bounds as in the proof of Theorem 16. Using the same calculations as in that proof, it follows now that $\mathbb{E}[\text{NN}] = O(n)$. □

Theorem 19 *For random shortest path metrics generated from sparse graphs with fast growing cut sizes we have $\mathbb{E}[\frac{\text{NN}}{\text{TSP}}] = O(1)$.*

The proof of this theorem is similar to that of Theorem 17, with the worst-case approximation ratio of the nearest neighbor heuristic on metric instances being $O(\ln(n))$ [27].

4.3 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. Usually the rule R prescribes

³ The expected distance of the k th added ‘edge’ is equal to the minimum of $n - k$ independent edge weights, and thus has an expected length of $1/(n - k)$. The final edge (connecting the n th vertex back to the starting point) has an $\text{Exp}(1)$ distribution, but is conditioned to be longer than the shortest edge from the starting point. Due to memorylessness, this edge has an expected value of $1 + 1/(n - 1)$.

that the initial tour is just some tour on three vertices or an edge (i.e., a tour on two vertices) or even a single vertex. Examples of rules used for choosing a vertex to insert in the tour are ‘nearest insertion’ (choose the vertex that has the shortest distance to a vertex already in the tour), ‘farthest insertion’ (choose the vertex whose minimal distance to a vertex already in the tour is maximal) and ‘cheapest insertion’ (choose the vertex whose insertion causes the smallest increase in the length of the tour) [24]. 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 $O(\ln(n))$ [27]. Moreover, for random Euclidean instances, some insertion rules R have an approximation ratio of $\Omega(\ln(n)/\ln \ln(n))$ [4]. 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 [8, 22]. We show that a similar result holds for random shortest path metrics generated from sparse graphs with fast growing cut sizes.

Theorem 20 *For random shortest path metrics generated from sparse graphs with fast growing cut sizes we have $\mathbb{E}[IN_R] = O(n)$.*

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 16. Observe that $Y_1 \leq n$. For $i > 1$, by Theorem 14, we can partition the vertices in an expected number of at most $O(1 + n/(i - 1)^{1/(1-\varepsilon)})$ clusters (where the constant hidden by the O -notation does not depend on 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 $\mathbb{E}[Y_i] \leq O(1 + n/(i - 1)^{1/(1-\varepsilon)})$ for $i > 1$. On the other hand, for values of i such that $8(i - 1) \geq 6n$, it follows from Lemma 15 that $\mathbb{E}[Y_i] \leq n \cdot \mathbb{P}(\Delta_{\max} \geq 4(i - 1)) \leq n^2 e^{-2(i-1)}$.

Using the same method as in the proof of Theorem 16 (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, we obtain

$$\begin{aligned} \mathbb{E}[IN_R] &\leq \mathbb{E}[T_R] + \sum_{i=1}^{\infty} 8i \cdot \mathbb{E}[X_i] = \mathbb{E}[T_R] + \sum_{i=1}^{\infty} 8 \cdot \mathbb{E}[Y_i] \\ &= \mathbb{E}[T_R] + 8 \cdot \mathbb{E}[Y_1] + \sum_{i=2}^{n-1} 8 \cdot \mathbb{E}[Y_i] + \sum_{i=n}^{\infty} 8 \cdot \mathbb{E}[Y_i] \\ &\leq O(n) + 8n + \sum_{i=2}^{n-1} O\left(1 + \frac{n}{(i - 1)^{1/(1-\varepsilon)}}\right) + \sum_{i=n}^{\infty} 8n^2 e^{-2(i-1)} = O(n), \end{aligned}$$

where we used Theorem 18 to bound the expected contribution of the initial tour by $\mathbb{E}[T_R] \leq \mathbb{E}[\text{TSP}] \leq \mathbb{E}[\text{NN}] = O(n)$. Observe that this proof is independent of the choice of rule R . \square

Theorem 21 *For random shortest path metrics generated from sparse graphs with fast growing cut sizes we have $\mathbb{E}[\frac{\text{IN}_R}{\text{TSP}}] = O(1)$.*

The proof of this theorem is similar to that of Theorem 17, with the worst-case approximation ratio of the insertion heuristic (with any rule R) on metric instances being $O(\ln(n))$ [27].

5 Analysis of 2-opt for TSP

In this section, we consider the arguably most famous local search heuristic for the TSP, the 2-opt heuristic, and show that it achieves a constant expected approximation ratio for random shortest path metrics generated from any sparse graph. Note that in this section we do not require the sparse graphs to have fast growing cut sizes.

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, the heuristic takes two ‘edges’ $\{u_1, v_1\}$ and $\{u_2, v_2\}$, where u_1, v_1, u_2, v_2 are visited in this order in the current solution, and replaces them 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. [9, 11]), 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 $O(\sqrt{n})$ [9]. Moreover, for Euclidean instances, the 2-opt heuristic has an expected approximation ratio of $O(1)$ [9]. For instances with independent edge lengths (thus not necessarily metric), the 2-opt heuristic has an expected approximation ratio of $O(\sqrt{n \ln(n)})$ [29]. For random shortest path metrics generated from complete graphs the expected approximation ratio of the 2-opt heuristic is $O(\ln(n))$, but it is an open problem whether this (almost) trivial bound can be improved or not [8]. 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 belong 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.

In this lemma and these theorems we consider the TSP tours as being directed and use the following notation. For each $i, j \in V$, let P_{ij} denote the set of all (directed) edges in the shortest i, j -path.

Lemma 22 *Let $G = (V, E)$ be an arbitrary connected graph and consider a random shortest path metric generated from this graph. Also, let \mathcal{S} denote an arbitrary 2-optimal solution for the TSP on this random metric. Moreover, let $x_{ij} := 1$ if this solution \mathcal{S} 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 $(i, j) = (k, l)$.)

We first look at the case where $|A| = 4$. 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 $\{j, l\}$. Note that $\delta \leq 0$ since \mathcal{S} is 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 clearly is a contradiction. Therefore, we must have $P_{ij} \cap P_{kl} = \emptyset$ in this case.

Now, we look at the case where $|A| = 3$. Since the x variables describe a solution to the TSP, this implies that either $j = k$ or $i = l$. These cases are analogous, 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 desired result follows. \square

Theorem 23 *For random shortest path metrics generated from arbitrary (connected) sparse graphs we have $\mathbb{E}[\text{WLO}] = O(n)$.*

Proof Let $x_{ij} = 1$ if WLO travels directly from vertex i to vertex j , and $x_{ij} = 0$ otherwise. From Lemma 22 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 in Sect. 3.2). Combining this with Corollary 7, it follows that

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

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

Theorem 24 *For random shortest path metrics generated from arbitrary (connected) sparse graphs we have $\mathbb{E}[\frac{WLO}{TSP}] = O(1)$.*

The proof of this theorem is similar to that of Theorem 17, with the worst-case approximation ratio of the 2-opt heuristic on metric instances being $O(\sqrt{n})$ [9].

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 sparse graphs with fast growing cut sizes (which includes many classes of sparse 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.

Declarations

Conflict of interest We have no conflicts of interest to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Auffinger, A., Damron, M., Hanson, J.: 50 years of first passage percolation. In: arXiv e-prints, [arXiv:1511.03262](https://arxiv.org/abs/1511.03262) [math.PR] (2015)
2. Aven, T.: Upper (lower) bounds on the mean of the maximum (minimum) of a number of random variables. *J. Appl. Probab.* **22**(3), 723–728 (1985). <https://doi.org/10.2307/3213876>

3. Avis, D., Davis, B., Steele, J.M.: Probabilistic analysis of a Greedy Heuristic for Euclidean matching. *Probab. Eng. Inf. Sci.* **2**(2), 143–156 (1988). <https://doi.org/10.1017/S0269964800000711>
4. Bafna, V., Kalyanasundaram, B., Pruhs, K.: Not all insertion methods yields constant approximate tours in the Euclidean plane. *Theor. Comput. Sci.* **125**(2), 345–353 (1994). [https://doi.org/10.1016/0304-3975\(94\)90257-7](https://doi.org/10.1016/0304-3975(94)90257-7)
5. Bentley, J. L., Saxe, J.B.: An analysis of two Heuristics for the Euclidean traveling salesman problem. In: *Proceedings of the Eighteenth Annual Allerton Conference on Communication, Control, and Computing*, October 8–10, 1980. Allerton House, Monticello, Illinois (1980) pp. 41–49
6. Bollobás, B., Leader, I.: Edge-isoperimetric inequalities in the grid. *Combinatorica* **11**(4), 299–314 (1991). <https://doi.org/10.1007/BF01275667>
7. Bon, J.-L., Păltănea, E.: Ordering properties of convolutions of exponential random variables. *Lifetime Data Anal.* **5**(2), 185–192 (1999). <https://doi.org/10.1023/A:1009605613222>
8. Bringmann, K., Engels, C., Manthey, B., Rao, B.V.R.: random shortest paths: non-Euclidean instances for metric optimization problems. *Algorithmica* **73**(1), 42–62 (2015). <https://doi.org/10.1007/s00453-014-9901-9>
9. Chandra, B., Karloff, H., Tovey, C.: New results on the old k-opt algorithm for the traveling salesman problem. *SIAM J. Comput.* **28**(6), 1998–2029 (1999). <https://doi.org/10.1137/S0097539793251244>
10. Davis, R., Prieditis, A.: The expected length of a shortest path. *Inf. Process. Lett.* **46**(3), 135–141 (1993). [https://doi.org/10.1016/0020-0190\(93\)90059-I](https://doi.org/10.1016/0020-0190(93)90059-I)
11. Engels, C., Manthey, B.: Average-case approximation ratio of the 2-opt algorithm for the TSP. *Oper. Res. Lett.* **37**(2), 83–84 (2009). <https://doi.org/10.1016/j.orl.2008.12.002>
12. Englert, M., Röglin, H., Vöcking, B.: Worst case and probabilistic analysis of the 2-Opt algorithm for the TSP. *Algorithmica* **68**(1), 190–264 (2014). <https://doi.org/10.1007/s00453-013-9801-4>
13. Frieze, A.M.: On random symmetric travelling salesman problems. *Math. Oper. Res.* **29**(4), 878–890 (2004). <https://doi.org/10.1287/moor.1040.0105>
14. Frieze, A. M., Yukich, J. E.: Probabilistic analysis of the TSP. In: Gutin, G. and Punnen, A. P. (eds) *The Traveling Salesman Problem and Its Variations*. Springer, Boston, MA, Chap. 7, pp. 257–307 (2007) https://doi.org/10.1007/0-306-48213-4_7
15. Hammersley, J. M., Welsh, D. J. A.: First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In: Neyman, J. and Le Cam, L. M. (eds) *Bernoulli 1713 Bayes 1763 Laplace 1813, Anniversary Volume, Proceedings of an International Research Seminar Statistical Laboratory, University of California, Berkeley 1963*. Springer Berlin Heidelberg, pp. 61–110 (1965) https://doi.org/10.1007/978-3-642-49750-6_7
16. Hassin, R., Zemel, E.: On shortest paths in graphs with random weights. *Math. Oper. Res.* **10**(4), 557–564 (1985). <https://doi.org/10.1287/moor.10.4.557>
17. Hoory, S., Linial, N., Wigderson, A.: Expander graphs and their applications. *Bull. Am. Math. Soc.* **43**(4), 439–561 (2006). <https://doi.org/10.1090/S0273-0979-06-01126-8>
18. Howard, C. D.: Models of first-passage percolation. In: Kesten, H. (ed) *Probability on Discrete Structures*. Springer, Berlin Heidelberg, pp. 125–173 (2004). https://doi.org/10.1007/978-3-662-09444-0_3
19. Janson, S.: One, two and three times $\log n/n$ for paths in a complete graph with random weights. *Combinat. Probab. Comput.* **8**(4), 347–361 (1999). <https://doi.org/10.1017/S0963548399003892>
20. Janson, S.: Tail bounds for sums of geometric and exponential variables. *Stat. Probab. Lett.* **135**, 1–6 (2018). <https://doi.org/10.1016/j.spl.2017.11.017>
21. Karp, R. M., Steele, J. M.: Probabilistic analysis of heuristics. In: Lawler, E. L., Lenstra, J. K., Rinnooy Kan, A. H. G., and Shmoys, D. B. (eds) *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*. Wiley, pp. 181–205 (1985)
22. Klootwijk, S., Manthey, B., Visser, S.K.: Probabilistic analysis of optimization problems on generalized random shortest path metrics. *Theoret. Comput. Sci.* **866**, 107–122 (2021). <https://doi.org/10.1016/j.tcs.2021.03.016>
23. Lawler, E.L.: *Combinatorial Optimization: Networks and Matroids*. Holt, Rinehart and Winston, New York (1976)
24. Reinelt, G.: *The Traveling Salesman: Computational Solutions for TSP Applications*. Lecture Notes in Computer Science 840. Berlin, Heidelberg: Springer-Verlag (1994) <https://doi.org/10.1007/3-540-48661-5>
25. Reingold, E.M., Tarjan, R.E.: On a greedy heuristic for complete matching. *SIAM J. Comput.* **10**(4), 676–681 (1981). <https://doi.org/10.1137/0210050>

26. Richardson, D.: Random growth in a tessellation. *Math. Proc. Camb. Philos. Soc.* **74**(3), 515–528 (1973). <https://doi.org/10.1017/S0305004100077288>
27. Rosenkrantz, D.J., Stearns, R.E., Lewis, P.M.: An Analysis of Several Heuristics for the Traveling Salesman Problem. *SIAM J. Comput.* **6**(3), 563–581 (1977). <https://doi.org/10.1137/0206041>
28. Ross, S.M.: *Introduction to Probability Models*. Academic Press, Burlington, MA (2010)
29. Slootbeek, J. J. A.: *Average-Case Analysis of the 2-opt Heuristic for the TSP*. Master Thesis. University of Twente (2017)
30. Walkup, D.W.: On the expected value of a random assignment problem. *SIAM J. Comput.* **8**(3), 440–442 (1979). <https://doi.org/10.1137/0208036>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.