

# An Asymptotic Determination of the Minimum Spanning Tree and Minimum Matching Constants in Geometrical Probability

Dimitris J. Bertsimas \*      Garrett van Ryzin †

August 1, 1989

## Abstract

Given  $n$  uniformly and independently distributed points in a ball of unit volume in dimension  $d$ , it is well established that the length of several combinatorial optimization problems (including the minimum spanning tree (MST), the minimum matching (M), the traveling salesman problem (TSP), etc.) on these  $n$  points is asymptotic to  $\beta(d)n^{(d-1)/d}$ , where the constant  $\beta(d)$  depends on the dimension  $d$  and the problem solved. It has been a long open problem to determine the constants  $\beta(d)$  for these problems. In this paper we make progress in establishing the constants  $\beta_{MST}(d), \beta_M(d)$  for the MST and the matching problem. By applying Crofton's method, an old method in geometrical probability, we prove that  $\beta_{MST}(d) \sim \sqrt{\frac{d}{2\pi e}}$ ,  $\beta_M(d) \sim \frac{1}{2}\sqrt{\frac{d}{2\pi e}}$  as  $d$  tends to infinity. Moreover, our method corresponds to heuristics for these problems, which are asymptotically exact as the dimension increases. Finally, we examine the asymptotics for the TSP constant and improve the best known bounds for  $d = 3, 4$ .

---

\*Dimitris Bertsimas, Sloan School of Management and Operations Research Center, MIT, Rm E53-359, Cambridge, MA 02139.

†Garrett van Ryzin, Operations Research Center, MIT, Cambridge, MA 02139.

*Key words.* Probabilistic Analysis, Euclidean Spaces, Crofton's Method, Minimum Spanning Tree, Matching.

# 1 Introduction

Research in the area of probabilistic analysis of combinatorial optimization problems in Euclidean spaces was initiated by the pioneering paper by Beardwood, Halton and Hammersley [1], where the authors prove the following remarkable result:

**Theorem ([1]):** If  $X_i$  are independent and uniformly distributed points in a region of  $R^d$  with volume  $a$ , then the length  $L_{TSP}$  of the traveling salesman tour (TSP) under the usual Euclidean metric through the points  $X_1, \dots, X_n$  almost surely satisfies

$$\lim_{n \rightarrow \infty} \frac{L_{TSP}}{n^{(d-1)/d}} = \beta_{TSP}(d)a^{1/d},$$

where  $\beta_{TSP}(d)$  is a constant that depends only on the dimension  $d$ .

This result was generalized to other combinatorial problems defined on Euclidean spaces, including the minimum spanning tree (MST) ([9]), the minimum matching (M) ([6]), the Steiner tree (ST) ([8]), the Held and Karp (HK) lower bound for the TSP ([4]) and other problems. Indeed, Steele [8] generalized the previous theorem for a class of combinatorial problems called subadditive Euclidean functionals. All these results say that there exist constants  $\beta_{MST}(d), \beta_M(d), \beta_{ST}(d), \beta_{HK}(d)$  such that the corresponding values for the MST, M, ST, HK over  $n^{(d-1)/d}$  tend almost surely to the corresponding constants. One of the important open problems in this area is the exact determination of these constants.

In this paper we make progress for two of these constants, namely  $\beta_{MST}(d), \beta_M(d)$ . We find upper and lower bounds for the  $\beta_{MST}(d), \beta_M(d)$  that are asymptotic to the same value as the dimension  $d$  increases. As a result, we prove that

$$\beta_{MST}(d) \sim \sqrt{\frac{d}{2\pi e}}, \quad \beta_M(d) \sim \frac{1}{2} \sqrt{\frac{d}{2\pi e}}.$$

We use Crofton's method, an old but surprisingly neglected method in geometrical probability, to compute the upper bounds. The bounds correspond to heuristics for these problems which have the interesting property that they are asymptotically

exact as both the dimension and the number of points increases. Furthermore, the lower bounds are based only on elementary arguments concerning nearest neighbors.

The paper is structured as follows. In the next section we find lower bounds for the two constants. In Section 3 we review Crofton's method and use it for the MST and the minimum matching problem in order to find upper bounds for  $\beta_{MST}(d)$  and  $\beta_M(d)$ . In addition, we analyze the connection of Crofton's method and heuristics for the MST and the matching problem. In Section 4 we state our central results on the asymptotics of  $\beta_{MST}(d)$  and  $\beta_M(d)$  and extend the results for more general MST functionals studied by Steele [9]. Finally, in Section 5 we review the best bounds for the  $\beta_{TSP}(d)$ , and observe that bounds based on our results improve the best known bounds for  $d = 3, 4$ . We also conjecture that

$$\beta_{TSP}(d) \sim \sqrt{\frac{d}{2\pi e}}.$$

## 2 Lower bounds

The lower bounds on  $\beta_{MST}(d), \beta_M(d)$  are found from the following elementary observation. The MST has larger length than the sum of the lengths of the nearest neighbor from each point. If  $L_n(a), M_n(a)$  are the expected length of the MST and minimum matching on  $n$  uniformly and independently distributed points in the  $d$ -ball of volume  $a$ , then

$$L_n(a) \geq (n-1)N(n, d, a), \quad (1)$$

where  $N(n, d, a)$  is the expected length of the nearest neighbor of a random point in the region. Similarly, since in every matching there are  $n/2$  (assume  $n$  is even) edges

$$M_n(a) \geq \frac{n}{2}N(n, d, a). \quad (2)$$

In the following lemma, which is well known, we find a lower bound on  $N(n, d, a)$ . Based on this bound Beardwood et. al. [1] find a lower bound for  $\beta_{TSP}(d)$ , although they did not include its proof. We present the proof for completeness.

**Lemma 1**

$$N(n, d, a) \geq \frac{1}{d} \left( \frac{a}{c_d n} \right)^{\frac{1}{d}} \Gamma\left(\frac{1}{d}\right), \quad (3)$$

where  $c_d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}$  is the volume of the ball of unit radius in dimension  $d$ .

**Proof**

If  $d(p)$  denotes the length of the nearest neighbor to a point  $p$ , then

$$\Pr\{d(p) \geq r\} = \left(1 - \frac{V(p, r)}{a}\right)^{n-1},$$

where  $V(p, r)$  is the volume of the intersection of the given  $d$ -ball of volume  $a$  and the ball of center  $p$  and radius  $r$ . Thus  $V(p, r) \leq \min(c_d r^d, a)$ . As a result,

$$N(n, d, a) = \int_0^{(a/c_d)^{1/d}} \Pr\{d(p) \geq r\} dr \geq \int_0^{(a/c_d)^{1/d}} \left(1 - \min(c_d r^d/a, 1)\right)^{n-1} dr.$$

If we make a change of variables to  $x = c_d r^d/a$ , we obtain

$$\begin{aligned} N(n, d, a) &\geq \frac{a^{1/d}}{dc_d^{1/d}} \int_0^1 x^{1/d-1} (1 - \min(x, 1))^{n-1} dx \\ &\geq \frac{a^{1/d}}{dc_d^{1/d}} \int_0^1 x^{1/d-1} (1 - x)^{n-1} dx \\ &= \frac{a^{1/d}}{dc_d^{1/d}} \frac{\Gamma(\frac{1}{d})\Gamma(n)}{\Gamma(n + \frac{1}{d})} \\ &\geq \frac{1}{d} \left( \frac{a}{c_d n} \right)^{\frac{1}{d}} \Gamma\left(\frac{1}{d}\right), \end{aligned}$$

where the last two relations follow from well known properties of the gamma function (see for example Rudin [7], p.192).  $\square$

From (1), (2), (3), and since  $\lim_{n \rightarrow \infty} \frac{L_n(a)}{n^{(d-1)/d}} = \beta_{MST}(d)a^{1/d}$ ,  $\lim_{n \rightarrow \infty} \frac{M_n(a)}{n^{(d-1)/d}} = \beta_M(d)a^{1/d}$ , we find the following lower bounds on  $\beta_{MST}(d), \beta_M(d)$ :

$$\beta_{MST}(d) \geq \frac{1}{dc_d^{1/d}} \Gamma\left(\frac{1}{d}\right), \quad (4)$$

$$\beta_M(d) \geq \frac{1}{2dc_d^{1/d}} \Gamma\left(\frac{1}{d}\right), \quad (5)$$

where  $c_d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}$ .

### 3 Crofton's Method Applied to MST and MATCHING

Crofton's Method is an old and specialized technique for determining mean values of random variables in certain geometrical probability problems (see for example [5]). It applies to problems in which  $n$  points are independently and uniformly distributed in a region of the  $d$  dimensional space and one would like to determine the mean value of some function defined on these points. The fundamental idea of the method is to view the expectation as a function of the size of the geometric region and then to derive a differential equation for this function by perturbing the region's size. In our case, the function is the length of the MST (or the matching problem) of  $n$  such points and the region is a  $d$ -dimensional ball of volume  $a$  centered at the origin.

Let  $L_n(a)$  denote the expected value of the MST as a function of  $n$  and  $a$ . Suppose we increase the volume of the ball incrementally from  $a$  to  $a + \delta a$  and distribute  $n$  points uniformly and independently in the enlarged ball. We consider two events:

- $E_1$ : The event all  $n$  points lie in original ball of volume  $a$ .
- $E_2$ : Exactly  $n - 1$  points lie in the original ball of volume  $a$  and one point lies in the infinitesimal, spherical shell of volume  $\delta a$ .

All other events have probability  $o(\delta a)$  and are ignored, since we will take  $\delta a \rightarrow 0$ .

Note that

$$P(E_1) = \left[ \frac{a}{a + \delta a} \right]^n = 1 - \frac{n}{a} \delta a + o(\delta a), \quad (6)$$

and

$$P(E_2) = n \frac{\delta a}{a + \delta a} \left[ \frac{a}{a + \delta a} \right]^{n-1} = \frac{n}{a} \delta a + o(\delta a). \quad (7)$$

The expected length of the MST in the enlarged ball given  $E_1$  is just  $L_n(a)$ . If we

let  $L_n(a|E_2)$  denote the expected length of the MST given  $E_2$ , then

$$L_n(a + \delta a) = L_n(a)P(E_1) + L_n(a|E_2)P(E_2) + o(\delta a),$$

which using (6) and (7) gives

$$\frac{L_n(a + \delta a) - L_n(a)}{\delta a} + \frac{n}{a}L_n(a) = \frac{n}{a}L_n(a|E_2) + \frac{o(\delta a)}{\delta a}.$$

Letting  $\delta a \rightarrow 0$ , we obtain

$$\frac{dL_n(a)}{da} + \frac{n}{a}L_n(a) = \frac{n}{a}L_n(a|E_2). \quad (8)$$

Observe that we can construct a feasible spanning for the case where one point lies on the exterior of the ball (event  $E_2$ ) by forming the MST of the  $n - 1$  points in the interior of the ball and then connecting the exterior point to the closest point on this tree. If we let  $R_d(n, a)$  denote the expected value of the distance from a point on the exterior of the  $d$ -dimensional ball of volume  $a$  to the nearest of  $n$  uniform points in the interior, then

$$L_n(a|E_2) \leq L_{n-1}(a) + R_d(n - 1, a).$$

Substituting this into (8) gives

$$\frac{dL_n(a)}{da} + \frac{n}{a}L_n(a) \leq \frac{n}{a}L_{n-1}(a) + \frac{n}{a}R_d(n - 1, a). \quad (9)$$

Note that if we defined  $L_n(a)$  has the expected length of the spanning tree that results from inserting points according to their distance from the origin, i.e. the *exodic tree* heuristic (see Section 3.1), then we could express (9) as an equation rather than an inequality; however, since we are ultimately concerned with the MST, we choose to work with the MST length and the resultant inequality (9) directly.

We are naturally led to find an upper bound on  $R_d(n, a)$ .

**Lemma 2**

$$\frac{1}{d} \left( \frac{2a}{c_d n} \right)^{\frac{1}{d}} \Gamma\left(\frac{1}{d}\right) \leq R_d(n, a) \leq \frac{1}{d} \left( \frac{2a}{c_d n} \right)^{\frac{1}{d}} \Gamma\left(\frac{1}{d}\right) + O(\delta^n), \quad (10)$$

where  $c_d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}$  is the volume of the ball of unit radius in dimension  $d$  and  $\delta$  is a constant such that  $0 < \delta < 1$ .

**Proof**

The complement of the distribution of the distance from a point on the exterior of the ball to the nearest point is

$$(1 - A(r)/a)^n, \quad 0 \leq r \leq \frac{2a}{c_d^{1/d}},$$

where  $A(r)$  is the volume of the intersection of a ball of radius  $r$  centered at the exterior of the ball of volume  $a$  with the interior of the ball of volume  $a$ . Then

$$R_d(n, a) = \int_0^{(2a/c_d)^{1/d}} (1 - A(r)/a)^n dr.$$

Intuitively the asymptotically important contribution in the above integral comes near  $r = 0$ , where  $A(r) \approx \frac{c_d r^d}{2}$ . In order to formalize this we make the following observation. For every  $b < \frac{1}{2}$  there exists a  $\gamma$ , such that

$$A(r) \geq \begin{cases} bc_d r^d & 0 \leq r \leq \gamma \\ bc_d \gamma^d & \gamma \leq r \leq (2a/c_d)^{1/d}. \end{cases}$$

Therefore,

$$\begin{aligned} R_d(n, a) &\leq \int_0^\gamma \left(1 - \frac{bc_d r^d}{a}\right)^n dr + \int_\gamma^{(2a/c_d)^{1/d}} \left(1 - \frac{bc_d \gamma^d}{a}\right)^n dr \\ &\leq \int_0^\gamma e^{-\frac{bnc_d r^d}{a}} dr + O(\delta^n) \\ &\leq \int_0^\infty e^{-\frac{bnc_d r^d}{a}} dr + O(\delta^n), \end{aligned}$$

where  $0 < \delta = 1 - \frac{bc_d \gamma^d}{a} < 1$ . If we make a change of variables  $x = \frac{bnc_d r^d}{a}$ , then

$$\begin{aligned} R_d(n, a) &\leq \frac{1}{d} \left( \frac{a}{bc_d n} \right)^{1/d} \int_0^\infty x^{1/d-1} e^{-x} dx + O(\delta^n) \\ &= \frac{1}{d} \left( \frac{a}{bc_d n} \right)^{1/d} \Gamma\left(\frac{1}{d}\right) + O(\delta^n). \end{aligned}$$



Since we can select any  $b < \frac{1}{2}$  the right inequality in (10) follows.

Since  $A(r) \leq \min(a, \frac{c_d r^d}{2})$  we can use exactly the same method as in Lemma 1 to prove the left inequality in (10).  $\square$

We next characterize the solution to the differential-difference equation (9) as  $n \rightarrow \infty$ . Upon substituting the results of Lemma 2 into (9) we note that the error term is  $O(n\delta^n)$ , i.e. exponentially small, and therefore can be neglected for large  $n$ . By a simple scaling argument,  $L_n(a) = L_n a^{1/d}$ , where  $L_n$  is the expected length of the MST in a unit  $d$ -ball. Substituting this into (9) does indeed solve the differential equation and yields the following difference equation for  $L_n$ ,

$$\frac{L_n}{d} + nL_n \leq nL_{n-1} + R_d n(n-1)^{-1/d} \quad (11)$$

where  $R_d = \frac{1}{d} \left(\frac{2}{c_d}\right)^{1/d} \Gamma(\frac{1}{d})$ .

To determine the leading behavior of the solution  $L_n$  for  $n \rightarrow \infty$  we note that the corresponding differential equation to (11) is

$$\frac{dy}{dx} = -\frac{y}{xd} + R_d(x-1)^{-1/d},$$

which has the following asymptotic behavior as  $x \rightarrow \infty$

$$y(x) \sim R_d x^{\frac{d-1}{d}}.$$

Because  $x = \infty$  is a regular singular point for the differential equation, the leading behaviors of the difference and corresponding differential equations are the same (see Bender and Orszag [2] pages 64 and 206).

To verify this we try a solution of the form  $L_n = \beta_{MST}(d)n^{(d-1)/d}$  and check that it is consistent with (11) for  $n \rightarrow \infty$ . Substituting  $L_n = \beta_{MST}(d)n^{(d-1)/d}$  into (11) yields

$$\frac{\beta_{MST}(d)n^{(d-1)/d}}{d} + \beta_{MST}(d)n^{(d-1)/d+1} \leq \beta_{MST}(d)n(n-1)^{(d-1)/d} + R_d n(n-1)^{-1/d}.$$

Using the expansion

$$(n-1)^{(d-1)/d} = n^{(d-1)/d} \left(1 - \frac{d-1}{dn} + O(n^{-2})\right),$$

and rearranging we obtain

$$\beta_{MST}(d) \leq R_d + O\left(\frac{1}{n}\right).$$

Therefore, the solution is indeed consistent for  $n \rightarrow \infty$  if  $\beta_{MST}(d) \leq R_d$ . Moreover, it is also consistent with our a priori knowledge of the asymptotic behavior of the MST length; therefore, we conclude that

$$\beta_{MST}(d) \leq \frac{2^{1/d}\Gamma(\frac{1}{d})}{dc_d^{1/d}}. \quad (12)$$

In addition, the leading behavior of the expected length of the exodic tree heuristic is  $R_d n^{(d-1)/d} a^{1/d}$ .

With some modification, a similar analysis applies to the minimum matching. Let  $M_n(a)$  be the expected length of the minimum matching. Applying the same technique to the matching problem we obtain

$$\frac{dM_n(a)}{da} + \frac{n}{a}M_n(a) \leq \frac{n}{a}\mathcal{M}_{n-2}(a - c_R) + \frac{n}{a}R_d(n - 1, a), \quad (13)$$

where  $\mathcal{M}_{n-2}(a - c_R)$  is the length of the matching of the  $n - 2$  points in the region consisting of the sphere of volume  $a$  minus those points within a radius  $R_d(n - 1, a)$  of the exterior point. Note that within this region, the  $n - 2$  point are uniformly and independently distributed; therefore, since the remaining  $n - 2$  points are uniformly and independently distributed in an area which is a subset of the original area  $\mathcal{M}_n(a - c_R) \leq M_n(a)$  and since  $c_R \rightarrow 0$  as  $n \rightarrow \infty$

$$\mathcal{M}_n(a - c_R) = M_n(a) \quad n \rightarrow \infty.$$

Therefore (13) becomes

$$\frac{dM_n(a)}{da} + \frac{n}{a}M_n(a) \leq \frac{n}{a}M_{n-2}(a) + \frac{n}{a}R_d(n - 1, a). \quad (14)$$

Again, it is straightforward to verify that a solution of the form  $M_n a^{1/d}$  solves (14) and yields

$$\frac{M_n}{a} + nM_n \leq nM_{n-2} + R_d n(n - 1)^{-1/d}. \quad (15)$$

The only difference between (11) and (15) is that in the matching problem given that we match two points there are  $n - 2$  remaining points to be matched, while in the MST given that we connect one point there are  $n - 1$  points remaining to be connected.

Following the same analysis as in the MST and using the expansion

$$(n - 2)^{(d-1)/d} = n^{(d-1)/d} \left(1 - \frac{2(d-1)}{dn} + O(n^{-2})\right),$$

we obtain

$$\beta_M(d) \leq \frac{2^{1/d} \Gamma(\frac{1}{2})}{(2d-1)c_d^{1/d}}. \quad (16)$$

### 3.1 Relation of the Crofton method with heuristic algorithms

Crofton's method produced an upper bound for the the values of the constants  $\beta_{MST}(d), \beta_M(d)$ . As mentioned, the method gives rise to heuristic algorithms for the MST and the matching problem. For the MST problem the heuristic algorithm is the following:

1. Given the points  $X_1, \dots, X_n$  sort the points such that  $|X_1| \leq |X_2| \leq \dots \leq |X_n|$ .
2. For  $i = 2, 3, \dots, n$  connect  $X_i$  to  $X_j$ , such that  $|X_i - X_j| = \min_{r < i} |X_i - X_r|$ .

It is easy to see by induction on  $i$  that these  $n - 1$  connections form a tree.

This construction was proposed by Gilbert [3], who gave it the name the exodic tree heuristic. He obtained the same constant (12) for the expected value of the exodic tree for  $d = 2$  through the use of generating functions and multidimensional integrals. Our analysis using the differential equation approach is considerably easier, has the advantage of naturally generalizing to every  $d$  and also generalizes to matching and possibly to other problems.

A similar connection holds for the matching problem. Namely our construction gives rise to a heuristic algorithm, which we call exodic matching to parallel the MST construction, as follows:

1. Given the points  $X_1, \dots, X_n$  sort the points such that  $|X_1| \leq |X_2| \leq \dots \leq |X_n|$ .
2. Starting with the outer point  $X_n$  connect  $X_n$  to its nearest neighbor, call it  $X_j$ . Delete  $X_n, X_j$  from the list of points. Repeat the procedure for the remaining points thus producing a matching.

As we show in the next section these heuristics have the surprising property of being asymptotically optimal as the dimension  $d$  increases. Furthermore, we find simple asymptotic values for  $\beta_{MST}(d), \beta_M(d)$  as the dimension increases.

## 4 The main result

In this section we combine the upper and lower bounds derived in the two previous sections and derive the central result of the paper.

**Theorem 3** *The constants  $\beta_{MST}(d), \beta_M(d)$  satisfy*

$$\frac{1}{\sqrt{\pi}} \Gamma\left(\frac{1}{d} + 1\right) \Gamma\left(\frac{d}{2} + 1\right)^{1/d} \leq \beta_{MST}(d) \leq \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{1}{d} + 1\right) \Gamma\left(\frac{d}{2} + 1\right)^{1/d} 2^{1/d}, \text{ and}$$

$$\frac{1}{2\sqrt{\pi}} \Gamma\left(\frac{1}{d} + 1\right) \Gamma\left(\frac{d}{2} + 1\right)^{1/d} \leq \beta_M(d) \leq \frac{d}{2d-1} \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{1}{d} + 1\right) \Gamma\left(\frac{d}{2} + 1\right)^{1/d} 2^{1/d}.$$

As  $d \rightarrow \infty$

$$\beta_{MST}(d) \sim \sqrt{\frac{d}{2\pi e}}, \quad \beta_M(d) \sim \frac{1}{2} \sqrt{\frac{d}{2\pi e}}. \quad (17)$$

*Proof*

The bounds for  $\beta_{MST}(d)$  follow by combining (4) with (12) and (5) with (16). Note that  $\Gamma(x+1) = x\Gamma(x)$ . (17) follows from the fact that  $\lim_{d \rightarrow \infty} 2^{\frac{1}{d}} = 1$ ,  $\lim_{d \rightarrow \infty} d^{\frac{1}{d}} = 1$  and the asymptotics of the gamma function, i.e.  $\Gamma(k+1) \sim \sqrt{2\pi} k^{k+\frac{1}{2}} e^{-k}$  as  $k \rightarrow \infty$ .  $\square$

From Theorem 3, we remark that for large dimensions almost every point in the optimum MST and matching is connected to its nearest neighbor, since we derived

the lower bound from exactly this observation. Moreover, the value of the MST is twice the value of the matching problem.

Although we only determined the two constants asymptotically for large dimensions, the upper and lower bounds are close to each other even for small dimensions. In Table 1 we calculated the upper and lower bounds and the asymptotic values from Theorem 3 for  $\beta_{MST}(d)$  as a function of  $d$ .

$d$	Lower bound	Upper bound	Asymptotics
10	0.866331	0.928511	0.765179
50	1.77978	1.80462	1.71099
100	2.47619	2.49342	2.41971
150	3.01347	3.02743	2.96352
200	3.46761	3.47965	3.42198
300	4.23107	4.24085	4.19106
400	4.87577	4.88422	4.83941
500	5.44433	5.45188	5.41063
1000	7.67823	7.68355	7.65179

Table 1: The bounds and the asymptotics for  $\beta_{MST}(d)$  as a function of  $d$ .

#### 4.1 Generalizations to more general MST functionals

Steele [9] analyzed the asymptotic behavior of the following MST functional. Let  $\phi(x)$  be a monotone function and  $S_n = \min_T \sum_{e \in T} \phi(|e|)$ , where the minimization is taken over all spanning trees. Note that when  $\phi(x) = x$  the problem reduces to the MST and also, since  $\phi(x)$  is monotone, the MST is the optimal tree for every  $\phi(x)$ . Furthermore, he proved that if  $\phi(x) \sim x^k$ ,  $k < d$  as  $x \rightarrow 0$  and the points are uniformly and independently distributed in a  $d$  ball of volume 1, then with probability 1

$$\lim_{n \rightarrow \infty} \frac{S_n}{n^{(d-k)/d}} = \beta(d, k).$$

In particular  $\beta(d, 1) = \beta_{MST}(d)$ . By applying exactly the same methods for the upper and lower bounds we can establish the following bounds for  $\beta(d, k)$ :

$$\frac{k}{dc_d^{k/d}} \Gamma\left(\frac{k}{d}\right) \leq \beta(d, k) \leq \frac{k}{dc_d^{k/d}} \Gamma\left(\frac{k}{d}\right) 2^{k/d}. \quad (18)$$

As a result, by examining the asymptotics of the upper and lower bounds we can easily obtain that for large  $d$

$$\beta(k, d) \sim \left(\frac{d}{2\pi e}\right)^{k/2}.$$

## 5 The best known bounds for the TSP

In this section we summarize the best known analytic bounds for the constant  $\beta_{TSP}(d)$  for the TSP. The goal of this section is to compare the bounds and to reveal that this constant has also an  $\Theta(\sqrt{d})$  behavior, which is consistent with the behavior of  $\beta_{MST}(d)$  and  $\beta_M(d)$  from the previous section.

**Theorem 4**  $\beta_{TSP}(d)$  satisfies

$$\frac{1}{dc_d^{1/d}} \Gamma\left(\frac{1}{d}\right) \left(1 + \frac{1}{2d}\right) \leq \beta_{TSP}(d) \leq \min\left[2^{\frac{1}{d}+1} \frac{1}{dc_d^{1/d}} \Gamma\left(\frac{1}{d}\right), \sqrt{\frac{d}{6}} 12^{1/2d}\right],$$

where  $c_d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}$ . Asymptotically,  $\beta_{TSP}(d) = \Theta(\sqrt{d})$ .

**Proof**

The lower bound follows by considering both the nearest and the second nearest neighbor for each point, since in the optimal tour there are two edges coming out from each point, whose length should be greater than the sum of the two closest neighbors. The upper bound follows from [1] by applying the strip heuristic and from the observation that twice the length of the MST is an upper bound of the optimum TSP tour, where we used the upper bound for the MST from (12). In fact, in [1] the authors prove a better bound for  $d = 2$ , i.e.  $\beta_{TSP}(2) \leq 0.9204$ . Note that

for  $d = 3, 4$  the bound from the MST is better than the one of [1] and for  $d \geq 5$  the bound of [1] dominates. Asymptotically, for large dimensions

$$\sqrt{\frac{d}{2\pi e}} \leq \beta_{TSP}(d) \leq \sqrt{\frac{d}{6}},$$

i.e.  $\beta_{TSP}(d) = \Theta(\sqrt{d})$ . □

In table 2 below we compute the best known upper and lower bounds for the TSP constant for dimensions  $d = 2, \dots, 10$ .

$d$	Lower bound	Upper bound
2	0.625000	0.92040
3	0.646287	1.39589
4	0.684158	1.44641
5	0.724529	1.50053
6	0.764356	1.51309
7	0.802857	1.54043
8	0.839871	1.57531
9	0.875436	1.61419
10	0.909648	1.65517

Table 2: The best known bounds for  $\beta_{TSP}(d)$ .

## 6 A closing conjecture

Our success with the MST and the matching constants, at least asymptotically, raises the natural question whether the upper bounding method can work for other combinatorial problems, in particular the TSP. Despite our attempts we were not able to generalize the method for the TSP. In light of Theorem 3 and our preliminary calculations we conjecture that the asymptotic behavior for large dimensions of  $\beta_{TSP}(d)$  is  $\sqrt{\frac{d}{2\pi e}}$ , which is consistent with the statements in [1].

## Acknowledgment

The research of both authors was partially supported by the National Science Foundation under grant ECS-8717970.

## References

- [1] J. Beardwood, J.H. Halton and J.M. Hammersley, "The shortest path through many points", *Proc. Cambridge Philos. Soc.*, 55, 299-327 (1959).
- [2] C.M. Bender and S.A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, New York, 1978.
- [3] E.N. Gilbert, "Random minimal trees", *J. Soc. Indust. Appl. Math.*, 13, 376-387 (1965).
- [4] M.X. Goemans and D. J. Bertsimas, "Probabilistic analysis of the Held-Karp relaxation for the Euclidean traveling salesman problem", MIT Sloan School Working Paper (1988), to appear in *Mathematics of Operations Research*.
- [5] R. Larson and A. Odoni, *Urban Operations Research*, Prentice-Hall, New Jersey (1981).
- [6] C.H. Papadimitriou, "The probabilistic analysis of matching heuristics", *Proc. 15th Annual Conference Comm. Contr. Computing*, Univ. Illinois (1978).
- [7] W. Rudin, *Principles of Mathematical Analysis*, third edition, McGraw-Hill, Tokyo (1976).
- [8] J.M. Steele, "Subadditive Euclidean Functionals and Nonlinear Growth in Geometric Probability", *Ann. Prob.*, 9, 365-376 (1981).
- [9] J.M. Steele, "Growth Rates of Euclidean Minimal Spanning Trees with Power Weighted Edges", Princeton University, technical report (1988).