Minimum spanning trees across dense cities

Ghurumuruhan Ganesan *

New York University, Abu Dhabi

Abstract

Consider $n$ nodes distributed independently across $N$ cities contained with the unit square $S$ according to a distribution $f$. Each city is modelled as an $r_n \times r_n$ square contained within $S$ and $MSTC_n$ denotes the length of the minimum spanning tree containing all the $n$ nodes. We use approximation methods to obtain variance estimates for $MSTC_n$ and prove that if the cities are well-connected and densely populated in a certain sense, then $MSTC_n$ appropriately centred and scaled converges to zero in probability.

Using the proof techniques, we alternately derive corresponding results for the length $MST_n$ of the minimum spanning tree for the usual case when the nodes are independently distributed throughout the unit square $S$. In particular, we obtain that the variance of $MST_n$ grows at most as a power of the logarithm of $n$ and use a subsequence argument to get almost sure convergence of $MST_n$ appropriately centred and scaled.

Key words: Minimum spanning tree, dense cities.

AMS 2000 Subject Classification: Primary: 60J10, 60K35; Secondary: 60C05, 62E10, 90B15, 91D30.

*E-Mail: gganesan82@gmail.com
1 Introduction

The study of minimum weight spanning trees of a graph arise in many applications and many analytical results have been derived regarding the weight of the minimum spanning tree (MST) for various types of weighted graphs. In this paper, we concern with Euclidean random graphs where nodes are distributed randomly across the unit square and the goal is to determine the overall length of the MST. Beardwood et al used subadditive ergodic type results to obtain that the minimum length of the MST \( \frac{\text{MST}_n}{\sqrt{n}} \) appropriately scaled converges to a constant a.s. as \( n \to \infty \). For more results on MST, we refer to Steele (1988, 1993), Alexander (1996), Kesten and Lee (1996).

Because of its practical importance, many algorithms have been proposed over the years to compute the MST for various kinds of graphs. For example, Kruskal’s algorithm (Cormen et al (2001)) iteratively adds edges to a sequence of increasing subtree of the original graph until a spanning tree is obtained. Much of the analytical literature is devoted to nodes distributed on regular shapes like circles or squares where subadditive techniques are applicable.

In the first part of this paper, we consider a slightly different problem where nodes are distributed across small cities distributed throughout the unit square \( S \). The cities are not necessarily regularly placed and therefore subadditive techniques are not directly applicable. We use approximation methods to obtain sharp bounds for the length of the minimum spanning tree and thereby deduce the corresponding convergence properties.

Model Description

Structure of the cities

For integer \( n \geq 1 \), let \( r_n \) and \( s_n \) be real numbers such that \( \frac{1-r_n}{r_n+s_n} \) is an integer. Tile the unit square \( S \) regularly into \( r_n \times r_n \) size squares in such a way that the distance between any two squares is at least \( s_n \) as shown in Figure 1. In Figure 1 the grey square is of size \( r_n \times r_n \), the segment \( AB \) has length \( r_n \) and the segment \( BC' \) has length \( s_n \). The \( r_n \times r_n \) squares are called cities and the term \( s_n \) denotes the intercity distance.

Label the \( r_n \times r_n \) squares (cities) as \( \{ S_l \} \) and identifying the centres of the squares \( \{ S_l \} \) with vertices in \( \mathbb{Z}^2 \), we obtain a corresponding subset of vertices \( \{ z_l \} \subset \mathbb{Z}^2 \). For example, in Figure 1 identify the centre of the
Figure 1: Tiling $S$ into $r_n \times r_n$ squares with an inter-square distance of $s_n$.

square labelled $S_1$ with $(0, 0)$, the centre of $S_2$ with $(1, 0)$, the centre of $S_3$ with $(0, 1)$ and so on. Two vertices $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$ are adjacent and connected by an edge if $|x_1 - x_2| + |y_1 - y_2| = 1$.

Fix $N = N(n)$ cities $\{S_{j_1}, \ldots, S_{j_N}\}$ and let $\{z_{j_1}, \ldots, z_{j_N}\}$ be the vertices in $\mathbb{Z}^2$ corresponding to the centres of $\{S_{j_i}\}$. We say that the cities $\{S_{j_1}, \ldots, S_{j_N}\}$ are well-connected if the corresponding set of vertices $\{z_{j_i}\}$ form a connected subgraph of $\mathbb{Z}^2$. Henceforth, we assume that $\{S_{j_1}, \ldots, S_{j_N}\}$ are well-connected and without loss of generality denote $S_{j_i}$ by $S_i$ for $1 \leq i \leq N$.

**Nodes in the cities**

Let $f$ be any density on the unit square $S$ satisfying the following conditions:

There are constants $0 < \epsilon_1 \leq \epsilon_2 < \infty$ such that

$$
\epsilon_1 \leq \inf_{x \in S} f(x) \leq \sup_{x \in S} f(x) \leq \epsilon_2 \quad (1.1)
$$

and

$$
\int_{x \in S} f(x)dx = 1. \quad (1.2)
$$

Define the density $g_N(.)$ on the $N$ cities $\bigcup_{1 \leq i \leq N} S_i$ as

$$
g_N(x) = \frac{f(x)}{\int_{\bigcup_{1 \leq i \leq N} S_i} f_j(x)dx} \quad (1.3)
$$

for all $x \in \bigcup_{1 \leq j \leq N} S_j$. 

3
Let \( X_1, X_2, \ldots, X_n \) be \( n \) nodes independently and identically distributed (i.i.d.) in the \( N \) cities \( \{S_j\}_{1 \leq j \leq N} \), each according to the density \( g_N \). Define the vector \((X_1, \ldots, X_n)\) on the probability space \((\Omega_X, \mathcal{F}_X, \mathbb{P})\). Let \( K_n = K(X_1, \ldots, X_n) \) be the complete graph whose edges are obtained by connecting each pair of nodes \( X_i \) and \( X_j \) by the straight line segment \((X_i, X_j)\) with \( X_i \) and \( X_j \) as endvertices. The line segment \( e_{ij} = (X_i, X_j) \) is the edge between the nodes \( X_i \) and \( X_j \) and \( d(e_{ij}) \) denotes the (Euclidean) length of the edge \((X_i, X_j)\).

Let \( Y_1, \ldots, Y_t \subset \{X_k\} \) be \( t \) distinct nodes. A path \( P = (Y_1, \ldots, Y_t) \) is a subgraph of \( K_n \) with vertex set \( \{Y_j\}_{1 \leq j \leq t} \) and edge set \( \{(Y_j, Y_{j+1})\}_{1 \leq j \leq t-1} \). The nodes \( Y_1 \) and \( Y_t \) are said to be connected by edges of the path \( P \). The subgraph \( C = (Y_1, Y_2, \ldots, Y_t, Y_1) \) with vertex set \( \{Y_j\}_{1 \leq j \leq t} \) and edge set \( \{(Y_j, Y_{j+1})\}_{1 \leq j \leq t-1} \cup \{(Y_t, Y_1)\} \) is said to be a cycle.

A subgraph \( T \) of \( K_n \) with vertex set \( \{Y_i\}_{1 \leq i \leq t} \) and edge set \( E_T \) is said to be a tree if the following two conditions hold:

1. The graph \( T \) is connected; i.e., any two nodes in \( T \) are connected by a path containing only edges in \( E_T \).
2. The graph \( T \) is acyclic; i.e., no subgraph of \( T \) is a cycle.

The length of the tree \( T \) is the sum of the lengths of the edges in \( T \); i.e.,

\[
L(T) = \sum_{e \in T} d(e) = \frac{1}{2} \sum_{i=1}^t l(Y_i, T),
\]

where \( l(Y_i, T) \) is the sum of lengths of edges in \( T \) containing \( Y_i \) as an endvertex.

The tree \( T \) is said to be a spanning tree if \( T \) contains all the \( n \) nodes \( \{X_k\}_{1 \leq k \leq n} \). Let \( T_n \) be a spanning tree satisfying

\[
MSTC_n = L(T_n) := \min_T L(T),
\]

where the minimum is taken over all spanning trees \( T \). If there is more than one choice for \( T_n \), choose one according to a deterministic rule. The tree \( T_n \) is defined to be the minimum spanning tree (MST) with corresponding length \( MSTC_n \).

Letting

\[
b_n := r_n \sqrt{nN},
\]

we have the following result.
Theorem 1. Suppose \( r_n, s_n \) and \( N = N(n) \) satisfy
\[
r_n^2 \geq \frac{M \log n}{n}, \quad \frac{n}{N^2} \to 0 \quad \text{and} \quad \frac{Ns_n}{b_n} \to 0 \tag{1.7}
\]
as \( n \to \infty \), for some constant \( M > 0 \). If \( M = M(\epsilon_1, \epsilon_2) > 0 \) is large, then
\[
\frac{1}{b_n} (\text{MST}_n - \mathbb{E} \text{MST}_n) \to 0 \quad \text{in probability} \tag{1.8}
\]
as \( n \to \infty \). In addition, there are positive constants \( \{\theta_i\}_{1 \leq i \leq 6} \) such that
\[
\theta_1 b_n \leq \mathbb{E} \text{MST}_n \leq \theta_2 b_n, \tag{1.9}
\]
\[
\mathbb{P}(\text{MST}_n \geq \theta_3 b_n) \geq 1 - e^{-\theta_4 N} \tag{1.10}
\]
and
\[
\mathbb{P}(\text{MST}_n \leq \theta_5 b_n) \geq 1 - \exp\left(-\theta_6 \frac{n}{N}\right) \tag{1.11}
\]
for all \( n \) large.

In words, if the cities are wide and dense enough, then the centred and scaled minimum length of the MST converges to zero in probability.

Unconstrained MST

There are \( n \) nodes \( \{X_k\}_{1 \leq k \leq n} \) independently distributed in the unit square \( S \), each according to the distribution \( f \) satisfying (1.1). Let \( T_n \) and \( \text{MST}_n \) denote the minimum spanning tree and its length, respectively, as defined in (1.5). Beardwood et al (1959) use subadditive techniques to study the convergence of the ratio \( \frac{\text{MST}_n}{\sqrt{n}} \to \beta \) for some constant \( \beta > 0 \), a.s. as \( n \to \infty \). Another approach involves the study of concentration of \( \text{MST}_n \) around its mean via concentration inequalities (see Steele (1993)). Here we use the techniques used in the proof of Theorem 1 to obtain the following result.

Theorem 2. The variance
\[
\mathbb{E} \left( \text{MST}_n - \mathbb{E} \text{MST}_n \right)^2 \leq C (\log n)^3 \tag{1.12}
\]
for some constant \( C > 0 \) and for all \( n \geq 1 \) and
\[
\frac{1}{\sqrt{n}} (\text{MST}_n - \mathbb{E} \text{MST}_n) \to 0 \quad \text{a.s.} \tag{1.13}
\]
as \( n \to \infty \). There are positive constants \( \{\theta_i\}_{1 \leq i \leq 3} \) such that

\[ \theta_1 \sqrt{n} \leq \mathbb{E} \text{MST}_n \leq 3 \sqrt{n}, \tag{1.14} \]

\[ \mathbb{P} \left( \text{MST}_n \leq 3 \sqrt{n} \right) = 1 \tag{1.15} \]

and

\[ \mathbb{P} \left( \text{MST}_n \geq \theta_2 \sqrt{n} \right) \geq 1 - \exp \left( - \frac{\theta_3 n}{\log n} \right) \tag{1.16} \]

for all \( n \) large.

Moreover, if the nodes are uniformly distributed in \( S \),

\[ \frac{\mathbb{E} \text{MST}_n}{\sqrt{n}} \to \beta \tag{1.17} \]

as \( n \to \infty \) for some constant \( \beta > 0 \).

The paper is organized as follows. In Section 2, we state the preliminary estimates needed for the proofs of the main Theorems. In Section 3, we prove Theorem 1 and in Section 4, we prove Theorem 2.

\section{Preliminary estimates}

We first derive a deterministic estimate based on the strips method used throughout.

\textbf{Strips estimate}

Suppose there are \( a \geq 3 \) nodes \( \{x_i\}_{1 \leq i \leq a} \) placed in a square \( R \) of side length \( b \) such that no two of the nodes share the same \( x \)- or \( y \)-coordinate. This is a mild condition since if \( \{X_j\}_{1 \leq j \leq n} \) are i.i.d. with density \( g_N \) as in (1.3), this condition is satisfied with probability one. For \( 3 \leq j \leq a \) let \( K(x_1, \ldots, x_j) \) be the complete graph with vertex set \( \{x_i\}_{1 \leq i \leq j} \) and let \( T_j \) be a spanning tree of \( K(x_1, \ldots, x_j) \) such that

\[ L(T_j) = \min_T L(T) =: \text{MST}(x_1, \ldots, x_j; R), \tag{2.1} \]

where the minimum is taken over all spanning trees of \( K(x_1, \ldots, x_j) \) and \( L(T) \) is the length of the tree \( T \) (see (1.4)).
We have that
\[ \text{MST}(x_1, \ldots, x_a; R) \leq 3b\sqrt{a}. \]  
\hspace{1cm} (2.2)

**Proof of (2.2):** Divide the square \( R \) into vertical rectangles (strips) each of size \( c \times b \) so that the number of strips is \( \frac{b}{c} \) as shown in Figure 2. Here \( a = 7 \) and without loss of generality suppose that \( P, Q, R, S, T, U \) and \( V \), are the nodes \( x_1, x_2, x_3, x_4, x_5, x_6 \) and \( x_7 \), respectively. The dotted line corresponds to a path containing all the nodes \( P, Q, R, S \) and \( T \). Starting from the top most node \( P \) in the first strip, vertically down in the strip and each time we are close to a node, we “reach” for the node by a slightly inclined line. In Figure 2 the vertical dotted line \( PA \) is joined to the node \( Q \) by the inclined line \( AQ \).

Continue vertically down from \( Q \) until we reach close to the bottom of the strip. Proceed along a horizontal line until we are directly below the lowest node in the second strip. In Figure 2 the point \( B \) is directly below the node \( R \). Continue vertically from \( B \), pass through \( R \) until we reach close to the next node \( S \). Join to \( S \) by a slightly inclined line and continue this procedure until all nodes in all strips have been exhausted.

The number of strips is \( \frac{b}{c} \) and the sum of the lengths of the vertical lines of \( P \) in a particular strip is at most the height of the strip \( b \). Therefore the total length of vertical lines in \( P \) is at most \( \frac{b}{c}b \).
The total length of the horizontal lines in $\mathcal{P}$ is at most $b$. Finally, each inclined line in $\mathcal{P}$ has length at most $\frac{a}{\sqrt{2}}$, since the corresponding slope is at most 45 degrees. Each of the $a$ nodes is attached to at most one inclined line and so the total length of the inclined lines in $\mathcal{P}$ is at most $ac\sqrt{2}$.

Summarizing, the total length of edges in $\mathcal{P}$ is at most $b^2 + ac\sqrt{2} + b$.

By construction, the path $\mathcal{P}$ encounters the nodes $x_1, \ldots, x_a$ in that order and so applying triangle inequality as before, the path $\mathcal{P}_0 = (x_1, x_2, \ldots, x_a)$ with edges being the straight lines $(x_1, x_2), (x_2, x_3), \ldots, (x_a, x_1)$, has total length no more than the sum of length of edges in $\mathcal{P}$. Thus

$$MST(x_1, \ldots, x_a; R) \leq L(\mathcal{P}_0) \leq \frac{b^2}{c} + \frac{ac}{\sqrt{2}} + b. \tag{2.3}$$

Setting $c = \frac{b}{\sqrt{2}}$ in (2.3), we get that $MST(x_1, \ldots, x_a; R)$ is bounded above by $b\sqrt{a} + \frac{b\sqrt{a}}{\sqrt{2}} + b \leq 3b\sqrt{a}$, since $a \geq 1$.

**Length of MST within cities**

Recall from discussion prior to (1.7) that $n \geq 1$ nodes $\{X_k\}_{1 \leq k \leq n}$ are distributed across the $r_n \times r_n$ squares $\{S_j\}_{1 \leq j \leq N}$ according to a Binomial process with intensity $g_N$ as defined in (1.3). In this subsection, we obtain estimates for the length $R_l$ of the MST containing all the nodes of the square $S_l$.

If $p_l$ denotes the probability that a node of $\{X_j\}$ occurs inside $S_l$, then

$$\frac{\eta_1}{N} \leq p_l := \frac{\int_{S_l} f(x) dx}{\int_{\bigcup_{j} S_j} f(x) dx} \leq \frac{\eta_2}{N}, \tag{2.4}$$

where $\eta_1 = \frac{a}{c_2} \leq \frac{a}{c_1} = \eta_2$ (see (1.1)). Therefore if

$$N_l = \sum_{i=1}^{n} 1(X_i \in S_l) \tag{2.5}$$

denotes the number of nodes of $\{X_j\}$ in the square $S_l$, then $N_l$ is Binomially distributed with parameters $n$ and $p_l$; i.e., for any $1 \leq k \leq n$,

$$\mathbb{P}(N_l = k) = B(k; n, p_l) := \binom{n}{k} p_l^k (1 - p_l)^{n-k}. \tag{2.6}$$
where \( \binom{n}{k} = \frac{n!}{k!(n-k)!} \) is the Binomial coefficient. Moreover,

\[
\frac{\eta_1 n}{N} \leq E N_l = n p_l \leq \frac{\eta_2 n}{N}
\]

by (2.4).

Let \( \{Y_j\}_{1 \leq j \leq N_l} \) be the nodes of \( \{X_j\} \) present in the square \( S_l \). Formally, if \( N_l = 0 \), set \( \{Y_j\}_{1 \leq j \leq N_l} := \emptyset \). If \( N_l \geq 1 \), define \( N_l \) indices \( j_1, \ldots, j_{N_l} \) as follows. Let

\[
j_1 = j_1(X_1, \ldots, X_n) := \min\{1 \leq k \leq n : X_k \in S_l\}
\]

be the least indexed node of \( \{X_k\} \) present in \( S_l \). Let

\[
j_2 = \min\{j_1 + 1 \leq k \leq n : X_k \in S_l\}
\]

be the next least indexed node of \( \{X_k\} \) present in \( S_l \) and so on. Set \( Y_i = X_{j_i} \) for \( 1 \leq j \leq N_l \).

Set \( R_l = 0 \) if \( N_l \leq 2 \) and if \( N_l \geq 3 \) set

\[
R_l := MST(Y_1, \ldots, Y_{N_l}; S_l)
\]

where \( MST(\cdot, \cdot) \) is as defined in (2.1). The following is the main lemma proved in this subsection.

Lemma 3. If \( M > 0 \) is arbitrary and (1.7) holds, the following is true: There are positive constants \( \{\delta_i\}_{1 \leq i \leq 3} \) such that for all \( n \geq 2 \) and for any \( 1 \leq l \leq N \),

\[
\delta_1 r_n \sqrt{\frac{n}{N}} \leq E R_l \leq \delta_2 r_n \sqrt{\frac{n}{N}} \quad \text{and} \quad E R_l^2 \leq \delta_3 \left( r_n \sqrt{\frac{n}{N}} \right)^2 .
\]

Moreover, if

\[
U_l = U_l(n) := \left\{ \frac{\eta_1 n}{2N} \leq N_l \leq \frac{2\eta_2 n}{N} \right\},
\]

where \( \eta_1 \) and \( \eta_2 \) are as in (2.4), then there are positive constants \( \{\delta_i\}_{i=4,5} \) such that for all \( n \geq 2 \) and for any \( 1 \leq l \leq N \),

\[
\Pr(U_l) \geq 1 - \exp\left(-\delta_4 \frac{n}{N}\right) \quad \text{and} \quad R_l 1(U_l) \leq \delta_5 r_n \sqrt{\frac{n}{N}} .
\]
To prove the above Lemma, we perform some preliminary computations. We first derive bounds for the total number of squares $N$. From (1.7) we have that $r_n^2 \geq \frac{M \log n}{n}$ and since all the $r_n \times r_n$ squares $\{S_l\}_{1 \leq l \leq N}$ are contained within the unit square $S$, we also have $N r_n^2 \leq 1$ and therefore $N \leq \frac{n}{M \log n}$. Similarly from (1.7) we also have that $\frac{n}{N} \to 0$ as $n \to \infty$ and so $N \geq \sqrt{n}$ for all $n$ large. Combining we get

$$\sqrt{n} \leq N \leq \frac{n}{M \log n} \quad \text{and} \quad \frac{n}{N} \geq M \log n$$

(2.12)

for all $n$ large.

For $k \geq 2$, let $D_l(k)$ be the expected minimum distance between the node $Y_k$ and every other node in $S_l$, given that there are $N_l = k$ nodes in $S_l$; i.e.,

$$D(k) = D_l(k) := \mathbb{E}(d(Y_k, \{Y_u\}_{1 \leq u \leq k-1}) | N_l = k),$$

(2.13)

where $d(A, B) = \min_{x \in A, y \in B} d(x, y)$ is the minimum distance between finite sets $A$ and $B$. We have the following properties.

(b1) For any $k \geq 2$ and $1 \leq l \leq N$, the term

$$D_l(k) \geq \int_0^{\frac{r_n}{k}} \left(1 - \pi \eta_2 \left(\frac{r}{r_n}\right)^2\right)^{k-1} dr$$

(2.14)

where $\eta_2 = \frac{\epsilon_2}{\epsilon_1}$ is as in (2.4).

(b2) There are positive constants $\gamma_i$, $1 \leq i \leq 3$ such that for any $k \geq 2$ and $1 \leq l \leq N$, the minimum distance

$$\gamma_1 \frac{r_n}{\sqrt{k}} \leq D_l(k) \leq \gamma_2 \frac{r_n}{\sqrt{k}} \quad \text{and} \quad \mathbb{E}(d^2(Y_k, \{Y_u\}_{1 \leq u \leq k-1}) | N_l = k) \leq \gamma_3 \frac{r_n^2}{k}.$$  

(2.15)

The proof of (b1) - (b2) uses the fact that given $N_l = k$, the nodes in $S_l$ are independently distributed in $S_l$ with distribution $f$; i.e.,

$$D_l(k) = \mathbb{E}d(Z_k, \{Z_j\}_{1 \leq j \leq k-1})$$

(2.16)

where $\{Z_i\}_{1 \leq i \leq k}$ are i.i.d. with distribution

$$\mathbb{P}(Z_1 \in A) = \frac{\int_{A \cap S_l} f(x)dx}{\int_{S_l} f(x)dx}.$$  

(2.17)
Use Fubini’s theorem and (2.17) to write
\[ D_l(k) = \frac{1}{\int_{S_l} f(x)dx} \int_{S_l} \mathbb{E}d(x, \{Z_j\}_{1 \leq j \leq k-1}) f(x) dx, \] (2.18)
where \( \mathbb{E}d(x, \{Z_j\}_{1 \leq j \leq k-1}) = \int_0^\infty \mathbb{P}(d(x, \{Z_j\}_{1 \leq j \leq k-1}) \geq r) dr. \) For any \( x \in S_l \), the minimum distance from \( x \) to \( \{Z_1, \ldots, Z_{k-1}\} \) is at least \( r \) if and only if \( B(x, r) \cap S_l \) contains no point of \( \{Z_j\}_{1 \leq j \leq k-1} \). Here \( B(x, r) \) is the ball of radius \( r \) centred at \( x \). Wherever the point \( x \in S_l \), the area of \( B(x, r) \cap S_l \) is at most \( \pi r^2 \) and so together with (1.1), we then get that
\[ \mathbb{P}(d(x, \{Z_j\}_{1 \leq j \leq k-1}) \geq r) = \left( 1 - \frac{\int_{B(x,r) \cap S_l} f(x) dx}{\int_{S_l} f(x) dx} \right)^{k-1} \]
is bounded below by \( \left( 1 - \pi \eta_2 \frac{r^2}{\pi n} \right)^{k-1} \), where \( \eta_2 = \frac{\alpha}{n} \) is as in (2.4). This proves (2.14).

To prove the lower bound for \( D_l(k) \) in (2.15) of (b2), fix \( k \geq 2 \) and use (2.14) to get that
\[ D_l(k) \geq \int_{\frac{r}{\sqrt{\delta k}}}^{\frac{r}{n}} \left( 1 - \delta \left( \frac{r}{r_n} \right)^2 \right)^{k-1} dr \geq \int_{0}^{\frac{r}{\sqrt{\delta k}}} \left( 1 - \frac{1}{k} \right)^{k-1} dr \geq \frac{e^{-1} r_n^2}{\sqrt{\delta k}} \]
for all \( n \) large. The final estimate is obtained by using \((1 - \frac{1}{r})^{r-1} \geq e^{-1}\) for all \( r \geq 2 \).

For the upper bound for \( D_l(k) \) in (2.15), again use (2.19) and the fact that \( B(x, r) \cap S_l \) has area at least \( \frac{\pi r^2}{4} \) no matter where the position of \( x \), to get
\[ \mathbb{P}(d(x, \{Z_j\}_{1 \leq j \leq k-1}) \geq r) \leq \left( 1 - \frac{\pi}{4 \epsilon_1} \left( \frac{r}{r_n} \right)^2 \right)^{k-1} \leq \exp \left( -\frac{\pi(k-1)}{4 \epsilon_1 r_n^2} r^2 \right) \]
and so \( D_l(k) \leq \int_0^\infty \exp \left( -\frac{\pi(k-1)}{4 \epsilon_1 r_n^2} r^2 \right) dr \leq \frac{C r_n}{\sqrt{k-1}} \leq \frac{2 C r_n}{\sqrt{k}} \) for all \( k \geq 2 \) and for some positive constant \( C \), not depending on \( k \) or \( l \).

Finally for the second moment estimate in (2.13), we argue analogous to (2.13) and get that the term \( \mathbb{E} \left( d^2(Y_k; \{Y_u\}_{1 \leq u \leq k-1}) | N_l = k \right) \) equals
\[ \mathbb{E}d^2(Z_k, \{Z_j\}_{1 \leq j \leq k-1}) = \frac{1}{\int_{S_l} f(x)dx} \int_{S_l} \mathbb{E}d^2(x, \{Z_j\}_{1 \leq j \leq k-1}) f(x) dx \] (2.19)
where \( \{Z_i\}_{1 \leq i \leq k} \) are i.i.d. with distribution as in (2.17). Arguing as in the previous paragraph we get that

\[
\mathbb{E}(d^2(x, \{Z_j\}_{1 \leq j \leq k-1})) = \int_0^\infty r \mathbb{P}(d(x, \{Z_j\}_{1 \leq j \leq k-1}) \geq r)dr
\]

is bounded above by \( \int_0^\infty r \exp\left(-\frac{\pi(k-1)r^2}{4e_1^2n}\right) dr \leq \frac{C_1^2}{k} \) for some positive constant \( C \), not depending on \( k \) or \( x \). This proves the desired bound for the second moment in (2.15).

\[\square\]

**Proof of Lemma 3** The proof of the first estimate in (2.11) follows from standard Binomial estimates and the estimate for \( \mathbb{E}N_i \) in (2.7) (see Corollary A.1.14, pp. 312, Alon and Spencer (2008)). The proof the second estimate in (2.11) follows from the strips estimate (2.2) with \( a = \frac{2b_2n}{N} \) and \( b = r_n \).

To prove the first estimate of (2.9) assume \( N_i \geq 3 \) and recall that \( \{Y_u\}_{1 \leq u \leq N_i} \) are the nodes of the Binomial process in the square \( S_i \) (see paragraph prior to (2.13)). Let \( R_i \) denote the MST of length \( R_i \) containing the nodes \( \{Y_u\}_{1 \leq u \leq N_i} \).

If \( l(Y_u, R_i), 1 \leq u \leq N_i \) is the sum of length of the edges containing \( Y_u \) as an endvertex then \( l(Y_u, R_i) \geq d(Y_u, \{Y_v\}_{v \neq u}) \), the minimum distance of \( Y_u \) from all the other nodes in \( S_i \) as defined in (2.13).

From (1.4), \( R_i = L(R_i) = \frac{1}{2} \left( \sum_{u=1}^{N_i} l(Y_u, R_i) \right) \geq \frac{1}{2} \left( \sum_{u=1}^{N_i} d(Y_u, \{Y_v\}_{v \neq u}) \right) \)

and so

\[
\mathbb{E}R_i = \sum_{k \geq 2} \mathbb{E}R_i^k \mathbb{1}(N_i = k) \geq \frac{1}{2} \mathbb{E} \sum_{k \geq 2} \sum_{u=1}^{k} d(Y_u, \{Y_v\}_{v \neq u}) \mathbb{1}(N_i = k). \quad (2.20)
\]

Recalling the definition of \( D_i(k) \) in (2.13) we then get

\[
\mathbb{E}R_i \geq \frac{1}{2} \sum_{k \geq 2} \mathbb{P}(N_i = k)kD_i(k) \geq \frac{1}{2} \sum_{\frac{\eta_1 n}{2N} \leq k \leq \frac{2b_2 n}{N}} \mathbb{P}(N_i = k)kD_i(k), \quad (2.21)
\]

provided \( n \) is large enough so that \( \frac{\eta_1 n}{2N} \geq \frac{n}{2} M \log n \geq 2 \), the middle estimate being true because of (2.12).

Using the estimate \( D_i(k) \geq \frac{\eta_1 r_n}{\sqrt{k}} \) (see (2.15)) in (2.21) we then get that \( \mathbb{E}R_i \) is bounded below by

\[
\gamma_1 r_n \sum_{\frac{\eta_1 n}{2N} \leq k \leq \frac{2b_2 n}{N}} \mathbb{P}(N_i = k)\sqrt{k} \geq \gamma_1 r_n \sum_{\frac{\eta_1 n}{2N} \leq k \leq \frac{2b_2 n}{N}} \mathbb{P}(N_i = k),
\]

12
which in turn is bounded below by $\gamma_i r_n \sqrt{\frac{2N}{n}} (1 - e^{-C_n})$ for some constant $C > 0$, by (2.11). Since $\frac{n}{N} \to \infty$ as $n \to \infty$, (see (2.12)), this proves the lower bound for $\mathbb{E}R_l$ in (2.9).

To prove the upper bound of $\mathbb{E}R_l$ in (2.9), we argue as follows. If the number of nodes $N_l \leq 2n^2n$, then from (2.11), $R_l \leq C r_n \sqrt{\frac{N}{n}}$ for some constant $C > 0$. If $N_l \geq 2n^2n$, then $R_l \leq N_l r_n \sqrt{2}$, since there are at most $N_l - 1 \leq N_l$ edges in the MST $R_l$ of length $R_l$ and each such edge has both endvertices in the $r_n \times r_n$ square $S_l$ and therefore has length at most $r_n \sqrt{2}$. Thus

$$
\mathbb{E}R_l \leq C r_n \sqrt{\frac{n}{N} + r_n \sqrt{2}} \mathbb{E} \left( N_l \left( N_l > \frac{2n^2n}{N} \right) \right) \leq C r_n \sqrt{\frac{n}{N} + r_n \sqrt{2}} \mathbb{E}(N_l 1(U_l^r)),
$$

(2.22)

where $U_l$ is as defined in (2.10).

Recall from discussion following (2.5) that $N_l$ is Binomially distributed with parameters $n$ and $p_l$ and so by standard Binomial estimates $\mathbb{E}N_l^2 \leq C n^2 n^2$ for some constant $C > 0$, by (2.4). Using Cauchy-Schwarz inequality and the estimate for $\mathbb{P}(U_l^c)$ in (2.11), we therefore get

$$
\mathbb{E}N_l 1(U_l^c) \leq (\mathbb{E}N_l^2)^{\frac{1}{2}} (\mathbb{P}(U_l^c))^\frac{1}{2} \leq C n \frac{n}{N} \exp \left( -C n \frac{n}{N} \right) \leq \sqrt{n} \frac{n}{N},
$$

(2.23)

for all $n$ large and for some positive constants $C_1, C_2$. The final inequality in (2.23) is true since $\frac{n}{N} \to \infty$ as $n \to \infty$ (see (2.12)). Substituting (2.23) into (2.22) gives the upper bound for $\mathbb{E}R_l$ in (2.9). The proof of the bound for $\mathbb{E}R_l^2$ is analogous as above.

Define the covariance between $R_{l_1}$ and $R_{l_2}$ for distinct $l_1$ and $l_2$ as

$$
cov(R_{l_1}, R_{l_2}) = \mathbb{E}R_{l_1}R_{l_2} - \mathbb{E}R_{l_1}\mathbb{E}R_{l_2}.
$$

(2.24)

We need the following result for future use. Recall the constants $\epsilon_1, \epsilon_2$ in (1.1).

**Lemma 4.** There is a positive constant $M_0 = M_0(\epsilon_1, \epsilon_2)$ large so that the following holds if (1.7) is satisfied with $M > M_0$: There are positive constants $C_1, C_2$ such that for all $n \geq 2$ and for any $1 \leq l_1 \neq l_2 \leq N$,

$$
|\text{cov}(R_{l_1}, R_{l_2})| \leq C_1 \left( \mathbb{E}R_{l_1}R_{l_2} \right) \frac{n}{N^2} \leq C_2 \frac{n^2 n^2}{N^3}.
$$

(2.25)

To prove Lemma 4, we use Poissonization described in the next subsection.
Poissonization

Recall from discussion prior to (1.7) that $n \geq 1$ nodes $\{X_k\}_{1 \leq k \leq n}$ are distributed across the $r_n \times r_n$ squares $\{S_j\}_{1 \leq j \leq N}$ according to a Binomial process with intensity $g_N(.)$ as defined in (1.3). Throughout, we use Poissonization as a tool to obtain estimates for probabilities of events for the corresponding Binomial process. We make precise the notions in this subsection.

Let $\mathcal{P}$ be a Poisson process on the squares $\bigcup_{j=1}^{N} S_j$ with intensity function $ng_N(.)$ defined on the probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$. If $N_{l}(\mathcal{P})$ be the number of nodes of $\mathcal{P}$ present in the square $S_l$, $1 \leq l \leq N$, then

$$\mathbb{P}_0(N_l(\mathcal{P}) = k) = \text{Poi}(k; np_l) := e^{-np_l} \frac{(np_l)^k}{k!}, \quad (2.26)$$

where $p_l$ is as defined in (2.4). Moreover,

$$\frac{\eta_1 n}{N} \leq \mathbb{E}_0 N_l(\mathcal{P}) = np_l \leq \frac{\eta_2 n}{N} \quad (2.27)$$

by (2.4).

Let $\{Y_j\}_{1 \leq j \leq N_l(\mathcal{P})}$ be the nodes of $\mathcal{P}$ present in the square $S_l$. Analogous to (2.8), set $R_l(\mathcal{P}) = 0$ if $N_l(\mathcal{P}) \leq 2$ and if $N_l(\mathcal{P}) \geq 3$ set

$$R_l(\mathcal{P}) := \text{MST}(Y_1, \ldots, Y_{N_l(\mathcal{P})}; S_l) \quad (2.28)$$

where $\text{MST}(.;.;)$ is as defined in (2.1). The following result is analogous to Lemma 3.

**Lemma 5.** If $M > 0$ is arbitrary and (1.7) holds, the following is true: There are positive constants $\{\delta_i\}_{1 \leq i \leq 5}$ such that for all $n \geq 2$ and for any $1 \leq l \leq N$,

$$\delta_1 r_n \sqrt{\frac{n}{N}} \leq \mathbb{E}_0 R_l(\mathcal{P}) \leq \delta_2 r_n \sqrt{\frac{n}{N}}, \quad \mathbb{E}_0 \left( R_l(\mathcal{P}) \right)^2 \leq \delta_3 \left( r_n \sqrt{\frac{n}{N}} \right)^2 \quad (2.29)$$

and

$$\mathbb{P}_0 \left( R_l(\mathcal{P}) \geq \delta_4 r_n \sqrt{\frac{n}{N}} \right) \geq \delta_5. \quad (2.30)$$

**Proof of Lemma 5.** The proof of (2.29) is analogous as in the Binomial case and proceeds as follows. Define

$$U_l(\mathcal{P}) = U_l(\mathcal{P})(n) := \left\{ \frac{\eta_1 n}{2N} \leq N_l(\mathcal{P}) \leq \frac{2\eta_2 n}{N} \right\}, \quad (2.31)$$
where \( \eta_1 \) and \( \eta_2 \) are as in (2.4). Analogous to (2.11), the following bound is obtained from standard Poisson distribution estimates (see Theorem A.1.15, pp. 313, Alon and Spencer (2008)): There is a positive constant \( \gamma \) such that for all \( n \geq 2 \) and for any \( 1 \leq l \leq N \),

\[
P_0 \left( U_l^{(P)} \right) \geq 1 - \exp \left( -\gamma \frac{n}{N} \right).
\] (2.32)

As in the Binomial case, given \( N_l^{(P)} = k \), the nodes of \( P \) are i.i.d. distributed according to distribution (2.17). Therefore for \( k \geq 2 \) we let

\[
D_l^{(P)}(k) = \mathbb{E}_0 \left( d(Y_k, \{ Y_j \}_{1 \leq j \leq k-1}) \mid N_l^{(P)} = k \right)
\]
and as in (2.13) obtain that

\[
D_l^{(P)}(k) = \mathbb{E}d(Z_k, \{ Z_j \}_{1 \leq j \leq k-1}) = D_l(k),
\] (2.33)

where \( D_l(k) \) is as defined in (2.13), the random variables \( \{ Z_j \}_{1 \leq j \leq k} \) are i.i.d. with distribution (2.17) and the final equality in (2.33) is true because of (2.16). Consequently \( D_l^{(P)}(k) \) also satisfies properties \((b1) - (b2)\) and the rest of the proof of (2.29) is analogous to the Binomial case.

Finally, the estimate in (2.30) is obtained by using (2.29) and the Paley-Zygmund inequality

\[
P_0 \left( R_l^{(P)} \geq \lambda \mathbb{E}_0 R_l^{(P)} \right) \geq (1 - \lambda)^2 \frac{\mathbb{E}_0 (R_l^{(P)})^2}{\mathbb{E}_0 (R_l^{(P)})^2}
\] (2.34)

for \( 0 < \lambda < 1 \).

We now use Poissonization and obtain intermediate estimates needed to prove Lemma 4. Recall from (2.8) and (2.28) that \( R_l \) and \( R_l^{(P)} \) are the lengths of the MSTs containing all the nodes in the \( r_n \times r_n \) square \( S_l, 1 \leq l \leq N \) in the Binomial and the Poisson process, respectively.

**Lemma 6.** There is a positive constant \( M_0 = M_0(\epsilon_1, \epsilon_2) \) large so that the following holds if (1.7) is satisfied with \( M > M_0 \): There are positive constants \( C_0, C_1 \) and \( C_2 \) not depending on \( l \) such that the following estimates hold for all \( n \geq C_0 \) : For \( 1 \leq l \leq N \),

\[
| \mathbb{E} R_l - \mathbb{E}_0 R_l^{(P)} | \leq C_1 \left( \mathbb{E} R_l \right) \left( \frac{n}{N^2} \right) \leq C_2 \left( \frac{r_n n^{3/2}}{N^{5/2}} \right).
\] (2.35)
For any $1 \leq l_1 \neq l_2 \leq N$

$$|\mathbb{E}(R_{l_1} R_{l_2}) - \mathbb{E}_0(R_{l_1}^{(P)} R_{l_2}^{(P)})| \leq C_1 \left( \mathbb{E} R_{l_1} \mathbb{E} R_{l_2} \right) \left( \frac{n}{N^2} \right) \leq C_2 \left( \frac{r^2 n^2}{N^3} \right). \quad (2.36)$$

To prove Lemma 6, we need estimates on the difference between Binomial and Poisson distributions. For $k, l \geq 1$ recall the Binomial distribution $B(k; n, p_l)$ and the Poisson distribution $Poi(k; np_l)$ as defined in (2.6) and (2.26), respectively. For $k_1, k_2, l_1, l_2 \geq 1$, let

$$B(k_1, k_2; n, p_{l_1}, p_{l_2}) := \binom{n}{k_1, k_2} p_{l_1}^{k_1} p_{l_2}^{k_2} (1 - p_{l_1} - p_{l_2})^{n-k_1-k_2}, \quad (2.37)$$

where $\binom{n}{k_1, k_2} = \frac{n!}{k_1! k_2! (n-k_1-k_2)!}$. We have the following properties.

(c1) There is a constant $C > 0$ such that for all $n \geq 3, 1 \leq l \leq N$ and $\frac{m n}{2N} \leq k \leq \frac{2m n}{N}$,

$$|B(k; n, p_l) - Poi(k; np_l)| \leq Poi(k; np_l) \left( 1 + \frac{C n}{N^2} \right). \quad (2.38)$$

(c2) There is a constant $C > 0$ such that for all $n \geq 3$, and for any $1 \leq l_1, l_2 \leq N$ and $\frac{m n}{2N} \leq k_1, k_2 \leq \frac{2m n}{N}$,

$$|B(k_1, k_2; n, p_{l_1}, p_{l_2}) - Poi(k_1; np_{l_1}) Poi(k_2; np_{l_2})| \leq Poi(k_1; np_{l_1}) Poi(k_2; np_{l_2}) \left( 1 + \frac{C n}{N^2} \right). \quad (2.39)$$

Proof of (c1) - (c2): To prove (2.38) in (c1), we write $p_l = p$ for simplicity. Use $\binom{n}{k} \leq \frac{n^k}{k!}$ and $1 - x \leq e^{-x}$ for $0 < x < 1$ to get

$$\binom{n}{k} p^k (1-p)^{n-k} \leq \frac{(np)^k}{k!} e^{-p(n-k)} = Poi(k; np) e^{kp}.$$ 

Using (2.4) and the fact that $k \leq \frac{2m n}{N}$ we get $e^{kp} \leq \exp \left( \frac{kmn}{N} \right) \leq \exp \left( 2\eta_2 \frac{mn}{N^2} \right)$ and since

$$e^x = 1 + x + \sum_{k \geq 2} \frac{x^k}{k!} \leq 1 + x + \sum_{k \geq 2} x^k \leq 1 + 2x \quad (2.40)$$

for all $x$ small, we get $e^{kp} \leq 1 + \frac{4\eta_2 mn}{N^2}$, proving the upper bound in (2.38).
To obtain a lower bound, we use the estimate
\[ 1 - x \geq e^{-x - x^2} \] (2.41)
for all \(0 < x < \frac{1}{2}\). To prove (2.41), write \( \log(1 - x) = -x - R(x) \) where
\[ R(x) = \sum_{k \geq 2} \frac{x^k}{k} \leq \frac{1}{2} \sum_{k \geq 2} x^k = \frac{x^2}{2(1 - x)} \leq x^2 \]
since \( x < \frac{1}{2} \). Use \( \binom{n}{k} \geq \frac{(n-k)^k}{k!} \) and (2.41) to get
\[ B(k; n, p) \geq \frac{1}{k!} (n - k)^k p^k e^{-p(n-k) - p^2(n-k)} = Poi(k; np) \left( 1 - \frac{k}{n} \right)^k e^{kp - (n-k)p^2} \] (2.42)
As before, using the fact that \( \frac{mn}{2N} \leq k \leq \frac{2mn}{N} \) we get
\[ \left( 1 - \frac{k}{n} \right)^k \geq 1 - \frac{k^2}{n} \geq 1 - \frac{4\eta_2^2 n}{N^2} \] (2.43)
and using (2.4) we get
\[ kp - (n-k)p^2 \geq kp - np^2 \geq \eta_1 \frac{n}{2N} \frac{\eta_1}{\eta_2} - n \left( \frac{\eta_2}{N} \right)^2 = -\eta_1 \frac{n}{N^2} \] (2.44)
where \( \eta = \eta_2^2 - \frac{\eta_1}{\eta_2} > 0 \), since \( \epsilon_1 \leq \epsilon_2 \) and so \( \eta_1 = \frac{\epsilon_2}{\epsilon_1} \leq \frac{\epsilon_2}{\epsilon_2} = \eta_2 \). Using (2.43) and (2.44) into (2.42) gives
\[ B(k; n, p) \geq Poi(k; np) \left( 1 - \frac{\eta_2^2 n}{4 N^2} \right) \exp \left( -\eta_1 \frac{n}{N^2} \right) \geq Poi(k; np) \left( 1 - \frac{\eta_2^2 n}{4 N^2} \right) \left( 1 - \eta_1 \frac{n}{N^2} \right) , \]
and using (2.4) we get
\[ B(k; n, p) \geq Poi(k; np) \left( 1 - \frac{\eta_2^2 n}{4 N^2} \right) \exp \left( -\eta_1 \frac{n}{N^2} \right) \]
\[ \geq Poi(k; np) \left( 1 - \frac{\eta_2^2 n}{4 N^2} \right) \left( 1 - \eta_1 \frac{n}{N^2} \right) , \]
since \( e^{-x} \geq 1 - x \) for \( 0 < x < 1 \). This proves (2.38).

To prove (2.39), write \( p_{11} = p_1, p_{12} = p_2 \) and \( B_{12} = B(k_1, k_2; n, p_1, p_2) \) for simplicity. Use
\[ \binom{n}{k_1, k_2} = \frac{1}{k_1! k_2!} n(n-1) \ldots (n-k_1-k_2+1) \leq \frac{n^{k_1+k_2}}{k_1! k_2!} \] (2.45)
to get
$$B_{12} \leq \frac{(np_1)^{k_1} (np_2)^{k_2}}{k_1! k_2!} e^{-(p_1+p_2)n} e^{(p_1+p_2)(k_1+k_2)}. \quad (2.46)$$

Using (2.4), we get $p_1 + p_2 \leq \frac{2np}{N}$ and since $k_1, k_2 \leq \frac{2np}{N}$, we get using (2.40) that
$$e^{(p_1+p_2)(k_1+k_2)} \leq \exp \left( \frac{4n^2 \eta n}{N^2} \right) \leq 1 + \frac{8n^2 \eta}{N^2} \quad (2.47)$$
for all $n$ large, since $\frac{\eta}{N^2} \to 0$ as $n \to \infty$ (see (1.7)). Substituting (2.47) into (2.46), we get the upper bound for $B_{12}$ in (2.39).

For the lower bound for $B_{12}$ again use (2.45) to get
$$\left( \frac{n}{k_1, k_2} \right) \geq \frac{1}{k_1! k_2!} (n - k_1 - k_2)^{k_1+k_2} = \frac{n^{k_1+k_2}}{k_1! k_2!} \left( 1 - \frac{k_1 + k_2}{n} \right)^{k_1+k_2}. \quad (2.48)$$

Using $$(1 - x)^r \geq 1 - rx$$ for $r, x > 0$ we further get
$$\left( \frac{n}{k_1, k_2} \right) \geq \frac{n^{k_1+k_2}}{k_1! k_2!} \left( 1 - \frac{(k_1 + k_2)^2}{n} \right) \geq \frac{n^{k_1+k_2}}{k_1! k_2!} \left( 1 - \frac{4n^2 \eta}{N^2} \right) \quad (2.49)$$
since $k_1, k_2 \leq \frac{2np}{N}$. Substituting (2.48) into (2.37) we get
$$B_{12} \geq \frac{(np_1)^{k_1} (np_2)^{k_2}}{k_1! k_2!} (1 - p_1 - p_2)^{n-k_1-k_2} \left( 1 - \frac{4n^2 \eta}{N^2} \right). \quad (2.49)$$

To evaluate $(1 - p_1 - p_2)^{n-k_1-k_2}$, we use the estimate (2.41) which is applicable since from (2.4), we have $p_1 + p_2 \leq \frac{2np}{N} \leq \frac{2np}{\sqrt{n}} \to 0$ as $n \to \infty$ (see (2.12)). Using (2.41), we get
$$e^{-(p_1+p_2)(n-k_1-k_2)-(p_1+p_2)^2(n-k_1-k_2)} = e^{-np_1} e^{-np_2} e^{I_1-I_2}, \quad (2.50)$$
where $I_1 = (p_1 + p_2)(k_1 + k_2) \geq 0$ and $I_2 = (p_1 + p_2)^2(n - k_1 - k_2) \leq n(p_1 + p_2)^2 \leq \frac{2n^2 \eta}{N^2}$ for some constant $C_1 > 0$, by (2.4). Using $e^{-x} \geq 1 - x$ we get $e^{I_1-I_2} \geq e^{-I_2} \geq 1 - \frac{2n^2 \eta}{N^2}$ and so from (2.50), we get
$$e^{-(p_1+p_2)(n-k_1-k_2)-(p_1+p_2)^2(n-k_1-k_2)} = e^{-np_1} e^{-np_2} \left( 1 - \frac{\eta^2 n}{N^2} \right). \quad (2.51)$$

Using (2.51) in (2.49), we get the lower bound for $B_{12}$ in (2.39).
Using properties (c1) – (c2) we prove Lemma 6.

Proof of (2.35) in Lemma 6: Recall from (2.5) that \( N_l \) is the number of nodes of the Binomial process \( \{X_k\} \) in the square \( S_l \) and let \( U_l \) be the event as defined in (2.10). Write \( \mathbb{E} R_l = I_1 + I_2 \) where

\[
I_1 = \mathbb{E} R_l \mathbf{1}(U_l) = \sum_{\frac{m n}{2N} \leq k \leq \frac{2m n}{N}} \mathbb{E} R_l \mathbf{1}(N_l = k), I_2 = \mathbb{E} R_l \mathbf{1}(U_l^c) \tag{2.52}
\]

and \( \eta_1, \eta_2 \) are as in (2.4). Similarly \( \mathbb{E} R_l^{(P)} = I_1^{(P)} + I_2^{(P)} \), where

\[
I_1^{(P)} = \mathbb{E} R_l^{(P)} \mathbf{1}(U_l^{(P)}), I_2^{(P)} = \mathbb{E} R_l^{(P)} \mathbf{1}(U_l^{(P)})^c, \tag{2.53}
\]

\( U_l^{(P)} = \left\{ \frac{m n}{2N} \leq N_l^{(P)} \leq \frac{2m n}{N} \right\} \) is as defined in (2.31) and \( N_l^{(P)} \) is the number of nodes of the Poisson process \( \mathcal{P} \) inside the square \( S_l \) (see discussion prior to (2.26)). From (2.52) and (2.53), we therefore get

\[
|\mathbb{E} R_l - \mathbb{E} R_l^{(P)}| \leq |I_1 - I_1^{(P)}| + I_2 + I_2^{(P)}. \tag{2.54}
\]

The remainder terms \( I_2 \) and \( I_2^{(P)} \) satisfy

\[
\max(I_2, I_2^{(P)}) \leq C (\mathbb{E} R_l) \frac{n}{N^2} \tag{2.55}
\]

for some constant \( C > 0 \). We prove (2.55) for \( I_2 \) and an analogous proof holds for \( I_2^{(P)} \). Indeed, every edge in the MST \( R_l \) containing all the nodes in the \( r_n \times r_n \) square \( S_l \) has both endvertices within \( S_l \) and so has length at most \( r_n \sqrt{2} \). Since there are \( N_l \) nodes in the square \( S_l \), there are \( N_l - 1 \leq N_l \) edges in \( R_l \) and so the length \( R_l \leq N_l r_n \sqrt{2} \) and

\[
I_2 = \mathbb{E} R_l \mathbf{1}(U_l^c) \leq r_n \sqrt{2} \mathbb{E} N_l \mathbf{1}(U_l^c). \tag{2.56}
\]

Using the third expression in (2.23) to estimate \( \mathbb{E} N_l \mathbf{1}(U_l^c) \) we get

\[
I_2 \leq C_1 r_n \sqrt{2} \frac{n}{N} \exp \left( -C_2 \frac{n}{N} \right) = C_1 \sqrt{2} \left( r_n \sqrt{\frac{n}{N}} \right) \left( \sqrt{\frac{n}{N}} \exp \left( -C_2 \frac{n}{N} \right) \right) \tag{2.57}
\]

for some constants \( C_1, C_2 > 0 \). From the lower bound in (2.9) we have \( \mathbb{E} R_l \geq C_3 r_n \sqrt{\frac{n}{N}} \) and so

\[
I_2 \leq C_4 (\mathbb{E} R_l) \left( \sqrt{\frac{n}{N}} \exp \left( -C_2 \frac{n}{N} \right) \right) = C_4 (\mathbb{E} R_l) \left( \frac{n}{N^2} \right) \delta_N \tag{2.58}
\]
where
\[\delta_N = \left(\frac{N^3}{n}\right) \exp \left(-\frac{C_2 n}{2N}\right) \leq \frac{n^2}{M^3(\log n)^3} \exp \left(-\frac{C_2 M}{2\log n}\right) \leq 1 \quad (2.59)\]
for all \(n\) large, provided \(M > 0\) large. The first estimate in (2.59) follows from the upper bound \(N \leq \frac{n}{M\log n}\) in (2.12). Fixing such an \(M\) and using (2.59) in (2.58), we get (2.55).

To estimate the difference \(I_1 - I_1^{(P)}\) in (2.54), recall that given \(N_l = k\), the nodes in \(S_l\) are independently distributed in \(S_l\) with distribution \(f(x)dx\) (see (2.17)) and so
\[I_1 = \sum_{\frac{n^2}{M} \leq k \leq 2n \log n} \mathbb{P}(N_l = k) \mathbb{E}(R_l|N_l = k) = \sum_{\frac{n^2}{M} \leq k \leq 2n \log n} B(k; n, p_l) \Delta(k, q_l) \quad (2.60)\]
where \(B(k; n, p_l)\) is the Binomial probability distribution as defined in (2.6).

Similarly, as argued in (2.33), given \(N_l^{(P)} = k\), the nodes of the Poisson process \(P\) are also distributed in \(S_l\) according to distribution \(f(x)dx\).

Therefore \(\mathbb{E}(R_l^{(P)}|N_l^{(P)} = k) = \Delta(k, q_l)\) as defined in (2.61) and so
\[I_1^{(P)} = \sum_{\frac{n^2}{M} \leq k \leq 2n \log n} \Delta(k, q_l) \text{Poi}(k; np_l), \quad (2.62)\]
where \(\text{Poi}(k; np_l)\) is the Poisson distribution as defined in (2.26). From (2.60) and (2.62), we therefore get
\[|I_1 - I_1^{(P)}| \leq \sum_{\frac{n^2}{M} \leq k \leq 2n \log n} \Delta(k, q_l) |B(k; n, p_l) - \text{Poi}(k; np_l)|. \quad (2.63)\]
Using estimate (2.38) of property (c1) to approximate the Binomial distribution with the Poisson distribution, we get

\[ |I_1 - I_1^{(P)}| \leq C_1 \left( \sum_{\frac{n}{N} \leq k \leq 2\frac{n}{N}} \text{Poi}(k; np_i) \Delta(k, q_i) \right) \frac{n}{N^2} \]

\[ \leq C_1 \left( \sum_{k \geq 0} \text{Poi}(k; np_i) \Delta(k, q_i) \right) \frac{n}{N^2} \]

\[ = C_1 \left( \mathbb{E}_0(R_i^{(P)}) \right) \frac{n}{N^2} \] (2.64)

for some constant \(C_1 > 0\). But \(\mathbb{E}_0(R_i^{(P)})\) and \(\mathbb{E}R_i\) both are bounded above and below by constant multiples of \(r_n \sqrt{\frac{n}{N}}\) (see (2.9) and (2.29)). From (2.64), we therefore get

\[ |I_1 - I_1^{(P)}| \leq C_3 \left( \mathbb{E}R_i \right) \frac{n}{N^2} \] (2.65)

for some constant \(C_3 > 0\). Substituting (2.65) and (2.55) into (2.54) gives

\[ |\mathbb{E}R_i - \mathbb{E}_0 R_i^{(P)}| \leq C_4 \left( \mathbb{E}R_i \right) \frac{n}{N^2} \leq C_5 \left( \frac{r_n^2 n^{3/2}}{N_5^5} \right) \]

for some positive constants \(C_4, C_5\), again using the upper bound for \(\mathbb{E}R_i\) from (2.9). This proves (2.35).

Proof of (2.36) of Lemma 6: The proof is analogous to (2.35).

Write \(\mathbb{E}R_i R_{i2} = J_1 + J_2\) where \(J_1 = \mathbb{E}R_i R_{i2} 1(U_i \cap U_{i2})\) and \(J_2 = \mathbb{E}R_i R_{i2} 1(U_i^c \cup U_{i2}^c)\). Similarly, for the Poisson case let \(U_i^{(P)}\) be the event defined in (2.31) and define analogous terms \(J_1^{(P)}\) and \(J_2^{(P)}\) so that \(\mathbb{E}_0 R_i^{(P)} R_{i2}^{(P)} = J_1^{(P)} + J_2^{(P)}\). The difference

\[ |\mathbb{E}R_i R_{i2} - \mathbb{E}_0 R_i^{(P)} R_{i2}^{(P)}| \leq |J_1 - J_1^{(P)}| + J_2 + J_2^{(P)}. \] (2.66)

Arguing as in (2.55), the remainder terms \(J_2\) and \(J_2^{(P)}\) satisfy

\[ \max(J_2, J_2^{(P)}) \leq C_1(\mathbb{E}R_i R_{i2}) \frac{n}{N^2} \leq C_2 \left( \frac{r_n^2 n^2}{N^3} \right) \] (2.67)

for some constants \(C_1, C_2 > 0\). We prove (2.67) for \(J_2\) and an analogous proof holds for \(J_2^{(P)}\). As argued in the proof of (2.55), every one of the \(N_i\) edges in
the MST $R_{i_k}$ of length $R_{i_k}$ has both endvertices within $S_i$ and so has length at most $r_n \sqrt{\frac{n}{N}}$. Therefore

\[
J_2 = \mathbb{E}R_{i_1}R_{i_2}1(U_{i_1}^c \cup U_{i_2}^c) \leq \left( r_n \sqrt{\frac{n}{2}} \right)^2 \mathbb{E}N_{i_1}N_{i_2}1(U_{i_1}^c \cup U_{i_2}^c). \tag{2.68}
\]

Using Cauchy-Schwarz inequality,

\[
\mathbb{E}N_{i_1}N_{i_2}1(U_{i_1}^c \cup U_{i_2}^c) \leq (\mathbb{E}N_{i_1}^2N_{i_2}^2)^\frac{1}{2} \mathbb{P} (U_{i_1}^c \cup U_{i_2}^c)^\frac{1}{2} \leq (\mathbb{E}N_{i_1}^2N_{i_2}^2)^\frac{1}{2} \exp \left( -2C \frac{n}{N} \right) \tag{2.69}
\]

for some constant $C > 0$ using the estimate (2.11).

To evaluate $\mathbb{E}N_{i_1}^2N_{i_2}^2$, use $ab \leq \frac{a^2+b^2}{2}$ to write $\mathbb{E}N_{i_1}^2N_{i_2}^2 \leq \frac{1}{2} (\mathbb{E}N_{i_1}^4 + \mathbb{E}N_{i_2}^4)$ and use the fact that the term $N_i$ is Binomially distributed with parameters $n$ and $p_i$, where $p_i \leq \frac{M}{n}$ (see (2.4)) and $\eta_i$ does not depend on $l$ or $n$. Therefore $\mathbb{E}N_i^4 \leq C_1 (np_i)^4 \leq C_2 \left( \frac{n}{N} \right)^4$ for some constants $C_1, C_2$ not depending on $l$ or $n$ and so $\mathbb{E}N_{i_1}^2N_{i_2}^2 \leq C_3 \left( \frac{n}{N} \right)^4$. Therefore $\mathbb{E}N_{i_1}N_{i_2}1(U_{i_1}^c \cup U_{i_2}^c) \leq C_4 \left( \frac{n}{N} \right)^2 e^{-2C \frac{n}{N}}$ (see (2.69)) and so from (2.68)

\[
J_2 \leq C_5 r_n^2 \left( \frac{n}{N} \right)^2 \exp \left( -2C \frac{n}{N} \right) = C_5 \left( \frac{r_n^2 n^2}{N^3} \right) N \exp \left( -2C \frac{n}{N} \right).
\]

Since $N \leq \frac{n}{M \log n}$ (see (2.72)) we have that $N e^{-2C \frac{n}{N}} \leq \frac{n}{M \log n} e^{-2CM \log n} \leq 1$ for all $n$ large provided $M > 0$ is large. Fixing such an $M$, we get (2.67).

To evaluate the difference $J_1 - J_1^{(P)}$, recall from discussion prior to (2.60) that given $N_i = k$, the nodes of the Binomial process are distributed in the square $S_i$ with distribution (2.17). Similarly, given $N_i^{(P)} = k$, the nodes of the Poisson process are also distributed according to (2.17). Therefore analogous to (2.63) we get

\[
|J_1 - J_1^{(P)}| = \sum_{\frac{n}{2N} \leq k_1, k_2 \leq \frac{2n}{N}} |B_{i_1,i_2} - Po(i_1; np_{i_1}) Po(i_2; np_{i_2})| \Delta(k_1, q_{i_1}) \Delta(k_2, q_{i_2}) \tag{2.70}
\]

where $q_{i_1}, q_{i_2}$ and $\Delta(\ldots)$ are as defined in (2.61) and $B_{i_1,i_2} = B(k_1, k_2; n, p_{i_1}, p_{i_2})$ is as defined in (2.37). Using (2.39) and arguing as in (2.64) we then get $|J_1 - J_1^{(P)}| \leq C \mathbb{E}0(R_{i_1}^{(P)}) \mathbb{E}0(R_{i_2}^{(P)}) \left( \frac{n}{N} \right)$ for some constant $C > 0$. Using the fact that bound $\mathbb{E}0(R_{i_1}^{(P)})$ and $\mathbb{E}R_{i_1}$ are both bounded above and below by constant multiples of $r_n \sqrt{\frac{n}{N}}$ (see (2.29) and (2.9)), we get (2.36).
Proof of Lemma 4: Since Poisson process is independent on disjoint subsets, we have
\[ \text{cov}_0(R_{l_1}^{(P)}, R_{l_2}^{(P)}) = \mathbb{E}_0(R_{l_1}^{(P)} R_{l_2}^{(P)}) - \mathbb{E}_0 R_{l_1}^{(P)} \mathbb{E}_0 R_{l_2}^{(P)} = 0. \]

Therefore write
\[ |\text{cov}(R_{l_1}, R_{l_2})| = |\text{cov}(R_{l_1}, R_{l_2}) - \text{cov}_0(R_{l_1}^{(P)}, R_{l_2}^{(P)})| \leq Z_1 + Z_2 + Z_3, \]
where
\[ Z_1 = |\mathbb{E}_0 R_{l_1} R_{l_2} - \mathbb{E}_0 R_{l_1}^{(P)} R_{l_2}^{(P)}| \leq C \left( \frac{r^2 n^2}{N^3} \right), \]
\[ Z_2 = |\mathbb{E}_0 R_{l_1}^{(P)} \mathbb{E}_0 R_{l_2}^{(P)} - \mathbb{E} R_{l_1} \mathbb{E} R_{l_2}| \leq Z_3 + Z_4, \]
\[ Z_3 = |\mathbb{E}_0 R_{l_1}^{(P)} - \mathbb{E} R_{l_1}| \mathbb{E}_0 R_{l_2}^{(P)} \leq C \left( \frac{r^2 n^{3/2}}{N^{5/2}} \right) \left( r_n \sqrt{\frac{n}{N}} \right) = C \left( \frac{r^2 n^2}{N^3} \right) \]
and similarly,
\[ Z_4 = \mathbb{E} R_{l_1} |\mathbb{E}_0 R_{l_2}^{(P)} - \mathbb{E} R_{l_2}| \leq C \frac{r^2 n^2}{N^3}, \]
for some constant \( C > 0 \). The estimate for \( Z_1 \) follows from (2.36) and the estimates for \( Z_3 \) and \( Z_4 \) follow from (2.35) and the estimates for \( \mathbb{E} R_{l_1} \) and \( \mathbb{E}_0 R_{l_1}^{(P)} \) in (2.9) and (2.29), respectively.

\[ 3 \quad \text{Proof of Theorem 1 } \]

For \( 1 \leq l \leq N \), recall that \( R_l \) is the length of the MST containing all the \( N_l \) nodes of \( \{X_k\} \) present in the square \( S_l \). The first step is to see that \( \text{MST}_C \) is well approximated by \( \sum_{i=1}^N R_i \). Recall that \( s_n \) denotes the intercity distance i.e., the minimum distance between squares in \( \{S_l\} \) (see paragraph prior to (1.7)).

We have the following bounds for \( \text{MST}_C \).

Lemma 7. We have that
\[ \text{MST}_C \leq (V_n + (N - 1)(s_n + 8r_n)) \mathbf{1}(U_{\text{tot}}(n)) + 3\sqrt{n} \mathbf{1}(U_{\text{tot}}^c(n)), \quad (3.1) \]
where
\[ V_n := \sum_{l=1}^{N} R_l, U_{tot} = U_{tot}(n) := \bigcap_{l=1}^{N} U_l \]  
and \( U_l \) is the event defined in (2.10). If the intercity distance \( s_n > r_n \sqrt{2} \), then
\[ MST C_n \geq V_n. \]  

Proof of (3.1) of Lemma 7: We construct a tree containing all the nodes \( \{X_k\}_{1 \leq k \leq n} \) and satisfying the upper bound in (3.1). Suppose that the event \( U_{tot} \) occurs so that each square \( S_l, 1 \leq l \leq N \) contains at least
\[ \frac{\eta_1 n}{2N} \geq \frac{\eta_1 M}{2} \log n \geq 2, \]  
nodes of \( \{X_k\}_{1 \leq k \leq n} \) for all \( n \) large, by (2.12). Let \( T(l) \neq \emptyset \) be the MST containing all the nodes of \( S_l \).

Recall from the discussion following (1.7) that the cities are well connected in the sense that the vertices \( \{z_l\} \) corresponding to the centres of the squares \( \{S_l\} \) is a connected graph \( G_{Z} \subset \mathbb{Z}_{2} \). The spanning tree \( T_{Z} \subset G_{Z} \) contains \( N - 1 \) edges \( f_k, 1 \leq k \leq N - 1 \). Let \( f_k \) have endvertices \( z_1, z_2 \in \mathbb{Z}_{2} \) and let \( S_l \) and \( S_l \) be the corresponding squares whose centres are associated with \( z_1 \) and \( z_2 \), respectively. Pick an edge \( e_k \) with one endvertex being a node of \( \{X_k\} \) in \( S_l \) and another endvertex being a node of \( \{X_k\} \) in \( S_l \). Performing this operation iteratively, we obtain \( N - 1 \) edges \( \{e_k\}_{1 \leq k \leq N-1} \).

The union of the MSTs and the edges
\[ T_{up} := \bigcup_{l=1}^{N} T(S_l) \bigcup_{k=1}^{N-1} e_k \]  
is a tree containing all the nodes \( \{X_k\}_{1 \leq k \leq n} \) and whose length is
\[ L(T_{up}) = \sum_{l=1}^{N} R_l + \sum_{k=1}^{N-1} l(e_k) \leq \sum_{l=1}^{N} R_l + (N - 1)(s_n + 8r_n), \]  
since each edge \( e_k \) has length at most \( s_n + 8r_n \), the sum of the intercity distance and the total perimeter of the two \( r_n \times r_n \) squares containing the endvertices of \( e_k \).
Figure 3: Modifying the path $P_{ab} = acfd$ to obtain a new tree $T_{\text{new}}$.

If the event $U_{\text{tot}}$ does not occur, then by the strips estimate (2.2), the minimum spanning tree containing all the nodes $\{X_k\}_{1 \leq k \leq n}$ has length at most $3\sqrt{n}$.

To prove the lower bound (3.3) in Lemma 7, we need additional properties. Recall from (1.4) that $T_n$ is the minimum spanning tree containing all the nodes $\{X_k\}_{1 \leq k \leq n}$. Suppose there are two nodes $a, b \in \{X_k\}$ in some square $S_l, 1 \leq l \leq N$. Since $T_n$ is a tree, there is a unique path $P_{ab} \subset T_n$ containing $a$ and $b$ as endvertices. The following crucial property also holds.

$(g1)$ Every node in $P_{ab}$ belongs to the square $S_l$.

Proof of $(g1)$: We prove by contradiction and suppose that the path $P_{ab}$ contains a node outside the square $S_l$. This means that $P_{ab}$ “exits” and “re-enters” the square $S_l$ at two distinct nodes. Without loss of generality, we assume that $a$ and $b$ are the exit and entry points; i.e., there are edges $e_a$ and $e_b$ both in $P_{ab}$ such that $e_a$ contains $a$ as an endvertex and $e_b$ contains $b$ as an endvertex.

If $c$ and $d$ are the other endvertices of $e_a$ and $e_b$ respectively, then $c$ and $d$ both lie outside $S_l$, as shown in Figure 3. Here, the path $P_{ab} = acfd$ is the union of the two edges $ac, bd$ and the wavy path $cfd$.

Since the distance between any two squares in $\{S_j\}$ is at least $s_n > r_n \sqrt{2}$, the edges $ac$ and $bd$ have length at least $s_n > r_n \sqrt{2}$, each. The edge $ab$ however has length at most $r_n \sqrt{2}$. Consider the new graph $T_{\text{new}}$ formed by deleting the edge $ac$ and adding the edge $ab$. The graph $T_{\text{new}}$ is a tree and
by construction, the sum of the length of edges in $T_{new}$ is strictly less than the sum of length of edges in the MST $T_n$. This is a contradiction and so all nodes of $P_{ab}$ are contained in the square $S_l$. ■

Proof of (3.3) in Lemma 7: For $1 \leq l \leq N$, let $H_n(l)$ be the subgraph of $T_n$ containing all the nodes of $S_l$ and all edges with both endvertices inside $S_l$. From property (g1), the graph $T_n(l)$ is connected and is therefore a tree. The length of $T_n(l)$ is at least $R_l$, the length of the MST containing all the nodes of $S_l$. Since the above statement is true for each $1 \leq l \leq N$, we obtain the lower bound in (3.3). ■

We use Lemma 7 to prove Theorem 1. From Lemma 7, we have that the overall minimum length $MST_{C_n}$ is bounded above and below by the sum of the local MST lengths $\sum_{l=1}^N R_l$ apart from some residual terms. From the bounds on $\mathbb{E}R_l$ in (2.29) of Lemma 3, we have that $\sum_{l=1}^N \mathbb{E}R_l$ is of the order of $N r_n \sqrt{n/N} = r_n \sqrt{nN} = b_n$ as defined in (1.6). We therefore study the convergence of $MST_{C_n}/b_n$. We henceforth fix $M > 0$ large so that (2.25) of Lemma 4 holds.

Proof of (1.8) in Theorem 1: From the upper and lower bounds (3.1) and (3.3) in Lemma 7, we have that

$$\frac{1}{b_n} (V_n - \mathbb{E}V_n) - \Delta_n \leq \frac{1}{b_n} (MST_{C_n} - \mathbb{E}MST_{C_n}) \leq \frac{1}{b_n} (V_n - \mathbb{E}V_n) + \Delta_n(3.5)$$

where $V_n = \sum_{l=1}^N R_l$ is as defined (3.2) and

$$\Delta_n = \frac{2(N-1)(s_n + 8r_n)}{b_n} 1(U_{tot}(n)) + \frac{4\sqrt{n}}{b_n} 1(U_{tot}^c(n)).$$

The variance of $V_n$ satisfies

$$\text{var}(V_n) \leq C \frac{r_n^2 n^2}{N} = C b_n^2 \left( \frac{n}{N^2} \right)$$

for some constant $C > 0$ and all $n$ large and since $\frac{n}{N^2} \to 0$ (see (1.7)), we get that

$$\frac{1}{b_n} (V_n - \mathbb{E}V_n) \to 0 \text{ in probability} \quad (3.7)$$

26
as \( n \to \infty \). Also
\[
\Delta_n \to 0 \text{ a.s.} \quad (3.8)
\]
as \( n \to \infty \). This proves (1.8) and we prove (3.6) and (3.8) separately below.

Proof of (3.6): Write
\[
\text{var}(V_n) = \sum_l \text{var}(R_l) + \sum_{l_1, l_2} \text{cov}(R_{l_1}, R_{l_2}) \leq \sum_l \mathbb{E}R_l^2 + \sum_{l_1, l_2} \text{cov}(R_{l_1}, R_{l_2}), \quad (3.9)
\]
where \( \text{cov}(X, Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y \). Using (2.29) of Lemma 3 to estimate \( \mathbb{E}R_l^2 \) we get
\[
\sum_{l=1}^N \mathbb{E}R_l^2 \leq NC_1 \left( r_n \sqrt{\frac{n}{N}} \right)^2 = C_1 r_n^2 n \quad (3.10)
\]
for some constant \( C_1 > 0 \). Similarly using estimate (2.25) of Lemma 4 for the covariance, we get
\[
\sum_{l_1, l_2} \text{cov}(R_{l_1}, R_{l_2}) \leq N^2 \left( C_2 \frac{r_n^2 n^2}{N^3} \right) = C_2 \frac{r_n^2 n^2}{N}. \quad (3.11)
\]
for some constants \( C > 0 \). Substituting (3.10) and (3.11) into (3.9), we get
\[
\text{var}(V_n) \leq C_1 r_n^2 n + C_2 \frac{r_n^2 n^2}{N} = \frac{r_n^2 n^2}{N} \left( C_1 \frac{N}{n} + C_2 \right).
\]
Since \( \frac{N}{n} \leq \frac{1}{M \log n} \leq 1 \) for all \( n \) large (see (2.12)), we get that \( \text{var}(V_n) \leq C_3 \frac{r_n^2 n^2}{N} \) for some positive constant \( C_3 \) and for all \( n \) large.

Proof of (3.8): From (3.6) and the fact that \( r_n < r_n \sqrt{2} < s_n \) (see statement of the Theorem), we get
\[
0 \leq \Delta_n \leq \frac{18Ns_n}{b_n} + \frac{4\sqrt{n}}{b_n} 1(U_{tot}^c(n)) \quad (3.12)
\]
and so
\[
0 \leq \lim_{n} \sup \Delta_n \leq \lim_{n} \sup \frac{4\sqrt{n}}{b_n} 1(U_{tot}^c(n)), \quad (3.13)
\]
since \( \frac{N_{s_n}}{b_n} \to 0 \) as \( n \to \infty \) by the statement of the Theorem. From the estimate for the event \( U_l \) in (2.10),
\[
\Pr(U_{tot}^c(n)) \leq \sum_{i=1}^{N} \Pr(U_i^c) \leq N \exp \left( -C \frac{n}{N} \right), \tag{3.14}
\]
for some constant \( C > 0 \). Using the fact that \( \frac{n}{N} \geq M \log n \) (see (2.12)), we get
\[
\Pr(U_{tot}^c(n)) \leq \frac{n}{M \log n} \frac{1}{nMC} \leq \frac{1}{n^2}, \tag{3.15}
\]
provided \( M > 0 \) is large. Fixing such an \( M \), we have from Borell-Cantelli lemma that \( \Pr(\limsup_n U_{tot}^c(n)) = 0 \) and so a.s. \( 1(U_{tot}^c(n)) = 0 \) for all large \( n \).

From (3.13), we therefore get (3.8).

**Proof of (1.9) in Theorem 1** Recalling that \( V_n = \sum_{i=1}^{N} R_l \) from (3.2), we use Lemma 7 to get
\[
\mathbb{E} V_n \leq \mathbb{E} MST C_n \leq \mathbb{E} V_n + b_n \mathbb{E} \Delta_n, \tag{3.16}
\]
where \( \Delta_n \) satisfies (see (3.12))
\[
\mathbb{E} \Delta_n \leq \frac{18N_{s_n}}{b_n} + 4\sqrt{n}\Pr(U_{tot}^c(n)) \leq 18 + \frac{4\sqrt{n}}{b_n}\Pr(U_{tot}^c(n)), \tag{3.17}
\]
since \( \frac{N_{s_n}}{b_n} \to 0 \) as \( n \to \infty \) (see statement of the Theorem). Using (3.15) for estimating the probability of the event \( U_{tot} \) we get
\[
\sqrt{n}\Pr(U_{tot}^c(n)) \leq \frac{\sqrt{n}}{n^2} \leq \sqrt{\frac{M \log n}{n}} \leq r_n \leq r_n \sqrt{nN} = b_n \tag{3.18}
\]
for all \( n \) large, where the second inequality is true by the condition for \( r_n \) in (1.7). Thus \( \frac{4\sqrt{n}}{b_n}\Pr(U_{tot}^c(n)) \leq 4 \) and so \( \mathbb{E} \Delta_n \leq 22 \) and
\[
\mathbb{E} V_n \leq \mathbb{E} MST C_n \leq \mathbb{E} V_n + 22b_n, \tag{3.19}
\]
by (3.17) and (3.16), respectively.

To estimate \( \mathbb{E} V_n \) use the bounds for \( \mathbb{E} R_l \) in (2.29) of Lemma 3 to get
\[
C_1b_n = N \left( C_1r_n \sqrt{\frac{n}{N}} \right) \leq \mathbb{E} V_n \leq N \left( C_2r_n \sqrt{\frac{n}{N}} \right) = C_2b_n \tag{3.20}
\]
for some constants $C_1, C_2 > 0$. From (3.20) and (3.19), we get the bounds for $\mathbb{E}MSTC_n$ in (1.9).

Proof of (1.10) of Theorem 1: We consider Poissonization and recall the Poisson process $\mathcal{P}$ on the squares $\{S_l\}_{1 \leq l \leq N}$, defined on the probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ (see paragraph prior to (2.26)). Analogous to $MSTC_n$ defined in (1.5), let $MSTC_n^{(P)}$ denote the length of the MST containing all the nodes of the Poisson process $\mathcal{P}$. Recall from (2.28) that $R_l^{(P)}$ denotes the length of the MST containing all the nodes of $\mathcal{P}$ in the square $S_l$.

Analogous to (3.3), we have that if the intercity distance $s_n > r_n \sqrt{2}$, then

$$MSTC_n^{(P)} \geq V_n^{(P)} = \sum_{l=1}^{N} R_l^{(P)}. \tag{3.21}$$

Define the event $E_l^{(P)} = \left\{ R_l^{(P)} \geq \delta_4 r_n \sqrt{2/N} \right\}$, where $\delta_4$ is the constant in (2.30) of Lemma 5. Since the Poisson process is independent on disjoint sets, the events $E_l^{(P)}$ are independent and each occurs with probability at least $\delta_5$, by (2.30). If $F_{\text{sum}}^{(P)} := \sum_{l=1}^{N} 1(E_l^{(P)})$ then $\mathbb{E}_0 \left( F_{\text{sum}}^{(P)} \right) \geq \delta_5 N$ and from the standard Chernoff bound estimate for sums of independent Bernoulli random variables (see Corollary A.1.14, pp. 312 of Alon and Spencer (2008)) we also have $\mathbb{P}_0 \left( F_{\text{sum}}^{(P)} \geq C_1 N \right) \geq 1 - e^{-2C_2 N}$ for some positive constants $C_1$ and $C_2$. If $F_{\text{sum}}^{(P)} \geq C_1 N$, then $\sum_{l=1}^{N} R_l^{(P)} \geq C_1 N \left( \delta_4 r_n \sqrt{2/N} \right) = C_3 b_n$ for some constant $C_3 > 0$ and so from (3.21),

$$\mathbb{P}_0(\text{MSTC}_n^{(P)} \geq C_3 b_n) \geq 1 - e^{-2C_2 N} \tag{3.22}$$

for all $n$ large.

To convert the probability estimates to the Binomial process, let

$$A_P = \{ TSPC_n^{(P)} \geq C_3 b_n \}, A = \{ TSPC_n \geq C_3 b_n \}$$

and use the dePoissonization formula

$$\mathbb{P}(A) \geq 1 - D \sqrt{n} \mathbb{P}(A^c_P) \tag{3.23}$$

for some constant $D > 0$ and (3.22) to get that

$$\mathbb{P}(\text{MSTC}_n \geq C_3 b_n) \geq 1 - D \sqrt{ne^{-2C_2 N}} = 1 - e^{-a N},$$

29
where \( \alpha_N = 2C_2N - \log D - \frac{1}{2}\log n \geq C_2N \) for all \( n \) large, since \( N \geq \sqrt{n} \) for all \( n \) large (see (2.12)). This proves (1.10) and it only remains to prove (3.23).

To prove (3.23), let \( N_P \) denote the random number of nodes of \( P \) in all the squares \( \cup_{j=1}^N S_j \) so that \( E_0N_P = n \) and \( P_0(N_P = n) = e^{-n} \frac{n^n}{n!} \geq D_1 \sqrt{n} \) for some constant \( D_1 > 0 \), using the Stirling formula. Given \( N_P = n \), the nodes of \( P \) are i.i.d. with distribution \( g_N \) as defined in (1.3); i.e., \( P_0(A^c P | N_P = n) = P(A^c) \) and so

\[
P_0(A^c P) \geq P_0(A^c | N_P = n)P_0(N_P = n) = P(A^c)P_0(N_P = n) \geq P(A^c) \frac{D_1}{\sqrt{n}},
\]

proving (3.23).

**Proof of (1.11) of Theorem 1** As in the proof of (1.10) above, we consider the Poisson process \( P \) on the squares \( \{S_l\}_{1 \leq l \leq N} \) defined in the paragraph prior to (2.26). As before, let \( \text{MST}_C(P)_n \) denote the length of the minimum length cycle containing all the nodes of the Poisson process \( P \).

Recall from (2.28) that \( R_l(P) \) denotes the length of the minimum length cycle containing all the nodes of \( P \) inside the square \( S_l \).

Analogous to (3.1) of Lemma 7, we have

\[
\text{MST}_C(P)_n \leq (V_n(P) + (N - 1)(s_n + 8r_n))1(U_{\text{tot}}(n)) + 4\sqrt{n}1(U_{\text{tot}}(n))^c,
\]

where

\[
V_n(P) := \sum_{l=1}^{N} R_l(P), \quad U_{\text{tot}}(P) = U_{\text{tot}}(n) := \bigcap_{l=1}^{N} U_l(P)
\]

and \( U_l(P) = \{ \frac{m}{2N} \leq N_l(P) \leq \frac{2m}{N} \} \) is the event defined in (2.31). Recall that \( N_l(P) \) is the total number of nodes of \( P \) inside the square \( S_l \).

Suppose now that the event \( U_{\text{tot}}(n) \) occurs so that

\[
\text{MST}_C(P)_n \leq V_n(P) + (N - 1)(s_n + 8r_n) = \sum_{l=1}^{N} R_l(P) + (N - 1)(s_n + 8r_n).
\]

Since \( U_l(P) \supseteq U_{\text{tot}}(P) \) occurs for every \( 1 \leq l \leq N \), we use the strips estimate (2.2) with \( a = \frac{2m}{N} \) and \( b = r_n \) to get that the corresponding minimum length \( R_l(P) \leq 4b\sqrt{a} \leq Cr_n\sqrt{\frac{m}{N}} \) for some constant \( C > 0 \) and for
every $1 \leq l \leq N$. Thus $V_n^{(P)} = \left( \sum_{l=1}^{N} R_l^{(P)} \right) \leq Cb_n$ and from (3.26) we therefore get

$$\text{MSTC}_n^{(P)} \leq Cb_n + 2(N-1)(s_n + 8r_n) \leq Cb_n + 18Ns_n \leq (C + 1)b_n, \quad (3.27)$$

for all $n$ large. The second inequality in (3.27) is true since $r_n < r_n \sqrt{2} < s_n$. The final inequality in (3.27) is true since $\frac{Ns_n}{b_n} \to 0$ and so $\frac{Ns_n}{b_n} \leq \frac{1}{18}$ for all $n$ large.

Summarizing, we have that if the event $U_{tot}^{(P)}$ occurs, then the overall minimum length $\text{MSTC}_n^{(P)} \leq C_1b_n$ for some constant $C_1 > 0$. To evaluate $\mathbb{P}(U_{tot}^{(P)})$, use the estimate (2.32) for the event $U_l^{(P)}$ to get

$$\mathbb{P}_0(U_{tot}^{(P)}) \geq 1 - N \exp\left(-2C \frac{n}{N}\right) \quad (3.28)$$

for some constant $C > 0$. Thus

$$\mathbb{P}_0(\text{MSTC}_n^{(P)} \leq C_1b_n) \geq \mathbb{P}(U_{tot}^{(P)}) \geq 1 - N \exp\left(-2C \frac{n}{N}\right). \quad (3.29)$$

To convert the probabilities to the Binomial process, we again use the dePoissonization formula (3.23) to get that

$$\mathbb{P}(\text{MSTC}_n \leq C_1b_n) \geq 1 - DN \sqrt{n} \exp\left(-2C \frac{n}{N}\right) = 1 - e^{-\delta_N}, \quad (3.30)$$

where $D > 0$ is as in (3.23) and $\delta_N = 2C \frac{n}{N} - \log D - \log N - \frac{1}{2} \log n$. Since $\frac{n}{N} \geq M \log n$ for all $n$ large (see (2.12)), we get

$$\log D + \log N + \frac{1}{2} \log n \leq \log D + \log\left(\frac{n}{M \log n}\right) + \frac{1}{2} \log n \leq 2 \log n \leq C \frac{n}{N},$$

provided $M > 0$ is large. Fixing such an $M$ we get that $\delta_N \geq C \frac{n}{N}$ and so (1.11) follows from (3.30).

4 Proof of Theorem 2

To prove Theorem 2, we need a preliminary estimate regarding the difference in the total length of the MSTs upon adding or deleting a single node.
For \( n \geq 1 \), divide the unit square \( S \) into \( r_n \times r_n \) squares \( \{S_i\}_{1 \leq i \leq N} \) each of side length \( r_n \) satisfying
\[
\frac{2M \log n}{n} \leq r_n^2 := \frac{2M \log n + c_n}{n} \leq \frac{3M \log n}{n},
\]
where \( M \) is a large integer to be determined later and \( c_n \in (0, 1) \) is chosen such that \( \frac{1}{4r_n} \) is an integer.

For \( 1 \leq i \leq N \), let \( N_i \) be the random number of nodes of \( \{X_k\}_{1 \leq k \leq n} \) in the square \( S_i \). Using (1.1), the average number of nodes
\[
\mathbb{E} N_i = n \int_{S_i} f(x)dx
\]
satisfies
\[
8 \leq 2\epsilon_1 M \log n \leq n\epsilon_1 r_n^2 \leq \mathbb{E} N_i \leq n\epsilon_2 r_n^2 \leq 3\epsilon_2 M \log n
\]
where \( \epsilon_1, \epsilon_2 > 0 \) is as in (1.1). The first estimate in (4.2) is true provided the constant \( M > 0 \) is large and we fix such an \( M \) henceforth. The other estimates in (4.2) follow from (4.1).

For \( 1 \leq j \leq n+1 \) and \( 1 \leq i \leq N \), let \( Z_j(i) \) be the event that the square \( S_i \) contains between \( \epsilon_1 M \log n \) and \( 4\epsilon_2 M \log n \) nodes of \( \{X_k\}_{1 \leq k \neq j \leq n+1} \) and define
\[
Z_{\text{tot}}(n+1) := \bigcap_{1 \leq j \leq n+1} \bigcap_{i=1}^{N} Z_j(i).
\]
By standard Binomial estimates and (4.2) (see Corollary A.1.14, pp. 312, Alon and Spencer (2008))
\[
\mathbb{P}(Z_j(i)) \geq 1 - e^{-2C_1 M \log n}
\]
for some positive constant \( C_1 \) not depending on \( i \) or \( j \). Thus
\[
\mathbb{P}(Z_{\text{tot}}(n+1)) \geq 1 - (n+1) \cdot N \cdot e^{-2C_1 M \log n}
\]
and since the number of squares is \( N = \frac{1}{r_n^2} \leq \frac{C_2 n}{\log n} \) for some constant \( C_2 > 0 \) (see (1.1)), we get
\[
\mathbb{P}(Z_{\text{tot}}(n+1)) \geq 1 - \frac{C_2 n}{\log n} (n+1) e^{-2C_1 M \log n} \geq 1 - e^{-C_1 M \log n}
\]
for all \( n \) large, provided \( M > 0 \) is large. Fix such a \( M \).

Recall that \( T_n \) is the MST containing all the \( n \) nodes \( \{X_k\}_{1 \leq k \leq n} \). The following Lemma estimates the edge lengths in the MSTs \( T_{n+1} \) and \( T_n \).
Lemma 8. For $1 \leq j \leq n + 1$, let $MST_n(j)$ be the length of the minimal spanning tree containing the nodes $\{X_k\}_{1 \leq k \neq j \leq n + 1}$. The difference

$$|MST_{n+1} - MST_n(j)| \leq C_1 r_n \log n 1(Z_{tot}(n + 1)) + n \sqrt{2} 1(Z_{tot}^c(n + 1)),$$

for some constant $C_1 > 0$ not depending on $j$. Also, if $M > 0$ is large then

$$\mathbb{E}|MST_{n+1} - MST_n| \leq C_2 \frac{(\log n)^{3/2}}{\sqrt{n}},$$

for some constant $C_2 > 0$.

We henceforth fix $M$ large enough so that (4.7) is also satisfied.

We first perform some preliminary computations. For a square $S_i$, let $N_1(S_i)$ be the set of all squares in $\{S_j\}$ sharing a corner with $S_i$. For $k \geq 2$, let $N_k(S_i)$ be the set of squares sharing a corner with some square in $N_{k-1}(S_i)$. We use the following property to prove Lemma 8.

(h1) Suppose the event $Z_{tot}(n + 1)$ occurs and suppose $X_j = v \in S_i$ for some $1 \leq i \leq N$. Let $e$ be any edge in the tree $T_{n+1}$ containing $v$ as an endvertex. If $u$ denotes the other endvertex of $e$, then $u \in S_k$ for some $S_k \in N_{20}(S_i)$ and the length of $e$ is at most $20r_n \sqrt{2}$.

Proof of (h1): The fact that the edge length is at most $20r_n \sqrt{2}$ is a consequence of the definition of $N_{20}(S_i)$.

We prove by contradiction and assume that $u$ does not lie in any square of $N_{20}(S_i)$. Let $S_k \in N_{10}(S_i)$ be a square whose centre is at a distance of at least $5r_n$ from the centre of $S_i$, intersecting the edge $(u, v)$. Since the event $Z_{tot}(n + 1)$ occurs, the square $S_k$ contains a vertex $z$ which also belongs to the MST $T_{n+1}$. The distance between $z$ and $u$ is strictly less than the distance between $v$ and $u$. Similarly the distance between $v$ and $z$ is strictly less than the distance between $v$ and $u$.

Let $P_{vz}$ be the unique path in the tree $T_{n+1}$ with endvertices $v$ and $z$. If the path $P_{vz}$ does not contain $u$ as shown in Figure 4(a), then the edge $(u, z)$ cannot be present in $T_{n+1}$ as this would create a cycle. Removing the edge $(u, v)$ and adding the edge $(u, z)$, we get a new tree $T_{new}$. By construction, the sum of length of edges in $T_{new}$ is strictly less than the sum of length of edges in the MST $T_{n+1}$, a contradiction.

If the path $P_{vz}$ contains the node $u$, then the edge $(u, v)$ necessarily belongs to $P_{vz}$ because $(u, v)$ is the unique path in the tree $T_{n+1}$ connecting $u$ and $v$. In this case, the edge $(v, z)$ cannot be in $T_{n+1}$ as this would create a
(a) The node \( u \notin P_{vz} = vQz \).  
(b) The node \( u \in P_{vz} = vuQz \).

Figure 4: Creating the new tree \( T_{\text{new}} \) depending on whether \( u \in P_{vz} \) or not.

cycle (see Figure 4(b)). Define \( T_{\text{new}} \) to be the graph obtained by deleting the edge \((u, v)\) and adding the edge \((v, z)\). The graph \( T_{\text{new}} \) is again a tree and the sum of length of edges in \( T_{\text{new}} \) is strictly less than the sum of length of edges in the MST \( T_{n+1} \), a contradiction.

Proof of Lemma 8: Suppose that the event \( Z_{\text{tot}}(n+1) \) defined in (4.3) occurs and suppose the node \( X_j = v \in S_i \) for some \( 1 \leq i \leq N \).

To find an upper bound for \( MST_{n+1} - MST_n(j) \), let \( T_n(j) \) be the MST containing the nodes \( \{X_k\}_{1 \leq k \neq j \leq n+1} \). Since the event \( Z_j(i) \supseteq Z_{\text{tot}}(n+1) \) occurs, the square \( S_i \) contains some node \( w \in \{X_k\}_{1 \leq k \neq j \leq n} \). Joining \( v \) and \( w \) by an edge, we get a new tree containing all the nodes \( \{X_k\}_{1 \leq k \leq n+1} \). The edge length between \( v \) and \( w \) is at most \( r_n \sqrt{2} \) and so

\[
MST_{n+1} - MST_n(j) \leq r_n \sqrt{2} \mathbf{1}(Z_{\text{tot}}(n+1)).
\]

To obtain a lower bound for \( MST_{n+1} - MST_n(j) \), we use property (h1) and estimate the difference in length of the MST obtained by removing the node \( X_j = v \) from the MST \( T_{n+1} \) containing all the nodes \( \{X_k\}_{1 \leq k \leq n+1} \). From property (h1), every edge in the MST \( T_{n+1} \) containing \( v \) as an endvertex, has its other endvertex in some square \( S_k \in \mathcal{N}_{20}(S_i) \). Since \( Z_{\text{tot}}(n+1) \) occurs, there are at most \( 4\epsilon_2 M \log n \) nodes of \( \{X_k\}_{1 \leq k \neq j \leq n} \) in every square \( S_k \in \mathcal{N}_{20}(S_i) \) (see definition of \( Z_{\text{tot}}(n+1) \) prior to (4.3)). There are at most \( 40^2 \) squares of \( \{S_k\} \) in \( \mathcal{N}_{20}(S_i) \) and so the degree \( d(v) \) of \( v \) in the tree \( T_{n+1} \) is at most

\[
d(v) \leq 40^2 \cdot (4\epsilon_2 M \log n) = C_2 \log n
\]

34
Suppose $\{v_k\}_{1 \leq k \leq d(v)}$ are the neighbours of $X_j = v$ in the tree $T_{n+1}$. Remove the node $v$ and the edges containing $v$ as an endvertex and add the edges $(v_k, v_{k+1})$ for $1 \leq k \leq d(v) - 1$ as shown in Figure 5. Here $d(v) = 3$ and the broken triangles represent the corresponding subtrees of $T_{n+1}$ attached to the nodes $v_1, v_2$ and $v_3$.

The resulting graph is a tree containing all the nodes $\{X_k\}_{1 \leq k \neq j \leq n+1}$. Each edge removed in the above process belongs to $T(n + 1)$ and so has length at most $20r_n\sqrt{2}$ (property (h1)). Using (4.9), the total length of the edges removed is then at most

$$d(v) \cdot (20r_n\sqrt{2}) \leq C_3 r_n \log n$$

for some constant $C_3 > 0$. Consequently

$$MST_n(j) \leq MST_{n+1} + C_3 r_n \log n \mathbf{1}(Z_{\text{tot}}(n+1)). \quad (4.10)$$

From (4.8) and (4.10), we obtain (4.6) for the case when $Z_{\text{tot}}(n + 1)$ occurs.

If $Z_{\text{tot}}(n+1)$ does not occur, we use the crude upper bound that any edge belonging to either of the spanning trees $T_{n+1}$ or $T_n(j)$ has length most $\sqrt{2}$ and there are $n$ edges in $T_{n+1}$ and $n - 1 \leq n$ edges in $T_n$. This proves (4.6).
To prove (4.7), let $M > 0$ be large so that $\mathbb{P}(Z_{\text{tot}}(n+1)) \geq 1 - \frac{1}{n^4}$. Setting $j = n$ in (4.6) and using the estimate for $r_n$ in (4.1), we then get

$$
\mathbb{E}|MST_{n+1} - MST_n| \leq C_2 \frac{(\log n)^{3/2}}{\sqrt{n}} + n\sqrt{\frac{1}{n^3}} \leq C_3 \frac{(\log n)^{3/2}}{\sqrt{n}}
$$

This proves (4.7).

**Proof of 1.12 of Theorem 2** We use the martingale difference method and for $1 \leq j \leq n+1$, let

$$
\mathcal{F}_j = \sigma(X_1, \ldots, X_j)
$$

denote the sigma field generated by the random variables $X_1, \ldots, X_i$. Defining the martingale difference

$$
G_j = \mathbb{E}(MST_{n+1}|\mathcal{F}_j) - \mathbb{E}(MST_{n+1}|\mathcal{F}_{j-1}),
$$

we have that

$$
MST_{n+1} - \mathbb{E}MST_{n+1} = \sum_{j=1}^{n+1} G_j
$$

and so by the martingale property

$$
\text{var}(MST_{n+1}) = \left(\sum_{j=1}^{n+1} G_j\right)^2 = \sum_{j=1}^{n+1} \mathbb{E}G_j^2.
$$

(4.12)

There is a constant $C > 0$ such that

$$
\max_{1 \leq j \leq n+1} \mathbb{E}G_j^2 \leq \frac{C(\log n)^3}{n}
$$

(4.13)

for all $n \geq 1$ and this proves (1.12).

To prove (4.13), we rewrite $G_j$ in a more convenient form. Let $\omega = (x_1, \ldots, x_{n+1})$ and $\omega' = (y_1, \ldots, y_{n+1})$ be two vectors in $(\mathbb{R}^2)^{n+1}$. We say that $\{x_k\}_{1 \leq k \leq n+1}$ are the nodes of $\omega$. Defining $\omega_j = (x_1, \ldots, x_j, y_{j+1}, \ldots, y_{n+1})$ for $1 \leq j \leq n+1$ and using Fubini’s theorem, we get

$$
|G_j| = \left|\int (M(\omega_j) - M(\omega_{j-1}))f(y_i) \ldots f(y_{n+1})dy_j \ldots dy_{n+1}\right| \leq H_j,
$$

(4.14)
where
\[
H_j := \int |M(\omega_j) - M(\omega_{j-1})| f(y_j) \ldots f(y_{n+1}) dy_j \ldots dy_{n+1},
\] (4.15)
and $M(\omega_j)$ is the length of the MST containing all the nodes in $\omega_j$.

**Proof of (4.13):** Let $Z_{tot}(n + 1)$ be the event defined in (4.13) prior to the proof of property (h2) above. From (4.15),
\[
H_j = I_1 + I_2,
\] (4.16)
where
\[
I_1 = \int |M(\omega_j) - M(\omega_{j-1})| 1(\omega_j \in Z_{tot}(n + 1)) 1(\omega_{j-1} \in Z_{tot}(n + 1)) f(y_j) \ldots f(y_{n+1}) dy_j \ldots dy_{n+1}
\] (4.17)
and $I_2 = I_1 - H_j$.

We have that
\[
\mathbb{E}I_2^2 \leq \frac{C(\log n)^3}{n} \quad \text{and} \quad \mathbb{E}I_2^2 \leq \frac{4}{n}
\] (4.18)
for some constant $C > 0$ and all $n$ large. Since $|G_j|^2 \leq H_j^2 = (I_1 + I_2)^2 \leq 2(I_1^2 + I_2^2)$, we get that
\[
\mathbb{E}(G_j^2) \leq 2 \left( \frac{C(\log n)^3}{n} + \frac{4}{n} \right) \leq \frac{3C(\log n)^3}{n},
\] proving (4.13).

We obtain the estimates for $\mathbb{E}I_1^2$ and $\mathbb{E}I_2^2$ in (4.18), separately below.

**Estimate for $I_1$:** Let $T_n(j)$ be the MST containing all the vertices $\{x_k\}_{1 \leq k \leq j-1} \cup \{y_k\}_{j+1 \leq k \leq n}$. If $L(T_n(j))$ is the length of $T_n(j)$, then from (4.16) we have for $t \in \{j-1, j\}$ that
\[
|M(\omega_t) - L(T_n(j))| 1(\omega_t \in Z_{tot}(n + 1)) \leq Cr_n \log n
\] (4.19)
for some constant $C > 0$. From (4.19), (4.17) and triangle inequality, we therefore have
\[
I_1 \leq 2Cr_n(\log n) \quad \text{and so} \quad \mathbb{E}(I_1^2) \leq 4C^2r_n^2(\log n)^2 \leq C_1 \frac{(\log n)^3}{n}
\] (4.20)
for some constant $C_1 > 0$. The final estimate in (4.20) follows from the expression for $r_n$ in (4.1).
Estimate for $I_2$: To estimate $I_2$, use the fact the MST containing all the nodes of $ω_t$, $t = j - 1, j$ has $n$ edges, each of which has length at most $\sqrt{2}$. Therefore

$$I_2 \leq \int n\sqrt{2} (1(ω_j \notin Z_{tot}(n + 1)) + 1(ω_{j-1} \notin Z_{tot}(n + 1)) f(y_j) \cdots f(y_{n+1})dy_j \cdots dy_{n+1}$$

$$= J_1 + J_2,$$  \hspace{1cm} (4.21)

where $J_1 = n\sqrt{2} \int 1(ω_j \notin Z_{tot}(n + 1)) f(y_j) \cdots f(y_n)dy_j \cdots dy_n$ and $J_2$ is the remaining term. Using Cauchy-Schwarz inequality,

$$J_1^2 \leq 2n^2 (\mathbb{E}(1(Z_{tot}(n + 1))|F_j))^2 \leq 2n^2 \mathbb{E}(1(Z_{tot}(n + 1))|F_j)$$

Similarly $J_2^2 \leq 2n^2 \mathbb{E}(1(Z_{tot}(n + 1))|F_{j-1})$. Using $I_2^2 \leq 2(J_1^2 + J_2^2)$ and the fact that $\mathbb{E}(\mathbb{E}(X|F_j)|F_{j-1}) = \mathbb{E}(X|F_{j-1})$, we get

$$\mathbb{E}(J_1^2 + J_2^2|F_{j-1}) \leq 4n^2 \mathbb{P}(Z_{tot}(n + 1)|F_{j-1}).$$

Since $I_2^2 \leq (J_1 + J_2)^2 \leq 2(J_1^2 + J_2^2)$, we get

$$\mathbb{E}(I_2^2) \leq 4n^2 \mathbb{P}(Z_{tot}(n + 1)^c) \leq 4n^2 e^{-CM\log n}$$

for some constant $C > 0$, using (1.13). Letting $M > 0$ large so that $e^{-CM\log n} \leq \frac{1}{n^2}$, we get the estimate for $I_2$ in (4.18). \hfill \blacksquare

Using the variance estimate (1.12), we prove the almost sure convergence result.

Proof of (1.13) in Theorem 2: From (1.12) and Borel-Cantelli lemma,

$$\frac{1}{n} (\text{MST}_{n^2} - \mathbb{E}MST_{n^2}) \longrightarrow 0 \text{ a.s.}$$  \hspace{1cm} (4.22)

as $n \to \infty$. For convergence along the sequence $a_n = n$, we use a subsequence argument and define

$$D_n := \max_{n^2 \leq k < (n+1)^2} |\text{MST}_k - \text{MST}_{n^2}|.$$  \hspace{1cm} (4.23)

Recalling the event $Z_{tot}(n + 1)$ defined in (4.3), let

$$Y_{tot}(n) := \bigcap_{n^2 \leq k < (n+1)^2} Z_{tot}(k + 1)$$  \hspace{1cm} (4.24)

38
so that from (4.6), the difference
\[ |MST_{k+1} - MST_k| \leq C_1 r_k (\log k) \mathbf{1}(Y_{\text{tot}}(n)) + k \sqrt{2} \mathbf{1}(Y_{\text{tot}}^c(n)) \]
for each \( n^2 \leq k < (n + 1)^2 \) and for some constants \( C_1, C_2 > 0 \) not depending on \( k \) or \( n \).

From (4.1) we have that
\[ r_k \leq C_2 \sqrt{\frac{\log k}{k}} \leq C_3 \frac{\log n}{n} \]
for some positive constants \( C_2, C_3 \) and so
\[ r_k \log k \leq C_3 \left( \frac{\log n}{n} \right)^{3/2} \] and \( k \sqrt{2} \leq (n + 1)^2 \sqrt{2} \) (4.25)
for some positive constants \( C_2, C_3, C_4 \) and for all \( n^2 \leq k < (n + 1)^2 \). Using (4.25) in (4.25) and adding telescopically, we get
\[ |MST_k - MST_{n^2}| \leq C_4 \left( \frac{\log n}{n} \right)^{3/2} (k - n^2) \mathbf{1}(Y_{\text{tot}}(n)) + (k - n^2)(n+1)^2 \sqrt{2} \mathbf{1}(Y_{\text{tot}}^c(n)) \] (4.26)
for \( n^2 \leq k < (n + 1)^2 \).

From (4.23), (4.26) and the fact that \( k - n^2 \leq (n + 1)^2 - n^2 \leq 4n \) for all \( n \) large, we get
\[ D_n \leq C_5 \left( \frac{\log n}{n} \right)^{3/2} \mathbf{1}(Y_{\text{tot}}(n)) + 4n(n + 1)^2 \mathbf{1}(Y_{\text{tot}}^c(n)) \] (4.27)
From the estimate for \( Z_{\text{tot}}(k) \) in (4.3)
\[ \mathbb{P}(Y_{\text{tot}}(n)) \geq 1 - \sum_{k=n^2}^{(n+1)^2-1} \mathbb{P}(Z_{\text{tot}}^c(k)) \]
\[ \geq 1 - \sum_{k=n^2}^{(n+1)^2-1} \exp(-CM \log k) \]
\[ \geq 1 - ((n + 1)^2 - n^2) \exp(-CM \log(n^2)) \]
for all \( n \) large and for some constant \( C > 0 \). Setting \( M > 0 \) large so that \( \exp(-CM \log(n^2)) \leq \frac{1}{n^6} \) we then get that
\[ \mathbb{P}(Y_{\text{tot}}(n)) \geq 1 - \frac{2n + 1}{n^9} \geq 1 - \frac{1}{n^7} \] (4.28)
for all \( n \) large.
From Borel-Cantelli lemma and (4.28) we get that \( P(\lim \inf_n Y_{tot}(n)) = 1 \) and so a.s. \( 1(Y_{tot}(n)) = 0 \) for all large \( n \). From (4.27) and (4.28), we therefore get

\[
\frac{D_n}{n} = C_5 \frac{\log n^{3/2}}{n} + 4n(n+1)^2 1(Y_{tot})(n) \rightarrow 0 \text{ a.s. (4.29)}
\]

and

\[
\frac{\mathbb{E}D_n}{n} = C_5 \frac{\log n^{3/2}}{n} + \frac{4n(n+1)^2}{n^2} \rightarrow 0 \text{ a.s. (4.30)}
\]

as \( n \to \infty \).

Finally for \( n^2 \leq k < (n+1)^2 \), write

\[
\frac{1}{\sqrt{k}} |MST_k - EMST_k| \leq \frac{1}{\sqrt{k}} |MST_k - MST_{n^2}| + \frac{1}{\sqrt{k}} \mathbb{E}|MST_k - MST_{n^2}|
\]

\[
\leq \frac{1}{n} |MST_k - MST_{n^2}| + \frac{1}{n} \mathbb{E}|MST_k - MST_{n^2}|
\]

\[
\leq \frac{D_n}{n} + \frac{\mathbb{E}D_n}{n}
\]

and use (4.29) and (4.30) to get that \( \frac{1}{\sqrt{k}} (MST_k - EMST_k) \) converges to zero a.s. as \( k \to \infty \). \( \blacksquare \)

**Proof of (1.14) and (1.15) in Theorem 2** The variance estimate (1.12) is proved above. The upper bound for \( EMST_n \) in (1.14) is obtained from the strips estimate (2.2) with \( a = n \) and \( b = 1 \). This also proves (1.15).

To prove the lower bound for \( EMST_n \) in (1.14), let \( l(X_j, T_n) \) denote the total length of the edges containing the node \( X_j \) in the MST \( T_n \). From (1.4),

\[ MST_n = \frac{1}{2} \sum_{j=1}^n l(X_j, T_n) \geq \frac{1}{2} \sum_{j=1}^n d(X_j, \{X_k\}_{k \neq j}), \]

where \( d(X_j, \{X_k\}_{k \neq j}) \) is the minimum distance of the node \( X_j \) from all the other nodes. Therefore \( EMST_n \geq \frac{a}{2} \mathbb{E}d(X_1, \{X_j\}_{2 \leq j \leq n}) \geq C_1 \sqrt{n} \) for some constant \( C_1 > 0 \), by arguing analogous to the proof of (2.15) in property \( (b2) \). \( \blacksquare \)

**Proof of (1.16) in Theorem 2** We perform Poissonization and construct a Poisson process \( \mathcal{P} \) in the unit square \( S \) with intensity \( nf(.) \) as follows. Let \( \{V_{i,k}\}_{1 \leq i \leq N, k \geq 1} \) be i.i.d. random vectors in \( \mathbb{R}^2 \) with density

\[
\int_{S_i} f(x) dx \mathbf{1}(x \in S_i).
\]

Let \( \{N(S_i)\}_{1 \leq i \leq N} \) be independent Poisson random variables such that \( N(S_i) \) has mean \( n \int_{S_i} f(x) dx \) for \( 1 \leq i \leq T \). The random variables \( \{N(S_i)\} \) are independent of \( \{V_{i,k}\} \) and we define \( (\{V_{i,k}\}, \{N(S_i)\}) \) on the probability space \( (\Omega_0, \mathcal{F}_0, \mathbb{P}_0) \).
For $1 \leq i \leq N$, if $N(S_i) \geq 1$, then we set $\{V_{i,k}\}_{1 \leq k \leq N(S_i)}$ to be the nodes of $\mathcal{P}$ in the square $S_i$. Analogous to (1.5), let $\mathcal{T}_n^{(P)}$ be the MST containing all the nodes of $\mathcal{P}$ in the unit square $S$ and as in (1.5) define $\text{MST}_n^{(P)} := L(\mathcal{T}_n^{(P)})$.

We find lower bounds for the length $\text{MST}_n^{(P)}$ in the Poisson process and then later convert the estimates to the Binomial process. We first need some preliminary definitions and computations. Analogous to (4.2), we have for every $1 \leq i \leq N$ that

$$2\epsilon_1 M \log n \leq n\epsilon_1 r_n^2 \leq \mathbb{E}_0 N(S_i) \leq n\epsilon_2 r_n^2 \leq 3\epsilon_2 M \log n \quad (4.31)$$

where $\epsilon_1, \epsilon_2 > 0$ is as in (1.1). Defining

$$Y_i := \{\epsilon_1 M \log n \leq N(S_i) \leq 4\epsilon_2 M \log n\} \quad (4.32)$$

we get by standard Poisson distribution estimates (Theorem A.1.15, pp. 313, Alon and Spencer (2008)) that

$$\mathbb{P}_0(Y_i) \geq 1 - e^{-2CM \log n} \quad (4.33)$$

for some constant $C > 0$ not depending on $M$ and for all $n$ large.

For $q \geq 1$, recall the definition of the $q$–neighbourhood $\mathcal{N}_q(S_i)$ of the square $S_i$, $1 \leq i \leq N$, from the discussion following Lemma 8. Let $W_1, \ldots, W_T$ be a maximal set of squares in $\{S_k\}$ such that $\mathcal{N}_{20}(W_i) \cap \mathcal{N}_{20}(W_j) = \emptyset$ for any $1 \leq i \neq j \leq T$. There are $(41)^2$ squares in $\mathcal{N}_{20}(W_i)$ for any square $W_i$ and so by our choice of $r_n$ in (1.1), we have that $\bigcup_{i=1}^T \mathcal{N}_{20}(W_i) = \bigcup_{k=1}^N S_k$.

Since there are a total $N = \left(\frac{1}{r_n}\right)^2$ squares in $\{S_k\}$, we must have

$$C_1 \frac{n}{\log n} \leq T = \frac{N}{(41)^2} = \left(\frac{1}{41r_n}\right)^2 \leq C_2 \frac{n}{\log n} \quad (4.34)$$

for some positive constants $C_1, C_2$, using the bounds for $r_n$ in (1.1). For $1 \leq i \leq T$, let

$$Q_i := \bigcap_{k:S_k \in \mathcal{N}_{20}(W_i)} Y_k, \quad (4.35)$$

so that from (4.33) we get

$$\mathbb{P}_0(Q_i) \geq 1 - (41)^2 e^{-2CM \log n} \quad (4.36)$$

for some constant $C > 0$. 
The event $Q_i$ is useful in the following way.

(h2) Suppose the event $Q_i$ occurs for some $1 \leq i \leq T$ and let $e$ be an edge of the MST $T_n^{(P)}$ containing a node $v \in W_i$. If $u$ denotes the other endvertex of $e$, then $u \in S_k$ for some $S_k \in N_{20}(W_i)$.

The proof of (h2) is analogous to the proof of property (h1) stated below Lemma 8.

Recall from paragraph prior to (4.31) that $N(W_i)$ is the number of nodes of the Poisson process $\mathcal{P}$ in the square $W_i$ and that $\{V_{i,k}\}_{1 \leq k \leq N(W_i)}$ are the nodes of $\mathcal{P}$ in $W_i$. Let $l(V_{i,k}, T_n^{(P)})$ be the sum of length of the edges containing the node $V_{i,k}$ as an endvertex in the MST $T_n^{(P)}$, with the notation that the sum length is zero if $N(W_i) = 0$. From (4.4) $MST_n^{(P)} = L(T_n^{(P)})$ satisfies

$$MST_n^{(P)} \geq \frac{1}{2} \sum_{i=1}^{T} \sum_{k=1}^{N(W_i)} l(V_{i,k}, T_n^{(P)}) \geq \frac{1}{2} \sum_{i=1}^{T} \sum_{k=1}^{N(W_i)} l(V_{i,k}, T_n^{(P)}) 1(Q_i). \quad (4.37)$$

If the event $Q_i$ occurs, the number of nodes $N(W_i) \geq \epsilon_1 M \log n$. Moreover, from property (h2) above, every edge containing $V_{i,k} \in W_i$ as an endvertex has its other endvertex in some square belonging to the neighbourhood $N_{20}(W_i)$. Therefore

$$l(V_{i,k}, T_n^{(P)}) 1(Q_i) \geq d_{i,k} 1(Q_i)$$

where $d_{i,k}$ is the minimum distance of the node $V_{i,k} \in W_i$ from all the nodes of $\mathcal{P}$ in $N_{20}(W_i)$.

Summarizing,

$$MST_n^{(P)} \geq \sum_{k=1}^{\epsilon_1 M \log n} \sum_{i=1}^{T} F_{i,k}, \quad (4.38)$$

where $F_{i,k} := d_{i,k} 1(Q_i)$. We need the following property regarding the moments of $F_{i,k}$.

(h3) There are positive constants $C_1, C_2$ and $C_3$ such that for any $1 \leq i \leq T$ and any $1 \leq k \leq M \log n$,

$$C_1 \frac{r_n}{\sqrt{\log n}} \leq \mathbb{E} F_{i,k} \leq C_2 \frac{r_n}{\sqrt{\log n}}$$

and

$$\mathbb{E} F_{i,k}^2 \leq C_3 \frac{r_n^2}{\log n}. \quad (4.39)$$
Proof of (h3): There are \( L = (41)^2 \) squares of \( \{S_k\} \) in \( \mathcal{N}_{20}(W_i) \) and if the event \( Q_i \) occurs, then each square \( S_k \in \mathcal{N}_{20}(W_i) \) has between \( \epsilon_1 M \log n \) and \( 4\epsilon_2 M \log n \) nodes of \( P \) (see (4.35) and (4.32)).

For positive integers \( l_1, \ldots, l_L \) define

\[
E(l_1, \ldots, l_L) = \bigcap_{S_k \in \mathcal{N}_{20}(W_i)} \{ N(S_k) = l_k \}
\]

and use the definition of \( Q_i \) in (4.35) to get that

\[
Q_i = \bigcup (l_1, \ldots, l_L) E(l_1, \ldots, l_L),
\]

where the union is over all \( L \)-tuples satisfying

\[
\epsilon_1 M \log n \leq l_k \leq 4\epsilon_2 M \log n, \quad 1 \leq k \leq L.
\]

If (4.40) holds, then arguing as in the proof of (2.15) in property (b2), we get

\[
C_4 \frac{r_n}{\sqrt{\log n}} \leq \mathbb{E}_0 (d_{i,k} | E(l_1, \ldots, l_L)) \leq C_5 \frac{r_n}{\sqrt{\log n}}
\]

and \( \mathbb{E}_0 (d_{i,k}^2 | E(l_1, \ldots, l_L)) \leq C_6 \frac{r_n^2}{\log n} \) for some positive constants \( C_4, C_5 \) and \( C_6 \), not depending on \( \{l_k\} \) or \( i \). Thus

\[
C_4 \frac{r_n}{\sqrt{\log n}} \mathbb{P}_0(Q_i) \leq \mathbb{E}_0 (d_{i,k} 1(Q_i)) \leq C_5 \frac{r_n}{\sqrt{\log n}} \mathbb{P}_0(Q_i).
\]

and \( \mathbb{E}_0 (d_{i,k}^2 1(Q_i)) \leq C_6 \frac{r_n^2}{\log n} \mathbb{P}_0(Q_i) \). Using the estimate for \( \mathbb{P}_0(Q_i) \) in (4.36) we then get (4.39).

From (4.39) and the Paley-Zygmund inequality (2.34), we have for \( 1 \leq i \leq T \) and \( 1 \leq k \leq \epsilon_1 M \log n \) that

\[
\mathbb{P}_0 \left( F_{i,k} \geq \delta_1 \frac{r_n}{\sqrt{\log n}} \right) \geq \delta_2
\]

for some positive constants \( \delta_1 \) and \( \delta_2 \), not depending on \( i \) or \( k \). We use (4.41) to lower bound \( \text{MST}^{(P)}_n \) in (4.38) as follows. Let \( G_{i,k} = \{ F_{i,k} \geq \delta_1 \frac{r_n}{\sqrt{\log n}} \} \) and use (4.38) to get

\[
\text{MST}^{(P)}_n \geq \sum_{k=1}^{\epsilon_1 M \log n} \sum_{i=1}^T F_{i,k} 1(G_{i,k}) \geq \delta_1 \frac{r_n}{\sqrt{\log n}} \sum_{k=1}^{\epsilon_1 M \log n} \sum_{i=1}^T 1(G_{i,k}).
\]

Since the Poisson process is independent on disjoint sets, the terms \( F_{i_1,k} \) and \( F_{i_2,k} \) are independent for distinct \( 1 \leq i_1 \neq i_2 \leq T \). Therefore we get
from (4.41) and standard Chernoff estimates for Bernoulli random variables that
\[
P_0 \left( \sum_{i=1}^{T} 1(G_{i,k}) \geq \delta_3 T \right) \geq 1 - e^{-\delta_4 T} \tag{4.42}
\]
for some positive constants \(\delta_3, \delta_4\) not depending on \(k\). Using the bounds for \(T\) in (4.34), we get
\[
P_0 \left( \sum_{i=1}^{T} 1(G_{i,k}) \geq \delta_5 \frac{n}{\log n} \right) \geq 1 - \exp \left( -\delta_6 \frac{n}{\log n} \right) \tag{4.43}
\]
for some positive constants \(\delta_5, \delta_6\). Consequently,
\[
P_0 \left( \sum_{k=1}^{T} \sum_{i=1}^{T} 1(G_{i,k}) \geq \left( \delta_5 \frac{n}{\log n} \right) \cdot \epsilon_1 M \log n \right)
\geq 1 - (\epsilon_1 M \log n) \exp \left( -\delta_6 \frac{n}{\log n} \right)
\geq 1 - \exp \left( -\delta_7 \frac{n}{\log n} \right) \tag{4.44}
\]
for all \(n\) large, for some constant \(\delta_7 > 0\).

Using (4.44) in (4.42) we get that with \(P_0\)-probability at least
\[1 - \exp \left( -\delta_7 \frac{n}{\log n} \right),\]
the term
\[
MST_n^{(P)} \geq \delta_1 \frac{r_n}{\sqrt{\log n}} \left( \delta_5 \frac{n}{\log n} \right) \cdot \epsilon_1 M \log n \geq C \sqrt{n}, \tag{4.45}
\]
for some constant \(C > 0\), using the lower bound \(r_n \geq \sqrt{M \log n} / n\) from (1.1).

Finally, to convert the estimates to the length \(MST_n\) of the MST in the Binomial process, we let
\[
A := \{ MST_n \geq C \sqrt{n} \}, A_P = \{ MST_n^{(P)} \geq C \sqrt{n} \}
\]
and use dePoissonization formula \(P(A) \geq 1 - D P_0(A_P) \sqrt{n}\) for some constant \(D > 0\) (see (3.23)). From (4.45) we then get (1.16).

Proof of (1.17): We need some preliminary definitions and estimates. For a set of nodes \(x_1, \ldots, x_n\) in the unit square \(S\), recall from Section II that
$K_n(x_1, \ldots, x_n)$ is the complete graph formed by joining all the nodes by straight line segments and $\text{MST}(x_1, \ldots, x_n)$ is the length of the minimum spanning tree of $K_n(x_1, \ldots, x_n)$.

For any $a > 0$, consider the graph $K_n(ax_1, \ldots, ax_n)$ where the length of the edge between the vertices $ax_1$ and $ax_2$ is simply $a$ times the length of the edge between $x_1$ and $x_2$ in the graph $K_n(x_1, \ldots, x_n)$ Using the definition of MST in (1.5) we then have

$$\text{MST}(ax_1, \ldots, ax_n) = a \text{MST}(x_1, \ldots, x_n).$$  \hspace{1cm} (4.46)

Therefore if $Y_1, \ldots, Y_n$ are $n$ nodes uniformly distributed in the square $aS$ of side length $a$, then we get from (4.46) that

$$\text{MST}(n; a) := \text{MST}(Y_1, \ldots, Y_n) = a \text{MST}(X_1, \ldots, X_n),$$

where $X_i = \frac{Y_i}{a}, 1 \leq i \leq n$ are i.i.d. uniformly distributed in $S$. Recalling the notation $\text{MST}_n = \text{MST}(X_1, \ldots, X_n)$ (see paragraph prior to Theorem 2) we therefore get

$$\mathbb{E} \text{MST}(n; a) = a \mathbb{E} \text{MST}_n. \hspace{1cm} (4.47)$$

The following property is also needed for future use.

$(t1)$ For any positive integers $n_1, n_2 \geq 1$ we have that

$$\text{MST}_{n_1 + n_2} \leq \text{MST}_{n_1} + 3\sqrt{n_2} + \sqrt{2}. \hspace{1cm} (4.48)$$

Proof of $(t1)$: Let $\mathcal{T}_1$ be the MST formed by the $n_1$ nodes $\{X_i\}_{1 \leq i \leq n_1}$ and let $\mathcal{T}_2$ be the MST formed by the remaining $n_2$ nodes. Joining $\mathcal{T}_1$ and $\mathcal{T}_2$ by an edge $e_{12}$, we get a tree containing all the $n_1 + n_2$ nodes. Since $e_{12}$ has length at most $\sqrt{2}$, we get

$$\text{MST}_{n_1 + n_2} \leq \text{MST}_{n_1} + \text{MST}(X_{n_1+1}, \ldots, X_{n_2}) + \sqrt{2}. \hspace{1cm} (4.49)$$

Using the strips estimate (2.2), the middle term in (4.49) is bounded above by $3n_2\sqrt{2}$.

To prove (1.17), it suffices to see that

$$\frac{\mathbb{E} \text{MST}_{n^2}}{n} \rightarrow \beta \hspace{1cm} (4.50)$$
as $n \to \infty$ for some constant $\beta > 0$. To see this is true, use the definition of $D_n = \max_{n^2 \leq k < (n+1)^2} |\text{MST}_k - \text{MST}_{n^2}|$ in (4.23) to get for $n^2 \leq k < (n+1)^2$ that

$$\frac{\text{EMST}_k}{\sqrt{k}} \leq \frac{\text{EMST}_k}{n} \leq \frac{\text{EMST}_{n^2}}{n} + \frac{ED_n}{n}$$

and

$$\frac{\text{EMST}_k}{\sqrt{k}} \geq \frac{\text{EMST}_k}{n+1} \geq \frac{\text{EMST}_{n^2}}{n+1} - \frac{ED_n}{n+1}$$

and then use the fact that $\frac{ED_n}{n} \to 0$ as $n \to \infty$ (see (4.29)).

In the first step in the proof of (4.50), we show that

$$\limsup_n \frac{\text{EMST}_{n^2}}{n} \leq \limsup_k \frac{\text{EMST}_{k^2m^2}}{km}$$

for any fixed integer $m \geq 1$.

**Proof of (4.51):** Fix an integer $m \geq 1$ and write $n = qm + s$ where $q = q(n) \geq 1$ and $0 \leq s = s(n) \leq m - 1$ are integers. As $n \to \infty$,

$$q(n) \to \infty \text{ and } \frac{n}{q(n)} \to m.$$  

(4.52)

Using property (t1),

$$\text{MST}_{n^2} = \text{MST}_{(qm+s)^2} = \text{MST}_{q^2m^2 + 2qms + s^2} \leq \text{MST}_{q^2m^2 + 3\sqrt{2qms + s^2} + \sqrt{2}}$$

and so

$$\limsup_n \frac{\text{EMST}_{n^2}}{n} \leq \limsup_n \frac{q^2m \text{EMST}_{q^2m^2}}{nm} + \limsup_n \frac{3\sqrt{2qms + s^2} + \sqrt{2}}{qm}.$$  

(4.53)

Since $s \leq m - 1 < m$, $3\sqrt{2qms + s^2} + \sqrt{2} \leq 4m\sqrt{2q + 1} + \sqrt{2}$ and so using (4.52), the second term in (4.53) is zero. Using (4.52) again, the first term in (4.53) equals

$$\limsup_n \frac{\text{EMST}_{q^2m^2}}{qm} \leq \limsup_k \frac{\text{EMST}_{k^2m^2}}{km}.$$  

(4.54)

This proves (4.51).

**Proof of (4.54):** Let $L_1 = \limsup_n \frac{\text{EMST}_{q^2m^2}}{qm}$ and $L_2 = \limsup_k \frac{\text{EMST}_{k^2m^2}}{km}$.

For $q = q(n)$ as defined prior to (4.52) and for all integers $l \geq 1$, we have

$$\sup_{n \geq l} \frac{\text{EMST}_{q^2m^2}}{qm} \geq L_1 \text{ and so } \sup_{n \geq lm + m} \frac{\text{EMST}_{q^2m^2}}{qm} \geq L_1.$$  

46
But \( n \geq lm + m \) implies that \( q(n) = \frac{lm + m - s}{m} \geq l \) since \( s = s(n) \leq m \) (see statement prior to (4.52)). Therefore

\[
L_1 \leq \sup_{n \geq lm+m} \frac{\mathbb{E} \text{MST}_{q^2m^2}}{qm} \leq \sup_{k \geq l} \frac{\mathbb{E} \text{MST}_{k^2m^2}}{km} \downarrow L_2
\]
as \( l \to \infty \).

If \( \lambda := \liminf_n \frac{\mathbb{E} \text{MST}_n}{n} \) then from (1.14), we have that \( \lambda > 0 \). Moreover,

\[
\limsup_k \frac{\mathbb{E} \text{MST}_{k^2m^2}}{km} \leq \lambda
\]
and so (1.17) follows from (4.51).

To prove (4.55), we proceed as follows. For positive integers \( k \) and \( m \), distribute \( k^2m^2 \) nodes \( \{X_i\}_{1 \leq i \leq k^2m^2} \) independently and uniformly in the unit square \( S \). Divide \( S \) into \( k^2 \) disjoint squares \( \{W_j\}_{1 \leq j \leq k^2m^2} \) each of size \( \frac{1}{k} \times \frac{1}{k} \) and let

\[
N_j = \sum_{i=1}^{k^2m^2} 1(X_i \in W_j)
\]
denote the number of nodes in the square \( W_j \).

(t2) If \( \text{MST}(N_j) \) denotes the length MST of the nodes in the square \( W_j \) then

\[
\text{MST}_{k^2m^2} \leq \sum_{j=1}^{k} \text{MST}(N_j) + 4k\sqrt{2}.
\]

Proof of (t2): For the proof of (1.57), we proceed as in the proof of the strips method (see (2.22)). Suppose the top left most square is labelled \( W_1 \), the square below \( W_1 \) is \( W_2 \) and so on until we reach the square \( W_k \) intersecting the bottom edge of the unit square \( S \). The square to the right of \( W_1 \) is then labelled \( W_{k+1} \) and the square below \( W_{k+1} \) is \( W_{k+2} \) and so on. For \( j \geq 1 \), let \( T(j) \) be the MST formed by the nodes of \( W_j \). We set \( T(j) = \emptyset \) if \( W_j \) contains no node. Suppose \( T(1) \neq \emptyset \) and let \( W_{j_1} \) be the “first” square below \( W_1 \) in the first column of squares \( \{W_i\}_{1 \leq i \leq k} \) also containing at least one node.

Join some node of \( A \in T(1) \) with some node \( B \in T(j_1) \) and call the resulting edge as an inclined extra edge (see Figure 6). Similarly let \( j_2 \geq j_1 + 1 \) be the least indexed square containing at least one node in the first column.
of squares \{W_i\}_{1 \leq i \leq k} and join some node of \( T(j_1) \) with some node of \( T(j_2) \) by an inclined extra edge.

Let \( j_{last} \) be the “last” square in \( \{W_i\}_{1 \leq i \leq k} \) containing at least one node and let \( i_1 \) be the “first” square in the second column of squares containing at least one node. Join some node \( C \in T(j_{last}) \) with some point \( D \) within the first square in the second column \( W_{k+1} \) by vertical and horizontal extra edges as shown in Figure 6. Join \( D \) to some node of \( E \in T(i_1) \) by an inclined extra edge as shown in Figure 6.

Continue the above procedure for the second column of squares and proceeding iteratively, we finally obtain a spanning tree containing all the \( k^2m^2 \) vertices. By construction, any extra edge (horizontal, vertical or inclined) intersecting a square \( W_j \) has length no more than \( \sqrt{2}/k \), the length of the diagonal of \( W_j \). Also, at most four extra edges intersect \( W_j \). Since there are \( k^2 \) squares in \( \{W_j\} \), the total length of the extra edges added is no more than \( 4\sqrt{2}/k \times k^2 = 4k\sqrt{2} \).

From property (t2) and (4.48), we get

\[ \mathbb{E}MST_{k^2m^2} \leq k^2\mathbb{E}MST(N_1) + 4k\sqrt{2} \]  \hspace{1cm} (4.58)

To evaluate \( MST(N_1) \), write \( \mathbb{E}MST(N_1) = I_1 + I_2 \), where

\[ I_1 = \mathbb{E}MST(N(1))\mathbb{1}(F_1), I_2 = \mathbb{E}MST(N(1))\mathbb{1}(F'_1) \]

and \( F_1 := \{m^2 - m \log m \leq N_1 \leq m^2 + m \log m\} \). Each node \( X_i, 1 \leq i \leq k^2m^2 \) has a probability \( \frac{1}{k^2} \) of being present in the \( \frac{1}{k} \times \frac{1}{k} \) square \( W_1 \). Therefore
the number of nodes $N_1$ in the square $W_1$ is binomially distributed with mean $E N_1 = m^2$ and $\text{var}(N_1) \leq k^2 m^2 \frac{1}{k^2} = m^2$ (see (4.56)). We therefore get from Chebychev’s inequality that

$$\mathbb{P}(F_1) \leq \frac{1}{(\log m)^2} \leq \epsilon$$

(4.59) for all $m \geq M_0$ large, not depending on $k$.

We evaluate $I_1$ and $I_2$ separately below.

**Evaluation of $I_1$:** Write $I_1 = \sum_{j=j_{\text{low}}}^{j_{\text{up}}} \mathbb{E} MST(N(1)) \mathbb{P}(N(1) = j)$, where $j_{\text{low}} := m^2 - m \log m \leq m^2 + m \log m =: j_{\text{up}}$. Given $N_1 = j$, the nodes in $W_1$ are uniformly distributed in $W_1$ and recall from discussion prior to (4.47) that $\mathbb{E} MST(j; \frac{1}{k})$ is the expected length of the MST containing $j$ nodes uniformly distributed in the $\frac{1}{k} \times \frac{1}{k}$ square $W_1$. Thus

$$I_1 = \sum_{j=j_{\text{low}}}^{j_{\text{up}}} \mathbb{E} MST \left( j; \frac{1}{k} \right) \mathbb{P}(N(1) = j) = \frac{1}{k} \sum_{j=j_{\text{low}}}^{j_{\text{up}}} \mathbb{E} MST_j \mathbb{P}(N(1) = j),$$

(4.60) by (4.47).

Using the difference estimate (4.7) from Lemma 8 we have for any $j_{\text{low}} \leq j_1, j_2 \leq j_{\text{up}}$ that

$$\mathbb{E}|MST_{j_2} - MST_{j_1}| \leq \sum_{u=j_{\text{low}}}^{j_{\text{up}}-1} \mathbb{E}|MST_{u+1} - MST_u| \leq \sum_{u=j_{\text{low}}}^{j_{\text{up}}-1} C \frac{(\log u)^{3/2}}{\sqrt{u}}$$

for some constant $C > 0$ not depending on $j_1$ or $j_2$. For all $j_{\text{low}} \leq u \leq j_{\text{up}}$, the term $\frac{(\log u)^{3/2}}{\sqrt{u}} \leq C_1 \frac{(\log m)^{3/2}}{m}$ for some positive constant $C_1$ and so the term $\mathbb{E}|MST_{j_2} - MST_{j_1}|$ is bounded above by

$$(j_{\text{up}} - j_{\text{low}}) C_1 \frac{(\log m)^{3/2}}{m} \leq (2m \log m) C_1 \frac{(\log m)^{3/2}}{m} = C_2 (\log m)^{5/2}$$

(4.61) for some constant $C_2 > 0$. Setting $j_1 = m^2$ and $j_2 = j$ and using (4.61) we get $MST_j \leq MST_{m^2} + C_2 (\log m)^{5/2}$ for all $j_{\text{low}} \leq j \leq j_{\text{up}}$. From (4.60) we therefore have that

$$I_1 \leq \frac{1}{k} \mathbb{E} MST_{m^2} + \frac{1}{k} C_2 (\log m)^{5/2}. \quad (4.62)$$
Evaluation of $I_2$: There are $N(1)$ nodes in the square $W_1$ and so from the strips estimate $\text{(2.2)}$, $\text{MST}(N(1)) \leq \frac{3}{k} \sqrt{N(1)}$. Thus

$$I_2 = \mathbb{E}_{\text{MST}}(N(1))1(F_1^c) \leq \frac{3}{k} \mathbb{E} \sqrt{N(1)}1(F_1^c) \leq \frac{3}{k} (\mathbb{E} N_1)^{\frac{1}{2}} (\mathbb{P}(F_1^c))^{\frac{1}{2}},$$

by the Cauchy-Schwarz inequality. Since $\mathbb{E}N_1 = m^2$ and $\mathbb{P}(F_1^c) \leq \epsilon$ for a fixed $\epsilon > 0$ and for all $m$ large (see $\text{(4.59)}$), we get

$$I_2 \leq \frac{3}{k} m \sqrt{\epsilon}.$$  \hspace{1cm} (4.64)

Substituting $\text{(4.64)}$ and $\text{(4.62)}$ into $\text{(4.58)}$ gives

$$\mathbb{E}_{\text{MST}} k^2 m^2 \leq k \mathbb{E}_{\text{MST}} m^2 + C_2 k (\log m)^{5/2} + 3 m k \sqrt{\epsilon} + 4 k \sqrt{2}$$

and so

$$\limsup_k \frac{\mathbb{E}_{\text{MST}} k^2 m^2}{km} \leq \frac{1}{m} \mathbb{E}_{\text{MST}} m^2 + C_2 \frac{(\log m)^{5/2}}{m} + 3 \sqrt{\epsilon} + 4 \sqrt{2}$$

for all $m$ large. Consequently, $\limsup_k \frac{\mathbb{E}_{\text{MST}} k^2 m^2}{km} \leq \lambda + 3 \sqrt{\epsilon}$ and since $\epsilon > 0$ is arbitrary, we get $\text{(4.55)}$. $\blacksquare$

Acknowledgement

I thank Professors Rahul Roy, Jacob van den Berg, Anish Sarkar and Federico Camia for crucial comments and for my fellowships.

References

[1] K. Alexander. (1996). The RSW theorem for continuum percolation and the CLT for Euclidean minimal spanning trees. *Annals of Applied Probability*, 6, 466–494.

[2] N. Alon and J. Spencer. (2008). *The probabilistic method*. Wiley.

[3] J. Beardwood, J. H. Halton and J. M. Hammersley. (1959). The shortest path through many points. *Proceedings Cambridge Philosophical Society*, 55, pp. 299–327.
[4] T. Cormen, C. E. Leiserson, R. R. Rivest and C. Stein. (2009). *Introduction to Algorithms*. MIT Press and McGraw-Hill.

[5] J. M. Steele. (1988). Growth rates of Euclidean minimal spanning trees with power weighted edges. *Annals of Probability, 16*, pp. 1767–1787.

[6] J. M. Steele. (1993). Probability and Problems in Euclidean Combinatorial Optimization. *Statistical Science, 8*, pp. 48–56.

[7] H. Kesten and S. Lee. (1996). The central limit theorem for weighted minimal spanning trees on random points. *Annals of Applied Probability, 6*, pp. 495–527.